Catarina Dutilh Novaes

LOGIC, EPISTEMOLOGY, AND THE UNITY OF SCIENCE 7

Formalizing Medieval Logical Theories

Suppositio, Consequentiae and Obligationes

Springer

FORMALIZING MEDIEVAL LOGICAL THEORIES

LOGIC, EPISTEMOLOGY, AND THE UNITY OF SCIENCE

VOLUME 7

Logic, Epistemology, and the Unity of Science aims to reconsider the question of the unity of science in light of recent developments in logic. At present, no single logical, semantical or methodological framework dominates the philosophy of science. However, the editors of this series believe that formal techniques like, for example, independence friendly logic, dialogical logics, multimodal logics, game theoretic semantics and linear logics, have the potential to cast new light no basic issues in the discussion of the unity of science.

This series provides a venue where philosophers and logicians can apply specific technical insights to fundamental philosophical problems. While the series is open to a wide variety of perspectives, including the study and analysis of argumentation and the critical discussion of the relationship between logic and the philosophy of science, the aim is to provide an integrated picture of the scientific enterprise in all its diversity.

Formalizing Medieval Logical Theories

Suppositio, Consequentiae and *Obligationes*

By

Catarina Dutilh Novaes
Leiden, The Netherlands

Springer

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN 978-1-4020-5852-3 (HB)
ISBN 978-1-4020-5853-0 (e-book)

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

www.springer.com

Printed on acid-free paper

Cover image: Adaptation of a Persian astrolabe (brass, 1712–13), from the collection of the Museum of the History of Science, Oxford. Reproduced by permission.

For Reinout, Marie and Little-one-on-its-way

CONTENTS

INTRODUCTION

Perhaps one of the most striking characteristics of later medieval philosophy and science is the remarkable unity with which the different fields of investigation were articulated to each other, in particular with respect to the methodology used. While it is fair to say that current science is characterized by a plurality of methodologies and by a high degree of specialization in each discipline, in the later medieval period there was one fundamental methodology being used across disciplines, namely logic. One can say without hesitation that logic provided unity to knowledge and science in the later medieval times. Logic (which was then understood more broadly than it is now, including semantics and formal epistemology) was one of the first subject-matters in the medieval curriculum; it was thought that the knowledge of logic was a necessary, methodological requirement for a student to move on to the other disciplines. And indeed, the widespread use of this logical and semantic methodology can be perceived in disciplines as diverse as natural philosophy (physics), theology, ethics and even medicine.

Besides the fact that medieval logic provided unity to science then, while modern logic does not play the same role now (if anything at all, it is mathematics that might be considered as the fundamental methodology for current investigations), it is also widely acknowledged that the medieval and modern traditions in logic are very dissimilar in many other respects. Of course, this holds of most domains of knowledge: Copernican astronomy also has little resemblance to current astrophysics; current chemistry came a long way from long-forgotten alchemy. Nevertheless, even if the main assumptions and methods are radically different, most present-time disciplines share at least a common subject matter with their predecessors; indeed, Copernican astronomy and astrophysics both have stars, planets and the universe as their subject matter.[1] But the same cannot be said of logic: at first sight, the subject matters of current logic seem to have no counterpart in, for example, Aristotelian or medieval logic, to name but two of its 'predecessors'. In fact, we may doubt whether

[1] Even though our conceptions of what planets, stars and the universe are have changed considerably, as these are essentially theory-laden concepts.

these past traditions should be viewed as predecessors of what we now call logic, or, alternatively, whether what is now known as logic deserves this name at all, in light of its history. In other words, can we really speak of a unified discipline – logic – or is each of these traditions a discipline in its own right? This seems a hard pill to swallow, but at the same time it is not evident what, if anything, would constitute the very nature of logic, that is, the traits common to all these different traditions.

This apparent lack of uniformity in logic lies at the origin of the main question driving the present investigation: in which senses (if any) can medieval logic be viewed as logic (in particular from the viewpoint of modern logic)? It is not so much that medieval logic is of interest to us only insofar as it satisfies modern criteria of what is to count as logic; rather, it is the quest for the common grounds of these two traditions that motivates the search for the senses in which medieval logic is to be seen as logic also by us, 21st century philosophers and logicians. In other words, this investigation seeks to outline unity in two main respects: the unity of medieval science and knowledge provided by medieval logic, and the diachronic unity of logic as a discipline, in spite of the apparent profound dissimilarity between the traditions of medieval logic and modern logic.

Of course, there is a fundamental disparity in their respective general approaches: while, for medieval logicians, their investigations were very closely related to the general study of language, logic is nowadays a part of mathematics. This, among other reasons, is held to justify the skepticism with which medieval logic and other past logical traditions are often viewed by modern logicians (not to mention the widespread positivistic credo to the effect that everything that is 'old' is necessarily obsolete). Notwithstanding (or because of?) these dissimilarities, the degree of sophistication attained by medieval logicians is impressive, just as much as what are, to my mind, significant resemblances (albeit not easily perceived at first sight) between the medieval investigations and current developments in logic and philosophy.

At the same time, it appears that many lessons can be learned from the medieval logicians, as they were aware of some of the intricacies of logic and language whose importance we seem to have forgotten. That is, while the quest for the common grounds of the two traditions is essentially motivated by an inquiry on the nature of logic, the aspects in which medieval logic differs from modern logic are just as significant, as they are a potential source of inspiration for new developments within the current tradition. At any rate, it is clear that to establish a dialogue between the two traditions can only be beneficial.

How can this be done? From a modern perspective, the medieval writings in logic are incomprehensible. Not only is the language (Latin) a barrier; medieval logic was embedded in a complex conceptual framework, with constant use of highly technical jargon. But the most serious obstacle may be the modern tendency to express logical theories in especially devised notations, and with a certain axiomatic structure, which are not to be found in the medieval writings. Either way, it is clear that one way of establishing such a dialogue between these two traditions is to *formalize* fragments of medieval logic. And this is precisely what I set out to do. In particular, the objects of formalization in the present study are three topics from medieval logic, namely supposition, *consequentia* and *obligationes*; each can be seen as a case study demonstrating the fruitfulness of formalizing medieval logic.

By the term 'formalization', one usually understands the 'translation' of something expressed in ordinary language into a symbolic counterpart. In fact, as I carried out the formalizations presented here, it became increasingly evident that, for an adequate formalization, more important than just the choice of symbols is a suitable conceptual analysis of the theory to be formalized. For this reason, the project presupposed an in-depth conceptual understanding of the topics and theories being formalized. In this sense, the present work is just as much a conceptual-historical examination of these topics as it is an attempt at formalization.

Moreover, the term 'formalization' obviously refers to the notion of the formal. This is a rather telling element; currently, formality is often thought to be what is distinctive about logic, so that, for a theory to deserve the attribute 'logical', it must be formal.[2] Therefore, to formalize a theory, that is, to render it (more) formal, is also to show that it is (or the extent to which it is) logical and/or essentially grounded on logical concepts.[3]

However, that formality is what is characteristic of logic is indeed a strong assumption, which must not be plainly taken for granted; in effect, one of the important upshots of examining other logical traditions is to put this assumption to test. Four views are possible: (i) the theoretical constructs of a given logical tradition do conform to the formality criterion; (ii) these theories do not conform to the formality criterion, and thus are not logical theories properly speaking; (iii) these obviously logical theories do not conform to the formality criterion, so the criterion may have to be modified; (iv) formality is irrelevant as a criterion demarcating what is to count as logic.

Obviously, the very notion of the formal demands careful consideration, as it is clear that distinct concepts of the formal are in play. I will argue that, according to some suitable notions of the formal, some of the medieval logical theories are (at least to some extent) formal – and this is made patent by means of the formalizations offered here – in particular if this notion is understood more broadly than it usually is in current developments (especially with respect to permutation invariance – cf. MacFarlane 2000). In other words, I defend view (iii) as defined above: I maintain that the notion of the formal is relevant at least as a necessary condition for what is to count as logic, but that it must go beyond the rather restricted concept of the formal as permutation invariance.

---

[2] It is disputable whether formality is a sufficient condition for what is to count as logic, but it seems to me that it is in any case a necessary condition.

[3] One may argue that this does not hold, as a formalization of a mathematical theory does not turn it into a logical theory. But it is not a coincidence that the usual practitioners of formalization in mathematics are advocates of mathematical logicism; the underlying idea seems to be that a formalization of a mathematical theory corroborates the view that mathematics ultimately rests on logical concepts.

Overall, the aims of the present investigation can be summarized as follows:

1. Historical aim: an investigation of some aspects of medieval logic and semantics, so as to obtain a better understanding of them. In particular, I investigate the extent to which these theories are *formal*, in such a way that they could play the methodological role ascribed to them in medieval science.
2. Pedagogical aim: the attempt to make these medieval theories more easily understandable from a modern vantage point.
3. Philosophical aim: the search for the common grounds underlying different logical traditions (medieval vs. modern), in order to explore the nature and unity of logic as such. The underlying assumption is that logic is formal, but that of itself does not say much as long as it is not clear what is meant by 'formal'.

Given these aims, the use of formalization as the main tool seemed to impose itself. Now, this decision is of itself not of much help, as one can hardly speak of well-defined guidelines as to how a formalization must be carried out. In fact, this is rather murky terrain; several different loose ideas seem to be associated with the concept of formalization, so it became clear that a philosophical reflection on this very concept was not only a necessary addendum to this project; it might also be a welcome contribution to the philosophy of logic in general. As a consequence, in addition to the three case studies on medieval logic, this work contains a fourth chapter on the philosophy of formalization. In this chapter, I argue that formalization corresponds to three distinct but related tasks, that is, axiomatization, symbolization and conceptual translation of a non-formalized theory into an already existing formal theory. A formalization may consist of one of these three procedures, or, more typically, of a combination of them.

HISTORICAL PRELUDE

A systematic overview of the history of later medieval logic is not to be found in the present work. For this, the reader is referred elsewhere.[4] Here, the main goal is that of conceptual analysis, presupposing familiarity with the medieval logical framework. But a few preliminary words on the history behind the authors that figure prominently in my investigation can certainly do no harm.

The later medieval period in (Christian) philosophy starts in the 12th century, with Abelard. This 12th century tradition is a world of its own, extremely complex and interesting, which requires separate attention. Therefore, in the present work, I have deliberately chosen not to deal with the 12th century tradition. It should be mentioned, though, that, while philosophy and theology were still essentially part of the same broad domain of investigation, it is in the 12th century that laymen such as Abelard

[4] The *Cambridge History of Later Medieval Philosophy* (Kretzmann, Kenny and Pinborg 1982) is particularly useful for this purpose, as is (Spade 1996).

(who became a cleric only later in life[5]) became important figures in the Christian intellectual environment.[6]

The 13th century witnessed the emergence of terminist logic, that is, the tradition marked by the study of the so-called properties of terms, such as signification, supposition etc. (cf. Read 2002, De Rijk 1962/67). Two authors from this period will often be referred to in the present work, namely William of Sherwood and Peter of Spain. Both wrote what we could call 'textbooks' in logic, which were then widely used for the study of logic.

But most of the authors considered here belong to the 14th century. At that time, there were two major traditions, namely the English tradition revolving around the University of Oxford, and the continental tradition, whose center was the University of Paris (cf. De Libera 1982). Burley, Ockham, Swyneshed and Strode all stem from the English tradition, while Buridan, Albert of Saxony and Marsilius of Inghen, among many others, belong to the continental tradition. For sure, there are points of contact and exchanges between these two traditions, but each has its own distinctive spirit.

That is, this work is mainly based on 14th century authors, predominantly from the English tradition. Earlier authors are considered only insofar as their writings offer elements for the conceptual understanding of the 14th century theories that are my object of analysis.

SUBJECT-MATTER

I have chosen three topics from medieval logic as objects of formalization: supposition, *consequentia* and *obligationes*. Why these topics, and not others? There is no principled answer to this question. Various contingent reasons led me to focus on these three topics.

The concept of supposition was already the topic of my master thesis, where I dealt with Ockham's truth conditions for the main propositional forms, leaving aside the different kinds of supposition that are my concern here. Besides, supposition is a crucial concept in the medieval semantic framework, so it seemed appropriate to treat of it in the present context – even more so since supposition remains an unfinished topic within medieval scholarship. My main tenet is that, contrary to the accepted view, theories of supposition should not be compared to modern theories of reference. Within the modern framework, they are best seen as theories of meaning, more specifically as theories for the algorithmic generation of the meanings that a certain body of propositions may carry. This insight came to me from a switch of perspective: theories of supposition should not be seen as static, but rather as procedural, in a sense that has recently become influential in logic.

---

[5] Cf. (King 2004).

[6] Notice that the (very rich) Arab and Jewish traditions of the time also fall out of the scope of this work.

As for *consequentia*, it was not obvious to me at first in which way the medieval discussions on the topic had something to add to the current state of affairs (notwithstanding the central position occupied by consequence and related notions in logic, then as well as now). But I quickly realized that these medieval discussions touched upon various important topics. In particular, Buridan's commitment to tokens as truth-value bearers leads him to inquiries that are strikingly similar to current investigations in two-dimensional semantics. Moreover, the distinction material vs. formal semantics as found in Buridan turns out to have important connections with the modern debate on logical consequence. That is, while, on the one hand, some of the modern apparatus of two-dimensional semantics is crucial for spelling out the details of Buridan's views, on the other hand, his notion of formal consequence offers an interesting vantage point for current discussions of the notion of logical consequence; that is to say, the dialogue seems to benefit both sides, as I show in part 2.

Lastly, *obligationes*. It is a doubly fashionable topic: at present, *obligationes* is a popular subject matter among medievalists, and the modern counterpart that I found for it, namely the application of the game-theoretical framework to logic, is equally popular among logicians. Of course, this is not the (only) reason why I chose *obligationes* to be one of my objects of formalization; in fact, it is a remarkable case of conceptual similarity between a medieval and a modern theoretical framework and, accordingly, one of the best examples of the fruitfulness of this kind of investigation. Most of all, recent research on *obligationes* has made important progress, but we are still a long way from totally understanding this genre. There is certainly room for further research on the topic, and with the formalization presented in part 3, I hope to offer further insight using the framework of logical games as point of vantage.

Moreover, these three topics are related to one another in many important ways. As already said, in the later medieval period, logic was a tool to be used for a wide variety of intellectual investigations; in particular, a given logical theory or topic was often used for the analysis of other logical theories or topics (that is, logic as a discipline was not articulated in a strict, foundational way). The notion of supposition was at the core of the medievals' machinery of semantic analysis, and thus was used virtually everywhere; the notion of *consequentia*, or entailment, was of course at the center of all investigations, since it permeates the all-crucial notion of inference of new knowledge from known premises; the obligational framework, which may seem to us a rather artificial and regimented construction, amply underlined the analysis of a variety of topics. The specific connections between each of these topics shall be pointed out in due course, but for now it is important that the organic character of the articulation of the different topics and theories in later medieval logic be borne in mind.

In sum, the present text is composed of four main parts: part 1 is dedicated to supposition theory, part 2 to the notion of *consequentia*, part 3 to *obligationes* and part 4 to the philosophy of formalization. Finally, in the conclusion, I draw some general remarks on the nature of logic, inspired by the foregoing analyses and formalizations.

PART 1

SUPPOSITION THEORY: ALGORITHMIC HERMENEUTICS*

1.0 INTRODUCTION

Over the last decades, medieval theories of supposition have attracted a great deal of interest from historians of philosophy as well as from theoretic philosophers. On the one hand, historians are mostly interested in the historical importance of such theories, and therefore try to explain their development against a historical background. Theoretic-oriented philosophers, on the other hand, judge that the conceptual content of these theories may contribute to the development of new, original semantic systems, and therefore evaluate them against the background of current investigations in philosophy of logic and language and semantics. These are of course not mutually exclusive alternatives, yet in practice it is rare to see both approaches combined.

In truth, theories of supposition have proven to be extremely resistant to comparisons with modern semantic theories. In fact, it is probably not an overstatement to say that we still do not know exactly what theories of supposition were supposed to accomplish. One view that is often put forward by both historians and philosophers is that theories of supposition are the medieval counterpart of theories of reference. But it seems to me that this approximation does more harm than good to our general understanding of these medieval theories. Hence, in what follows I spend quite some time discussing this hypothesis; I examine why it is such a recurring view, but I also argue that there are problems with this view by outlining significant dissimilarities between the concepts of reference and *suppositio*. If they are to be compared to any group of modern theories at all, I argue that it may be more fruitful to view supposition theories as theories of (sentential/propositional[7]) meaning.

However, the very dichotomy meaning/reference is a product of 20th century philosophy of logic and language, and might be inadequate when applied to medieval semantics. Ideally, one should try to understand medieval theories, of supposition

* My views on Ockham's supposition theory can also be found in (Dutilh Novaes 2008).

[7] Throughout the present work, I will use the term 'proposition' in its Latin acceptation, that is, corresponding to what are now known as declarative (meaningful) sentence-tokens, unless otherwise stated.

or otherwise, in their own terms. But precisely here lies the dilemma of studies of this kind: on the one hand, a 21st century philosopher is inexorably prisoner of a certain conceptual paradigm, so any attempt to understand theories belonging to a different paradigm always remains a (mis?)projection; on the other hand, there may not even be much of a point in understanding a medieval theory in its own terms, since this would not really advance the knowledge of this given theory here and now. In sum, analyzing a medieval theory in its own terms may be not only impossible, but (were it to be possible) also uninformative. Still, my main assumption here is that the endeavor of examining these theories and avoiding misprojections for as much as possible – but at the same time offering intelligible (to the modern audience) accounts of a group of theories such as supposition theories – is worth undertaking.

The conclusion I will draw from the historical and conceptual analysis to follow will be that theories of supposition, in particular Ockham's theory of supposition, can be seen as a formal method for the semantic analysis of propositions, which generates their possible readings. The theory as formulated by Ockham is essentially intended to interpret propositions, a sort of hermeneutics, but not a hermeneutics that relies on the subjective skills of the interpreter; rather, it is a set of rules that could *prima facie* be applied mechanically, in a way similar to what is now known as computational semantics (cf. Blackburn and Kohlhase 2004); very important in my interpretation of supposition theory is its procedural nature, insofar as these rules are instructions for (interpretational) procedures to be carried out. Therefore, perhaps the best description of Ockham's theory of supposition is that it is an **algorithmic hermeneutics** – this is in any case what I intend to argue for in the coming pages.

1.1 THEORIES ON THEORIES OF SUPPOSITION

1.1.1 Two kinds of approach: projections

As just said, there seem to be two traditional lines of investigation on medieval logic: the historical line and the systematic line. The historical line is primarily concerned with the establishment of reliable editions of the original Latin texts, with the identification of historical threads of influence among the different authors, and so forth. The systematic line is often undertaken by philosophers of logic, frequently with no formal training in medieval logic, who for one reason or another estimate that some of the theories and ideas developed by medieval logicians can be fruitfully applied to current problems of philosophy of logic and language. In order to do so, they take up the task of 'reconstructing' medieval logical systems so that the latter acquire the form to which philosophers and logicians of the 20th and 21st centuries are accustomed.

Both traditions are of course concerned with the content of these medieval systems, but in quite different ways; whereas the historian is deeply involved with the medieval intellectual universe, the systematic-oriented philosopher is not committed to it, but typically to the post-Fregean, 20th century universe. The historian looks upon medieval logical systems from within, so to say, and therefore with a non-critical eye; the systematic-oriented philosopher usually judges the quality of medieval systems based on his own, anachronistic criteria.

And yet, surprisingly enough, practitioners of both trends usually agree with the claim that *suppositio* is roughly the same as reference. How can this be? My hypothesis is that both sides project their own original conceptual framework into the other; crudely stated, the historian knows what *suppositio* is, but not exactly what reference is, and therefore assumes that reference is what he knows *suppositio* to be, whereas the systematic-oriented philosopher knows what reference is, but not what *suppositio* is, and therefore assumes that *suppositio* is what he knows reference to be, insofar as both concepts are related to the same phenomenon (words standing for things). More specifically, the medievalist may not be entirely familiar with the details of 20th century discussions on reference; he knows that 'reference' has become the standard name for the relation between words and things, and may overlook the fact that the term is in practice very theory-laden. By contrast, the systematic-oriented philosopher has usually been brought up within the 'reference paradigm', and therefore assumes that what this paradigm says the relation between words and things must be is the only acceptable view on this matter.

The result of this imbroglio is that both seem to project their paradigms into the other. As an attempt to make supposition theories palatable to the current audience, medievalists often say that they are theories of reference, probably assuming that this approximation is harmless.[8] But the harm it does is that it encourages people to look upon supposition theories with the same set of criteria that are applied to theories of reference. As a result, the understanding of supposition theories is very much hampered by this approach.

In fact, I am not alone in challenging the adequacy of attributing the title of 'theory of reference' to pre-Fregean theories:

> Discussion on the relation between words and world go back about as far as philosophy itself, and are detailed and sophisticated in Aristotle. The debate is whether just any relation between words and world counts as the reference relation. Some hold that it is constitutive of reference that it should be a relation distinct from that between a common noun and the things of which it is true. Some hear it as trivially false that 'man' *refers to* Socrates, among others, as opposed to merely *denoting* or *suppositing for* each man. If this is right, it will be natural to focus on proper names and to regard them as paradigm possessors of reference, a focus which would be hard to find in ancient, medieval, or early modern discussions. Indeed, if this view about what constitutes reference is correct, then as far as I know there are no discussions of reference before Frege: the history of the topic is short indeed. (Sainsbury 2005, 1)[9]

Misprojection can also occur in the 'reconstructions' of medieval systems in an axiomatic fashion. In itself this is a legitimate and interesting enterprise: the difficulty lies in using an anachronistic (yet fruitful) framework and at the same time

[8] For example: 'To a first (but pretty good) approximation, supposition in this first part of the theory is what nowadays we call "reference". It is the relation between the terms used in a proposition and the things those terms are used to talk about in that proposition'. (Spade 1996, 243)

[9] I owe the reference to this passage to Dean Buckner.

avoiding undue conceptual projections. Sometimes philosophers of logic try to use the apparatus of supposition theories to account for issues that are in fact alien to the medieval framework. Sometimes this attempt leads to interesting results;[10] but at times its failure is seen as a sign of the 'flaws' of supposition theories.[11] Well, try to use a hammer to do the job of a screwdriver: it may just as well work out, but if it does not, does it make sense to blame the hammer? Or, to use a Wittgensteinian metaphor, does it make sense to try to play chess on a backgammon board?

In conclusion, it seems that, even though some particular results have been obtained, the general panorama of what theories of supposition were meant to accomplish still leaves scope for further analysis.

1.1.2 Commentators

The trend of historical studies on medieval logic and semantics began quite recently. Until not so long ago, there were very few modern editions of the original texts, and the manuscripts were scattered around the world. This situation began to change during the 1930s and 1940s, under the pioneering efforts of Ph. Boehner and E. Moody to establish reliable editions of the texts and to deepen our conceptual knowledge of medieval logic in general and of supposition theories in particular. Still, until the 1960s, this was a rather limited field, with hardly any influence upon the general philosophical scene.

In 1962 a remarkable and original work appeared: Peter Geach's *Reference and Generality*. Its main theme was the emerging concept of reference, which had been acquiring new contours with the publication of the English translation of Frege's *Über Sinn und Bedeutung*.[12] Geach's book is seminal in many respects (it launched e.g., a whole new approach to the use of formal logical tools for the analysis of natural language), but the point of concern for the present discussion is that Geach was the first to view medieval logic as capable of contributing to 20th century philosophical and semantic debates. In this sense, as much as Boehner and Moody can be seen as the initiators of the so-called historical trend in medieval semantics studies, Geach is perhaps the pioneer of the systematic-oriented approach.

To my knowledge, Geach was the first to claim that theories of supposition were theories of reference:

> The medieval term for what I call the mode of reference of a referring phrase was '*suppositio*'. Apparently in origin this is a legal term meaning 'going proxy for': Aquinas and Ockham say quite indifferently that a

---

[10] See for example (Klima and Sandu 1991).

[11] Cf. the debate between Swiniarski and Matthews (Swiniarski 1970), (Matthews 1973): both identify a problem with respect to the supposition of the predicate of 'Some *a* is not *b*' propositions. But while Swiniarski maintains that it is a flaw inherent to supposition theory, Matthews contents that the problem lies in a given interpretation of it.

[12] (Frege 1948). It is debatable, though, whether the 20th century concept of reference can be attributed to Frege. Indeed, Geach himself later urged that Frege's term *Bedeutung* be translated as 'meaning', changing his earlier use of 'reference' in later editions of the Frege translations.

term has *suppositio* for (*supponit pro*) and that it stands for (*stat pro*) one or more objects. In paraphrasing medieval writers I shall quite often tacitly use 'mode of reference' for their '*suppositio*'. (Geach 1962, 84)

The impact of *Reference and Generality* was double-sided: On the one hand the book attracted the attention of many scholars to whom it had never occurred that medieval philosophy could be relevant at all to present-time philosophy; on the other hand, Geach was claiming (i) that supposition theories were theories of reference and (ii) that they were **bad** theories of reference, that is, they failed miserably in their task of accounting for the phenomenon of reference.

After that, many studies were devoted to the comparison of theories of supposition to modern theories of quantification, in which scholars debated whether they were similar or in fact essentially different (actually, this discussion concerned only a fragment of supposition theories, namely the modes of personal supposition). The general project of comparing medieval logic to modern logic (in particular by means of more or less formal reconstructions) was quite popular for some time, with different verdicts.[13] Yet nobody seems to have questioned the common assumption that theories of supposition are theories of reference.

For example, P.V. Spade and C. Panaccio seem to endorse this assumption in the passages below.

The difference between these two theories can perhaps best be seen by asking what question each one tries to answer. The first theory, the doctrine of supposition proper, is in effect a theory of reference. It answers the question: What thing or things are referred to by a given term-occurrence in a given sentence? Thus the question *what* a term refers to or supposits for in a given instance is completely answered by the doctrine of supposition proper. The second theory, the doctrine of 'modes of supposition', has nothing to do with that question – at least not in the fourteenth century. (Spade 1988, 190)

[...] la théorie de la supposition permet de déterminer quels sont leurs [des termes] référents dans les divers contextes propositionnels où ils peuvent apparaître. (Panaccio 1992, 58)

Spade seems to be making a loose use of the term 'reference', as the as-of-now official name for the relation between a word and the thing(s) it stands for. As for Panaccio, he seems to be claiming that theories of supposition provide a procedure to determine which things a term stands for in a given propositional context:[14] this is, as we shall see, the core of most theories of reference, but it does not seem to be the main (or only) purpose of theories of supposition (in a way that should be clarified subsequently).

Interestingly, no commentator appears to have realized that an adequate definition of 'reference' ought to be provided to justify the approximation of this concept to that of supposition. An exception to this rule is P. King, in the introduction to his translation of Buridan's *Treatise of Supposition* and *Treatise of Consequences* (King 1985). He presents two (according to him) paradigmatic views on reference

[13] Cf. (Priest and Read 1977), (Karger 1976), (Mullick 1971), (Cocchiarella 2001), (Klima 1993a), (Klima 1993b), (Spade 1978), (Bird 1961).

[14] To be fair, it must be said that the core of Panaccio's interpretation of Ockham's supposition theory is that it is an explanation of the **cognitive**, psychological mechanisms of the semantic relations between words and things, and not a logico-semantic theory offering a procedure for reference determination.

(Davidson's and Quine's) and concludes that there is enough similarity between the two concepts (reference and supposition) to justify the claim of identity between them.

The theory of supposition should not be assimilated to formal logic, but to the philosophy of logic; it is the mediaeval theory of reference. (King 1985a, 35)

Unfortunately, his analysis of the concept of reference is far too brief, and a more extensive examination thereof may indicate that the very opposite claim, namely that reference and supposition are in fact two quite dissimilar concepts, is a more faithful picture of the situation. Hence, in order to avoid misprojection from either side, I shall present a brief conceptual analysis of both 'paradigms', and attempt to expose some of their underlying assumptions.

1.1.3 Theories of reference

If one wants to challenge this conflation, the first important step is to define the notions that are (albeit tacitly) associated with the concept of a theory of reference. Obviously, in order to make perfectly general claims, one would have to go through the entire literature on the topic of the last 50 years, a virtually impossible endeavor. Moreover, there is an array of different views of what a theory of reference is, to the point that there appears to be no set of characteristics common to all these views – the concept of a theory of reference would be something like a vague concept, or else a concept whose unity is at best based on a family resemblance, and not on features common to all its members. If it is so, then it is indeed impossible to provide a reconstruction of a 'standard' theory of reference; there would always be counterexamples of 'theories of reference' not fitting the 'standard' profile.

But for the present purposes, and at the risk of being accused of building a strawman, I will focus on what I consider to be paradigmatic versions of a theory of reference, namely A. Church's logic of 'sense and denotation'[15] and T. Parsons 'Fregean theory of truth and meaning';[16] they exemplify what I believe to be two most relevant aspects of 20th century theories of reference.[17] These are the idea that a theory of reference provides sufficient conditions for the **determination** of the (unique) referent of a term, and the emphasis on a **many-one mapping** between expressions and their referents. But first, I will briefly examine the descriptive aspect of some theories of reference, that is, the attempt to explain the **mechanisms** according to which words come to refer to the things they refer, insofar as this descriptive aspect also plays a major role in several chief theories of reference.

---

[15] (Church 1956). Ironically, Church himself did not use the term 'reference' as a translation for *Bedeutung*, but rather the more old-fashioned 'denotation'.

[16] (Parsons 1996).

[17] The fact that I select two 'Fregean' theories of reference does not compromise the generality of my claim. What I say about then holds *mutatis mutandi* for the so-called 'new theories of reference' as well (but I shall not justify this latter claim in the present text).

1.1.3.1 The mechanisms of reference

The entry named 'Reference' in a well-known encyclopedia of philosophy begins as follows:

Reference is a relation that obtains between expressions and what speakers use expressions to talk about. […] The central question concerning reference is: How do words refer? What, in other words, is the 'mechanism' of reference? (Reimer 2003)

Indeed, a very important aspect of the debate on the concept of reference concerns the so-called 'mechanisms of reference'. If reference is the relation holding between an expression and what it stands for, one is naturally led to wonder how and why a given expression can stand for a given entity (and not for another): in virtue of what does an expression have its reference? In other words, a significant part of many theories of reference is the descriptive task of elucidating the mechanisms involved in the actual occurrences of referential relations. Questions such as 'So what *does* make my use of "Cicero" into a name of *him*?' (Kripke 1980, 91) are among the main concerns of philosophers dealing with the issue.

The debates on the concept of reference of roughly the last 50 years revolved mostly around the semantics of proper names, with two main opposing factions: those arguing in favor of the mediated nature of the mechanism of reference, and those defending the view that, at least for an important group of expressions, the relation between term and referent is direct. According to the former, a word refers to a thing in virtue of a third entity mediating this relation, corresponding to the neo-Fregean picture of the relation between a term and a referent as being established by means of a sense.

The regular connection between a sign, its sense, and its referent is of such a kind that to the sign there corresponds a definite sense and to that in turn a definite referent, while to a given referent (an object) there does not belong only a single sign. (Frege 1948, 211)

One of the most notable versions of this view is the so-called descriptive theory of proper names (cf. Reimer 2003), which holds that a proper name refers to its bearer in virtue of the fact that the latter satisfies the description associated to the proper name in question (the 'third entity' being in this case the description).

The other main view on the mechanism of reference is a clear reaction to the neo-Fregean-Russellian position (which had been influential for many decades); accordingly, its partisans often call it 'the new theory of reference'. Its fundamental claim is that the relation of reference between a proper name and its bearer is not of a mediated nature, but rather of a direct, causal nature. Kripke's *Naming and Necessity* is the seminal text for this view.

Someone, let's say, a baby, is born; his parents call him by a certain name. They talk about him to their friends. Other people meet him. Through various sorts of talk the name is spread from link to link as if by a chain. (Kripke 1980, 91)

The opposition between the two views can be summarized as follows:

By contrast [to the 'Fregean-Carnapian-Churchian' theory of reference], it is often said that the theory of singular reference espoused by Donnellan and Kripke is the theory that denies that proper names have a sense, contending instead that proper names are no more than 'empty tags' which merely label objects, and for which reference is determined not by way of any conceptual content in the name but by tracing back along some sort of causal chain leading originally from the object to a speaker's use of its name. Here, this theory of reference is often called *the causal theory of reference*. (Salmon 1982, 11)

These are two very influential positions, but other accounts of the mechanisms of reference have also been proposed, sometimes hybrid versions of both (cf. Reimer 2003). Another important trend is what can be called the pragmatic explanation for this phenomenon: the reference of a term is determined solely by the intention of the speaker. Partisans of this view are, among others, Strawson (Strawson 1950) and Donnellan (cf. Donnellan 1966).

But regardless of which position one is inclined to advocate, it is manifest that an account of the mechanisms of reference is an indispensable part of any professed respectable theory of reference. In this sense, there is a clear unity among these different theories of reference; different explanations are given to the relation between proper names (and other singular terms) and their bearers, but the phenomenon at stake – reference – is the same. Now, as we shall see, in medieval semantics, the issue of the foundations for the relation between words and things is also tackled,[18] but not within the realm of theories of supposition; this is for the medieval authors the domain of theories of signification.

1.1.3.2 Determination

A widespread interpretation of Frege's distinction between sense and reference (*Sinn* and *Bedeutung*) is the view that the sense of an expression is an intensional object whose content corresponds to an efficient procedure for the determination of the referent of the given expression. The view that the sense should and does provide sufficient conditions for reference determination is often attributed to Frege himself, but it has been argued[19] that this view is a product of later interpretations of Frege's ideas, in particular by Church and Carnap.[20] However, it is clear that in Church's theory of sense and denotation (reference), the idea of reference determination[21] plays a crucial role.

We shall say that a name *denotes* or names its denotation and *expresses* its sense. Or less explicitly we may speak of a name just as *having* a certain denotation and *having* a certain sense. Of the sense we say that it *determines* the denotation, or is a *concept* of the denotation. (Church 1956, 6)

The denotation of a name (if there is one) *is a function of* the sense of the name […], that is, given the sense, the existence and identity of the denotation are thereby fixed, though they may not necessarily therefore be known to everyone who knows the sense. (Church 1956, 9)

---

[18] It must also be mentioned that, while theories of reference usually take proper names and other singular terms as their paradigmatic cases, the same does not hold of medieval theories of signification, as we shall see shortly.

[19] Cf. (Korte 2001) and (Parsons 1996).

[20] Cf. (Korte 2001).

[21] I use here the term 'reference' for the **relation** between an expression and a thing, and the term 'referent' for the thing itself, that is, the second *relatum* of the relation of reference.

Under Church's formulation, the procedure whereby the referent of an expression is determined (by means of the content of the sense) has a distinguished 'mathematical' flavor, insofar as senses are seen as functional operators that, when applied to an expression, yield its referent. Perhaps owing to Church's influential role, the dominant notion of what a theory of reference ought to be seems to have become inseparable from the idea of determination. The latter is found in many prominent theories of reference of the second half of the 20th century, also in those where the functional approach isn't emphasized as much as in Church's. Again, *Reference and Generality* seems to capture perfectly the central elements of the concept of reference.

> The view that in an assertion of the form 'Some man is P' 'some man' refers to some man seems to make sense because as regards any assertion of this form the question 'Which man?' is in order and if the assertion is true the question can be answered by naming a man who is P. (Geach 1962, 31)

It is implied thus that the relation of reference between an expression 'α' and an object α properly occurs only when it makes sense to ask for the determination (specification) of the object α in question. It is expected that the interpreter possesses all the necessary conditions to perform this determination and to give an answer to the 'which?' question. According to this view, it is as though a purported theory of reference that is not able to provide a systematic answer to the 'which?' question is not a good theory of reference.

1.1.3.3 Many-one mapping

Related to the idea that reference should be of a functional nature is the emphasis on the fact that this relation is (or in any case should be) a many-one mapping between expressions and what they refer to. It is possible that more than one word refers to the same object (in which case these words are usually said to be synonyms), and it is a trivial fact that there are different names for the same things in different languages. But there can be no equivocation concerning the second relatum of the relation of reference; it must be unique. This way, when a word is used, its referent is always unambiguous. Clearly, a many-one mapping between words and things is the best way to secure the *desideratum* of reference determination aforementioned (although determination has an epistemological dimension that is not necessarily guaranteed by a many-one mapping, namely if the relations between the different words and things are not **known** to the interpreter).

But it is a known fact that most languages are prone to ambiguity and equivocation. Some terms – that is, general terms such as 'dog', 'man' etc. – can be used to talk about a multitude of different individuals. Moreover, it sometimes occurs that words are used not to refer to their usual referent, but to some other entity, such as in 'My name is John', where 'John' is used to refer not to its usual referent – the bearer of the proper name 'John' – but rather to the term 'John' itself. To eradicate the first source of ambiguity, the usual strategy of most theories of reference propounded in the last century has been to stipulate 'artificial' entities as the unique referent of general terms. This procedure is exemplified by the following example: (Parsons 1996,

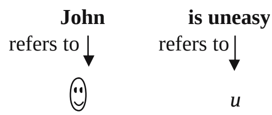


Figure 1.1.3.3.1. The referents of ‘John is uneasy’

373: ‘*u* is the function that maps uneasy things to truth and all other things to falsehood’) (Figure 1.1.3.3.1).

The preferred candidates for the post of unique referent of general terms are usually one of the following: *functions*, *classes* composed of the individuals of which the term can be predicated, or *intensional contents*. It is crucial that the phenomenon of multiple denotation be avoided, and this is accomplished by an expansion of the ontology with entities such as these.

As for the other source of ambiguity, related to words standing for things other than their usual referents, it is dealt with by an expansion of the language: new symbols are added to the language in order to ensure that overburdened terms be relieved of the task of naming more than one individual. The most conspicuous case is that of quotation marks; in order to refer to a term ‘X’ itself and not to its usual referent, the usual procedure is, as I have just done, to add quotation marks to the term in question. The string ‘X’ is thus the **name** of the string X. This general procedure can be expanded to similar cases of ambiguity, for example, the ambiguity between the reference to a thing or to a picture of this thing (the new term being in this case ‘picture of X’). Geach spells out these strategies in the following passage:

> Obviously, though, the double use of ‘man’ (say) in subject and in predicate position is not a casual ambiguity, like the use of ‘beetle’ for a mallet and for an insect; it is a systematic ambiguity, like the way that a common noun may be used to label either a thing of a given kind or a picture of such a thing, or again like the way that a word may be used to refer to that word itself. These systematic ambiguities **are removable** by the use of special signs, for example, the modifying words ‘picture of a’, or quotation marks; and similarly, if we have a logical sign (‘thing that —’) by prefixing which to a predicable we generate (something like) a name, then we may eliminate the subject-predicate ambiguity of ‘man’ by taking subject occurrences of ‘man’ as short for ‘thing-that is-a-man’ (where the copula, I have argued, is logically superfluous). (Geach 1962, 173)

In sum, in order to secure that each expression refers to at most one entity, modern theorists of reference usually proceed by expanding the ontology and the language. In fact, over the last 100 years or so, very few systems have been proposed in which the imperative of unique denotation was not present and/or in which the idea of multiple denotation was taken seriously.[22] Interestingly, and for as far as I know, it is usually not **argued** in favor of singular denotation; it is simply assumed that every semantic

[22] Notice though that, while certainly idiosyncratic, some work has been done on plural reference/multiple denotation in the last decades. Examples: R.M. Martin 1953, ‘On Truth and Multiple Denotation’. *Journal of Symbolic Logic* 18(1); P. Simons 1982, ‘Plural Reference and set theory’. In B. Smith (ed.), *Parts and Moments*. Munich, Philosophia, 199–256.

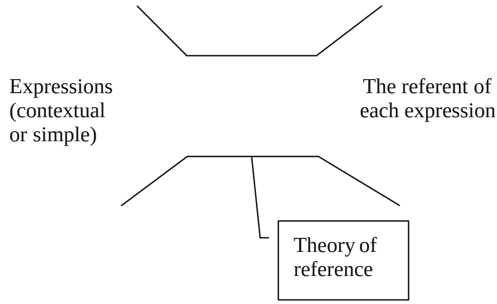


Figure 1.1.3.3.2. Input-output of a theory of reference

system should be based on the idea of many-one mappings between expressions and entities. At the origin of this assumption is perhaps the central role occupied by mathematics in the logic of this period, and in particular by the concept of **function**.

From this brief description emerges the following picture: a theory of reference ought to be something like a machine that, when given expressions as input, outputs their (unique) referents (Figure 1.1.3.3.2).

It could be argued that the picture here presented offers too narrow a view of the concept of reference, and that there are 'theories of reference' significantly different from this sketchy description. But in this case, my claim is that theories of supposition are unlike the theories of reference that do fit this description, and these are, I believe, an extremely significant and influential group of theories of reference.

1.2 WHAT SUPPOSITION THEORIES DO NOT DO?

Based on these three key aspects, which, I claim, underlie most of the received theories of reference, I will argue that theories of supposition are very different from theories of reference thus characterized. It will be shown that these aspects do not play a major role in the standard versions of supposition theory. This is true in particular of the determination aspect of reference and of the emphasis on many-one mappings between expressions and things; supposition theory, it will be argued, does not aim at the determination of referents, and it does not consist of procedures to establish many-one mappings between words and things, quite the contrary. Supposition theories are meant to establish the **range** of possible *supposita* of a term in a given propositional context.

As for the descriptive aspect of theories of reference – the description of the mechanisms in virtue of which the relation of reference occurs – medieval logicians also devoted their attention to this topic, but it did not pertain to the realm of supposition theory; it was in fact the subject matter of their theories of signification.

1.2.1 Theories of supposition do not explain the mechanisms of reference

The history of logic and philosophy seems to evolve in cycles; very often, that questions raised in one given period often recur after a certain time, within the specificities of each period. This is the case for instance of the issue of how and why words stand for things: the debate about the mechanisms of reference in the 20th century bears striking resemblance to discussions in the 13th and 14th century. This should not come as a surprise, as the issue of the grounds for the relation between words and the things they stand for is simply one of the crucial issues in the philosophy of language and logic of all times.

However, in the Middle Ages this debate took place on the *pre-propositional*, term level, whereas modern discussions usually take place against the still predominant background of propositional atomism. Therefore, while for philosophers of the 20th and 21st centuries this issue pertains to the general theme of reference (more often than not discussed with respect to propositional contexts), for medieval logicians it did not pertain to their theory of propositional contexts – supposition theory – but rather to their theory of **signification**. In the later Middle Ages, the issues as to whether the semantic relation between words and what they stand for is of a mediated or direct nature, as to the grounds for the relation between sign and signified, and similar issues all depended on one's position with respect to the notion of signification.

> *Supposition* is a semantic relation, holding between term(s) and thing(s). The relation of signification, however, is also a relation of term(s) and thing(s). Yet it is one matter to assign certain terms to certain things, so that a language may be set up in the first place; this is the contribution of signification. It is quite another matter to actually use that language to talk about things; this is explained by supposition […]. (King 1985, 35)

Thus, aside from signification pertaining to a pre-propositional level of semantics and supposition to a propositional level,[23] another major difference between the two notions is the following: the theory of signification is supposed to tackle the process by means of which words come to entertain a special relation to some things; in King's (slightly anachronistic) terms, the assignment of certain terms to certain things.[24] Supposition, however, only occurs when this first step has been accomplished and words already have a meaning – for example, entertain a special relation with given things. So, discussions such as 'what makes "Cicero" a name of Cicero?' (cf. section 1.1.3.1) pertain, in medieval philosophy, to the realm of signification.

---

[23] A nuance must be made, though: it is true that in the 14th century supposition was exclusively a property of terms in propositions, but in the 13th century, some authors – for example, Peter of Spain – talk about the supposition of a term outside a proposition. Cf. (Spade 1996, 247); De Rijk, 'The Development of *Suppositio naturalis* in Medieval Logic I: Natural Supposition as non-contextual Supposition'. *Vivarium* 9, 1971.

[24] The term 'assign' is what is somehow anachronistic here. It echoes current procedures in the modeling of artificial languages; for medieval philosophers, however, words acquired meaning in a very natural, causal way.

There are many medieval theories on how words acquire the meaning they have, and it is impossible to discuss them all in the present context. However, as argued by U. Eco et al. (Eco 1989), the variations between them can be understood in terms of the place they occupy with respect to the following dichotomy: (a) words are signs just as much as natural signs; (b) words are very special kinds of signs, namely *voces*. Theories geared towards view (a) would emphasize the semiotic relation between sign and signified, in particular the relation between words and what they signify, and would often insist on its causal nature – the symbol being able to signify something because the latter would be the cause of the former, just as much as smoke signifies fire. Alternatively, theories closer to the Aristotelian-Boethian position (b) would stress the cognitive and psychological process by means of which words come to signify things, and the intentionality involved in the act of talking about something by means of certain words.

Another point of disagreement among medieval logicians with respect to the notion of signification concerns the nature of the entities with which words bear a relation of signification. There seemed to be consensus with respect to the definition of the concept of signification, but disagreement as to what should fill up the place of the second relatum. There was

> [...] a basic agreement about the psychologico-epistemological criterion of signification (to signify $x =_{\text{df.}}$ to establish an understanding of x), and at the same time a disagreement over what spoken words immediately signify. Is it the corresponding concept? Or is it some external object? (Spade 1996, 136)

Words entertain a special relation to certain things, but which are those? The influential Aristotelian-Boethian tradition held that words primarily signify concepts, and only insofar as certain things bear a special relation to certain concepts are words said to relate to those things – their ultimate *significata*, as Buridan would say. In sum, this notion of signification depends on mediating (abstract) entities, and it bears resemblance with, for example, some of the views expounded by Frege in *Uber Sinn und Bedeutung*.

This tradition seems to have dominated the philosophical scenario well into the 13th century, when Roger Bacon's *De signis* appeared.[25] As the title suggests, he incorporated elements of aforementioned view (a) to his semantic views, basically by treating words essentially as signs. Roger Bacon contended that words have a special, *direct* relation to things, namely the very relation of signification – that is, signification for him is not the relation between words and concepts.

> Until Bacon in fact – and one thinks in particular of the semiotics of Abelard – the *vox* signifies directly the concept, or the species, and indirectly the thing or also, it denotes or designates (*denotat, designat*) its meaning (*sententia*) and names (*nominat* or *appellat*) the thing or the state of the world (*res*). In a few words, until Bacon, and along the Aristotelian track, words 'signify' the passions of the soul (concepts, or universal species), species have a relation with things and words, through the mediation of the species, serve to name things. With *De signis* however, words begin to point *directly* towards individual things, of which the *species intelligibiles* are the mental equivalents: but the relation between words and species becomes secondary, and is reduced to a purely symptomatic relation. (Eco 1989, 21)

[25] (Rosier-Catach 1994).

Therefore, Bacon is something of a predecessor of 20th century's theories of direct reference, – the 'new theories of reference' – insofar as his rupture with the tradition is strikingly similar to the rupture of direct reference theories with the Fregean tradition.

Roger Bacon's 'new' semiotic triangle had a lasting influence on the philosophical panorama of the time, especially in Oxford. In particular, influential philosophers such as William of Ockham adhered to this semiotic model and incorporated it to the very core of their theories.

> I say that spoken words are signs subordinated to concepts or intentions of the soul not because in the strict sense of 'signify' they always signify the concepts of the soul primarily and properly. The point is rather that spoken words are used to signify the very things that are signified by concepts of the mind, so that a concept primarily and naturally signifies something and a spoken word signifies the same thing secondarily.[26] (William of Ockham 1998, 50)

But the traditional view, according to which the relation between words and things is mediated (by a concept or a universal), remained predominant: Walter Burley clearly maintained it (cf. Burley 2000, 80–82), and so did Buridan, through the distinction between the immediate and the ultimate *significata* of a term (cf. Buridan 2001, section 4.1.2)

The present analysis of medieval semiotics and theory of signification is admittedly brief, but it is nonetheless sufficient for the present purpose, which is to stress the fact that medieval discussions on the concept of signification are similar to some aspects of 20th century discussions concerning reference. But the concept of supposition has no place in the explanation of how certain words come to stand for certain things. From this conclusion emerges the first element of disparity between theories of reference and theories of supposition.

1.2.2 Supposition theories do not determine referent: only the will of the speaker does

If the discussions concerning the notion of reference analyzed in section 1.1.3.1 do bear similarities to medieval discussions pertaining to the notion of signification, the same cannot be said of discussions analyzed in section 1.1.3.2, namely those concerning the determination aspect of theories of reference. Better put, there is no medieval theory concerning the determination of what terms stand for in a proposition, at least not with the same degree of expected efficiency of modern theories of reference.

In the relevant medieval texts, one encounters discussions as to whether the meaning of a term in a given context is defined by intrinsic linguistic properties of the context (*de virtute sermonis*), or whether it is the customary usage (*usus*) of a term

[26] Dico autem voces esse signa subordinata conceptibus seu intentionibus animae, non quia proprie accipiendo hoc vocabulum 'signa' ipsae voces semper significant ipsos conceptus animae primo et proprie, sed quia voces imponuntur ad significandum illa eadem quae per conceptus mentis significantur, ita quos conceptus primo naturaliter significant aliquid et secundario vox significant illud idem. Ockham, *Summa Logicae I*, chap. 1, 26–31 (p. 7)((Ockham 1974a) for all passages of his *Summa Logicae*). Throughout this work, I will refer, whenever available, to published translations of the Latin texts, and, when noted, with modifications. The Latin text will be at the footnote.

that must be taken into account, or else whether the *user* of the term has the freedom to choose the meaning intended for it (cf. Buridan 2001, section 4.3.2, a passage that will be discussed in detail below). But these discussions do not come even near the idea of an effective procedure to determine the *suppositum* of terms in propositions, if 'to determine' is taken in a strict acceptation (as it appears to be the case for most theoreticians of reference). However, the idea that supposition theory's goal is to define a procedure for the determination of reference relations seems to be remarkably widespread, also among historians of medieval logic. Take this passage by C. Panaccio:

On posera donc, en guise de conclusion, que la théorie de la supposition, considerée dans sa totalité, fournit une procédure générale pour déterminer l'ensemble des supposita d'un terme catégorematique simple pris comme sujet ou comme prédicat d'une proposition donnée. Cette procédure tient compte de certaines particularités du contexte propositionnel (la position syntaxique du terme consideré, le temps ou la modalité de la copule, la signification – métalinguistique ou non – de l'extrême opposé) et de certaines relations non propositionnelles que le terme pris en lui-même entretient avec d'autres objets individuels: la signification surtout, mais parfois aussi la ressemblance morphologique ou l'équivalence sémantique. (Panaccio 1992, 43)

What Panaccio describes in this passage is indeed very similar to the gist of many significant theories of reference: a theory of reference defines a general procedure for the determination of the object that a term in a given proposition stands for. However, shortly after Panaccio adds that, in actual cases, the procedure is not entirely effective.

Appliquée aux discours réels, la procédure n'est pas totalement effective puisqu'elle laisse subsister certaines ambivalences systématiques, celles en particulier qui sont suscitées par le temps ou la modalité du verbe ou celles encore qui tiennent à la présence d'expressions métalinguistiques [supposition matérielle ou simple]. Les cas de ce genre représentent la limite de l'analyse purement sémantique: la désambiguïsation exige des lors le recours à des facteurs pragmatiques. (Panaccio 1992, 43)

Indeed, the purely semantic procedures defined within supposition theory are not sufficient for the determination of the exact *supposita* of a term in a proposition; as we shall see below, what in fact fixes the *suppositum(a)* of a term in a sentence is the intention of the speaker. From this perspective, Panaccio's assimilation of supposition theories to theories of reference is only warranted if the notion of reference is taken in its pragmatic variant. To be sure, pragmatic views on the phenomenon of reference have indeed been proposed, for example, in (Strawson 1950) and (Donnellan 1966), but this is arguably not a dominant trend. In an influential paper, Kripke (1977) has in fact explicitly argued against pragmatic views on reference, and has stated his preference for what he calls 'unitary theories' as opposed to theories that postulate an ambiguity (cf. Kripke 1977, 384). Hence, even though Panaccio recognizes that, as a purely semantic theory of reference, supposition theory is not entirely effective, and therefore seems to be comparing it to **pragmatic** theories of reference, the fact that the latter are a minor trend among modern theories of reference does mean that his comparison may be somehow misleading to the uninformed reader. If the reader simply assumes that a theory of reference ought to be purely semantic (which he is likely to do), then the association of the notion of supposition to that of reference might be deceptive; it might lead him to believe that theories of supposition are defective theories of reference, since they do not offer a procedure to fix the referent of a term on purely semantic grounds – a judgement which, as I will argue, is not fair on the medieval theories.

In effect, to the best of my knowledge, no medieval author ever said that the aim of supposition theories was to fix the referent (*suppositum*) of a term in a proposition. Moreover, the disambiguation of semantic relations does not seem to concern supposition theories; rather, an author such as Ockham insists that ambiguous propositions 'must be distinguished', that is, its possible readings must be brought into light. But it is not the task of (Ockham's) supposition theory to determine, among the alternative readings, which one is the 'right' one (the idea behind the notion of disambiguation). In this sense, instead of viewing a given theory as a defective theory of something, critics of supposition theory may do well to conjecture that this theory is a theory of something **else**.

To be precise, some commentators have voiced similar concerns, for instance P. King, in the passage below (although I find his choice of terms somewhat convoluted):

> Hence it is a mistake to suppose that supposition theory will say exactly which things a term in fact supposits for. Rather, supposition theory will specify what things a term semantically supposits for, and then it is a separate question whether the supposition is successful. (King 1985, 37)

The convoluted expression would be 'semantically supposits for' (is there any kind of supposition that is not semantic?). The point King seems to be making is exactly that supposition theory will specify the range of things a term could supposit for, the possible *supposita*, but will not decide which possible *suppositum* is the actual one. This is all the more the case since, according to most medieval authors, only what is now known as the intention[27] of the speaker[28] really determines the *suppositum* of a term in a proposition that he produces (as we shall see shortly). This fact is also acknowledged by Panaccio, a few pages after the passage quoted above:

> Lorsqu'elles [les règles] ne sont pas décisives, la phrase en cause doit être considerée comme sémantiquement ambivalente: seule la reconnaissance – pragmatique, dirait-on aujourd'hui – de l'intention du locuteur autorisera alors une interprétation unique. (Panaccio 1992, 58)

Certainly, supposition theory alone does not allow the interpreter to 'select' the actual (intended) *suppositum* (referent) of a term in a proposition. Therefore, the 'which?' question, the core of theories of reference, is indeed out of place when it comes to theories of supposition. This being given, the passage below, by P.V. Spade is all the more puzzling because it seems to suggest awareness of the inappropriateness of

[27] However, to prevent misunderstandings, the use of this term here must be avoided, since the Latin term *intentio* has a very different meaning from 'intention'. I shall often use the term 'will' instead, although this term is also not an ideal choice – the supposition of a term is not a matter of the utterer's volition, as the term 'will' may be taken to imply, but rather of his decision that it be so.

[28] Roughly, what Kripke has termed 'speaker's reference' (as opposed to 'semantic reference') in (Kripke 1977).

the 'which?' question with respect to the notion of supposition, but at the same time attribution of a crucial role to this question.[29]

> Determinate supposition appears to have been originally thought of as reference to (at least) one determinate thing. When someone said 'Some man is running', for instance, it was appropriate to ask 'Which one?' Unlike discrete supposition, in which also a term supposited for a single thing, in a case of determinate supposition one could not immediately pick out the referent; any one of several would suffice. (Spade 1988, 208)

These three passages, from distinguished medievalists, show in different ways how easy it is to assume that the general purpose of supposition theories is to be compared to that of theories of reference; clearly, Geach's 1962 statement has remained very influential. But there is something awkward in this approximation, since supposition theories do not seem to make sense entirely if viewed as theories of reference, in particular with respect to the 'which?' question – the determination of the referent (*suppositum*).

Surprisingly, this kind of misunderstanding was anticipated by one of supposition theory's main exponents, William of Sherwood. His clarification concerns the risk that the technical term 'determinate supposition' would induce the idea that the *suppositum* in question is determined.

> There is some doubt regarding the division of personal into determinate and confused supposition, for it seems that when I say 'A man is running' the term 'man' does not supposit determinately, since [*A*] the proposition is indefinite, and [*B*], it is uncertain for whom the term 'man' supposits. Therefore it supposits [*A*] indefinitely and [*B*] uncertainly; therefore indeterminately.
>
> In response we must point out that there is a respect in which determinateness is opposed to uncertainty, and in that respect we can say that 'man' *supposits* indeterminately, as the objection has it. There is, however, another respect in which determinateness is opposed to plurality (*multitudini*) and whatever is single is determinate; and in this respect 'man' does stand determinately. For by virtue of the expression the sentence 'A man is running' says that the predicate is in some one individual, not in many, even though the predicate is in many – for [a sentence] some time permits this but does not signify it. And therefore 'man' supposits determinately, not confusedly, since by virtue of the expression [in which it occurs] it supposits for one and not for many.[30] (William of Sherwood 1966, 115) (with modifications)

Hence, for a term to have determinate supposition does not mean that the *suppositum* is determined, but rather that it concerns one *suppositum* instead of many. The same holds for all other kinds of supposition (except for discrete supposition): a determination of the *suppositum* does not occur.

---

[29] G. Matthews is also under the impression that this passage makes conflicting suggestions, cf. (Matthews 1997, 37).

[30] Dubitatur de divisione personalis in determinatam et confusam. Videtur enim, quod iste terminus 'homo', cum dico 'homo currit', non supponit determinate. Est enim indefinita. Et iterum incertum est, pro quo supponit. Ergo supponit incerte et indefinite, ergo indeterminate.

Dicendum, quod determinatio uno modo opponitur incertinudini. Et secundum hoc potest dici, quod indeterminate supponit, ut obiectum est. Alio modo opponitur multitudini, et est determinatum, quod unum est. Et sic stat determinate. Vis enim huius sermonis vult predicatum inesse alicui uni et non multis, lict praedicatum insit multis. Hoc enim patitur aliquando, sed non significant. Et ideo supponit determinate et non confuse, quia de virtute locutionis supponit pro uno et non pro multis. (Sherwood 1995, 148)

But then, in virtue of what does a term stand for one given thing in a proposition and for no other? If the rules of supposition theory are not able to retrace the actual *suppositum* of a term, what determines the actual *suppositum* of a term? A discussion of this issue can be found in John Buridan's treatise on supposition. He presents three opinions concerning what defines the 'right' supposition of a term in a given context:

> Some people have said that 'Man is species' is false by virtue of the expression. For the principal sort of supposition is personal supposition.[31] (Buridan 2001, 255) (with modifications)

According to the first opinion, the very constraints of language – namely the primacy accorded to personal supposition – force a certain *suppositum* upon the terms of the proposition. In this case, '*homo*' must supposit for a man (for it must have personal supposition), in which case the proposition is false, for no individual man is a species. The second opinion goes as follows:

> Another opinion states the contrary, namely, that 'Man is species' is true by virtue of the expression. For speech, in signifying and suppositing, only has any import because of conventional imposition and usage, and we cannot know which was the imposition except from the usage of authors.[32] (Buridan 2001, 255) (with modifications)

The second opinion ascribes the burden of supposition (and signification) determination entirely to the will and usage of the utterer, a sort of linguistic pragmatism *avant la lettre*. According to this opinion, there are no linguistic constraints to the intention of the utterer. But Buridan himself endorses a third opinion:

> The third opinion, with which I agree, is that an utterance does not have proper import in signifying and suppositing, except from ourselves. So by an agreement of the disputing parties, as in obligational disputations, we can impose on it a new signification and not use it according to its common signification.[33] (Buridan 2001, 256)

Hence, according to Buridan, the property of signifying and suppositing that words have does not stem from their proper meaning (*de virtute sermonis*), and also not (solely) from the intention of their producer: rather, it is a matter of convention, established among the participants of a given linguistic exchange – an interesting case being that of disputations and *obligationes*. In such circumstances, the usual signification of a term can be shifted by means of mutual agreement among the participants (the 'ourselves' in question). In other cases, for example, textual interpretation, the interpreter has the power to fix the supposition of a term with respect to his own interpretation of the text. Yet, for this purpose, first he has to establish the range of possible interpretations in order to choose the adequate one. So it would seem that,

---

[31] Quidam enim dixerunt istam esse falsam de virtute sermonis 'homo est species', quia principalis suppositio est personalis. (Buridan 1998, 41)

[32] Alia opinio ponit contrarium, scilicet quod de virtute sermonis haec est vera 'homo est species' quia sermo in significando et supponendo non habet virtutem nisi ex impositione ad placitum et usu, et non possumus scire quae fuit impositio nisi per usum auctorum. (Buridan 1998, 41)

[33] Tertia opinio est, cui ego assentio, quod vox in supponendo vel significando non habet virtutem propriam nisi ex nobis. Unde per conventum disputantium, ut in obligationibus, sibi possumus imponere novam significationem et non uti communi eius significatione. (Buridan 1998, 41)

in contexts of speech (disputations), quite some freedom exists with respect to 'new' conventional significations for words, because it is assumed that common agreement legitimates the convention.[34] In textual contexts, however, legitimacy must come from the accurate applications of the rules of supposition by the interpreter.

> [...] We would say that 'Man is species' is true insofar as it is put forward by authors, but it is not true by virtue of the expression, that is, it would not be true if it were put forward as a proper locution [...].[35] (Buridan 2001, 256)

So Buridan's final verdict is in the spirit of Ockham's claim that propositions in which terms can have more than one supposition 'must be distinguished' (see section 1.3.2.2 below). One must spell out the reading(s) that the given proposition ought to receive in case the analysis would limit itself to purely linguistic constraints, but one must also investigate whether other readings are possible, perhaps more in accordance to the author's original intention – or to the interpretation favored by the interpreter.

Indeed, this procedure makes sense in particular against the background of textual interpretation, which dominated intellectual activity at the time, as it was of course virtually never possible to 'ask' the author what his original intentions were. Moreover, many of the statements contained in authoritative texts (the Bible being the best example) are strictly speaking false if literally interpreted. 'Man is a species', for example, if literally interpreted, is false (according to Buridan). However, if one considers the possibility of 'man' not suppositing for its usual *supposita*, but rather for the universal (concept) 'man' – thereby having simple supposition – then the proposition may turn out to be true. Since one usually prefers to make true statements instead of false ones, especially in the case of authoritative sources, it is to be assumed that this second reading was the one intended by the author.

A similar view on the topic was held by William of Ockham. He seems to maintain that there are strictly linguistic constraints with respect to the supposition of terms, such as the priority of personal supposition, but these can be overruled by a special imposition of the author. In other words, it is not always imperative to turn to the will of the author in order to interpret and understand a given proposition, as there are linguistic constraints to be taken into account.

> It should also be noted that a term, in any proposition in which it is placed, can always have personal supposition, unless those who use it limit it to other forms of supposition. In the same way an equivocal term can, in any proposition, supposit for any one of its *significata*, unless those who use it limit it to one particular *significatum*.[36] (Ockham 1998, 191)

---

[34] Indeed, one sort of obligational disputations is called *impositio*, and consists precisely in giving a word or phrase a new meaning and responding accordingly. Cf. (Spade 1978, part III).

[35] … diceremus quod ista 'homo est species' est vera sicut ponitur ab auctoribus, sed non est vera de virtute sermonis, idest non esset vera si poneretur secundum propriam locutionem… (Buridan 1998, 42)

[36] Notandum etiam est, quod semper terminus, in quacumque propositione ponatur, potest habere suppositionem personalem, nisi ex voluntate utentium arctetur ad aliam. Sicut terminus aequivocus in quacumque propositione potest supponere pro quolibet significato suo, nisi ex voluntate utentium arctetur ad certum significatum. (Ockham, *Summa Logicae I*, chap. 65, 5–7 (p. 197))

The limitation of the supposition to one particular *significatum* may be a matter of contingency or of doctrinary choice, but in any case it falls out of the scope of supposition theory properly speaking; supposition theory only establishes the possible readings of a proposition.

In conclusion to this section, one may say that, in the end (and at any rate for authors such as Buridan and Ockham), the actual supposition of a term in a proposition is the one accepted (imposed) by its author. But in order to interpret a (written) proposition, it is neither always necessary nor possible to consult with the author as to his intentions: therefore, the interpreter shall attempt not only to retrieve the actual supposition of a term, but also to reconstruct the range of possible *supposita* that could be the case (and possibly favor one of them). In other words, supposition theory does not have the means to determine the referent (*suppositum*) of terms; neither is it its purpose.

1.2.3 General and ambiguous designation: one-many correspondence

Perhaps the most interesting point of divergence between theories of reference and theories of supposition concerns the general position with respect to ambiguities of the kind discussed in section 1.1.3.3. While theories of reference usually introduce *ad hoc* procedures to guarantee uniqueness of reference, theories of supposition seem to embrace semantic ambiguities as their very reason to exist. It is because many propositions allow for more than one reading that a formal theory that generates the possible readings of propositions is necessary, so that one can distinguish legitimate interpretations from illegitimate ones.

Semantic ambiguities – let us call them 'variations', a more neutral term – can be said to be of two kinds, let us call them 'horizontal' and 'vertical' variations. Horizontal variation occurs when a term can stand for several entities that belong to the same ontological level, so to say. This occurs in the case of general terms, such as 'man', which are predicated of many entities, and therefore can supposit for several entities in a proposition, namely any man. By contrast, vertical variation concerns the fact that terms can supposit for entities of different ontological levels, namely either for extra-mental (physical) objects, of for concepts or linguistic occurrences. Again, the word 'man' can undergo this kind of variation insofar as it can supposit for a man, but also for the concept of man or for an occurrence of the word 'man'.

Supposition theories do not try to rule out these variations, the idea is rather to bring them under control and to establish the rules that determine when they can or cannot occur. Once again, one must bear in mind that, although it deals with variation, it is not a case of 'anything goes', that is, a theory of supposition is not trivial, since it distinguishes legitimate possible *supposita* for a term in a proposition from illegitimate ones.

Let us take the case of general (common) terms. The fact that they can stand for a plurality of entities is not viewed as something undesired by medieval logicians, rather this is the very basis of their account of generality. While the modern procedure is to obtain generality by means of quantifiers that range over a certain domain and are connected to functional predicates by variables, the medieval approach to generality is to allow meaningful general terms to range over a certain domain. On the one hand, uniqueness of reference is not preserved, but on the other hand, the nature of general terms seems to be rendered more faithfully. In other words, medieval logicians do not see the potential to supposit for more than one individual as a semantic flaw, but rather as a desired property of certain terms.

Of course, multiple denotation alone is not a sufficient account of generality: if general terms signify many entities, and, therefore, can stand for them in a proposition, it might not be possible to refer to only some of the *significata* of a term at a time, instead of to all of them. In other words, it would only be possible to talk about *all* men by means of the word 'man', and never about some men only. Now, there come in play the theories of modes of personal supposition: they account for the fact that it is possible to talk about some men only and not about all of them every time that the word 'man' is used.[37] That is, a restriction on the range of possible *supposita* of a term occurs if it has personal determinate supposition – better put, a restriction on the quantity of *supposita*, even though these can be picked among any of the *significata* of the given term.

> The supposition is called determinate not because a term suppositing determinately in this way supposits for one *suppositum* and not for another. Rather the supposition is called determinate because for the truth of a proposition in which a common term supposits determinately it is required that the proposition be made true for some determinate *suppositum*.[38] (Walter Burley 2000, 102/3)

It is the absence, or presence, of certain syncategorematic terms such as 'some' (*aliquod*) or 'all' (*omnis*) that will indicate how many *significata* of a term in a given proposition are required for the truth of the proposition. Still, determination of **which** of the *significata* of a term, that is, its actual *suppositum* in a proposition, does not occur.

Thus, general terms are able to supposit for any of their *significata* indistinctively, and how many of those are at stake is defined by auxiliary terms (currently called quantifying terms). Medieval authors did not experience the semantic phenomenon of (anachronistically put) multiple denotation as a threat, but rather as the very basis for generality.

So much for 'horizontal variation'. Let us now turn to 'vertical variation', that is, the fact that a term can stand for things of different ontological levels. As suggested by the passage by Geach quoted above, the modern attitude towards ambiguities of the use/mention kind, reference to concept/reference to physical entities, and so forth, is to introduce new terms in the language so that a one-one mapping between terms and things can be established. Medieval logicians also recognized that a term can

---

[37] According to G. Matthews, this is in fact the very purpose of the modes of personal supposition; cf. (Matthews 1997, 39).

[38] Et dicitur suppositio determinata, non quia terminus sic supponens determinate supponit pro uno et non pro alio, sed dicitur suppositio determinata, quia ad veritatem propositionis, in qua terminus communis supponit determinate, requiritur quod verificetur pro aliquo supposito determinato. (Burley 1955, 20 (32–36))

stand for these different things, but instead of using syntactic procedures to eliminate the ambiguities, they saw it as a legitimate semantic phenomenon, namely that of terms being interpreted in different ways – and therefore standing for different things, what was accounted for by the familiar distinction between the kinds of supposition (cf. Ockham's *Summa Logicae I*, chap. 64). The list of kinds of supposition varies per author: Ockham, for example, recognized three basic kinds of supposition, namely personal, simple and material. By contrast, Buridan accepted only personal and material supposition (and takes simple supposition to be a special type of material supposition) (cf. Buridan 1998, 39; Buridan 2001, section 4.3.2).

The fundamental principle behind these distinctions is similar to what is sometimes called 'referential shift', a concept whose (re)introduction can be attributed to Frege: words customarily stand for a given entity – their customary referent – but in some contexts they may be forced to stand for something else (cf. Frege 1948). Paradigmatic cases are the so-called intensional contexts: when they are the principal verb of a sentence with an embedded sentence, verbs such as 'believe', 'think', among others, force a referential shift upon the terms of the embedded sentence. So in 'Mary believes that John is uneasy',[39] 'John' does not refer to its usual referent (John) but rather to its usual sense, that is, something like the concept of John.

It seems to me that the different kinds of supposition can be successfully explained by means of a comparison with the shifts of reference identified by Frege (the usual comparison between material supposition and quotation marks seems to be more misleading than helpful). That is, in some propositional contexts, a term may be interpreted as not having its usual *suppositum* (most authors agree on the priority given to personal supposition), but as having a different kind of supposition. The crucial difference between Fregean shifts of reference and medieval kinds of supposition is that, at least according to some authors, among whom Ockham, a given context does not force another kind of supposition upon its term, but rather allows for that possibility **alongside** with the usual (personal) supposition. This difference is all the more significant precisely because it reveals a fundamental disparity between these two groups of theories – namely that, in theories of reference, the unique referent of a term in a given context must always be determined, whereas in theories of supposition, a term in one and the same proposition can have either personal and/or simple and/or material supposition.

Let us examine an example, the familiar 'man is a species'. Because the predicate is a term that signifies concepts, the subject can have simple supposition (cf. section 1.4.1.1 below); but this does not overrule the possibility of it having personal supposition as well. It is up to the interpreter to decide which supposition (and therefore which *suppositum*) is to be preferred (which he may do based on his own theoretical preferences or as an attempted reconstruction of the meaning intended by the original author of the proposition). So personal and simple supposition coexist as potential interpretations for the term 'man' – what is tantamount to saying that 'man'

---

[39] Example from (Parsons 1996, 374).

in this proposition can supposit either for a man or for the concept of man. In modern theories of reference, such coexistence of potential referents is usually not permitted.

Again, rather than considering such ambiguities (or ‘equivocations’, to use the Aristotelian vocabulary) as flaws of language, medieval authors seemed to view them as a sign of the resourcefulness of language, of its flexibility to express different ideas. If well mastered, these vertical variations would only contribute to a welcome multiplication of possible readings that a proposition might receive. The ability to establish such readings would show the degree of logical acumen of an interpreter, in particular in the case of difficult propositions (*sophismata* and *insolubiles*); the challenge they presented was an important stimulus for developments in logic and semantics.

1.2.4 Similarities

Still, one might argue, it is undeniable that theories of reference and theories of supposition are basically about the same thing, namely the fact that words can stand for things, so that we can talk about things by means of words. This much must of course be granted; but then, will any philosophical investigation of this phenomenon necessarily be a theory of reference? If so, one must speak of a Platonic theory of reference, of an Aristotelian theory of reference, of a Hobbesian theory of reference, to name but a few. But this seems far-fetched; the issue of the relation between words and things is simply a very old and recurring philosophical topic, which may or may not receive the general term ‘reference’ – as far as I am concerned, preferably not, since the term is too theory-laden.[40]

Yet, under the light of historical developments in philosophy, the comparison between theories of reference and theories of supposition is not as uninformative as I may have seemed to imply. After the decline of scholastic philosophy, so-called ‘logical investigations’ were mainly concerned with the manipulation of *concepts*, and interest in the linguistic level of analysis was limited (although, as always, there were exceptions). Frege’s criticism of the psychologistic approach to logic, at the end of the 19th century, redirected logic towards the relation between expressions and things, or, more radically, to the relation of expressions to expressions within a logical structure – and not of concepts to concepts, as had been the case roughly for four centuries.

So, it is understandable that medieval logical theories may strike the modern scholar as much closer to post-Frege discussions than, say, Kant’s work in logic (cf. Caygill 1995). It is thus easy to make the assumption that medieval theories must roughly be concerned with the same problems that we, logicians and philosophers of the 20th/21st century, think are the relevant semantic and logical issues. But this assumption can result in the aforementioned undue projections.

---

[40] Cf. the passage by Sainsbury quoted in section 1.1.1 above.

However, exactly because of this similarity, medieval discussions are more likely to contribute to modern logical discussions than most of post-Descartes, pre-Frege philosophy, provided that the due precautions are taken, and that a proper conceptual and historical analysis is made.

1.3 WHAT ARE SUPPOSITION THEORIES THEN?

In sum, there seem to be compelling arguments against the view that supposition theories are theories of reference. What are they theories of, then? If one is looking for the closest modern counterpart of these medieval theories, an immediate reaction is to turn one's attention to the other major group of modern semantic theories, namely theories of meaning. Indeed, according to Quine, modern semantics can be divided into two main groups:

> When the cleavage between meaning and reference is properly heeded, the problem of what is loosely called semantics becomes separated into two provinces so fundamentally distinct as not to deserve a joint appellation at all. They may be called the *theory of meaning* and *theory of reference*. 'Semantics' would be a good name for the theory of meaning, were it not for the fact that some of the best work in so-called semantics, notably Tarski's, belongs to the theory of reference. The main concepts in the theory of meaning, apart from meaning itself, are *synonymy* (or sameness of meaning), *significance* (or possession of meaning) and *analyticity* (or truth by virtue of meaning). Another is entailment, or analyticity of the conditional. The main concepts in the theory of reference are *naming, truth, denotation* (truth-of) and *extension*. Another is the notion of values of variables. In 'Notes on the theory of reference', in (Quine 1961, 130).

If it does not belong to the territory of theories of reference, then, according to Quine's description, supposition theory should belong to the territory of theories of meaning. Of course, it may be argued that the very division of semantics in these two areas does not apply to medieval semantics, and that there are probably just as many arguments against the ascription of supposition theory to the general group of theories of meaning as there are against its ascription to the group of theories of reference. Yet, it seems to me that the claim that supposition theory is a theory of meaning puts us in a more promising path towards a better grasp of this medieval group of theories than its assimilation to theories of reference.

Hence, in this section I argue that supposition theory is better seen as a theory of propositional meaning, primarily intended to provide a procedure of analysis for the establishment of what can be asserted by a given proposition, rather than the establishment of the entities that the proposition is about (as would be the case if it was a theory of reference). In other words, supposition theory is a theory of interpretation, of semantic analysis – in sum, a theory of hermeneutics. As the title of this chapter gives away, the main thread to be followed here is the idea that the corresponding procedure can be applied mechanically, that is, they are theories of algorithmic hermeneutics; in this section I will argue for the hermeneutic part of the claim, while the formalization presented in section V is intended to substantiate my attribution of an algorithmic, formal character to these theories.

The claim that Ockham's supposition theory is a theory of propositional meaning is bound to be seen with suspicion. The very appeal of this theory from a modern point of view seemed to be its strictly extensional and causal character, insofar as the signification of a term are the things it can be predicated of, instead of awkward abstract entities, and as the basic semantic relations between words and things are immediate and defined causally. But the extensional nature of Ockham's supposition theory seems to be lost if it is viewed as a theory of meaning (given that the notion of meaning is traditionally seen as an essentially intensional notion).

But even if it is a theory of intensions – if one views the meaning of a piece of language as an intension – Ockham's theory can be said to be an *extensional theory of intensions* (meaning). In other words, the intension(s) corresponding to a complex phrase, and in particular to propositions, is (are) determined by the extensions of its terms; to use Ockham's own terms, what is asserted by a proposition [*denotatur*] is determined by the supposition of its terms. Moreover, one of the most interesting aspects of Ockham's theory is the recognition that many propositions allow for more than one reading, and thus that different assertions can be made by means of such propositions; among those, many are cases of propositions with terms which can supposit in different ways and for different things, and precisely for this reason such propositions yield more than one reading. The phenomenon of multiplicity of propositional meanings is indeed one of the main focal points of the present analysis, insofar as it is (it seems to me) the very heart of Ockham's theory of supposition.

1.3.1 Historical arguments

Although philosophers tend to be resistant to the incorporation of historical elements into philosophical analyses, I think that it would be a mistake not to take into account how theories of supposition were actually used in the later Middle Ages. We know that they were extensively used for a variety of purposes – in philosophy, theology and other disciplines – so the effort to understand these theories will surely benefit from a brief survey of some historical elements. Here, I am particularly interested in the crucial role played by textual commentary and interpretation in the intellectual world of that time.

Moreover, it is a well-known fact that supposition theory bears close connections with theories of fallacies, whose main source is Aristotle's *On Sophistical Refutations*. It has even been claimed that theories of fallacies are the very source for the historical development of theories of supposition (cf. De Rijk 1967). Even if one disagrees with this claim, it would be hard to deny that there are indeed significant links between these two groups of theories. Therefore, an examination of some of these links should also shed some light on supposition theory.

1.3.1.1 Fallacies

The story I will tell now contains nothing novel; quite to the contrary, it is common knowledge among medievalists. It is nevertheless worth being told, as it may reveal a bit of the essence of supposition theories.

Aristotle's work in logic was 'edited' into six books by Andronicus of Rhodes in the first century BC: *Categories, On Interpretation, Prior Analytics, Posterior Analytics,* *Topics* and *On Sophistical Refutations*; he gave the name *Organon* to the collection of these six texts (cf. Spade 1996, 11). Some 850 years after Aristotle, Boethius (presumably) translated all of them into Latin, but somehow most of this material got lost or remained very little known; only the translations of the first two texts were well known in the western Christian world until the year 1100.

Then, in the 12th century a growing interest in the Aristotelian corpus came about, in particular in his logical texts. New translations were made, but also the old Boethian translations were somehow recovered (cf. Dod 1982, 46).[41] Due to several circumstances, the *Topics* and in particular *On Sophistical Refutations* had a decisive influence on medieval developments in logic.[42]

The aim of *Sophistical Refutations*, usually referred to under its Latin name, *Sophistici Elenchi* (henceforth SE), is to enable the recognition of pseudo-refutations. It is a compilation of fallacies, that is, ways in which reasoning may appear to be correct even though it is not; this often involves phenomena such as words being taken in different ways. In sum, it is, among other things, a manual on how complex expressions can be interpreted in several ways, giving rise to sophistical refutations – refutations that seem correct but are not. One of the underlying ideas of SE is thus that language is prone to conflicting interpretations. So, probably influenced by this idea, medieval masters started to develop a semantic system codifying these ambiguities. This was, according to de Rijk, the beginning of supposition theories.

> Much more evident, however, is the influence of the *Sophistici Elenchis* of Aristotle. Thus the different types of supposition are not only illustrated by means of fallacies, but, as we shall see, derived from the determination of those fallacies. (De Rijk 1962, 22)

Indeed, as far as in the 14th century, the existence of different kinds of supposition was still justified in terms of the concepts presented in SE.

> Further, with this proposition we must, in accordance with the *third mode of equivocation*, distinguish between the case where the subject may have simple supposition and the case where it may have personal supposition.[43] (Ockham 1998, 192) (with modifications – emphasis added)

In sum, supposition and fallacy theories are both theories about how terms can be interpreted in different ways: sometimes they stand for something other than what they usually stand for, or may stand for different, alternative things. This phenomenon is explained in terms of the different kinds of supposition that a term may have, or in terms of the different kinds of fallacies that can occur. In particular, supposition theories spell out in which contexts such variations can occur, and for which different,

---

[41] 'One of the results of the quickening interest in logic in the early twelfth century was the recovery, from about 1120 onwards, of the rest of Boethius' translations of [Aristotle's] logic: the *Prior Analytics, Topics* and *Sophistici elenchi*. How and when these translations, made some six centuries earlier, were found is not known'. (Dod 1982, 46)

[42] 'This little work [SE] began to circulate in Latin translations sometime after about 1120, and its appearance was a crucial event in the history of mediaeval logic'. (Spade 1996, 38)

[43] Et est propositio distinguenda penes tertium modum aequivocationis, eo quod subiectum potest habere suppositionem simplicem vel personalem. (Ockham, *Summa Logicae I*, chap. 65, 15–17 (p. 198))

alternative things a term can stand for. By means of this apparatus, an interpreter or hearer shall not be trapped in sophistical refutations, since he will be able to anticipate all possible readings a given proposition can receive.

There is abundant textual evidence to the effect that supposition theories developed in close proximity with the doctrine of fallacies (cf. De Rijk 1967); this fact suggests that supposition theories are particularly appropriate for dealing with semantic ambiguities and fluctuation of meaning – in sum, with problematic cases. This seems to be often forgotten; it is sometimes said that the concept of material supposition, for example, is trivial, since it only came to being because 'medievals did not have quotation marks'.[44] This kind of remark misses the point, since the interesting cases are exactly those of terms in propositions that can be interpreted either way, that is, as standing for themselves or for equiform strings or as standing for one of their usual *significata*. As argued in section 1.2.3, the whole point is not to rule out ambiguities, but to bring them under control. Now, this was certainly one of the purposes of the techniques presented in SE as well. Why the need for an alternative semantic system was felt (wasn't the fallacy framework sufficient?), and the conceptual and historical relations between these two groups of theories are delicate issues (cf. Dutilh Novaes forth coming); nevertheless, emphasizing the similarities between theories of supposition and theories of fallacies does seem to help clarifying the general nature of the former.

1.3.1.2 Commentaries

It is a well-known fact that academic activity in the Middle Ages was to a great extent articulated around textual commentaries. In this section, I argue that the nature of this activity gives us clues to understand supposition theories in general, insofar as it all seems to indicate that these theories were developed, among other things, in order to respond to intellectual needs related to the performance of textual commentary.

As a part of his academic education, any master of arts to-be had to write commentaries on some inescapable texts, in particular of the Aristotelian corpus. However, authoritative texts were not to be contested *verbatim*. So commentators had to interpret the original texts rather freely in order to express their own views,[45] which could be done by making use of the supposition apparatus; often, the result was that two different authors were able to extract two very different if not opposite views from the same text.[46] In order to be able to extract 'novelties' from authoritative texts, medieval

---

[44] Which is not even true; they made use of the particle *li* or *ly* preceding a word to indicate that the very word was being talked about, and not its usual *significata*.

[45] 'However, since new doctrines were often disguised as merely new interpretations of the authorities, it is difficult to know where the borderlines between the exposition of the authority and the author's personal philosophy may be drawn'. (Kenny and Pinborg 1982, 29)

[46] 'It is also interesting to follow the 14th-century debate between Ockham and his followers on one side, who brought their own doctrine into harmony with Aristotle by claiming that Aristotle often spoke metaphorically, and more traditional Aristotelians on the other side, who could not concede that the philosopher ever spoke metaphorically, since he himself claimed to proceed demonstratively'. (Kenny and Pinborg 1982, 30)

masters could certainly benefit from a flexible semantic and interpretational system, and supposition theories appear to be exactly what was needed. If the basic idea of supposition theories is that words can stand for things in different ways, and thus for different things, it is not too farfetched to imagine that, given a certain authoritative statement, one commentator would attribute, say, personal supposition to its subject, whereas another commentator, of a different intellectual trend, would see a case of simple supposition, so that two different positions would find confirmation in the very same authoritative statement.

Admittedly, this is only a hypothesis concerning the connection between textual interpretation and supposition theory: among authors of the 13th and 14th century an explicit statement of the purposes of supposition theories is virtually not to be found. But in later authors, for example, in a commentary to Marsilius of Inghen's treatise on supposition, one encounters this sort of meta-analysis:

> Supposition was invented in order to save the truth and falsity of propositions that are simultaneously granted and denied.[47]

As I interpret this passage, the author is contending that the apparatus of supposition theory can, on the one hand, 'save' a proposition that seems false and yet must be true, on account of being an authoritative statement (or the other way round, if it is attributed to a presumably unreliable source), and that, on the other hand, this apparatus is meant to distinguish the different readings (possibly with different truth values) of propositions that (in virtue of having more than one reading) must be granted and denied at the same time. (Notice also the use of the obligational framework, to be discussed in part 3 of the present work.)

But even in earlier authors, references to the relation between textual commentary and supposition theory can be found. Tellingly, in his treatise on supposition, while discussing what determines which supposition a term has in a proposition, Buridan is led to examine the issue of how to interpret authoritative statements:

> So it seems to me that wherever it is evident that an author has put forward a proposition in a true sense, although not as a proper locution, then to deny that proposition without qualification would be cantankerous and insolent. But to avoid error, it should be properly pointed out that the proposition is not true in the proper sense, or by virtue of its proper meaning, and then it has to be shown in which sense it is true.[48] (Buridan 2001, 256)

Not only for the interpretation of philosophical texts was supposition theory a good interpretational device; it is possible that its use in theology was even more important.

---

[47] 'Suppositio est adinuenta propter salvare veritates et falsitates propositionum comiter concessarum et negatarum' (my translation). In *Clarissimi philosophi Marsilii de Inguen textus dialectics de suppositionibus, ampliationibus, appellationibus, restrictionibus alienationibus et duabus consequentiarum partibus abbreviatus (...)*. Vienne 1512 (per Hieronymum Victorem et Johannem Singrenium calcographos excusum), p. 159v. E.P. Bos drew my attention to this passage.

[48] Videtur ergo mihi quod, ubi apparet auctorem posuisse aliquam propositionem ad aliquem sensum verum, licet non secundum propriam locutionem, negare simpliciter propositionem esset dyscolum et protervum. Sed ne erretur debet bene dici quod propositio non est vera secundum proprium sensum sive de proprietate sermonis, et debet ostendi secundum quem sensum est vera. (Buridan 1998, 42)

Many difficult passages of the Bible, or problematic dogmas of the Christian faith, were often accounted for by means of the distinction of different kinds of supposition. In fact (although this line of research has not been conducted extensively hitherto[49]), it is possible that the hermeneutic activity of biblical interpretation would have had significant influence in the development of logical methods such as those related to the concept of supposition. At any rate, in practice, the framework of supposition theory was extensively used for theological discussions: consider Ockham's commentary on the first book of the *Sentences* of Peter Lombard, discussing whether 'God generates God' should be conceded (Distinction IV, question 1). After examining and criticizing other opinions, he presents his own.

> Therefore I have another opinion about the problem. First we should consider supposition in general, secondly supposition of this term 'God' in particular.[50]

Of course, the historical relation between textual commentary and supposition theories is not of necessity, for many commentary traditions did not rely on anything like supposition theory and handled the possibility of alternative readings in different ways. Yet, the prominent role of textual commentaries in medieval academic activity and the use of supposition theories therein offer an interesting vantage point for a better understanding of these theories.

1.3.2 Conceptual arguments

While the historical arguments just presented do not concern any medieval author in particular, the conceptual arguments below are taken primarily from Ockham's supposition theory, and evolve around two key phrases, which occur countless times in Ockham's *Summa Logicae*: '*denotatur*' and '*propositio est distinguenda*'.

1.3.2.1 Denotatur

The expression '*denotatur*' occurs numerous times in the *Summa Logicae*, in particular in the chapters discussing supposition theory (but also in *Summa Logicae* III-4, where the theory of fallacies is expounded). It is remarkable though that few medievalists seem to have recognized the significance of this expression for Ockham's theory of supposition (notable exception being (Marmo 1984), only in Italian).

What does '*denotatur*' mean? If it is related to our current sense of 'denotation', then this suggests that Ockham's theory of supposition is indeed something like a theory of reference (see Quine's quote, where denotation is said to pertain to the

---

[49] Some of the few texts on the topic are: (Brown 1993) and (Bos 2004). Brown says in particular: 'We will be looking at one particular theological case and showing how a small, number of Chatton's predecessors dealt with that case and *how they tried to develop a theory of supposition to handle it*'. (p. 124, my emphasis)

[50] Ideo dico aliter ad questionem. Et primo videndum est aliquid de suppositione in communi, secundo in speciali de suppositione huius termini 'Deus'. (Ockham 1974b, p. 7, 11–13). The English translation was made by E.P. Bos, who also drew my attention to this passage.

general topic of reference). But the meaning of the phrase '*denotatur*' is much richer (cf. Eco 1989 – where Eco (wrongly?) concludes that '*denotatur*' has an extensional meaning for Ockham). In Ockham's text, the expression '*denotatur*' usually appears in phrases of the form '*per istam propositionem 'p' denotatur quod* ...' or similar ones. The expression is always used in the passive form, and applies to propositions, not to terms (as in the current meaning of 'denotation'). Roughly, it means: '[by this proposition] it is asserted that ... / it is indicated that ... / [this proposition] means that ...'.

In the very first chapter in which the notion of supposition is explained, Ockham makes extensive use of the expression '*denotatur*'.

> More generally, if the suppositing term is a subject, it supposits for the thing of which (or of the pronoun referring to which) it is asserted by the containing proposition that the predicate is to be predicated. [...] Thus, by the proposition 'Man is an animal' it is asserted that Socrates is an animal, so that were the proposition 'This is an animal' (pointing at Socrates) formed, it would be true.[51] (Ockham 1998, 189) (with modifications)

> In the former case [if the subject has simple supposition] we have a true proposition asserting that a concept or intention of the soul is a species, and that is true. In the latter case [if the subject has personal supposition], we have the simply false proposition asserting that something signified by 'man' is a species, which is clearly false.[52] (Ockham 1998, 192) (with modifications)

That '*denotatur*' is the key concept to understand supposition theory becomes particularly clear in the case of empty terms (in propositions) and of false propositions.

> To the second doubt, it must be said that, by virtue of the expression, it must be conceded that, if no man is white and no man sings the mass and if God does not create, then in the aforementioned propositions, the subject term does not supposit for anything. And yet it is taken significatively, since 'taken significatively' or 'supposit personally' can be understood in two ways: either that the term supposits for one of its

---

[51] Here I quote the whole passage in Latin: Et sic universaliter terminus supponit pro illo de quo – vel de pronomine demonstrante ipsum – per propositionem denotatur praedicatum praedicari, si terminus supponens sit subiectum; si autem terminus supponens sit praedicatum, denotatur quod subiectum subicitur respectu illius, vel respectu pronomis demonstrantis ipsum, si propositio formetur. Sicut per istam 'homo est animal' denotatur quod Sortes vere est animal, ita quod haec sit vera si formetur 'hoc est animal', demonstrando Sortem. Per istam autem 'homo est nomen' denotatur quod haec vox 'homo' sit nomen, ideo in ista supponit 'homo' per illa voce. Similiter per istam 'album est animal' denotatur quod illa res quae est alba sit animal, ita quod haec sit vera 'hoc est animal' demonstrando illam rem quae est alba; et propter hoc pro illa subiectum supponit. Et sic, proportionaliter, dicendum est de praedicato : nam per istam 'Sortes est albus' denotatur quod Sortes est illa res quae habet albedinem; et si nulla res haberet albedinem nisi Sortes, tunc praedicatum praecise supponeret pro Sorte. (Ockham, *Summa Logicae I*, chap. 63, 16–32 (p. 194))

[52] Primo modo est propositio vera, quia tunc denotatur quod una intentio animae sive conceptus sit species, et hoc est verum. Secundo modo est propositio simpliciter falsa, quia tunc denotatur quod aliqua res significata per hominem sit species, quod est manifeste falsum. (Ockham, *Summa Logicae I*, chap. 65, 17–21 (p. 198))

Another interesting passage: Et ideo dicitur suppositio determinata quia per talem suppositionem denotatur quod talis propositio sit vera pro aliqua singulari determinata; quae singularis determinata sola, sine veritate alterius singularis, sufficit ad verificandam talem propositionem. (Ockham, *Summa Logicae I*, chap. 70, 21–25 (p. 210)) *significata*, or that it is asserted to supposit for something or that it is asserted not to supposit for anything. Thus, in all such affirmative propositions, it is asserted that the term supposits for something, and therefore if it supposits for nothing, the proposition is false. Now, in negative propositions, it is asserted that the term does not supposit for anything, or that it supposits for something of which the predicate is denied, and therefore such negative [proposition] has two causes of truth. […] In 'a white man is white', if no man is white, the subject is taken significatively and personally, not because the subject supposits for something, but because it is asserted to supposit for something; and since it supposits for nothing, and yet it is presumed to supposit for something, the proposition is simply false.[53] (Ockham 1998, 206) (with modifications).

Let us take the case of an affirmative proposition whose subject is not empty, but whose intersection of *supposita* of subject and predicate is empty – a false proposition, thus: 'A man is an donkey'. One might object: since there are no men who are donkeys, for which men does the subject supposit? Apparently, for none. The reply to this objection relies on the second (intensional) sense of 'supposit personally' as defined by Ockham in the quotation above: it is asserted that 'man' supposits for men and that 'donkey' supposits for donkeys, but since there is no common *supposita* between subject and predicate, and it is asserted that there is one (since the proposition is affirmative), then the proposition is simply false.

A similar phenomenon occurs with affirmative propositions whose subject is an empty term, such as 'A chimera is white': what does 'chimera' supposit for, since there are no chimeras? Again, the reply to this objection is based on the concept of '*denotatur*'. It is asserted that 'chimera' supposits for something (since the proposition is affirmative), but since there are no chimeras, the proposition is false.

In sum, the purely extensional interpretation of supposition theory – supposition being only the relation between a word and the thing(s) it stands for – simply cannot account for false and vacuous propositions. It has been claimed that this is a sign of the inconsistency of Ockham's theory (cf. Karger 1978), but it seems manifest to me that Ockham was very well aware of these problems, and therefore formulated his theory of supposition in what may be called intensional terms:

One might contend that the notions of 'to supposit' and 'to supposit for nothing' are incompatible since the following is a valid inference: the term supposits, therefore it supposits for something. The response is

[53] Ad secundum dubium dicendum est quod de virtute sermonis est concedendum, si nullus homo est albus et si nullus homo cantat missam et si Deus non creat, quod in praedictis propositionibus subiecta pro nullo supponunt. Et tamen sumuntur significative, quia 'sumi significative' vel 'supponere personaliter' potest dupliciter contingere: vel quia pro aliquo significato terminus supponit, vel quia denotatur supponere pro aliquo vel quia denotatur non supponere pro aliquo. Nam semper in propositionibus talibus affirmativis denotatur terminus supponere pro aliquo, et ideo si pro nullo supponit est propositio falsa. In propositionibus autem negativis denotatur terminus non supponere pro aliquo, vel supponere pro aliquo a quo vere negatur praedicatum, et ideo talis negativa habet duas causas veritatis. […] In ista autem propositione 'homo albus est homo', si nullus homo sit albus subiectum sumitur significative et personaliter, non quia supponit pro aliquo, sed quia denotatur supponere pro aliquo; et ideo quia pro nullo supponit, cum tamen denotetur supponere pro aliquo, est propositio simpliciter falsa. (Ockham, *Summa Logicae I*, chap. 72, 113–130 (p. 218/9))

that the inference is not valid. The following, however, is valid: the term supposits; therefore it is asserted either to supposit for something or not to supposit for anything.[54] (Ockham 1998, 206)

Other authors handled this problem with different concepts, but the basic intuition seems to be the same. In the thirteenth century already, Peter of Spain was aware of this possible misinterpretation of the concept of supposition, and insisted on the difference between the concepts of supposition and of verification (what makes a proposition true):

Determinate supposition is the acceptance of a common term taken indefinitely or of a common taken with a sign of particularity, as in ‘A man runs’ or ‘Some man runs’. The supposition in both these [propositions] is said to be determinate because, although in both the term man supposits for every man, both those running and those not running, they are true when just one man is running. For suppositing is one thing and it is another thing for a locution to be true of something, since in those propositions, as was said, the term man supposits for every man, those running and those not running, although it renders a locution true only of the running ones.[55] (Peter of Spain 1945, 8) (with modifications)

The distinction between supposition and verification is later taken over by Buridan, whose *Summulae* is, as we know, a commentary (albeit heavily modified) of Peter of Spain’s *Tractatus*.

(2) Again, it is possible that terms have supposition in a proposition without the verification of the proposition, in affirmatives as well as in negatives, as in ‘A man is a donkey’ or ‘A man is not an animal’. (3) Furthermore, there can be no verification without supposition in negatives, as in ‘A chimera is not a goat-stag’. But in the case of true affirmatives it is necessary that the proposition be verified of some thing or of some things for which its terms supposit.[56] (Buridan 2001, 224)

Clearly, it is verification that opponents of supposition theory have in mind when they argue that, in ‘A man is a donkey’, ‘man’ cannot have supposition because there are no men that are donkeys. But if the concepts of supposition and verification are kept apart, or if the notion of supposition is assimilated to the intensional concept of ‘*denotatur*’, this argument obviously does not hold.

These passages also dissociate the semantic notion of supposition from the issue of existence or non-existence of the *suppositum* in question. It is true that medieval

[54] Et si dicatur: ista non stant simul ‘supponit’ et ‘pro nullo supponit’, quia sequitur ‘supponit, igitur pro aliquo supponit’, dicendum est quod non sequitur, sed sequitur ‘supponit, igitur denotatur pro aliquo supponere, vel denotatur pro nullo supponere’. (Ockham, *Summa Logicae I*, chap. 72, 135–139 (p. 219))

[55] Determinata suppositio dicitur quam habet terminus communis indefinite sumptus vel cum signo particulari, ut ‘homo currit’ vel ‘aliquis homo currit’. Et dicitur utraque istarum determinata, quia, licet in utraque illarum iste terminus ‘homo’ supponat pro omni homine tam currente quam non currente, tamen uno solo homine currente vera sunt. Aliud enim est supponere et aliud est reddere locutionem veram pro aliquot. In predictis enim, ut dictum est, iste terminus ‘homo’ supponit pro omni homine, tam currente quam non currente, sed reddit locutionem veram pro currente. (Peter of Spain 1972, 82 (14–22))

[56] Item, possible est esse suppositionem terminorum in propositione sine verificatione propositionis, tam in affirmatives quam in negativis, ut ‘homo est asinus’ vel ‘homo non est animal’. Item, potest esse in negativis verificatio sine suppositione, ut ‘chimaera non est hircocervus’; sed necesse est in affimativis veris verificationem propositionis esse pro aliquo vel aliquibus pro quo vel quibus termini supponunt. (Buridan 1998, 10)

logic attributes existential import to the subject of all categorical affirmative propositions, including universal propositions; but one must know that existential import occurs with respect to the **truth** of an affirmative proposition, and not with respect to supposition. In sum, the foregoing considerations suggest that supposition is less a theory of a proposition's aboutness than a method to establish the possible readings a proposition may have.

As I see it, the generation of possible readings takes place on a pre-assertoric level: the generated readings are assertion-apt, but not asserted. For the proposition under a certain reading to be asserted, the intentional choice of one's preferred reading is required. So, 'Man is a species' does not assert anything by itself; it is only under one of its possible readings that something is asserted [*denotatur*] by it. That is, if a proposition is ambiguous – prone to receiving more than one reading –, it has a potential for asserting any of its possible readings, but it does not assert any of them by itself. It only comes to assert when an interpreter associates it to a mental proposition,[57] thus choosing among the possible interpretations of the written/spoken proposition.

Hence, by means of the concept of supposition and its machinery, it is possible to retrace the assertions potentially contained in certain propositions.[58] The procedure of generating possible readings for a proposition consists of determining the range of possible *supposita* for its terms, and then combining each possible *suppositum* of the subject to each possible *suppositum* of the predicate. So, if n is the number of possible *supposita* for the subject and m is the number of possible *supposita* for the predicate, then the proposition in question will have $n \times m$ possible readings. The choice of one of these readings as the favored one (and, accordingly, the preferred kind of supposition for each term) occurs out of the scope of supposition theory, and takes into account matters such as truth, faithfulness to the author's intention, theoretical background etc.

These considerations lead us to the problem of actual/potential supposition, which became acute only in later stages of the development of supposition theory. While 13th and 14th century authors did not discuss this topic explicitly, 15th century authors were well aware of its relevance. E.J. Ashworth mentions three main views on the issue, upheld by different authors:

> [...] does the acceptance of the term have to be actual, or merely potential? At the end of the fifteenth century Johannes Magister outlined three possibilities. First is the belief of Dorp and some of the nominalists that a term in a closed book can have supposition if somebody would take it to refer if they were to take notice of it. Tartaretus and Eckius both echoed this view. Second is the belief of Marsilius and most of

---

[57] Presumably, mental propositions should never be equivocal. This is a delicate point in Ockham's system: if there are different kinds of supposition in mental language, then there are mental propositions that are ambiguous, which is an undesired result cf. (Spade 1980; Normore 1997; Dutilh Novaes 2004b). Buridan, on the other hand, avoids this problem by contending that there is only personal supposition in mental language – what corroborates the view that the machinery of supposition really is meant for the interpretation of written and spoken propositions.

[58] Admittedly, supposition theory cannot handle every proposition that is ambiguous, but rather only those whose ambiguity is related to the possibility of their terms standing for different things.

the nominalists that a term can only supposit in relation to a mental term, so that if 'God exists' is written in a closed book, it is not a proposition, and the word 'God' does not have supposition. This view was adopted by Hieronymus of St. Mark, who argued that a term had to be accepted by somebody as standing for something before it could be said to have supposition. The last view is that of realists, who drew a distinction between potential and actual supposition, and thus were able to combine the two other views. (Ashworth 1974, 78)

Even though earlier authors such as Ockham, Buridan and Peter of Spain did not explicitly discuss the problem of actual and potential supposition, the intensional interpretation of the concept of supposition (based on the notion of '*denotatur*' or on the distinction between supposition and verification) suggests that a term in a proposition does not have supposition unless it is assigned by the interpreter to stand for some of its possible *supposita* – indeed, as discussed in section 1.2.2, the actual supposition of a term is essentially determined by its user. Moreover, it is clear that the strictly extensional interpretation of supposition theory ('a term supposits, thus it supposits for something') leads to inconsistency, as shown in (Karger 1978) and by the cases of false affirmative propositions; to my mind, these purported 'inconsistencies' substantiate more than anything else the claim that the correct interpretation for supposition theory is the second, intensional interpretation.

1.3.2.2 Propositio est distinguenda

Is it possible that, with one and the same proposition (or tokens of the same type), different statements can be made? According to Ockham, and to many 14th century logicians, it is not only possible but also very frequent.[59] Indeed, the debate around such propositions – which in the medieval jargon are referred to as 'propositions that must be distinguished' (*propositio est distinguenda*) – was a lively one in that century; at that time the mainstream position was favorable to the 'distinction' of propositions (the 1340 statute of the University of Paris forbids the view that *nulla propositio est distinguenda* – cf. (Ashworth 1991), (Van der Lecq and Braakhuis 1994). Another term for '*propositio est distinguenda*' is '*propositio est multiplex*' (used by Burley too).[60]

To distinguish a proposition is to make its possible readings explicit (but not necessarily to choose among these possibilities). Now, the apparatus of supposition theory can generate the possible, legitimate readings of certain propositions (namely those whose ambiguity is related to the possibility of their terms standing for different things), on the basis of elements such as the propositional context and the signification of their terms; in effect, much of the discussions concerning the distinction of propositions is to be found in treatises on supposition.

This issue is not only historically but also conceptually very important. While 13th century authors usually maintained that the supposition of the subject term is defined

[59] When explaining the distinction of propositions, Ockham adds: 'Et sic de multis talibus'. (Ockham, *Summa Logicae I*, chap. 65, 27 (p. 198))

[60] Et non est idem dicere 'animal praedicatur de homine' et 'homo est animal', quia una est multiplex et alia non. (Ockham, *Summa Logicae I*, chap. 65, 13–21 (p. 198))

by the predicate (see below), early 14th century authors agreed almost unanimously that numerous propositions had to be distinguished, in particular those where a given term could have more than one kind of supposition; each kind of supposition of the subject would yield a different reading for the proposition. However, this view was later explicitly rejected by authors such as Paul of Venice (see below) and Albert of Saxony (cf. Ashworth 1991).

In particular the 14th century trio – Burley, Ockham and Buridan – very much insisted on the existence of alternative, equally legitimate readings (possibly with different truth values) for certain propositions, especially in connection with supposition theory.

> 'Species' signifies an intention of the soul; therefore, in the proposition 'Man is a species' the term 'man' can have simple supposition. With this proposition we must, in accordance with the third mode of equivocation, distinguish between the case where we have simple supposition and the case where we have personal supposition. In the former case we have a true proposition asserting that a concept or intention of the soul is a species, and this is true: in the latter case the proposition is simply false, because it asserts that something signified by 'man' is a species, which is clearly false.[61] (Ockham 1998, 192) (with modifications)

Of '*Homo est species*', two readings are possible. According to the first reading it is asserted [*denotatur*] that a man is a species, in which case '*homo*' has personal supposition: with the second reading it is asserted that a concept is a species, in which case '*homo*' has simple supposition. Under the first reading, the proposition is false; under the second, it is true. However, the first one is none the less legitimate than the second one (in spite of its truth-value). This point is of utmost importance, it concerns the separation of the establishment of legitimate readings of propositions – what can be asserted by means of them – from the issue of truth. We shall see later that the application of this theoretical apparatus to the analysis of a proposition consists of first establishing its possible readings, and only then does the question of truth come into play. Note also that the rules of supposition establish that only these two readings are possible, and exclude the possibility of '*homo*' having material supposition, for example.

Note again the presence of SE vocabulary: Ockham says that the ambiguity of this proposition follows from the third mode of equivocation.[62] Interestingly, Burley will say that the very same proposition is equivocal according to the second mode of equivocation, indicating that medieval authors did not always agree on their interpretation

---

[61] Sed in ista propositione 'homo est species' quia 'species' significat intentionem animae ideo potest habere suppositionem simplicem. Et est propositio distinguenda penes tertium modum aequivocationis, eo quod subiectum potest habere suppositionem simplicem vel personalem. Primo modo est propositio vera, quia tunc denotatur quod una intentio animae sive conceptus sit species, et hoc est verum. Secundo modo est propositio simpliciter falsa, quia tunc denotatur quod aliqua res significata per hominem sit species, quod est manifeste falsum. (Ockham, *Summa Logicae I*, chap. 65, 13–21 (p. 198))

[62] An interesting discussion of Ockham's distinction of propositions is to be found in (Priest and Read 1981).

of the fallacy framework.[63] Here is the passage where Burley tackles this point: his view is (perhaps surprisingly, given their philosophical divergences) very similar to Ockham's.

And just as an analogical term matched with something participating in it according to the secondary significate is equivocal insofar as it can be taken for its primary or secondary significate, so a term that can have different kinds of supposition and is matched with something participating in it according to its secondary supposition is ambiguous insofar as it can have one kind of supposition or the other, namely, the primary or the secondary.

Thus 'A man runs' is not ambiguous, and neither is 'A man is an animal', because in these propositions the subject supposits personally [only]. But propositions like 'Man is a species' and 'Man is a monosyllable' are ambiguous according to the second mode of equivocation insofar as the term 'man' can have personal or simple or material supposition. For the proposition 'Man is a species' is ambiguous insofar as the term 'man' can have personal or simple supposition. 'Man is a monosyllable' has to be distinguished too according to the second mode of equivocation insofar as the term 'man' can have personal or material supposition.[64] (Walter Burley 2000, 91)

In both passages, it is patent that the ambiguity of propositions such as 'Man is a species' and 'Man is a monosyllabe' lies in the flexibility with which the term 'man' can be interpreted – that is, it can have either personal or simple or material supposition. Each reading corresponds to the assignment of one of these kinds of supposition to the term in question.

Especially during dialectical disputations, such as *obligationes*, it was crucial for the respondent to be able to identify propositions prone to more than one interpretation, otherwise his accepting or denying it might be taken with respect to a reading other than the one intended by him (which might result in contradiction and in his defeat). For such cases, Burley gives the following practical advice:

During the time of the obligation one should not give a definite answer to a question requiring a distinction.[65] (Burley 1988, 385)

---

[63] For more on the relations between the fallacy framework and the supposition apparatus, cf. (Dutilh Novaes forthcoming).

[64] Et sicut terminus analogus comparatus alicui participanti sibi pro significato secundario est aequivocus penes secundum modum aequivocationis, ex eo quod potest accipi pro significtao primario vel secondario, ita terminus potens habere diversas suppositions comparatus alicui participanti sibi secundum suppositionem secundariam est multiplex, ex eo quod potest habere unam suppositionem vel aliam, videlicet primariam vel secundariam.

Unde haec non est multiplex: 'Homo currit', vel: 'homo est animal', quia in istis subiectum supponit personaliter. Sed tamen huiusmodi propositiones: 'Homo est species', 'Homo est disyllabum', sunt multiplices penes secundum modum aequivocationis, ex eo quod ille terminus 'homo'potest habere suppositionem personalem vel simplicem vel materialem. Ista enim propositio: 'Homo est species' est multiplex, ex eo quod iste terminus 'homo' potest habere suppositionem personalem vel simplicem. Et ista 'Homo est dissyllabum', est distinguenda penes secundum modum aequivocationis, ex eo quo ille terminus 'homo' potest habere suppositionem personalem vel materialem. (Burley 1955, 10 (12–34))

[65] It is controversial, though, whether the meaning intended by Burley really is the one attributed to him by the translators. The Green edition of the Latin text has the following: 'Alia regula est ista: durante tempore obligationis, non est certificanda quaestio <disciplinaris>; quod patet sic. Sit in rei veritate solum Socratem loqui, et ponatur Socratem tacere. Et proponatur: aliquis homo loquitur. Haec est vera et impertinens, igitur concedenda. Et si quaeratur "quis est iste homo", si tu respondeas, oportet concedere oppositum positi vel

Indeed, in Ockham's version of obligational disputations, '*distinguo*' is one of the replies available to respondent, along with '*concedo*', '*nego*' and '*dubito*'.[66] '*Distinguo*' is namely the correct reply to propositions proposed by the opponent that are ambiguous and therefore must be distinguished. Moreover, propositions that must be distinguished play a prominent role not only in Ockham's theory of supposition and in his rules for obligational disputations, but also in his theory of fallacies (*Summa Logicae* III-4). According to him, the key to many kinds of fallacies (in particular equivocation, amphiboly, division and composition) is that one or more propositions in a given fallacious argument are prone to more than one reading (cf. Dutilh Novaes forthcoming), that is, they must be distinguished, and the failure to do so yields a fallacious argument.

The need to 'distinguish propositions' also arises in connection with the interpretation of authoritative statements, as shown in the following passage by Buridan[67] (already quoted):

> So if we want, by agreement, to call such an appropriateness of a locution its 'proper meaning' [*virtutem sermonis*], then I would say that 'Man is a species' is true insofar as it is put forth by authors, but it is not true according to its proper meaning, that is, it would not be true if it were put forward as a proper locution; this is because utterances primarily and principally were imposed to signify so as to stand for their ultimate significata, and not for themselves, as was noted before. So it absolutely seems to me that wherever it is evident that an author puts forward a proposition in a true sense, although not as a proper locution, then to deny that proposition without qualification would be cantankerous and insolent. But to avoid error, it should be properly pointed out that the proposition is not true in the proper sense, or by virtue of its proper meaning, and then it has to be shown in which sense it is true.[68] (Buridan 2001, 256)

From Buridan's passage, it is manifest that there exists a close relation between the need to distinguish propositions and the traditions of commentary on the authorities and of disputations. In particular, when it comes to interpreting and commenting, the issue of truth is not primary; the first step is to extract every possible assertion that can be made by means of the proposition in question. Being able to do so is a skill

---

falsum et impertinens. Quia, sit tu dicas: "Socrates loquitur, concedes repugnans posito". Et sit tu dicas: "Plato vel Cicero (vel sic de aliis) loquitur", concedes falsum et impertinens'. (Green 1963, 52 (18–27))

[66] Et consistit ars ista in hoc quod in principio debet aliqua propositio poni, deinde debent respondens respondere concedendo vel negando vel dubitando vel distinguendo. (Ockham, *Summa Logicae I*, chap. 41, 28–32 (p. 736)). But notice that doubts have been voiced concerning the authorship of the chapter on obligations attributed to Ockham (included in his *Summa Logicae*), cf. (Weber 2003, p. 150, footnote 64).

[67] But notice that Buridan never states the view that propositions must be distinguished as clearly as Ockham and Burley do. In fact, at an earlier stage of his career, he seems to have been opposed to this view. (Cf. Van der Lecq and Braakhuis 1994)

[68] Si ergo huiusmodi proprietatem locutionis volumus ad placitum vocare 'virtutem sermonis', tunc diceremus quod ista 'homo est species' est vera sicut ponitur ab auctoribus, sed non est vera de virtute sermonis, idest non esset vera si poneretur secundum propriam locutionem, quia primo et principaliter voces fuerunt impositae ad significatum ut starent pro suis significatis et non pro se, ut dicebatur prius. Videtur ergo mihi quod, ubi apparet auctorem posuisse aliquam propositionem ad aliquem sensum verum, licet non secundum propriam locutionem, negare simpliciter propositionem esset dyscolum et protervum. Sed ne erretur debet bene dici quod propositio non est vera secundum proprium sensum sive de proprietate sermonis, et debet ostendi secundum quem sensum est vera. (Buridan 1998, 42)

in its own right; in first instance, one's theoretical preferences should not interfere with the interpretational process (although, of course, the interpretational apparatus is never entirely theoretically neutral). Similarly, in disputations such as *obligationes*, it is not the truth-value of propositions that is at stake, but rather their logical relations with other propositions.

In order to distinguish propositions, Ockham relies essentially on two different conceptual frameworks, namely these of supposition theory and of fallacy theory. The supposition apparatus is mostly used to distinguish propositions whose ambiguity is related to the possibility of their terms standing for different things (what is accounted for by the distinction between personal, simple and material supposition), but it can also handle modal propositions that must be distinguished[69] (i.e., modal propositions *cum dicto* – cf. *Summa Logicae* II, chap. 9) and propositions with a tensed copula (cf. *Summa Logicae* II, chap. 7). Obviously, there are many more propositions that must be distinguished, besides these; thus, presumably, there are many propositions that must be distinguished which cannot be handled by supposition theory (e.g., Ockham uses the second mode of amphiboly to handle propositions that must be distinguished in virtue of a possible metaphorical reading – cf. *Summa Logicae* III-4, chap. 6). Nevertheless, the fact remains that one of the main tools used by Ockham to distinguish propositions is his supposition theory, and therefore a proper understanding of supposition theory must take into account its application to distinguish propositions, that is, to generate their possible readings.

But, to be fair, it must be said that not all authors dissociated the generation of possible readings for a proposition from the issue of truth, which might mean that the claim that supposition theories are to a great extent theories for the interpretation of propositions (both in oral and written contexts) does not hold irrespectively. To the best of my knowledge, earlier authors such as Peter of Spain and William of Sherwood do not mention the need to 'distinguish' propositions: their focus is more often than not syntactical, as they emphasize the role of syncategorematic terms and other sentential constraints inducing a certain kind of supposition, in detriment of others, for a given term. Thus, these authors seem to hold that the kind of supposition that a term has in a proposition is completely determined by elements of the latter. In fact, William of Sherwood acknowledges that the diversity of kinds of supposition may give rise to equivocation,[70] but a few pages further he presents the following general rule:

> The subject, on the other hand [in opposition to the predicate], sometimes supposits absolutely for its form, sometimes not, and this according to the demands of the predicate according to this [rule]: The subjects are such as the predicates have allowed.[71] (Sherwood 1966, 113) (with modifications)

In other words, if the predicate is a term of second intention or of second imposition – that is, a term whose ultimate *significata* are concepts, universals or words –, then

---

[69] Cf. (Dutilh Novaes 2004a).

[70] (Sherwood 1995, 140) (109–113).

[71] Subiectum autem quandoque supponit formam absolute, quandoque autem non, et hoc secundum exigentiam praedicati secundum illud: Talia sunt subiecta, qualia permiserint praedicata. (Sherwood 1995, 144) (167–170)

it will impose on the subject a supposition other than personal supposition. Hence, while Ockham and Burley uphold that in such cases the subject can have either simple/material supposition or personal supposition, it seems that, for Sherwood, it can no longer have personal supposition, since the predicate term determines the kind of supposition the subject must have – presumably, the kind of supposition in question is the one that yields a true reading of the proposition. Apparently, for William of Sherwood and Peter of Spain, there was no need to 'distinguish' a proposition because syncategoremata and other aspects would induce only one kind of supposition for each term of a proposition, leaving no scope for ambiguity.

Interestingly, after the predominance of the Burley-Ockham-Buridan position in the first half of the 14th century, by the end of the 14th century Paul of Venice felt compelled to return to the earlier position according to which either the propositional context or the (desired) truth of a proposition determine the supposition of its terms.

> If someone says 'proposition is a name', the subject has either significate and the predicate only a formal significate: however, the subject stands materially because otherwise the proposition would not be true. Further, if it is said 'proposition is an incomplex term', both extremes have both significates: however, the subject stands materially, otherwise the proposition is false.[72] (Paul of Venice 1984, 146)

Problems with this account arise when a proposition yields two readings (due to different suppositions of a term), both of which are true. Consider the proposition: 'Noun has four letters'. The two possible readings are that some noun has four letters (true) and that the very noun 'noun' has four letters (also true). One wonders which reading Paul of Venice would accept as the right one, and based on which criteria, since he cannot resort to truth.

In view of this problem, the Burley-Ockham-Buridan position, according to which the kinds of supposition a term may have in a proposition are defined by certain rules, but are not related to truth-values, seems more accurate and felicitous. In effect, the task of generating the possible readings of a proposition is fully compatible with their theories of supposition, besides being motivated by the academic practices of the time.

In sum, it is to be expected that the kind of supposition that a term may have should be determined by formal features of the proposition in question, and not by its truth-value. If supposition is determined by truth, then there shall be no false propositions, or better said, no false statements, since one can always 'jiggle' around with the supposition of the terms.

Hence, even if the view that supposition theories are intended to generate the possible readings of a significant corpus of propositions is not an accurate description of some specific versions of these theories, this description does seem to apply to

---

[72] Sed si dicitur 'Propositio est nomen' subiectum habet utrumque significatum et praedicatum solum formale, et tamen subiectum supponit materialiter quia aliter non esset proposition vera. Item si dicitur 'propositio est terminus incomplexus' utrumque extremum habet ambo significata; et tamen subiectum supponit materialiter quia aliter proposition [35] esset falsa. (Paul of Venice 2002, 27)

the most interesting versions of supposition theories. If these theories were not all according to this description, they should have been.[73]

1.4 THE STRUCTURE OF OCKHAM'S THEORY OF SUPPOSITION

Finally, a closer exam of Ockham's theory offers further confirmation to the claims made so far. Admittedly, his theory is the one that fits best the foregoing general description, and it can be argued that other theories of supposition are not as clearly theories for the generation of a proposition's possible readings. Still, it is symptomatic that one of the most influential and well-known theories of supposition, now as well as then, seems to conform remarkably well to this characterization.

The concept of supposition is crucial in many aspects of Ockham's logical system, so it is important to define what is meant here by 'Ockham's supposition theory': it is the system put forward in the last chapters (63–77) of the first part of his *Summa Logicae*.[74] It has been argued that these chapters present not one, but two theories of supposition:[75] the theory of supposition 'proper' and the theory of the 'modes of personal supposition', the latter being structured in terms of ascent and descent. However, if my views are correct, there is no such division, as I shall argue.[76]

What follows is a systematic description of Ockham's theory, based on the relevant texts, which will serve as the foundation for the formalization presented in the next section.

There are two kinds of rules in Ockham's theory: the (quasi-syntactic) rules that define in which contexts a term can have such-and-such supposition, and the (semantic) rules that define what it means, with respect to the *supposita* (on the level of entities), for a term to have such-and-such supposition. The successive application of these two groups of rules yields the possible *supposita* of a term, possibly more than one alternative. When the latter are combined to the outcome of the same procedure

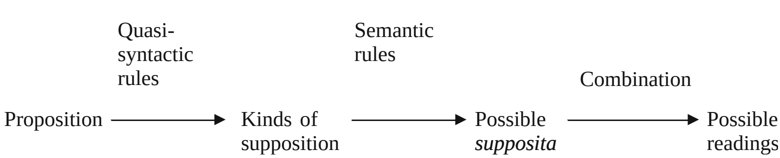


Figure 1.4. Structure of Ockham's theory of supposition

[73] But note that Ockham's supposition theory is certainly not free of problems. One of its problems is the existence of mental propositions that must be distinguished; in particular, if distinguishing written and spoken propositions amounts to mapping them into more than one mental proposition, then mental language must be free of ambiguities. However, Ockham claims that there are different kinds of supposition (personal, simple and material) also in mental language. (cf. Dutilh Novaes 2004b)

[74] Thus, it is **not** his theory of truth conditions of the first chapters of the second part of the *Summa Logicae*.

[75] Cf. (Spade 1988).

[76] The view that these are not two separate theories is also defended in (Matthews 1997).

applied to the other term of a categorical proposition, the possible readings of the proposition are generated.[77] This process is depicted in Figure 1.4.

1.4.1 Quasi-syntactic rules

A few lines above, I attributed a 'quasi-syntactic' character to the rules defining the kind(s) of supposition that a term has in a given context. The qualification 'quasi-' is necessary, because these rules are not syntactic in the case of the distinction between personal, simple and material supposition, as will be shown in what follows. Rather, in the latter case what determines the kind of supposition of a term is the semantic type of the other term in the proposition. This is not a syntactic feature of the proposition, but it is a **formal** feature thereof (in some senses analogous to the basic idea of sorted logics). By contrast, the rules for the modes of personal supposition are syntactic in a very straightforward way.

1.4.1.1 Personal, simple and material supposition

The rules for the division between personal, simple and material supposition are in fact quite simple; any term can always have personal supposition, and a term can have simple or material supposition only in some contexts, namely if the other extreme is a term of second imposition and/or of second intention.

> [...] a term, in any proposition in which it is placed, can always have personal supposition [...].[78] (Ockham 1998, 191) (with modifications)

> But a term cannot have simple or material supposition in any proposition, but only in propositions where it is compared with an extreme that relates to an intention of the soul or a spoken or written expression.[79] (Ockham 1998, 192) (with modifications)

1.4.1.2 Modes of personal supposition

As for the 'modes of personal supposition', what defines which kind of supposition a term may have is the presence or absence of syncategorematic terms such as 'some' [*aliquid*] or 'every' [*omnis*], and the position of each term in the proposition. As to be expected, these rules are much more complex than the ones for personal, simple and material supposition, because they must cover as many different logical forms for propositions as possible. Still, it is impossible to cover every possible case in a non-axiomatic language such as medieval Latin. In truth, the set of such rules is more

[77] For the sake of simplicity, I am not considering the case of modal and temporal propositions, and focusing only on propositions whose copula is present-tensed.

[78] Semper terminus, in qualcumque propositione ponatur, potest habere suppositionem personalem. (Ockham, *Summa Logicae I*, chap. 65, 3–4 (p. 197))

[79] Sed terminus non in omni propositione potest habere suppositionem simplicem vel materialem, sed tunc tantum, quando terminus talis comparatur alteri extremo, quod respicit intentionem animae vel vocem vel scriptum. (Ockham, *Summa Logicae I*, chap. 65, 7–10 (p. 197))

like an open system, to which new rules can be added in order to cover more and more cases.

In chapters 71, 73 and 74 of the first part of the *Summa Logicae*, Ockham discusses some of the logical forms determining which kind of personal supposition the terms of a proposition have.

> First it should be noted that when in a categorical proposition no universal sign distributing the whole extreme of a proposition is added to a term, either mediately or immediately (i.e., either on the part of the same extreme or on the part of the preceding extreme), and when no negation or any expression equivalent to a negative or a universal sign is added to a common term, that common term always supposits determinately. […] The same should be said in the case of 'some man runs'; for whether the sign of particularity is added or not does not alter the personal supposition of the term, such that it is frequent for a term to have personal supposition.[80] (William of Ockham 1998, 202) (with modifications)

The formulation of this rule may seem a bit convoluted at first sight, due to the ample use of technical terms, but the idea behind this rule is fairly simple: the default supposition of a term – if no universal or negative sign is added – is determinate supposition.

> First, where a common term mediately follows an affirmative sign of universality, it has merely confused supposition. That is, in an affirmative universal proposition the predicate always has merely confused supposition.[81] (William of Ockham 1998, 211) (with modifications)

> The first rule is that in every universal affirmative and universal negative proposition that is neither exclusive not exceptive, the subject has confused and distributive mobile supposition.[82] (William of Ockham 1998, 214)

> The second rule is that in every such universal negative proposition the predicate stands confusedly and distributively.[83] (William of Ockham 1998, 214)

> The third rule is that when a negation determining the principal composition precedes the predicate, the predicate stands confusedly and distributively. Thus the word 'animal' in 'Man is not an animal' stands confusedly and distributively. 'Man' however stands determinately.[84] (William of Ockham 1998, 214)

---

[80] Est ergo primo sciendum quod quando in categorica nullum signum universale distribuens totum extremum propositionis additur termino, nec mediate nec immediate, hoc est nec a parte euisdem extremi nec a parte extremi praecedentis, nec negatio praecedit nec aliqua dictio includens aequivalenter negationem vel signum universale, semper talis terminus communis supponit determinate. […] Idem est dicendum de ista 'aliquis homo currit', quia signum particulare additum vel non additum non variat suppositionem personalem, quamvis faciat frequenter terminum stare personaliter. (Ockham, *Summa Logicae I*, chap. 73, 7–18 (p. 212))

[81] Una est quod quando terminus communis sequitur signum universale affirmativum mediate, tunc stat confuse tantum, hoc est semper in universali affirmativa praedicatum supponit confuse tantum. (Ockham, *Summa Logicae I*, chap. 73, 5–8 (p. 226))

[82] Una est quod in omni propositione universali affirmativa et negativa, quae non est exclusiva nec exceptive, stat subiectum confuse et distributive mobiliter. (Ockham, *Summa Logicae I*, chap. 74, 6–8 (p. 228))

[83] Secunda regula: quod in omni tali universali negativa praedicatum stat confuse et distributive. (Ockham, *Summa Logicae I*, chap. 74, 10–11 (p. 229))

[84] Tertia regula est quod quando negatio determinans compositionem principalem praecedit, praedicatum stat confuse et distributive, sicut in ista 'homo non est animal' li animal stat confuse et distributive, sed 'homo' stat determinate. (Ockham, *Summa Logicae I*, chap. 74, 12–15 (p. 229))

A general rule is that if anything makes a term stand confusedly and distributively, it is either a sign of universality, a negation or an expression equivalent to a negation.[85] (William of Ockham 1998, 214)

After the application of these rules, the interpreter is able to identify (with the exception of 'unusual' logical forms) which kind(s) of supposition a term in a given proposition may have – whether personal, simple or material supposition, and in the case of personal supposition, which kind of personal supposition.

1.4.2 Semantic rules

The next step is to establish what it means in terms of *supposita* to have such-and-such supposition. This is determined by the semantic rules of supposition.

1.4.2.1 Personal, simple and material supposition

Personal, simple and material suppositions are defined by the kind of the *suppositum* and by whether the term is taken significatively or not.[86]

Thus, whenever the subject or predicate of a proposition supposits for its significate so that it is taken significatively, we always have personal supposition.[87] (Ockham 1998, 190)

Clearly, the notion of *significatum* of a term is the key element of this definition. As previously mentioned, for Ockham, signification is a relation between a (categorematic) term and each thing of which it can be predicated. 'Man', for example, signifies each and every man indistinctively. So, personal supposition is the kind of supposition that a term has iff it supposits for one of its *significata* (taken significatively). The other kinds of supposition are defined in a similar fashion:

Simple supposition occurs when a term supposits for an intention of the soul and is not functioning significatively.[88] (Ockham 1998, 190)

Material supposition occurs when a term does not supposit significatively, but supposits for a spoken or written expression.[89] (Ockham 1998, 190) (with modifications)

Simple supposition concerns the supposition of a term for a **concept** (not taken significatively) and material supposition, the supposition of a term for a written/spoken

[85] Hoc igitur universaliter est dicendum quod quidquid facit terminum stare confuse et distributive vel est signum universale vel negatio vel aliquid aequivalens negationi. (Ockham, *Summa Logicae I*, chap. 74, 45–47 (p. 230))

[86] The clause 'taken significatively' is rather delicate. Briefly put, if a term such as 'noun' supposits for the term 'noun' insofar as 'noun' **is** a noun, then the term is taken significatively. But if 'noun' supposits for 'noun' because they are equiform strings, then it is not taken significatively.

[87] … quandocumque subiectum vel praedicatum propositionis supponit pro suo significato, ita quod significative tenetur, semper est suppositio personalis. (Ockham, *Summa Logicae I*, chap. 64, 7–9 (p. 195)).

[88] Suppositio simplex est, quando terminus supponit pro intentione animae, sed non tenetur significative. (Ockham, *Summa Logicae I*, chap. 64, 26–27 (p. 196))

[89] Suppositio materialis est, quando terminus non supponit significative, sed supponit pro voce vel pro scripto. (Ockham, *Summa Logicae I*, chap. 64, 38–39 (p. 196))

expression[90] (not taken significatively). In sum, if it has been established that, in a given proposition, a term can have such-and-such supposition, by the application of the definitions of supposition, the interpreter determines which kinds of things can be a *suppositum* for the term.

1.4.2.2 Modes of personal supposition

As for the subdivisions of personal supposition, in Ockham's theory, what it means for a term to have determinate supposition, confused and distributive supposition or merely confused supposition is established in terms of ascent-and-descent (inferential) relations between the propositions in question and singular propositions of the form 'This *a* is *b*'.[91]

> Determinate supposition occurs when one can descend to singular [propositions] by some disjunctive proposition, such that this is a valid inference: 'a man is running, therefore this man is running, or that man etc.' The name 'determinate supposition' is employed because by such supposition, it is asserted that the proposition is true for some singular [proposition].[92] (Ockham 1998, 200) (with modifications)

> Merely confused personal supposition occurs when a common term supposits personally and one cannot descend to singular propositions by means of a disjunctive proposition with no change made on the part of the other extreme, but by means of a proposition having a disjunctive predicate, and one can infer [the proposition] from any of the singulars. For example, in the proposition 'Every man is an animal', the word 'animal' has merely confused supposition; for one cannot descend to particular under 'animal' to its contents by means of a disjunctive proposition.[93] (Ockham 1998, 201) (with modifications)

> Confused and distributive supposition occurs when one can descend in some way copulatively, if the term has many contents, and from no one of them is [the original proposition] formally inferred. For example, in 'Every man is an animal', the subject supposits confusedly and distributively. For it follows: 'Every man is an animal; therefore, *this* man is an animal and *that* man is an animal' and so on.[94] (Ockham 1998, 201)

[90] To say that, in material supposition, a term supposits for a word would be too narrow, as material supposition can also occur with respect to phrases, sentences or other complex expressions.

[91] In section 1.5.2.1 I discuss briefly different views on the purpose of these relations of ascent and descent, which have been proposed in the secondary literature.

[92] Suppositio determinata est quando contingit descendere per aliquam disiunctivam ad singularia; sicut bene sequitur 'homo currit, igitur iste homo currit, vel ille', et sic de singulis. Et ideo dicitur suppositio determinata quia per talem suppositionem denotatur quod talis propositio sit vera pro aliqua singulari determinata. (Ockham, *Summa Logicae I*, chap. 70, 19–27 (p. 210))

[93] Suppositio personalis confusa tantum est quando terminus communis supponit personaliter et non contingit descendere ad singularia per disiunctivam, nulla variatione facta a parte alterius extremi, sed per propositionem de disiuncto predicato, et contingit eam inferri ex quocumque singulari. Verbi gratia in ista 'omnis homo est animal', li animal supponit confuse tantum, quia non contingit descendere sub animali ad sua contenta per disiunctivam. (Ockham, *Summa Logicae I*, chap. 70, 44–50 (p. 211))

[94] Suppositio confusa et distributiva est quando contingit aliquo modo descendere copulative, si habeat multa contenta et ex nullo uno formaliter infertur. Sicut est in ista 'omnis homo est animal', cuius subiectum supponit confuse et distributive: sequitur enim 'omnis homo est animal, igitur iste homo est animal et ille homo est animal', et sic de singulis […]. (Ockham, *Summa Logicae I*, chap. 70, 62–68 (p. 211)). The translation as quoted here is by P.V. Spade, available at www.pvspade.com/Logic/docs/ockham.pdf

1.4.3 Combination

Finally, when the (alternative) possible *supposita* for each term of a proposition have been established, each alternative for one of the terms is combined to each alternative for the other term, and thereby the possible readings of the proposition in question are generated – that is, the proposition is 'distinguished'. In sum, the so-called 'theory of supposition proper' and 'theory of modes of personal supposition' seem to be different steps of application of the same machinery, which is applied to obtain the complete semantic analysis of a proposition.

To illustrate the procedure, let us take Ockham's recurrent example 'A man is a species' [*Homo est species*]. In the passage quoted above, he says that this proposition 'is to be distinguished'. Let us apply the machinery here described to this proposition and see whether it yields the desired results.

We begin by analyzing the term 'man': it can have personal supposition based on the rule that all terms can have personal supposition, but it can also have simple supposition, since the other term, 'species', signifies an intention of the soul. If it has personal supposition, then it has determinate personal supposition, for there are no syncategorematic signs of confusion. If it has simple supposition, then the term 'man' supposits for the concept <man>, according to the definition of simple supposition. If it has determinate personal supposition, then it supposits for a man.

As for the term 'species', it can only have personal supposition, since the other term 'man' does not signify intentions of the soul or written/spoken terms. According to the syntactic rules, it would have determinate supposition. If 'species' has personal determinate supposition, it supposits for its *significata* – that is, for the several mental terms that are species, and in particular for the mental term <man>.[95]

Finally, the combination of each alternative two by two yields the possible readings of this proposition.

Reading 1: A man is a species (a concept).
Reading 2: The concept *man* is a species.

Clearly, it is only under reading 2 that this proposition is true, but reading 1 is equally legitimate. More importantly, these are precisely the two readings recognized by Ockham for the original proposition.

1.4.4 Conclusion

In sum, Ockham's theory of supposition can be seen as a piece of machinery which, when given propositions as input, outputs their possible readings, as depicted in Figure 1.4.4.

The contrast between theories of supposition, thus described, and theories of reference is even more patent if one compares this picture to the picture representing a

[95] Recall Ockham's nominalistic claim that species, as much as other universals, have no extra-mental existence, they are only mental terms.

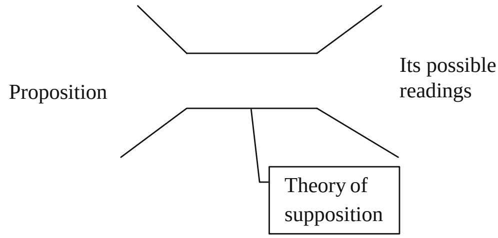


Figure 1.4.4. Input-output of Ockham's theory of supposition

theory of reference in section 1.1.3.3. The two 'pieces of machinery' receive similar input, but their respective outputs are quite different. The output of a theory of reference is the referent of the input-term – typically 'physical' things –, whereas the output of a theory of supposition (thus viewed) are the possible readings of the input-proposition; it is in this sense that a theory of reference is said to be 'extensional', while a theory of supposition, thus reconstructed, would be 'intensional'.

Perhaps the most important contribution of the present analysis is to show that challenging the well-entrenched view that theories of supposition are theories of reference sheds new light on the medieval theories, in a way that seems to clarify some of their obscure aspects, such as those raised in (Spade 2000). I suggest that supposition theories are best seen as formal theories of semantic analysis – of algorithmic hermeneutics –, and that is true particularly of William of Ockham's. But once the dogma 'supposition equals reference' is questioned, it is to be hoped that other, perhaps better theories about supposition theories may be proposed.

1.5 FORMALIZATION

The arguments hitherto presented were intended to substantiate the claim that theories of supposition, in particular Ockham's, are theories for the interpretation of propositions, that is, (loosely put) hermeneutic theories. In this section I intend to show that Ockham's theory is, besides being hermeneutic, also essentially formal, insofar as it provides a procedure for the analysis of propositions that could be mechanically applied by anybody (or even by a machine), always yielding the same results.[96] In this sense, this theory contrasts with other hermeneutical methods, which typically rely on the subjective skills of the interpreter. Moreover, Ockham's supposition theory abstracts from the specific content of each particular proposition (since the meaning

[96] In part 4 of this text I define six senses of formality, which seem relevant for the present investigation. As will become clear in the following pages, supposition theory can be said to be formal mainly according to the notions of formality as abstraction from content, of algorithmic formality and of substitutional formality (as the use of schematic letters in the formalization below shows).

of propositions is precisely what is to be established), as it is based on a finite, small number of propositional schemata; what determines to which schema a given proposition belongs – and thus the reading(s) it allows for – are formal features such as the semantic kind of its terms and the presence/absence of syncategorematic terms, and not the particular meanings of its terms.[97]

In order to spell out the formal, algorithmic character of Ockham's theory, I present here a formalization of it. It differs radically from previous formalizations of supposition theory in the literature;[98] the latter were basically 'translations' of supposition theory into first-order predicate logic (even though it was also sometimes stressed that these two theories were not equivalent and that adjustments had to be made to the first-order language being used – in particular with respect to merely confused supposition and disjunctive predicates). By contrast, the formalization presented here is essentially a project of formal semantics: it presents a **procedure** to generate the possible readings of some propositions, based on the relation between words and what they stand for. In sum, we will be dealing primarily with different ontological levels and with the semantic relations between them (the level of words and the level of things they stand for), and not with the logical relations between different bits of a given language.

It is noteworthy that, while the modes of personal supposition have been the object of several formalization attempts by modern commentators, (for as far as I know) nobody seems to have been aware of, or interested in, the formal character of the fragment of Ockham's supposition theory dealing with the division between personal, simple and material supposition. In this sense (and in a number of other senses), the present formalization presents a new perspective on medieval supposition theories in general and on Ockham's theory in particular.

It must be noted, of course, that any formalization of a (formal but not formalized) theory is nothing but an interpretation of the latter, an approximation, with different degrees of accuracy with respect to the theory being formalized. In the present case, we have the extra risk of anachronism, since the symbolic tools to be used are entirely alien to the medieval framework. Nonetheless, in accordance with the general assumption underlying this investigation – that such formalizations can be exceptionally illuminating –, it is assumed that the risk of anachronism is worth taking.

1.5.1 Personal, simple, and material supposition

The first part of Ockham's supposition theory concerns the division between personal, simple and material supposition; accordingly, the first step for the formalization is to account for this distinction. It will be shown that this portion of the theory is formal not because of an emphasis on the syntactic elements of propositions, but in virtue of the finite and effective character of the procedure presented here – it does not rely

[97] However, as we shall see, the semantic kind of a term is defined by the kind of entities it signifies.
[98] Cf. (Priest and Read 1977), (Karger 1976).

on the specific meanings of terms, but rather on their semantic type, a well-defined property of them.

1.5.1.1 Preliminary notions

The division between object-language and meta-language seems at first sight entirely out of place in a study of medieval logic and semantics. Nevertheless, in the present case, we clearly have a separation between bits of language being analyzed (categorical propositions such as 'Man is a species', 'Man is white' etc.[99]) and bits of language being used in the analysis. In particular, to the latter group belong the key terms describing the supposition conceptual apparatus (personal, simple, material supposition etc.). Occasionally, terms essentially belonging to the meta-Language can occur in propositions being analyzed (e.g., 'Concept is a term of second intention'), but this fact does not affect the basic division between the portions of language being analyzed and the portions of language being used for the analysis.

Object-language. The object language is composed of categorical propositions of the form [subject-copula-predicate]. The copula is present-tensed, non-modal. The subject and predicate are typically categorematic terms, possibly with the addition of quantifying syncategorematic terms, such as 'every' and 'some'.

Notation – terms: $a, b \ldots$
copula: $\circ$

Thus, basic categorical propositions are represented as $a \circ b$, where a and b stand for simple categorematic terms. At a later stage, the object-language will be enriched with other syncategorematic terms: 'quantifying terms' and negating expressions.[100]

Meta-language. It consists of the object language (in the meta-language, terms of the object-language stand for their counterparts in the latter), plus additional expressions for the concepts used in the analysis.

There are two basic sorts of schematic letters, those standing for terms (imported from the object-language) and those standing for things, for elements of the ontology. (Naturally, terms are also elements of the ontology).

Schematic letters for terms:	a, b …
Schematic letters for things:	**t**, **s** …
Abbreviation for propositions:	**P**, **Q** for proposition-types (equiform occurrences); $\mathbf{P_1}$, $\mathbf{P_2}$ …, $\mathbf{Q_1}$, $\mathbf{Q_2}$ … for individual occurrences.

[99] Obviously, in the original texts these are propositions in Latin, but this fact is mostly irrelevant for the present analysis, except for the fact that the examples being used are very often strange, almost ungrammatical English sentences. But for the purpose of presenting the theory, this is of no importance.

[100] For the moment, only indefinite propositions are dealt with, because quantifying terms force personal supposition upon the categorematic terms of the propositions cf. (*Summa Logicae II*, chap. 3, 5–9 (p. 255)). For now we are interested in propositions whose terms may have more than one kind of supposition.

Relational predicates. The two basic notions in Ockham's theory of supposition, those of signification and of supposition, shall be expressed by two-place relational predicates, associating a term to an object.

SIG(a, **t**) <=> The term a signifies object **t**.

SUP(a, **t**)$_{P_1}$ <=> The term a supposits for object **t** in proposition $\mathbf{P_1}$.

A common term usually signifies many different objects (multiple denotation); similarly, a common term typically supposits for many different objects at once, or it may supposit for alternative objects. Therefore, both supposition and signification cannot be represented by functions, uniquely associating a term to one single *suppositum* or *significatum*; they are relations between a term and (typically) several entities. In fact, there are three kinds of suppositional relations SUP(a, **t**)$_{P_1}$, which will be defined below.

Typing devices – Objects. A well-entrenched assumption in Ockham's framework (and in the general medieval framework, for that matter – cf. Dufour 1989, chap. II) is the division of the ontology in three levels: linguistic entities, mental entities, and 'physical' entities. (Here, it is important to notice that not only actually existent entities can be classified according to this taxonomy: past, future and possible entities can also be classified.)

$\mathbf{t}/\varkappa$ <=> Entity **t** belongs to the physical realm (denoted $\varkappa$).

$\mathbf{t}/\lambda$ <=> Entity **t** belongs to the linguistic realm (denoted λ).

$\mathbf{t}/\mu$ <=> Entity **t** belongs to the mental realm (denoted μ).

Typing devices – Terms. In the same way, terms can be classified according to the kind of entities they signify: terms that signify entities of the physical realm are said to be of first intention and imposition; terms that signify entities of the linguistic realm are said to be of second imposition, and terms that signify entities of the mental realm are said to be of second intention. Notice again that the relation of signification does not carry the presupposition of existence: 'chimera' is a term of first imposition and intention, even though there are no actual chimeras; but if there were chimeras, they would be entities belonging to the physical realm.[101]

a: $\varkappa$ <=> $\exists\mathbf{t}$ (SIG(a, **t**) and $\mathbf{t}/\varkappa$).

a: λ <=> $\exists\mathbf{t}$ (SIG(a, **t**) and $\mathbf{t}/\lambda$).

a: μ <=> $\exists\mathbf{t}$ (SIG(a, **t**) and $\mathbf{t}/\mu$).

[101] In *Summa Logicae I*, chap. 33, Ockham says that 'signification' can be understood as concerning only actually existent objects, or as concerning past, present, future and possible objects. Clearly, it is this second sense of 'signification' that must be adopted here.

Terms taken significatively and non-significatively. The definitions of 'taken significatively' and 'taken non-significatively' must rely on the attitude (intention) of the interpreter, and are expressed with a notoriously difficult connective: 'insofar as' or '*qua*'. Moreover, it is grounded on the concept of 'interpretation' of a proposition-token. Notice that these are two distinct concepts, for the very same proposition-token can receive more than one interpretation.[102]

$In<{*}a>_{\mathrm{P}_1} <=>$ the term a is taken significatively in $\mathbf{P_1}$ under interpretation In (where n is a natural number indexing the given interpretation) – that is, it supposits for a object **t** such that SIG(a, **t**), insofar as **t** is a *significatum* of a.

$In<\#a>_{\mathrm{P}_1} <=>$ the term a is taken non-significatively in $\mathbf{P_1}$ under interpretation In (where n is a natural number indexing the given interpretation) – that is, it supposits for an object **t** such that $\sim$SIG(a, **t**), or for an object **t** such that SIG(a, **t**), but not insofar as **t** is a *significatum* of a.

***Denotatur* – sign of assertion.** The importance of the concept of '*denotatur*', which is loosely translated here as 'it is asserted that', has been extensively analyzed in section 1.3.2.1. In effect, for the formalization we also need a symbol representing this concept; it seems suitable to use '$\vdash$', which is the Fregean symbol for assertive force.[103] Clearly, the concept of *denotatur* must be present in the formalization, as it is this concept that guarantees that the definitions of the kinds of supposition do not carry the presupposition of existence of the *suppositum* – this presupposition would be, as already shown, a mistake. By means of the sign of assertion, it becomes evident that, if the presupposition of existence of the *suppositum* fails, it is not the relation of supposition that fails, but rather the truth of the proposition in question, which is thus false if it is affirmative.

In other words, the relation of supposition seems to be what we could call an intensional relation. For instance, one can compare it to the relation of promising: if I promise you a horse, this establishes a relation between me and what I promise (and the person to whom I promised), but what has been promised does not necessarily exist (I can promise a horse even if there are none). Here, we have a similar situation: the relation of supposition holds even if the *suppositum* does not exist, in which case the proposition in question is false, if affirmative, or true, if negative. Now, the intensional approach to the concept of supposition adopted here must have a counterpart also in the formalization: the notion of '*denotatur*', represented by an assertion operator.[104]

---

[102] Usually, the notion of interpretation is treated within the realm of pragmatics, as it involves the intentions and actions of an agent. Still, current developments in formal pragmatics show that formal tools can also be successfully applied to the study of pragmatic phenomena (the seminal text for this trend is Stalnaker 1970, reprinted as chap. 1 of Stalnaker 1999).

[103] The same symbol is also used for the notion of provability/derivability, but this is not the case here.

[104] By contrast, the relation of signification is purely extensional and presupposes the existence of the *significatum*.

Notice that the assertion operator $\vdash$ belongs to the meta-language of the theory, and applies to sentences of the meta-theory – either the typing sentences of the form $\mathbf{t}/\varkappa$, $a{:}\ \varkappa$; or the relational sentences of the form $\mathrm{SIG}(a, \mathbf{t})$ or $\mathrm{SUP}(a, \mathbf{t})_{\mathrm{P}_1}$.

1.5.1.2 Definitions of the three kinds of supposition

If I were to follow the order of application of the procedures in Ockham's supposition theory, I would first have to introduce what I termed 'rules of supposition', that is, the rules determining when a term can have such-and-such supposition. But for explanatory purposes, it seems more convenient to introduce what I termed 'definitions of supposition' before, so that the concepts of personal/simple/material supposition are right away clear to the reader. So, based on Ockham's definitions of personal, material and simple suppositions, these concepts can be defined as follows:

Personal supposition: For some **t**,

$$\mathrm{PSUP}(a, \mathbf{t})_{\mathrm{P}_1} <=> \mathit{In}{<}{}^{*}a{>}_{\mathrm{P}_1} \text{ and } \mathrm{SIG}(a, \mathbf{t})$$

Therefore, for some type Ω, if $a{:}\ \Omega$ and $\mathrm{PSUP}(a, \mathbf{t})_{\mathrm{P}_1}$, then $\mathbf{t}/\Omega$.

Simple supposition: For some **t**,

$$\mathrm{SSUP}(a, \mathbf{t})_{\mathrm{P}} <=> \mathit{In}{<}\#a{>}_{\mathrm{P}_1} \text{ and } \mathbf{t}/\mu.$$

Material supposition: For some **t**,

$$\mathrm{MSUP}(a, \mathbf{t})_{\mathrm{P}} <=> \mathit{In}{<}\#a{>}_{\mathrm{P}_1} \text{ and } \mathbf{t}/\lambda.$$

Actual and potential supposition. Ockham's supposition theory defines the range of legitimate kinds of supposition for a term in a given propositional context; it does not define the kind of supposition that a term actually has in a proposition (only the intention of the user does – cf. section 1.2.2). So we must distinguish the cases in which a term actually has such-and-such supposition in a given proposition (and no other), in virtue of the intention of a user or interpreter, from the cases in which it is legitimate for a term to have such-and-such supposition in a proposition, but it does not necessarily have such-and-such supposition in that proposition (cf. section 1.3.2.1).

In effect, the actual supposition of a term is relative to one specific token $\mathbf{P_1}$ of a proposition **and** to one specific interpretation *In*. The most straightforward reading of a proposition is the meaning intended by the one who forms it; but as we have seen, in particular in the case of written propositions, it is often not possible to retrace the exact meaning intended by the author, so the interpreter must give his own interpretation to the proposition (which may or may not coincide with the intended meaning).[105]

---

[105] In practice, this kind of ambiguity and misunderstanding often occurs in actual situations as well. For a modern analysis of such phenomena, cf. (van Rooij 2004).

By contrast, the potential suppositions of a term in a proposition are relative only to the form of the proposition (presence or absence of syncategorematic terms and the type of the other extreme); now, every token of a proposition obviously shares this form with its other tokens. So the potential supposition of a term in a proposition is relative to a proposition-type **P**.

Therefore, the definitions of these notions are as follows: for some **t**,

$@_{\mathrm{In}}<\mathrm{PSUP}(a, \mathbf{t})_{\mathrm{Pn}}> \; <=>$ *a* actually has personal supposition in token *Pn*, under interpretation *In*, and no other kind of supposition.

$@_{\mathrm{In}}<\mathrm{SSUP}(a, \mathbf{t})_{\mathrm{Pn}}> \; <=>$ *a* actually has simple supposition in token *Pn*, under interpretation *In*, and no other kind of supposition.

$@_{\mathrm{In}}<\mathrm{MSUP}(a, \mathbf{t})_{\mathrm{Pn}}> \; <=>$ *a* actually has material supposition in token *Pn*, under interpretation *In*, and no other kind of supposition.

$+<\mathrm{PSUP}(a, \mathbf{t})_{\mathrm{P}}> \; <=>$ *a* possibly has personal supposition in **P**, and possibly other kind(s) of supposition as well.

$+<\mathrm{SSUP}(a, \mathbf{t})_{\mathrm{P}}> \; <=>$ *a* possibly has simple supposition in **P**, and possibly other kind(s) of supposition as well.

$+<\mathrm{MSUP}(a, \mathbf{t})_{\mathrm{P}}> \; <=>$ *a* possibly has material supposition in **P**, and possibly other kind(s) of supposition as well.

Supposition theory in itself cannot establish whether the actual relation of supposition is the case in **Pn**, but only whether such-and-such supposition is possible. Again, what defines the actual relation of supposition is only the intention of the speaker or of the interpreter, or a convention between the speakers, in such a way that a term can in fact supposit for anything, even going beyond the constraints and rules of supposition theory.

1.5.1.3 Quasi-syntactical rules for personal, simple, and material supposition

In the case of personal, simple and material supposition, the rules defining in which propositional context a term can have such-and-such supposition are quite simple (in contrast with the rules for the modes of personal supposition): a term can always have personal supposition, in whichever proposition it occurs, and it can have simple or material supposition if the other extreme of the proposition in which it appears is a term of second intention (a name of a concept) or of second imposition (a name of a word), respectively. These rules (which amount to interpretational procedures/instructions) can be formulated as follows:

For any term *a* in any proposition P, for some **t**,

$b: \lambda => +<\mathrm{PSUP}(a, \mathbf{t})_{\mathrm{P}}>$

$b: \mu => +<\mathrm{PSUP}(a, \mathbf{t})_{\mathrm{P}}>$

$b: \varkappa => +<\mathrm{PSUP}(a, \mathbf{t})_{\mathrm{P}}>$

For any term b in any proposition P, for some $\mathbf{t}$,

$$a\text{: } \lambda => +<\mathrm{PSUP}(a, \mathbf{t})_{\mathrm{P}}>$$
$$a\text{: } \mu => +<\mathrm{PSUP}(a, \mathbf{t})_{\mathrm{P}}>$$
$$a\text{: } \varkappa => +<\mathrm{PSUP}(a, \mathbf{t})_{\mathrm{P}}>$$

For any term a in any proposition P, for some $\mathbf{t}$,

$$b\text{: } \lambda => +<\mathrm{MSUP}(a, \mathbf{t})_{\mathrm{P}}>$$
$$b\text{: } \mu => +<\mathrm{SSUP}(a, \mathbf{t})_{\mathrm{P}}>$$

1.5.1.4 Table

These rules and definitions can all be represented in one single table (Figure 1.5.1.4). Since there are two terms in a categorical proposition and three kinds of terms (ϰ, λ, μ), clearly there are nine propositional schemata, corresponding to the combinatorial of the three kinds of terms to the two terms of the proposition. Notice also that some propositional schemata (indexed by numbers) correspond to more than one interpretational schema[106] (indexed by letters). These are represented in the left columns; the

	a	b	$\mathrm{SUP}(a, \mathbf{t}_1)_{\mathrm{P}}$	$\mathrm{SUP}(b, \mathbf{t}_2)_{\mathrm{P}}$	$\mathbf{t}_1$	$\mathbf{t}_2$
1	a: ϰ	b: ϰ	$+<\mathrm{PSUP}(a, \mathbf{t}_1)_{\mathrm{P}}>$	$+<\mathrm{PSUP}(b, \mathbf{t}_2)_{\mathrm{P}}>$	$\mathbf{t}_1/\varkappa$	$\mathbf{t}_2/\varkappa$
2.a 2.b	a: ϰ	b: λ	$+<\mathrm{PSUP}(a, \mathbf{t}_1)_{\mathrm{P}}>$ $+<\mathrm{MSUP}(a, \mathbf{t}_1)_{\mathrm{P}}>$	$+<\mathrm{PSUP}(b, \mathbf{t}_2)_{\mathrm{P}}>$ $+<\mathrm{PSUP}(b, \mathbf{t}_2)_{\mathrm{P}}>$	$\mathbf{t}_1/\varkappa$ $\mathbf{t}_1/\lambda$	$\mathbf{t}_2/\lambda$ $\mathbf{t}_2/\lambda$
3.a 3.b	a: ϰ	b: μ	$+<\mathrm{PSUP}(a, \mathbf{t}_1)_{\mathrm{P}}>$ $+<\mathrm{SSUP}(a, \mathbf{t}_1)_{\mathrm{P}}>$	$+<\mathrm{PSUP}(b, \mathbf{t}_2)_{\mathrm{P}}>$ $+<\mathrm{PSUP}(b, \mathbf{t}_2)_{\mathrm{P}}>$	$\mathbf{t}_1/\varkappa$ $\mathbf{t}_1/\mu$	$\mathbf{t}_2/\mu$ $\mathbf{t}_2/\mu$
4	a: λ	b: ϰ	$+<\mathrm{PSUP}(a, \mathbf{t}_1)_{\mathrm{P}}>$	$+<\mathrm{PSUP}(b, \mathbf{t}_2)_{\mathrm{P}}>$	$\mathbf{t}_1/\lambda$	$\mathbf{t}_2/\varkappa$
5.a 5.b	a: λ	b: λ	$+<\mathrm{PSUP}(a, \mathbf{t}_1)_{\mathrm{P}}>$ $+<\mathrm{MSUP}(a, \mathbf{t}_1)_{\mathrm{P}}>$	$+<\mathrm{PSUP}(b, \mathbf{t}_2)_{\mathrm{P}}>$ $+<\mathrm{PSUP}(b, \mathbf{t}_2)_{\mathrm{P}}>$	$\mathbf{t}_1/\lambda$ $\mathbf{t}_1/\lambda$	$\mathbf{t}_2/\lambda$ $\mathbf{t}_2/\lambda$
6.a 6.b	a: λ	b: μ	$+<\mathrm{PSUP}(a, \mathbf{t}_1)_{\mathrm{P}}>$ $+<\mathrm{SSUP}(a, \mathbf{t}_1)_{\mathrm{P}}>$	$+<\mathrm{PSUP}(b, \mathbf{t}_2)_{\mathrm{P}}>$ $+<\mathrm{PSUP}(b, \mathbf{t}_2)_{\mathrm{P}}>$	$\mathbf{t}_1/\lambda$ $\mathbf{t}_1/\mu$	$\mathbf{t}_2/\mu$ $\mathbf{t}_2/\mu$
7	a: μ	b: ϰ	$+<\mathrm{PSUP}(a, \mathbf{t}_1)_{\mathrm{P}}>$	$+<\mathrm{PSUP}(b, \mathbf{t}_2)_{\mathrm{P}}>$	$\mathbf{t}_1/\mu$	$\mathbf{t}_2/\varkappa$
8.a 8.b	a: μ	b: λ	$+<\mathrm{PSUP}(a, \mathbf{t}_1)_{\mathrm{P}}>$ $+<\mathrm{MSUP}(a, \mathbf{t}_1)_{\mathrm{P}}>$	$+<\mathrm{PSUP}(b, \mathbf{t}_2)_{\mathrm{P}}>$ $+<\mathrm{PSUP}(b, \mathbf{t}_2)_{\mathrm{P}}>$	$\mathbf{t}_1/\mu$ $\mathbf{t}_1/\lambda$	$\mathbf{t}_2/\lambda$ $\mathbf{t}_2/\lambda$
9.a 9.b	a: μ	b: μ	$+<\mathrm{PSUP}(a, \mathbf{t}_1)_{\mathrm{P}}>$ $+<\mathrm{SSUP}(a, \mathbf{t}_1)_{\mathrm{P}}>$	$+<\mathrm{PSUP}(b, \mathbf{t}_2)_{\mathrm{P}}>$ $+<\mathrm{PSUP}(b, \mathbf{t}_2)_{\mathrm{P}}>$	$\mathbf{t}_1/\mu$ $\mathbf{t}_1/\mu$	$\mathbf{t}_2/\mu$ $\mathbf{t}_2/\mu$

Figure 1.5.1.4. Table representing Ockham's theory of supposition

---

[106] I use the term 'propositional schema' to refer to the form of the proposition, including the semantic types of its terms, and the term 'interpretational schema' to the attribution of such-and-such supposition to its terms, on the basis of the interpretational constraints proper to each propositional schema.

kinds of supposition possible for each schema are represented in the middle columns; and the types of the *supposita* in question are represented in the right columns of the table.

Truth. As is well known, Ockham maintains that a proposition is true if there is coincidence of *supposita*, that is, if all the *supposita* of the subject are also *supposita* of the predicate.[107] Therefore, a necessary (but not sufficient) condition for the truth of a proposition is that the *supposita* of subject and predicate be of the same type (as sameness of type is obviously a necessary condition for identity). On the basis of the table above, it is easy to see which interpretational schemata can yield a true proposition (under a given interpretation) and which cannot. Those that fulfill the necessary condition to be true are those where $\mathbf{t_1}$ and $\mathbf{t_2}$ have the same type, namely 1, 2.b, 3.b, 5.a, 5.b, 6.b, 8.b, 9.a, 9.b. Notice that propositional schemata 4 and 7 can never be interpreted as true propositions (to these schemata belong awkward propositions such as 'a name is an animal', for example).[108] Other propositional schemata can yield true as well as false interpretations (2, 3, 6, 8), and yet other schemata yield two interpretations both of which satisfy the necessary condition for truth (5, 9).

1.5.2 Modes of personal supposition

The part of Ockham's supposition theory dealing with the modes of personal supposition is significantly more complex than the one dealing with personal, simple and material supposition, with respect to both the quasi-syntactical and the semantic rules of supposition. The problem with the rules of personal supposition is that, contrasting with the other part of the theory, here the enumeration of possible propositional forms simply cannot be as exhaustive. In the previous case, since a term can only be of one out of three kinds (a term signifying things, a term signifying concepts or a term signifying words), there are only nine possible propositional schemata; here, however, it is virtually impossible to provide an exhaustive enumeration of all possible schemata related to the form of propositions (i.e. the presence or absence of syncategorema and word order).

Another significant difference with respect to the other part of the theory is that no emphasis is placed on the need to 'distinguish' propositions whose both terms have personal supposition. In other words, the tacit assumption seems to be that such propositions can always be univocally interpreted, on the basis of its propositional form. In current semantics, the ambiguity inherent to propositions such as 'Every man loves a woman' (related to the scope of the quantifiers) is widely recognized, but interestingly, the regimentation of philosophical medieval Latin included conventions

---

[107] Cf. *Summa Logicae II*, chap. 2–3.

[108] One could argue though that a proposition such as 'a name is a sound' also belongs to this schema, which can be true in a straightforward way. To this it can be replied that it really depends on one's conception of a sound, and accordingly on whether 'sound' is indeed a $\varkappa$-term.

concerning word order that virtually excluded the possibility of such ambiguities. Hence, while the notion of '*denotatur*' is just as important for the modes of personal supposition as it is for the division of the three kinds of supposition, the notion of propositions that must be distinguished is not relevant here.

The problem concerning the definitions of the modes of personal supposition is of a different nature. There are three modes of personal supposition,[109] and therefore there can be only nine combinations of modes of personal supposition for subject and predicate, respectively – in fact, we shall see that only six combinations can actually occur, in virtue of the syntactical rules. So the problem at stake is not that of the impossibility of covering all cases with a few schemata; but the issue here is probably the (in the secondary literature) most discussed aspect of medieval theories of supposition, namely the relations of ascent and descent. What are these relations supposed to accomplish? Is this part of a theory of inference concerning the relation between different propositions, or does it concern the semantic relations between language and things? Are these supposed to provide the truth conditions of propositions?

The main assumption of the formalization presented here is that the ascents and descents are Ockham's manner of defining what it means, in terms of its *supposita*, for a term to have such-and-such personal supposition. If the structure of the theory is indeed as I describe, then the form of a proposition is analyzed to determine which supposition(s) a term can have (the rules of supposition), and subsequently meaning is ascribed to the proposition by the definitions of modes and kinds of supposition. If however the definitions based on ascent and descent are used to determine which kind of supposition a term has (inverting thus what I claim to be the correct order of application of the rules), then there is a primacy of content over form, and therefore the theory is no longer a **formal** theory of semantic analysis.[110]

Here, I defend the view that these inferential relations are meant to determine the meaning of such propositions, since they define what is being asserted by means of a proposition with respect to the *supposita*, just as much as the definitions of personal, simple and material supposition did. More specifically, this view is very much inspired by the inferentialist views defended by R. Brandom (1994, 2000), among others, according to whom the meaning of a proposition is defined by the inferential relations it entertains with other propositions.

Another important issue is the asymmetry between personal supposition, on the one hand, and simple and material supposition, on the other hand. Why are there different modes of personal supposition, but not of simple or material supposition? This fact is all the more puzzling because Ockham's ontology not only can easily accommodate but in fact even seems to require that the same distinctions be applied

---

[109] It has been discussed in the literature whether these three modes exhaust all possibilities, and whether there is or isn't a fourth missing mode, cf. (Spade 1996, 288–290).

[110] For more on why there only appears to be two distinctive sets of rules for the modes of personal supposition, but how in fact each set has a different purpose, see (Dutilh Novaes 2004c, footnote 8) and chapter 2.4 below.

to the other kinds of supposition as well. In his ontology composed of individuals, there are only individual occurrences of words as well as of concepts, every time that one of them is formed (spoken, written or conceived) – what are currently called tokens, as opposed to types. Now, if one considers a proposition such as 'Man is written on a piece of paper', clearly it is not being asserted that all occurrences of the word 'man' are written on a piece of paper, but rather that some are, whereas with 'Man has three letters', what is likely to be asserted is that all occurrences of the type 'man' have three letters. Ockham would probably not want to refer to the 'type "man",' since types very much resemble the (extra-mental) universals to which he denies existence, so a possible solution would be to attribute confused and distributive material supposition to 'man' in 'Man has three letters' in order to account for the intended meaning.

Hence, even though Ockham himself does not recognize different modes of supposition (determinate, confused and distributive and merely confused supposition) for simple and material supposition, this would be entirely compatible with his doctrines. In fact, Marsilius of Inghen, writing only some decades after Ockham, seems to have been the first to attribute different modes of supposition for material supposition (simple supposition being for him a special case of material supposition).

> Common determinate supposition is subdivided into material and personal.
>
> Common determinate material supposition is the acceptance of a term in a propositions for its non-ultimate significate, or its non-ultimate significates disjunctively, of which, or of which things, the term is verified through the copula of the proposition. As in the proposition *'to love' is a verb*, the verb '*to love*' stands disjunctively for each such term, viz. to love. Therefore, the following descent is allowed: *to love is a verb, therefore this case of 'to love' is a verb, or that case*, and so on.[111] (Marsilius of Inghen 1983, 57)

This is indeed a very natural expansion of the supposition framework; in practice, what will be presented below as modes of personal supposition could easily be applied to the other kinds of supposition as well, *mutatis mutandi*.

As in the previous section, for convenience of exposition I will first present the definitions of the modes of personal supposition, and then the rules defining which modes of personal supposition a term has in a proposition (again inverting the actual order of application of the procedure). I will only deal with common personal supposition; that is, discrete supposition, which is the supposition of proper names and demonstrative pronouns, will not be treated.

---

[111] Et divitur suppositio communis determinata, quia quedam est suppositio communis materialis determinata et quedam suppositio personalis determinata. Suppositio communis materialis determinata est acceptio termini in propositione pro sua significato non ultimato, sive pro suis significatis non ultimatis disiunctive, de quo, vel de quibus, talis terminus verificatur mediante copula talis propositionis. Ut hec 'amo' est verbum, li amo stat disiunctive pro quolibet termino tali, scilicet amo. Unde debet descendi sic 'Amo est verbum, ergo hoc 'amo' est verbum vel hoc et sic de singularis. (Marsilius of Inghen 1983, 56)

1.5.2.1 The semantic rules for the modes of personal supposition[112]

As already said, I maintain that the relations of ascent and descent characterizing the three modes of personal common supposition account for the semantic properties of the corresponding 'quantified' terms in a given proposition – that is, what it means, in terms of *supposita*, to have such-and-such supposition. It is noteworthy that earlier theories of supposition did not rely solely on ascent and descent for these definitions (in particular Peter of Spain's[113] and William of Sherwood's[114]); by contrast, in 14th century theories such as Ockham's, these inferential relations are at the core of the definitions of modes of personal supposition. One of the problems with this 14th century approach is, as has been argued in the literature,[115] that the definitions so cast simply do not work properly, since there isn't always logical equivalence between the original proposition and the chains of disjunction and conjunction.[116] Often, a given mode of personal supposition is defined in terms of which ascents and descents are **not** possible, which is obviously not sufficient for a full-fledged definition.

However, even if not good enough for definitions of logical equivalences or of truth conditions, the inferential relations of ascent and descent do seem to point towards what is asserted by means of a proposition. In fact, the idea that the meaning of a proposition is defined in terms of the inferential relations it entertains with other propositions has been given an elaborate theoretical status recently.[117] Now, it is very well conceivable that, within the supposition framework, it is the meaning of a given proposition that is established in terms of its inferential relations with chains of conjunctions and disjunctions of propositions of the form 'This *a* is *b*'. Hence, one could say that the moves of ascent and descent are not so much operational definitions as they are heuristic explanations of meaning.[118]

---

[112] The content of this section is very similar to section 3.2 of the present work, where Buridan's theory of inference between categorical propositions is treated. Here, the ideas are virtually the same (since Ockham's and Buridan's views on the modes of personal supposition are virtually the same), but the presentation is adapted to the notation being used for each formalization.

[113] (Peter of Spain 1972, 82–83).

[114] (William of Sherwood 1995, 136).

[115] Some of the studies on supposition theory are: (Swiniarski 1970); (Priest and Read 1977); (Priest and Read 1980); (Spade 1988); (Karger 1984), (Matthews 1997). For an informal but insightful overview, cf. (Spade 1996, chap. 9); notice though that I disagree with many of Spade's views on the modes of personal supposition.

[116] Cf. (Spade 1996, p. 289), (Spade 1988).

[117] Cf. (Brandom 2000).

[118] That is, even though I agree with Matthews' contention that there are not two theories of supposition, my view of the ultimate purpose of the theory differs from his: 'As I see it, the theory was not intended to provide, by itself, an account of truth conditions, at least until well after Ockham's time. It was intended as a way of making sense of the idea that general terms in categorical propositions can refer to individuals that fall under them' (Matthews 1997, 39). As I have stated many times, I do not think that supposition theories should be viewed as an account of the relation of reference, but I do agree with Matthews insofar as he holds that modes of personal supposition are related to the phenomenon of multiple denotation of general terms, cf. (section 1.2.3).

Following the tradition, Ockham maintains that personal supposition is divided into common and discrete supposition, and that common supposition is of three modes: determinate, confused and distributed, and merely confused personal supposition. As can be seen from the quotations in section 1.4.2.2, these modes of supposition are defined by the inferential relations of ascent and descent. In practice, what these relations illustrate are the situations that are asserted to be the case by means of this proposition; a proposition makes a claim about what is the case. Therefore, what can be determined by the relations of ascent and descent are the **models** underlying the different interpretational schemata defined in terms of the personal supposition of the terms in a proposition. Therefore, it seems that what is asserted [*denotatur*] by means of a proposition whose both terms have personal supposition can be rendered by some basic model theoretic tools – again, the use of modern tools does involve the risk of anachronism, but again, it is a risk worth being taken.

Of the three modes of personal supposition, it would seem that determinate supposition and confused and distributive supposition can be spelled out even if the terms are taken in isolation, that is, out of a propositional context, and that, by contrast, merely confused supposition can only be explained with respect to the supposition of the other term in the proposition. But in practice, one has to resort to the whole original proposition even in order to define the other two modes of personal supposition. As we shall see, the best way to approach the three modes of personal proposition seems to be through the (six, as we shall see) possible interpretational schemata that are formed by the attribution of a mode of personal supposition to subject and to predicate, respectively.

Nevertheless, I will first present characterizations of each mode of personal supposition, in order to reach the final goal of handling the different interpretational schemata (i.e., considering the supposition of subject and of predicate at once) and their underlying models. Based on the quotations in section 1.4.2.2, it is clear that one must start with the inferential relations that categorical propositions of the form '(S) *a* is (S) *b*' (where 'S' stands for syncategorema and is a placeholder for syncategorematic terms such as '*omnis*', '*aliquid*' or '*nullus*', which may or may not be filled) have with proposition of the form 'This *a* is (S) *b*' or '(S) *a* is this *b*'.

Notice that here I only treat the case of affirmative present-tense propositions whose verb is the appropriate conjugation of the verb 'to be' (*esse*); presumably, this account can be expanded so as to include negative propositions, past- or future-tense propositions, and propositions whose main verbs are other than the verb 'to be'.[119]

Notice also that, in the formulations of the kinds of personal supposition as well as of the different interpretational schemata, the implication holds in only one direction, from the kind of supposition(s) to the semantic interpretation of the proposition, and not the other way round. This is because, in the spirit of the procedural interpretation of supposition theory proposed here (cf. Figure 1.4), it is not the semantic properties of a proposition that determine the supposition of its terms, but rather the supposition of its terms that determines its semantic properties.

[119] The case of propositions having verbs other than 'to be' is treated in section 2.4 of the present work.

Determinate supposition. A term *a* has determinate supposition in $\mathbf{P_1}$ => A disjunction of propositions of the form 'This *a* is (S) *b*' can be inferred from $\mathbf{P_1}$ but a conjunction of propositions of the form 'This *a* is (S) *b*' cannot be inferred from $\mathbf{P_1}$ => With $\mathbf{P_1}$, it is asserted that (at least) one proposition of the form 'This *a* is (S) *b*' is true.

Confused and distributive supposition. A term *a* has confused and distributive supposition in $\mathbf{P_1}$ => A conjunction of propositions of the form 'This *a* is (S) *b*' can be inferred from $\mathbf{P_1}$[120] => With $\mathbf{P_1}$, it is asserted that every proposition of the form 'This *a* is (S) *b*' is true.

Merely confused supposition. A term *b* has merely confused supposition in $\mathbf{P_1}$ (where *a* has confused and distributive supposition) => A proposition with a disjunctive predicate of the form 'this *b*, or that *b* etc.' can be inferred from $\mathbf{P_1}$ but neither a disjunction nor a conjunction of propositions of the form '(S) *a* is this *b*' can be inferred from $\mathbf{P_1}$ => With $\mathbf{P_1}$, it is asserted that a proposition of the form '(S) *a* is this *b*, or that *b* etc.' is true.

Interpretational schemata. Since there are three modes of personal supposition, and we are taking into consideration only the basic (categorical) forms[121] of propositions (with two terms), one would expect there to be nine interpretational schemata related to the modes of personal supposition. But three of the interpretational schemata cannot occur, given the rules of personal supposition to be presented below – in fact, there do not even seem to be meaningful contents that could be associated to them. These three 'impossible' schemata are: the subject having determinate supposition and the predicate having merely confused supposition, the subject having merely confused supposition and the predicate having determinate supposition, and both terms having merely confused supposition.

Thus, the interpretational schemata to be considered are:

(11) $\text{PcdSUP}(a, \mathbf{t_1})_\text{P}$ and $\text{PcdSUP}(b, \mathbf{t_2})_\text{P}$
(12) $\text{PcdSUP}(a, \mathbf{t_1})_\text{P}$ and $\text{PdSUP}(b, \mathbf{t_2})_\text{P}$
(12′) $\text{PdSUP}(a, \mathbf{t_1})_\text{P}$ and $\text{PcdSUP}(b, \mathbf{t_2})_\text{P}$
(13) $\text{PcdSUP}(a, \mathbf{t_1})_\text{P}$ and $\text{PmcSUP}(b, \mathbf{t_2})_\text{P}$
(13′) $\text{PmcSUP}(a, \mathbf{t_1})_\text{P}$ and $\text{PcdSUP}(b, \mathbf{t_2})_\text{P}$
(14) $\text{PdSUP}(a, \mathbf{t_1})_\text{P}$ and $\text{PdSUP}(b, \mathbf{t_2})_\text{P}$

(12) and (12′), and (13) and (13′), respectively, are symmetric, so in effect there are only four schemata that must be treated.

[120] Here, the exclusion clause does not apply, since obviously from such a conjunction the appropriate disjunction can be inferred. Indeed, as shown in section 2.4.4, from a proposition where a given term has distributive and confused supposition, the corresponding proposition where the same term has determinate supposition can be inferred.
[121] According to the usual medieval procedure.

Schema (11)

By means of a proposition **P**, whose both terms *a* and *b* have confused and distributive supposition, it is asserted that **all** *significata* of *a* are (or not, if the proposition is negative) related by the relation **R** expressed by the copula or verb to **all** *significata* of *b*. Figure 1.5.2.1.1 is a diagram representing what is asserted to be the case by a proposition whose both terms have confused and distributive supposition ('*' represents entities, the circles represent the class of *significata* of each term, and the lines the relation between them).

In sum, according to this interpretational schema, the Cartesian product of the *significata* of *a* and *b* occurs. This is a rather awkward content in the case of an affirmative identity statement with the verb 'to be' as copula: it states that all *significata* of *a* are identical to all *significata b* – it is not obvious how an entity can be identical to several (presumably) different things. However, in the case of propositions with transitive verbs, having the subject-verb-object structure, and whose terms both have confused and distributed supposition, they simply state that all *significata* of the subject are related by the relation expressed by the verb to all *significata* of the object. Moreover, the propositional form 'No *a* is *b*', traditionally considered to be one of the four main propositional forms, belongs to this interpretational schema; with propositions of this form it is asserted that the relation of non-identity holds between every *suppositum* of *a* and every *suppositum* of *b*.

> PcdSUP(a, $\mathbf{t_1}$)$_P$ and PcdSUP(b, $\mathbf{t_2}$)$_P$ => $\vdash$ For all $\mathbf{t_1}$ and $\mathbf{t_2}$ such that SIG(a, $\mathbf{t_1}$) and SIG(b, $\mathbf{t_2}$), $\mathbf{t_1 R t_2}$.[122]

Example: *Nullo homo est asinus.*

By means of this proposition it is asserted that every man is in a relation of not being identical with **all** donkeys – that every man in not identical to each and every donkey, a situation corresponding to the diagram below.

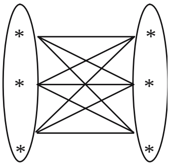


Figure 1.5.2.1.1. Schema (11)

[122] If either term is empty, such as 'chimera', then the assertion that every *significatum* of the subject is identical to every *significatum* of the predicate fails to be true (since what does not exist cannot be identical to whatever else), and therefore the proposition is false (if affirmative), but the supposition in itself does not fail. This is why the assertion operator must have wide scope, to avoid presuppositions of existence.

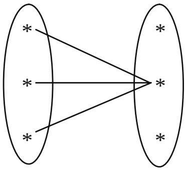


Figure 1.5.2.1.2. Schema (12)

Schema (12)
By means of a proposition **P**, whose term *a* has confused and distributive supposition and whose term *b* has determinate supposition, it is asserted that **all** *significata* of *a* are (or not, if the proposition is negative) related by the relation **R** expressed by the copula or verb to **one and the same** *significatum* of *b*. Figure 1.5.2.1.2 is a diagram representing what is asserted to be the case by a proposition whose subject has confused and distributive supposition and whose predicate has determinate supposition.

Again, this is not an immediately intuitive content in the case of an identity statements with the verb 'to be' as copula, but it makes perfect sense if the verb of the proposition is a transitive verb, or with a negated copula ('Some *a* is not *b*', one of the traditional propositional forms, belongs to interpretational schema (12′)).

> PcdSUP(a, $\mathbf{t_1}$)$_P$ and PdSUP(b, $\mathbf{t_2}$)$_P$ => ⊢ There is a $\mathbf{t_2}$ such that SIG (b, $\mathbf{t_2}$), and for all $\mathbf{t_1}$ such that SIG(a, $\mathbf{t_1}$), $\mathbf{t_1 R t_2}$.

Example: *Asinum omnis homo videt.*

According to the regimented use of Latin at Ockham's time,[123] by means of this proposition it is asserted that every man sees one and the same donkey, that is, is in a relation of seeing with the same donkey, as in the situation depicted in the diagram.

The definition of what is asserted by a proposition belonging to schema (12′) is easily obtained inverting the order of subject and predicate and following the same reasoning.

Schema (13)
By means of a proposition **P**, whose term *a* has confused and distributive supposition and whose term *b* has merely confused supposition, it is asserted that **all** *significata* of *a* are (or not, if the proposition is negative) related by the relation **R** expressed by the copula or verb to **some** *significatum* of *b*, forming ordered pairs, since merely confused supposition 'assigns' a *significatum* of the term in question to each *significatum*

[123] In this case, it regards word order. If the predicate is a general term and precedes the subject with a distributive sign of universality, then what is expressed is that one single member of the domain of the predicate is related to all members of the domain of the subject. This regimentation of the language displays a strong similarity with the idea of **scope** in current quantified logic.

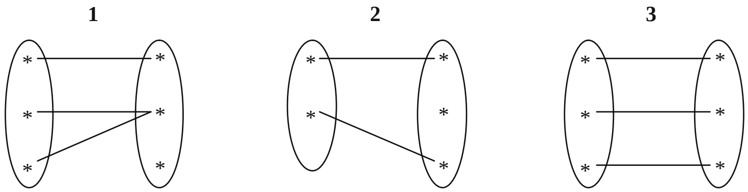


Figure 1.5.2.1.3. Schema (13)

of the other term. Thus, all *significata* of *a* are related to a *significatum* of *b*, but the latter is any of the *significata* of *b*. Three different contents can be asserted (Figure 1.5.2.1.3).

1 – Surjection. All *significata* of *a* are mapped into some *significatum* of *b*, but some *significatum* of *b* is related to more than one *significatum* of *a*.

Example: *Omnis homo videt asinum.*

This proposition is true, for example, if each *significatum* of '*homo*' sees one individual among the *significata* of '*asinus*', but it is possible that a certain donkey is seen by more than one man, as much as it is possible that another donkey is not seen by any man.

2 – Injection. All *significata* of *a* are mapped into a *significatum* of *b*, and a *significatum* of *b* is related to at most one *significatum* of *a*, but there are more *significata* of *b* than *significata* of *a*, so some *significata* of *b* are not related to any *significatum* of *a*.

Example: *Omnis homo est animal.*

All men are animals, but not all animals are men.

3 – Bijection. All *significata* of *a* are mapped into a *significatum* of *b*, and all *significata* of *b* are related to exactly one *significatum* of *a*.

Example: *Omnis homo est animal rationale.*[124]

If a proposition **P** belongs to this interpretational schema, it is not possible to determine from its form alone which of the three cases applies: as far as its form goes, any of these three situations would verify it. Only by means of an analysis of the **content** of the proposition is it possible to establish what exactly is being asserted, among these

[124] Indeed, a bijection is what is established by definitions of terms, implying equivalence between the two sides of the definition.

three possibilities. This can be tested for example, by means of the inferential relations between the proposition in question and similar ones. For example, the converse of 'Every man is a rational animal', 'Every rational animal is a man', is entailed by the original sentence, but 'Every animal is a man' is not entailed by 'Every man is an animal'. In sum:

> PcdSUP(a, $\mathbf{t_1}$)$_P$ and PmcSUP(b, $\mathbf{t_2}$)$_P$ => ⊢ For each $\mathbf{t_1}$ such that SIG (a, $\mathbf{t_1}$), there is a $\mathbf{t_2}$ such that SIG(b, $\mathbf{t_2}$) and $\mathbf{t_1 R t_2}$.

The definition of what is asserted by a proposition belonging to schema (13′) is easily obtained inverting the order of subject and predicate and following the same reasoning.

Schema (14)
By means of a proposition **P**, whose terms a and b have determinate supposition, it is asserted that **one** *significatum* of a is related by relation **R** expressed by the copula or verb to **one** *significatum* of b. Thus, one *significatum* of a is related to one *significatum* of b. Figure 1.5.2.1.4 is a diagram representing what is asserted to be the case by a proposition whose both terms have determinate supposition.

> PdSUP(a, $\mathbf{t_1}$)$_P$ and PdSUP(b, $\mathbf{t_2}$)$_P$ => ⊢ There is a $\mathbf{t_1}$ such that SIG (a, $\mathbf{t_1}$), and a $\mathbf{t_2}$ such that SIG(b, $\mathbf{t_2}$), such that $\mathbf{t_1 R t_2}$.

Example: *Homo est albus.*

This sentence asserts that least one of the *significata* of '*homo*' has a relation of identity with one of the *significata* of '*albus*'.

Polarity. So far, I have not explicitly discussed the case of negative propositions, that is, with negated verbs or copulas. While the syntactical aspects related to negation will be extensively discussed in the next section, I will now only hint at the semantic aspects thereof – but admittedly, more work on this should be done.

The issue here concerns the semantic effect of a proposition's negative polarity – that is, what this fact amounts to in terms of the models defined above. *Prima facie*, it seems to me that this is a straightforward matter: while by an affirmative proposition it is asserted that the relation expressed by the verb or copula **does exist** between the

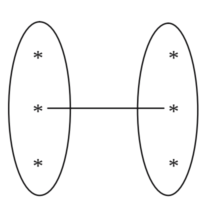


Figure 1.5.2.1.4. Schema (14)

relevant entities, by a negative proposition, that is, a proposition with one occurrence of a negating expression (in any position), it is simply asserted that the relation expressed by the verb or copula **does not exist** between any of the relevant entities.

Worth noting is the fact that negative propositions are also verified when the presupposition of existence of the *supposita* fails. In this case, the relation expressed by the verb or the copula obviously fails, since some or all of the 'entities' involved simply do not exist.

This matter becomes slightly more complicated with iterated occurrences of negating expressions. I would tend to think that, from a syntactical perspective, an affirmative proposition would be a proposition with 0 or an even number of negating expressions, while a negative proposition would be a proposition with an odd number of negations. Moreover, it would seem that the semantic definitions of polarity just sketched are sound with respect to the syntactical definitions of affirmative and negative propositions with negation iteration. However, intuitive as though they may seem, these claims would have to receive further corroboration.

1.5.2.2 Quasi-syntactical rules for the modes of personal supposition

Recall that in the case of personal, material and simple supposition, formal features of the propositional context define which kind(s) of supposition a term may have, namely the type of the predicate term. Since there are only three kinds of terms, the enumeration of the nine cases was exhaustive.

With respect to the modes of personal supposition, an exhaustive enumeration of all possible logical forms cannot occur. The syntactical structure of a proposition – namely the presence/absence of syncategorematic terms and the position of each term in the proposition – determines the kind of personal supposition of its terms. But it is virtually impossible to enumerate all cases: medieval logicians worked with a 'natural language'[125] (according to the modern acceptation of the expression), and not with an artificial, inductively generated language, as logicians now do.[126] So the range of possibilities – that is, well-formed propositions – is beyond enumeration.

It may seem that this hinders the generality of supposition theory, since no matter how long the enumeration of cases may be, there will always be well-formed propositions that do not display one of the logical forms recognized by the theory, and whose terms therefore do not have a previously defined kind of supposition. However, one must bear in mind that, since medieval logicians were dealing with object-languages in constant expansion (in the sense that, gradually, more and more complex propositional forms became a part of the fragment of the language that the theoretical apparatus could handle), the corresponding 'meta-theory' should also be in constant expansion. Current logical systems are almost always previously defined, closed languages, therefore, in this respect, a comparison with supposition theories would not be a fair one.

---

[125] In section 4.3 I question the use of this term to refer to 'regular' spoken and written language.

[126] However, medieval academic Latin is one of the best examples of how a 'natural language' can be regimented to such a point that it starts resembling an artificial language.

But there are other logical systems, not as 'old' as theories of supposition, that also make extensive use of the possibility of introducing new terms and definitions to the logical language, turning it into a language in expansion. A good example is the systems developed by the Polish logician S. Lesniewski (Lesniewski 1992). Much more recently, some of the linguists and semanticists involved in the project of formalizing 'natural' languages – in particular by means of first-order predicate logic – became aware of its limitations for this purpose, and gradually developed what was to be called the 'generalized quantifying theory' (cf. de Swart 1998, chap. 7 and 8). In particular, formalizations of determiners (a broader version of the notion of quantifiers), such as 'most', 'at least two', 'many', 'half' etc. (cf. van Benthem and ter Meulen 1985) display a certain similarity to the general idea behind supposition theory: at every occurrence of a 'new' (i.e., not yet defined) quantifying expression, the idea is to determine its semantics by means of the theory. In the case of the modern determiners, this is usually done in terms of set theory; in the case of the medieval theory, this was done in terms of the concept of supposition. For example, this concept was also used to account for the logic of relative pronouns (cf. *Summa Logicae* I, chap. 76), which is also a topic of major interest for semanticists nowadays.

Medieval logicians supply enumerations of logical forms of propositions and of the corresponding kinds of supposition that attempt to be as complete as possible, though obviously none of them is exhaustive. In the present reconstruction, I shall treat some of the rules presented by Ockham, in particular those related to the most often occurring propositional forms, which shall be sufficient to give the reader a hint of the general principle at stake. On the basis of the same general structure, the procedure could in principle be expanded so as to include other propositional forms.

Once again, when dealing with the syntactical structure of a proposition, what I call its propositional form, it is obvious that what is said about $\mathbf{P_1}$ holds for every equiform occurrence of $\mathbf{P_1}$ (in modern terms, for all tokens of the given type). So we will be dealing exclusively with types in this section.

On the basis of the rules quoted in section 1.4.1.2, I now present an account of the possible logical forms of propositions, and of the appropriate modes of personal supposition for each of them.

The scope of a syncategorema. Given the threefold logical form of categorical propositions, a syncategorematic term may be introduced.

– before the subject
– before the copula[127]
– before the predicate

[127] Grammatical detail: in Latin the negation typically comes before the copula, that is '… non est…', unlike the English structure '… is not…'. It is also possible in Latin that the negation comes after the copula, '… est non-…', but this structure is equivalent in English to '… is non-…'.

We shall define that scope is maximal with respect to left-hand association. So a syncategorema placed before the subject has both subject and predicate under its scope. A syncategorema placed before the copula or the predicate has only the predicate under its scope. Let (S) be any syncategorema. '**P**' names the proposition of which the relation of scope in question holds.

Definition 1.5.2.2.1 Scope of syncategorema

$$\mathbf{P}\text{: } (\mathrm{S})a \circ b => <(\mathrm{S});\ a, b>_{\mathrm{P}}$$

$$\mathbf{P}\text{: } a(\mathrm{S}) \circ b => <(\mathrm{S});\ b>_{\mathrm{P}}$$

$$\mathbf{P}\text{: } a \circ (\mathrm{S})b => <(\mathrm{S});\ b>_{\mathrm{P}}$$

We shall say that a categorema is immediately under the scope of a syncategorema if the former immediately follows the latter, and that a categorema is mediately under the scope of a syncategorema if the former mediately follows the later (i.e., when either the copula or another term stand between them).

$$\mathbf{P}\text{: } (\mathrm{S})a \circ b => \mathbf{I}<(\mathrm{S});\ a>_{\mathrm{P}} \text{ and } \mathbf{M}<(\mathrm{S});\ b>_{\mathrm{P}}$$

$$\mathbf{P}\text{: } a(\mathrm{S}) \circ b => \mathbf{M}<(\mathrm{S});\ b>_{\mathrm{P}}$$

$$\mathbf{P}\text{: } a \circ (\mathrm{S})b => \mathbf{I}<(\mathrm{S});\ b>_{\mathrm{P}}$$

Moreover, let us add the abbreviation

$$\mathbf{O}<(\mathrm{S});\ a>_{\mathrm{P}}$$

meaning that a is **only** under the scope of (S). This is important to safeguard the effect of 'weak' syncategoremata, in particular $\exists$. In first instance, we shall treat the basic types of syncategoremata: signs of universality – represented by '$\forall$' – signs of particularity – represented by '$\exists$' –, and the negation – represented by '$\sim$'. The absence of any sign shall be represented as '$\{\}$'. We shall start by the positive fragment of the language, that is, negation will be dealt with only at a later stage.

Let a be an incomplex (i.e., composed only of categorematic terms) common term. Then the rules are:

(rule 1) $\mathbf{O}<\{\}: a>_{\mathrm{P}} => \mathrm{PdSUP}(a, \mathbf{t})_{\mathrm{P}}$
(rule 2) $\mathbf{IO}<\exists: a>_{\mathrm{P}} => \mathrm{PdSUP}(a, \mathbf{t})_{\mathrm{P}}$
(rule 3) $\mathbf{IO}<\forall: a>_{\mathrm{P}} => \mathrm{PcdSUP}(a, \mathbf{t})_{\mathrm{P}}$
(rule 4) $\mathbf{MO}<\forall: a>_{\mathrm{P}} => \mathrm{PmcSUP}(a, \mathbf{t})_{\mathrm{P}}$

Negation: As we shall see, the treatment of the negation requires more ingenuity than the treatment of the positive sincategoremata. In fact, Ockham does not offer an explicit rule concerning the effect of the negation over a term that, without the negation, would have determinate supposition. But an explicit formulation thereof can be found in Buridan:

> A negating negation distributes every common term following it that without it would not be distributed and does not distribute anything that precedes it.[128] (Buridan 2001, 269)

Let A be any term, either complex or incomplex (but with no negation signs): '$\sim$' stands for any negation sign. Given a proposition **P**, let **P*** be the proposition resulting of the introduction of one negation sign in any position of P, such that A is in its scope. Buridan's rule can be formulated as follows:

(rule 5) $\text{PdSUP}(A, \mathbf{t})_{\text{P}}\ \&\ \langle\sim;\ A\rangle_{\text{P}*} \Rightarrow \text{PcdSUP}(A, \mathbf{t})_{\text{P}*}$
(rule 6) $\text{PmcSUP}(A, \mathbf{t})_{\text{P}}\ \&\ \langle\sim;\ A\rangle_{\text{P}*} \Rightarrow \text{PcdSUP}(A, \mathbf{t})_{\text{P}*}$

There is, however, a serious problem concerning the effect of a negating sign upon a term that, without the negation, would have confused and distributive supposition. Given the structure of the theory, it might seem impossible to provide a general rule for negation and confused and distributive supposition, for the following reason. Consider the four traditional kinds of categorical propositions (Figure 1.5.2.2).

'Some *a* is *b*' (1) should be equivalent to 'Not: No *a* is *b*' (2′) (the contradictory of (2)) and 'Every *a* is *b*' (3) should be equivalent to 'Not: Some *a* is not *b*' (4′) (the contradictory of (4)). If these equivalences hold, then the supposition of the terms in (1) and (2′) should be the same: *a* and *b* have determinate supposition in (1), so they should have the same kind of supposition in (2′).

For this to happen, the effect of the negation in (2′) should be to turn the confused and distributive supposition of *a* and *b* in 'No *a* is *b*' into determinate supposition. This is indeed the rule proposed by Ockham:

> Nevertheless, it should be noted that the aforementioned rules hold only in the case where the term in question would not stand confusedly and distributively if the negation sign or the relevant verb or name were taken away. For if the term were to stand confusedly and distributively when one of these expressions [negation sign] were taken away, then with the addition of such an expression it would stand determinately. This is clear in 'Socrates is every man'. Here the predicate 'man' stands confusedly and distributively;

Some *a* is *b* (1) —is the contradictory of→ No *a* is *b* (2) (contradiction 1)

Every *a* is *b* (3) —is the contradictory of→ Some *a* is not *b* (4) (contradiction 2)

Figure 1.5.2.2. Four traditional kinds of categorical propositions

[128] Negatio negans distribuit omnem terminum communem sequentem eam qui ea remota non esset distributus et nihil distribuit quod praecedit eam. (Buridan 1998, 57)

therefore, if a negation precedes it, it stands determinately, as in ‘Socrates is not every man’; for if Socrates is not that man (where any man is referred to), it follows that he is not every man.[129] (Ockham 1998, 214)

Ockham’s rule can be formulated as follows. Let $\sqsupset$ be a distributive sign (i.e., either a negation or a universal sign). Given a proposition P, let P# be the proposition resulting of the introduction of one distributive sign at the beginning of P, so that ‘*A*’ be under its scope. Then:

(rule 7o) $\mathrm{PcdSUP}(A, \mathbf{t})_{\mathrm{P}}\ \&\ <\sqsupset;\ A>_{\mathrm{P\#}} => \mathrm{PdSUP}(A, \mathbf{t})_{\mathrm{P\#}}$

But what about the equivalence between (3) and (4′)? In (3) *a* has confused and distributed supposition and *b* has merely confused supposition. So the same should occur in (4′). However, in ‘Some *a* is not *b*’, *a* has determinate supposition and *b* has confused and distributive supposition. According to rule 5, the negation would make *a* have confused and distributive supposition in (4′), that is, the same supposition of *a* in (3). But what about *b*? According to the rule proposed by Ockham, since it has confused and distributive supposition in ‘Some *a* is not *b*’, it would have **determinate** supposition in (4′), under the effect of the negation. But in fact it ought to have **merely confused** supposition, because of the equivalence between (3) and (4′). So the rule stated by Ockham does not safeguard this equivalence.

Buridan, on the other hand, presents a rule that does safeguard the equivalence between (3) and (4′):

A common term is confused nondistributively by two distributive [parts of speech] preceding it, either of which would distribute it without the other.[130] (Buridan 2001, 275)

Buridan’s rule can be formulated as follows:

(rule 7b) $\mathrm{PcdSUP}(A, \mathbf{t})_{\mathrm{P}}\ \&\ <\sqsupset;\ A>_{\mathrm{P\#}} => \mathrm{PmcSUP}(A, \mathbf{t})_{\mathrm{P\#}}$

That is, under the effect of two negations, *b* in (4′) would have merely confused supposition, which is the desired result. But then the equivalence between (1) and (2′) would not be preserved anymore: *b* would have merely confused supposition in (2′), whereas it ought to have determinate supposition, as in (1).

Thus, each of these rules is unable to preserve both equivalences: a rule concerning the effect of the negation upon a term originally with confused and distributive supposition either preserves the equivalence between (1) and (2′) (Ockham’s rule) or it preserves the equivalence between (3) and (4′) (Buridan’s rule). This is due to the following asymmetry: in the case of contradiction 1, the opposition in the supposition of

[129] Verumtamen sciendum est quod praedicta regulae verae sunt quando sine negatione, vel tali verbo vel nomine dempto, praedictus terminus non staret confuse et distributive, quia si aliquo praedictorum dempto terminus staret confuse et distributive, tunc per adventum talis dictionis idem terminus staret determinate. Sicut patet in ista ‘Sortes est omnis homo’, hoc praedicatum ‘homo’ stat confuse et distributive. Ideo si praecedat negatio, stabit determinate, sicut patet sic dicendo ‘Sortes non est omnis homo’; nam sequitur ‘Sortes non est iste homo’, quocumque homine demonstrato, ‘igitur Sortes non est omnis homo’. (Ockham, *Summa Logicae I*, chap. 74 (23–31))

[130] Terminus communis confunditur non distributive per duplex distributivum antecedens ipsum quorum utrumque distribueret ipsum sine reliquo. (Buridan 1998, 63)

b in each proposition is between determinate supposition and confused and distributive supposition, whereas in contradiction 2 the same opposition is between merely confused supposition and confused and distributive supposition. Therefore, it would seem impossible to provide a homogeneous account of the effect of the negation (or other distributive term) upon terms with confused and distributive supposition. However, in the writings of the early 17th century philosopher and theologian John of St. Thomas, one finds a precise and correct account of the effect of distributive terms upon terms already having confused and distributive supposition. For this purpose, one has to consider the whole propositional context, that is, the supposition of the other term in the proposition. Here is how John formulates it:

> If two universal signs simultaneously affect the same term, then you must see how it remains after the first negation or universal sign is removed; and if it remains distributed with reference to a term having determinate supposition, then it originally had confused supposition; if however the term remains distributive with reference to a term having confused supposition, it originally was determinate. For example, if I said, *No man is not an animal*, then when the first negative, i.e. the *no*, is taken away, *animal* becomes distributed with reference to *man*, which is determinate. Thus originally *animal* had confused supposition. However, if I said, *Not every man is an animal*, then when I take the *not* away, *man* becomes distributed with reference to *animal* which is confused. And thus *man* originally had determinate supposition.[131] (John of St. Thomas 1955, 69)

Making use of the symbolism introduced here, these rules can be formulated as:

(rule 7o′) $\mathrm{PcdSUP}(A, \mathbf{t_1})_{\mathrm{P}} \;\&\; \mathrm{PmcSUP}(B, \mathbf{t_2})_{\mathrm{P}} \;\&\; <\sqsupset; A>_{\mathrm{P\#}} => \mathrm{PdSUP}(A, \mathbf{t_1})_{\mathrm{P\#}}$

(rule 7b′) $\mathrm{PcdSUP}(A, \mathbf{t_1})_{\mathrm{P}} \;\&\; \mathrm{PdSUP}(B, \mathbf{t_2})_{\mathrm{P}} \;\&\; \sqsupset; A>_{\mathrm{P\#}} => \mathrm{PmcSUP}(A, \mathbf{t_1})_{\mathrm{P\#}}$

If a uniform account of the effect of distributive terms upon terms already having distributive and confused supposition could not be provided, this would be a grave drawback for the theory of supposition as a whole. Apparently, at the time of Ockham and Buridan a solution for this issue had not yet been found; however, later authors such as John Dorp (cf. Karger 1993) and John of St. Thomas were clearly aware of the problem, and succeeded in finding appropriate rules to deal with it. Clearly, many other cases may seem problematic and appear to be, at first sight, unaccountable for within supposition theory; but the reformulation of the rules for confused and distributive supposition above shows that the supposition framework is more resourceful than one might expect at first sight, allowing for constant refinement.[132]

---

[131] Si concurrant duo signa universalia super eundem terminum, videndum est, quomodo, dempta prima negatione aut signo universali, remaneat; et si remaneat distributus in ordine ad terminum supponentem determinata, antea supponebat confuse; si autem remanet distributive in ordine ad supponentem confuse, antea stabat determinate. Ut si dicam: ‘Nullus homo non est animal’, dempta prima negatione, scilicet ly nullus, ly animal remanet distributum in ordine ad ly homo, qui stat determinate, et sic antea ly animal supponebat confuse. Si autem dicam: ‘Non omnis homo est animal’, dempto ly non remanet ‘homo’ distributus in ordine ad ly animal stans confuse, et sic antea supponebat determinate. (John of St. Thomas 1929, 35 (6–26))

[132] See (Klima and Sandu 1991) for the use of supposition theory to account for complex quantificational cases.

1.5.3 Examples

How this machinery is applied to the analysis of propositions can be made clearer by means of examples.

1 – 'Noun has four letters' = 'Noun is four-lettered'

The subject is of type λ, and so is the predicate. According to the table above, this proposition belongs to propositional schema 5. Accordingly, two interpretational schemata are possible: (5.a) both terms have personal supposition; (5.b) subject has material supposition and predicate has personal supposition.

(5.a) By rule 1, both terms have personal determinate supposition, thus the proposition belongs to interpretational schema (14). Therefore, it asserts that there is a $\mathbf{t_1}$ such that SIG('noun', $\mathbf{t_1}$), and a $\mathbf{t_2}$ such that SIG('four-lettered', $\mathbf{t_2}$), such that $\mathbf{t_1} = \mathbf{t_2}$.

In other words, by means of this proposition, and under this interpretation, it is asserted that some noun has four letters (indeed, this is true of many nouns, including the noun 'noun').

(5.b) On this interpretational schema, the predicate has personal supposition. Again according to rule 1, it has personal determinate supposition. As for the subject, it supposits for a word, but not insofar as words are the *significata* of 'noun' (which is the case). In fact, the most obvious possibility is that it supposits for (occurrences of) the word 'noun'.[133] Therefore, what the proposition asserts under this interpretation is that it there is a $\mathbf{t_2}$ such that SIG('four-lettered', $\mathbf{t_2}$), such that 'noun' is identical to $\mathbf{t_2}$ – in other words, that the word 'noun' has four letters (which is indeed the case).[134]

2 – 'Not: every man is white.'

The subject and predicate are of type $\varkappa$. Therefore, the proposition belongs to propositional schema 1, that is, both subject and predicate can only have personal supposition.

Let us now see which modes of personal supposition the terms in this proposition have. First, we consider the supposition of the terms disregarding the negation, that is, in the proposition 'Every man is white': according to rules 3 and 4, respectively, the subject has confused and distributive supposition and the predicate, merely confused supposition. Now, with the negation added, the subject has then determinate supposition (rule 7o′), and the predicate has confused and distributive supposition (rule 6).

Therefore, the proposition belongs to the interpretational schema (12′). Hence, what the proposition asserts is that there is a $\mathbf{t_1}$ such that SIG('man', $\mathbf{t_1}$), and for all

[133] But notice that material supposition is far more complicated than this. See (Normore 1997), (Read 1999) and (Panaccio and Perini-Santos 2004).

[134] But notice that if the sentence was 'Noun has three letters', both readings would be equally legitimate, but the proposition would be true only under the first reading, as there exist indeed nouns with three letters.

$\mathbf{t_2}$ such that SIG('white', $\mathbf{t_2}$), $\mathbf{t_1} \neq \mathbf{t_2}$.[135] In other words, it asserts that one man is distinct from all white things, that is, he is not white – indeed, this proposition is equivalent to 'Some man is not white', according to the square of opposites; this fact is thus confirmed by the application of the suppositional machinery.

1.6 CONCLUSION

I hope to have made a convincing case for my main claim, namely that (Ockham's) supposition theory is best viewed as a theory of algorithmic hermeneutics. The historical and conceptual analysis aimed at arguing that it is a theory for the interpretation of propositions, that is, for the generation of their meanings, while the formalization was intended to show that this theory is essentially formal. While syntactic, semantic and pragmatic elements are involved, they can nevertheless all be treated with recognizably formal tools and procedures, in such a way that the apparatus of the theory thus described could, in principle, be applied by a machine.

It is not entirely clear the extent to which this description applies to other theories of supposition. For instance, as already mentioned, 13th century theories of supposition were not geared towards the distinction of propositions, as Ockham's is; other theories, such as Buridan's, seem to rely rather on non-formal, pragmatic elements. But I take this to be a sign of superiority of Ockham's theory *vis-à-vis* the others, the fact that it can be shown to be formal in this sense.

In any case, and this holds of supposition theories in general, it seems to me that the first step towards a better grasp of them is the emphasis on their procedural nature: they are essentially sets of instructions on how to interpret propositions, a point that seems to have been overlooked in previous formalizations.

[135] Notice that the negation also changes the polarity of the relation of equality; it becomes a relation of inequality.

PART 2

BURIDAN'S NOTION OF *CONSEQUENTIA**

2.1 INTRODUCTION AND HISTORY

2.1.1 Introduction

The concept of *consequentia* is among the most interesting medieval contributions to the development of logic. Even though investigations on the relations of entailment between propositions can be found in earlier theories (i.e., in Aristotle's *Prior Analytics* and in Stoic logic), these investigations did not reach the level of generality and systematicity of medieval theories of *consequentia*. True enough, after roughly two centuries of remarkable developments, this concept virtually disappeared from the philosophical scene for many centuries, to be rediscovered only in the 19th century by Bernard Bolzano. However, in spite of the lack of continuity between medieval and modern discussions on the concept of consequence, these medieval discussions seem to have much to contribute to the modern discussions; there is a great deal of similarity between them, but also of dissimilarity, and precisely for this reason the medieval theories may offer a refreshing vantage point for modern analyses of the concept of consequence.

In this part, I will focus on Buridan's theory of *consequentia*, one of the most significant and influential theories of this sort. Two aspects of this theory seem particularly worth examining: how he incorporates his commitment to tokens as truth-value bearers to his definition of consequence, and his substitutional notion of formal consequence. As a matter of fact, each of these aspects is strikingly similar to some recent lines of research, namely investigations on two-dimensional semantics and on the concept of logical consequence, respectively. For this reason, in what follows, some of the two-dimensional semantics apparatus will be used to account for Buridan's theory of consequence. Moreover, the comparison between Buridan's notion of formal consequence and some of the views on logical consequence that have been proposed recently turns out to be fruitful for both ends. Hence, in section 2.3 a comparative assessment of Buridan's theory and these modern theories is drawn. Finally, the last

* This part is largely based on (Dutilh Novaes 2004c), (Dutilh Novaes 2005a) and (Dutilh Novaes 2005c).

part of this chapter consists of an application of Buridan's concept of consequence to a fragment of his logical system, namely the inferential relations between some categorical propositions defined on the basis of the modes of personal supposition of their terms. With some rudiments of model theory, it will be shown that this fragment of Buridan's system is sound, that is, the rules syntactically defined are also semantically valid.

In sum, the aim of this part is again double-sided: by means of formalizations, I intend to attain a deeper understanding of Buridan's theory, but also to show its relevance to current investigations on token-based semantics and on the concept of logical consequence.

2.1.2 History of the notion of consequence

By now, the rough lines of the development of the medieval notion of consequence have been established (cf. Green-Pedersen 1984, chap. E), but some of its details remain to be clarified. The story begins, as almost always in logic, with Aristotle. For Aristotle and the Aristotelian tradition, the model of logical validity was that of syllogistic validity, according to the patterns described in the *Prior Analytics*. But the syllogistic system is a clear case of undergeneration[136] w.r.t. the intuitive notion of logical validity:[137] all valid syllogistic patterns are indeed intuitively valid, but the group of valid arguments described by syllogistic is but a very small subset of the set of intuitively valid logical arguments.

Besides syllogistic, Aristotle also produced what came to be known as the theory of Topics. *Topics* was the fifth book of Aristotle's *Organon*, a book that can be described as a rather loose collection of rules for the conduction of non-demonstrative reasoning and argumentation, whose logical status is very much inferior to that of syllogisms; for Aristotle, their application could at best generate reasonable dialectical arguments. In fact, the theory presented in the *Topics* was considered to be an art, not having the scientific status attributed to syllogisms. Yet, in the post-Aristotelian tradition, in particular among the Latin rhetoricians, the Topics remained an important object of study, cf. (Green-Pedersen 1984).

An addition to the Aristotelian framework introduced by Boethius (in the 5th–6th centuries), which was to be influential later in the Middle Ages, was the idea of hypothetical syllogisms[138] (the original Aristotelian system only dealt with categorical syllogisms). Indeed, the first medieval discussions related to the notions of consequence and entailment were prompted by reflections concerning hypothetical

[136] The concepts of undergeneration and overgeneration are borrowed from (Etchemendy 1990), and will be particularly important in section 2.3.2.1 below.
[137] Although this did not appear to be Aristotle's opinion; he seemed to think that every valid argument could be reduced to one of the valid syllogistic forms. Cf. (Smith 1995, 30)
[138] Arguably under the influence of Stoic logic, cf. (Boh 1982, 303).

syllogisms, and, to a minor extent, by the Topics.[139] Abelard (early 12th century) is probably the first medieval logician to have developed an early version of what we could call 'propositional logic' (cf. Tweedale 1982); moreover, he also introduced the famous distinction between perfect and imperfect inferences.[140] Two aspects of this distinction are worth noting: the autonomy of syllogisms and the idea that there are two 'sorts' of inferences. The first aspect would not always be unanimously accepted in later developments, but the division of consequences in two kinds is one of the most important traits of the evolution of this concept.

In the 13th century, there were no independent treatises on consequences yet,[141] and the precise historical origin of 14th century theories of consequence is still controversial among specialists. It is still something of a mystery why and how, all of a sudden, at the beginning of the 14th century, treatises bearing the title *De consequentiis* ((Burley 1980) is one of the first of such treatises) or the like began to appear. Why then, and not before? Naturally, the subject itself, that is, the logical and inferential relations between propositions, was very often discussed by earlier authors; the very term '*consequentia*' was in constant usage (in the same sense) since at least the 12th century, and dates as far back as Boethius in the 5th–6th century (cf. Boh 1982, 302). But no treatises or chapters were specifically dedicated to the topic or bore such titles before the 14th century.

According to an influential hypothesis concerning the origin of theories of consequences, they stemmed essentially from the tradition on the Topics (cf. Bird 1961; Stump 1982a). At first sight, this hypothesis makes good sense: in the tradition of Aristotelian logic, the role of the Topics was often that of accounting for the patterns of (correct) inference and reasoning that did not fit into the syllogistic system presented in the *Prior Analytics*. While it is a wonder of systematicity and formality, syllogistic is not a very wide-ranging theory in that it accounts for only a small portion of the patterns of reasoning that we are prepared to accept as valid. The Topics, even though not as rock-solid as syllogistic patterns, provided an account of many more of such patterns of reasoning. So, conceptually, it would seem quite natural that the tradition on the Topics would be at the origin of theories of consequences, as these are essentially theories about the relations between propositions that go beyond the patterns recognized by syllogistic. Moreover, some earlier investigations on the notion of consequence were made explicitly within the context of an analysis of the Topics; Abelard's sophisticated theory of the logical relations between propositions is to be found precisely in the part of his *Dialectica* dedicated to the Topics.

However, this hypothesis did not receive the historical confirmation that one could have expected. It has been shown (Green-Pedersen 1984, 270) that the late 13th century literature on the Topics, that is, the period immediately preceding the emergence

---

[139] This is the case in particular of Garlandus Compotista and Peter Abelard. 'Garlandus is interested in the Topics because he thinks they are useful in the study of hypothetical syllogisms, which appear to be his main interest in the *Dialectica*'. (Stump 1982a, 276)

[140] (Abelard, *Dialectica* 253/4, 31–1); (Abelard, *Dialectica* 256–7, 35–1); (Martin 1986, 566).

[141] Cf. (Green-Pedersen 1984, chap. E).

of treatises on consequences, gives absolutely no clue of what was to come; that is, there is no significant similarity between the contents of these 13th century treatises on the topics and 14th century treatises on consequences. Therefore, it has been concluded that the topics could not have been the main source for 14th century theories of consequences.[142]

Although our current state of knowledge on the matter still does not allow for a conclusive account of these developments, the picture that at this point seems more plausible is that different strands of traditional Aristotelian logic converged in order to give rise to the 14th century theories of consequences. It seems that at least three other traditions contributed to the development of theories of consequence: treatises on *syncategoremata*, especially in connection with the syncategorema '*si*' (corresponding to the 'if … then' structure in English); the analysis of hypothetical syllogisms (Boethius's *De hypotheticis syllogismis* is referred to 6 times in Burley's *De puritate artis logicae* – cf. Green-Pedersen 1984); and commentaries on the *Prior Analyticts* – indeed, it is in the *Prior Analytics* that Aristotle explicitly states a formulation of the notion of 'following' that is arguably the (remote) source for the most fundamental definition of consequence in the 14th century.[143]

Be that as it may, the importance of the Topics for the development of 14th century theories of consequences should also not be altogether dismissed. It is worth noticing that two of the first authors having written explicitly on consequence, Ockham and Burley, are both in some way or another influenced by the Topics. Burley explicitly says that all valid consequences are based on dialectical Topics (Burley 2000, 158). By contrast, the relation of Ockham's theory of consequence to the Topics is more convoluted; Green-Pedersen argues convincingly that Bird's reconstruction of Ockham's theory within the framework of the Topics (Bird 1961) is not satisfactory (Green-Pedersen 1984, 268), but he also confirms that Ockham's 'intrinsic' and 'extrinsic' middles, crucial concepts for his theory of consequence, are concepts essentially taken (albeit heavily modified) from the topical framework.

One important distinction introduced by Ockham (although related to Abelard's distinction between perfect and imperfect inferences) is that between formal and material consequences.[144] The criterion used by him to draw the distinction was still essentially based on the (heavily modified) topical framework (the notions of intrinsic and extrinsic middles), and was soon discarded by authors such as Buridan and Pseudo-Scotus. But the idea that there are basically two kinds of consequences, and that what distinguishes them is something related to their form, was to remain influential throughout the Middle Ages, and is still influential now (cf. Read 1994, Brandom 2000).

---

[142] Chapter E of (Green-Pedersen 1984), on Topics and the theory of consequence, is the most comprehensive survey of these developments that I am aware of.

[143] 'A deduction is a discourse in which, certain things being stated, something other than what is stated follows of necessity from their being so'. *Prior Analytics* 24^{b}19–20.

[144] Cf. (Ockham, *Summa Logicae*, III-3, chap. 1, 45–58 (p. 589)).

In any case, the presence of topical notions did not last for long: just a few years later, in Buridan's treatise[145] (written in the 1330s) and in Pseudo-Scotus's commentary on the *Prior Analytics* (written around 1350 – Pseudo-Scotus 2001), the validity of *consequentiae* was no longer related to Topics;[146] it finally acquired a status of its own, independent of the Aristotelian notions of syllogism and Topics.

In sum, while we are not yet able to reconstruct a complete history of the development of these theories, at this point it seems that the most plausible hypothesis is that at least these four traditions – Topics, *syncategoremata*, hypothetical syllogisms and *Prior Analytics* – must be taken into account to explain the rise of theories of consequences in the 14th century. Different aspects of each of these traditions contributed to the development of different aspects of the theories of consequence.[147] Green-Pedersen (1984, 295) argues for example, that the late 13th century treatises that most resemble early 14th century treatises on consequences are 'the treatises on syncategorematic words and a number of sophism-collections arranged after syncategoremes'.

For most medieval authors after Ockham, the primitive notion of consequence was that of a material consequence, usually expressed in terms of the 'modal', or incompatibility criterion: a consequence holds iff it is impossible for the antecedent to be the case while the consequent is not the case. Moreover, a sub-group of the (materially) valid consequences was thought to satisfy not only the modal criterion, but also another, stricter criterion, such that its members were said to be formally valid consequences. The different authors did diverge, however, with respect to the criterion differentiating a formal from a material consequence. Some defined formal consequences on the basis of semantic criteria,[148] others in terms of epistemic notions.[149]

Indeed, the second half of the 14th century, especially in England, witnessed the predominance of what we could call an epistemic view of logic,[150] and in particular of the notion of consequence. According to authors such as Strode and Billingham, a consequence was formal if the understanding of the consequent was somehow included in the understanding of the antecedent, in such a way that whoever acknowledged the truth of the antecedent could not but acknowledge the truth of the consequent.[151]

Eventually, as argued by C. Normore (1993), this epistemic notion of formal consequence paved the way to the Cartesian views on inference, no longer based on formal and logical features but rather on psychological elements such as intuitions and perceptions. Ironically, even though its original inspiration may have been the

[145] (Hubien 1976) – henceforth TC.
[146] Cf. (Boh 1982, 307–310).
[147] For several of these tentative connections, see chap. E of (Green-Pedersen 1984).
[148] The early tradition, represented by Ockham and Burley.
[149] The later, predominantly English, tradition, represented by Ralph Strode and Richard Billingham, cf. (Boh 2001); (Billingham 2003).
[150] Cf. (Boh 2000).
[151] Cf. (Boh 2001).

medieval epistemic notion of formal consequence, the influential status of Descartes' philosophy in the centuries to follow appears to have been one of the reasons why the notion of consequence as independent from the mental acts related to it remained virtually forgotten until the 19th century.

But the notion of formal consequence that will be of interest for the present discussion, the one advocated by Buridan, is what we could call a substitutional notion: a consequence is formal iff all substitutional instances of its categorematic (i.e., non-logical) terms yield material consequences. This notion was influential during the first half of the 14th century, but eventually lost terrain to the epistemic notion; as just said, it was only with Bolzano in the 19th century that it was rediscovered. Now, however, its descendants (in particular the model-theoretic notion of logical consequence) are almost unanimously viewed as the most accurate account of (logical) consequence.

2.2 INFERENCE AND CONSEQUENCE

Clearly, the concepts of consequence and inference are among the most important issues in the philosophy of logic of all times – after all, what is logic if not the science regulating the correct inference from known contents to new contents? Currently, discussions on them abound, but these discussions usually assume proposition-types to be the bearers of truth-values. In order to challenge this assumption, in what follows I propose to examine the concepts of consequence and inference within the framework of Buridan's token-based semantics; it will turn out that significant amendments to the usual definitions of consequence and inference are made necessary by the switch to the token perspective, insofar as pragmatic elements play a crucial role in the behavior of proposition-tokens. In fact, even if in modern logic we still observe the predominance of the type perspective, logicians and especially philosophers of language are increasingly aware of the fruitfulness of ascribing truth-values to tokens and not to types. In particular, phenomena such as the Liar paradox have made it patent that the token perspective is, at least in such cases, imperative – see (Gaifman 1992), (Kripke 1975).

In other words, beyond the historical interest, Buridan's theory of *consequentia* is indeed a fruitful starting point for the analysis of the concepts of inference and consequence also from a modern perspective. At first sight, Buridan's modal definition of *consequentia*[152] appears to be very similar to the (modal) model-theoretic notion of consequence,[153] based on the concept of truth/satisfiability in a model: the consequent follows from the antecedent iff it is impossible for the antecedent to be true while the consequent is false, that is, iff there is no situation (model) in which the antecedent is true and the consequent is false. On the basis of this

[152] A terminological clarification: I will be using the terms '*consequentia*' for the medieval notion and 'consequence' for the modern notion, except when there is no risk of ambiguity, in which case I may use the term 'consequence' while referring to Buridan's notion too.

[153] Whether Buridan advocates a representational or an interpretational notion of consequence, to use Etchemendy's terminology in (Etchemendy 1990), will be discussed below.

definition and the modes of personal supposition Buridan can, for example, give an account of the inferential relations between what we could call doubly-quantified propositions.[154]

However, Buridan advocates a strict commitment to what we now call proposition-tokens as the bearers of truth-value – more specifically, actually formed occurrences of propositions (either spoken, written or mental),[155] and this fact obliges him to refine his formulation of the (modal) definition of consequence. Roughly, the main sources of difficulty are propositions such as 'No proposition is negative', 'Some proposition is affirmative', which, when formed, interfere with their own truth-value. For this reason, Buridan has to eventually abandon the standard definition of consequence based on the truth-values of the propositions involved.

Inspired by Buridan's writings,[156] A. Prior (1969) has approached the phenomenon of propositions interfering with their own truth-value by coining the notions of possibly-true, necessarily-true, etc., as opposed to merely possible, necessary, impossible etc. Although Buridan's notion of consequence was not Prior's main concern in that article, one of his conclusions was that it cannot be correctly formulated in terms of the notions of impossibly/possibly-true, since some criterion other than their truth-values is needed to establish the (actual) modal value of propositions (with respect to each other) in order to assess the validity of the relation of consequence between them.[157] Prior's main idea was that it was necessary to differentiate the situation **in** which a proposition is true or false (namely, a situation in which it is actually formed) from the situation **of** which a proposition is true or false (namely, a situation in which it is not necessarily formed).

Here, the ideas of Prior and the general framework of two-dimensional semantics (cf. Chalmers 2003) – in particular views put forward by Kaplan and Stalnaker – serve as inspiration for the treatment of Buridan's notion of consequence within the framework of his token-based semantics. From Kaplan (1989), I borrow the distinction between the context of formation and the context of evaluation of a proposition, which I represent in two-dimensional matrices. Inspired by Stalnaker's work (cf. Stalnaker 1999), I focus on the mutual relations between contents and contexts.

I will be examining Buridan's treatise on consequences[158] and some parts of the eighth chapter of his treatise on *Sophismata*. Moreover, insofar as Buridan's notion of consequence depends on modal concepts, I will incidentally offer an account of modalities within Buridan's token-based semantics, even though the main topic of this part is his notion of consequence. Beyond the historical endeavor of analyzing Buridan's theory, I also intend to give a general idea of what a theory of consequence

---

[154] Cf. section 2.4 below.
[155] Cf. (Klima 2004a).
[156] He was mainly inspired by Buridan's remarks on the first and second sophisms in the eighth chapter of Buridan's *Summulae de Dialectica*. (Buridan 2001)
[157] See (Buridan 2001, 954–958).
[158] (Hubien 1976). The translations of this text quoted here are, when available, by G. Klima, in (Klima 2004a), otherwise they are my own. See also my review of Klima (2004a), (Dutilh Novaes 2004d).

and inference must be like within a two-dimensional semantics, where tokens are the truth-value bearers (section 2.3.1).

2.2.1 Fundamental notions: tokens and types, inference, (formal) consequence, *consequentia*

I start with the notions of type and token. Strictly speaking, this distinction is at odds with Buridan's nominalistic ontology, since the only entities whose existence he accepts are individual entities, and types would be somewhat obscure universal entities. In effect, for Buridan, only concrete occurrences of spoken, written or mental expressions exist.[159] Thus, Buridan's semantics is token-based not only insofar as tokens are the truth-value bearers of his system, but also because types understood as universal entities do not even exist according to his ontological principles.

Moreover, it is not clear whether the modern distinction between types and tokens was fully available to Buridan.[160] But for the present purposes, it is nonetheless necessary to formulate the notion of type in terms of the logical apparatus and ontology at our disposal (i.e., Buridan's). Types cannot be seen as abstract, independent entities, but it seems acceptable to define an expression-type as the class of all occurrences (tokens) that are equiform. Equiformity is certainly not an unproblematic notion; it is not evident how the written expression 'dog' is 'equiform' to utterances of the same word, or to mental occurrences of the term, but for now we will have to rely on an intuitive and vague understanding of the notion of equiformity.[161]

In sum, what we have are particular occurrences of propositions – in modern terms, proposition-tokens – and classes of equiform occurrences of propositions – in modern

---

[159] Buridan's nominalism does not consist of denying the existence of any non-physical entity, as some 20th century versions of nominalism do. Mental propositions are simply individual accidents of the individual who produces them; their ontological status is not problematic as far as Buridan is concerned. In fact, mental propositions are more fundamental to Buridan's semantics than written or spoken propositions, but an extensive analysis of the issue of mental language falls out of the scope of the present investigation.

[160] Nor to any other medieval author, for that matter. Here is what Nuchelmans says of this distinction with respect to Ockham: 'That the borderline between a *propositio* in the token-sense and a *propositio* in the type-sense was rather vague is further shown by SL II, 9'. (Nuchelmans 1973). In this book, Nuchelmans makes extensive use of this distinction to investigate the ancient and medieval theories on propositions; but it seems that the distinction belongs to the conceptual apparatus used by Nuchelmans for these analyses rather than to the very framework of the theories being analyzed.

Notice though that Marsilius of Inghen's view that there are different kinds of material supposition (as much as there are different kinds of personal supposition) (cf. section 1.5.2 of the present work) can be seen as a forerunner of the type-token distinction.

[161] Actually, this notion can be defined within Buridan's framework in terms of the concept of **subordination** to mental propositions: a proposition-type would be the class of all written, spoken and mental propositions that are subordinated to the same mental propositions (for convenience, assume that subordination is a reflexive property). The problem then becomes that of finding an individuation principle for mental propositions. The general problem of determining criteria according to which tokens belong to the same type is indeed delicate, as indicated by the list of possibilities considered by Chalmers (cf. Chalmers 2003): the criteria could be orthographic, linguistic, semantic or epistemic.

terms, proposition-types. Every time I use the phrase 'proposition-type', it should be understood as the class of all equiform occurrences of a given proposition.

As for the concepts of consequence and inference, it is important to notice that they are often conflated in recent philosophy and logic. Insofar as a valid inference is (usually) related to a valid consequence, it is not difficult to see why such confusion would arise. But these are in fact different concepts, and their conflation can be dangerous – see (Haack 1982), (Sundholm 1998a), (Sundholm 1998b).

There seem to be two fundamental differences between the concepts of inference and of consequence: (i) the relata in a relation of inference are two or more asserted propositions, or judgements,[162] whereas a relation of consequence concerns unasserted contents; (ii) inference is an action, the act of inferring a certain piece of knowledge (the conclusion), usually linguistically expressed, from previously possessed knowledge (the premises), whereas a consequence does not depend on human actions or states of mind to exist: it is seen as an objectual relation between contents.

> The relata in logical consequence are propositions, whereas an inference effects a passage from known judgements to a novel judgement that becomes known in virtue of the inference in question. (Sundholm 1998b, 27)

The conflation between the notions of inference and consequence is a result of the general predominance of what we could call the 'realist' tradition in logic:

> In the realist tradition from Bolzano onwards, through Wittgenstein's *Tractatus* and Tarski, up to modern model theory, the validity of [inferences] has been reduced to the – logical – holding of the consequence between the propositions that serve as contents of the judgements that are, respectively, the premises and the conclusion of the inference. (Sundholm 1998a, 184)

But the discrepancies between the concepts of consequence and inference become almost impossible to ignore within a token-based semantics. Most importantly, inference concerns actually formed tokens – more precisely, asserted propositions – whereas consequence seems prima facie to be independent of the actual formation of the propositions involved. Indeed, in the present framework, consequence is understood as a relation between two or more proposition-types, that is, equiform classes of proposition-tokens.

Moreover, according to many authors, there are two kinds of consequence, material consequence and formal/logical consequence. Formal consequences would be those that are valid in virtue of their form, while material consequences are consequences that are valid but not in virtue of their form (according to most authors, among whom Buridan, the criterion of validity for a consequence other than its form is the modal criterion just mentioned); thus stated, this definition is not very informative, for

[162] I here adopt a rather strict, 'Fregean' notion of inference, as concerning exclusively judgements, that is, actually asserted propositions. Admittedly, in other contexts, the term 'inference' can be used to cover cases in which the assertion of premises, properly speaking, does not occur, as in *per impossibile* inferences of natural deduction. (I owe this point to an anonymous referee).

an appropriate definition of 'form' is still needed.[163] But assuming that an appropriate notion of form is available, one of the main issues in recent debates on the notion of (logical) consequence has been as to whether 'mere' material consequences could also be said to be consequences, or whether the criterion of formality was in fact a necessary condition for a consequence to hold. I here adopt the view that there are such things as 'material consequences', which are in fact the primitive kind of consequences, and that formal consequences are a particular subset of the set of valid consequences.[164] Besides, this view is certainly in conformity with Buridan's notion of consequence (and with that of medieval logicians in general, for that matter), as will become clear below.

As for Buridan's notion of *consequentia*, this notion can seem ambiguous from a modern perspective, corresponding to the modern notions of (material or strict) implication as well as to those of (logical) consequence and of inference between sentences.

> [...] I reply that a *consequentia* is never true or false unless it is; and thus the validity or truth of a consequence requires that its antecedent and consequent exist.[165] (Buridan 2001, 957)

Notice that Buridan mentions 'the validity **or** truth of a *consequentia*'.[166] A *consequentia* can only be said to be true if it is a proposition, since only propositions are truth-value bearers. In other words, we have here the notion of a conditional sentence 'If A then B', which is itself true or false, but properly speaking its elements 'A' and 'B' are formed but not asserted – it is the relation between them that is asserted.[167]

Besides being true or false, a *consequentia* can be valid or invalid. Thus, other than being seen as a proposition, a *consequentia* is also seen as a relation between propositions. It is, however, still unclear whether the notion of *consequentia* should correspond to the modern notion of inference or to the modern notion of consequence.

According to the definitions of inference and consequence that I have adopted, an inference is said to be valid or invalid – related to the act of drawing new information from previously known information – whereas a relation of consequence (understood as objectual) does or does not hold between propositions.[168] Given Buridan's commitment to tokens, it seems more natural that his notion of *consequentia* should

---

[163] Related to this issue is the notoriously complex problem of the boundaries between logical and nonlogical terms.

[164] This view is also defended, among others, by (Read 1994) and (Brandom 2000, chap. 1).

[165] … dico primo quod numquam consequentia est vera vel etiam false nisi ipsa sit; et sic ad hoc quod consequentia sit bona aut vera, oportet quod antecedens et consequens illius sint. (Buridan 2004, 145 (4–6)).

[166] P. King (2001, 121) claims that there are only a handful of passages in which consequences are called true or false, this passage being one of them. He uses this fact to support his view that *consequentiae* should not be identified with modern conditional sentences. I am however not concerned with this identification, since I focus on the difference between the notions of consequence and of inference with respect to *consequentia*, a distinction not discussed by King.

[167] Buridan was very well aware of the so-called 'Frege point', the distinction between force and content. See (Klima 2004b).

[168] Medieval authors '… say that consequences are 'legitimate (*bona*)', or that they 'hold (*tenet*)' or 'are valid (*ualet*)'. (King 2001, 122). This terminology reinforces the ambiguity with respect to the notions of

correspond to the modern notion of inference, insofar as, in both cases, the relata are actually formed propositions (tokens).

However, C. Normore (1993) has argued that the concept of *consequentia* only acquired the meaning of a mental performance in later developments (in particular with Ralph Strode, who seems to have written his treatise roughly 10 years after Buridan's).[169] By contrast, Buridan's notion of *consequentia* seems to embody a clear objectual character. Hence, his concept of *consequentia* seems on the one hand similar to the notion of inference as defined with respect to the relata involved – actually formed tokens – but, on the other hand, the objectual nature of the relation – with the term '*consequentia*' Buridan seems to refer to the very relation between propositions, and not to the corresponding act – points in the direction of the modern notion of consequence.

In sum, Buridan's use of the term *consequentia* is seemingly equivocal, and is related to at least those three conceptions – implication, inference and consequence. This makes the interpreter's task more complicated, but that does not mean that Buridan confuses the concepts themselves; Buridan's use of the term '*consequentia*' seems to be a case of terminological equivocation, but not of conceptual conflation. Indeed, '*consequentia*' is simply the general term used by Buridan and other medieval logicians to refer to different kinds of relations between (asserted and unasserted) propositions.

A *consequentia* is a hypothetical proposition; for it is constituted of several propositions united by the expression 'if' or by the expression 'thus', or equivalent ones.[170] (TC, 21, 8–11)

However, the divisions of the different kinds of *consequentia* – formal consequence, material consequence, *ut nunc* consequence – allowed for the distinction of the different concepts being referred to by the same general term. It is surprising that, in modern times, while we do have different names for each concept, recent developments in logic have conflated them to a much greater extent than what appears to be the case with medieval logicians, in spite of their equivocal use of the term *consequentia*.

2.2.2 Buridan's definition of consequence

In his treatise on consequences, Buridan makes successive attempts to find the correct definition of the notion of consequences, which are then falsified by counterexamples. I now examine each one of these attempts.

---

inference and of consequence, since 'valid' is the term usually used for inferences, whereas 'to hold' is usually used for consequences.

[169] 'Thus, Strode's view […] puts in play the idea that deduction is not an objective relation between abstract objects or sentences but a mental operation performed on the basis of what can be understood or imagined'. (Normore 1993, 450). Notice though that the intensional notion of consequence is not a 14th century invention: Abelard and Kilwardby already defended the view according to which, in a valid consequence, the consequent is 'contained' in the antecedent. (cf. Ashworth 2002)

[170] Consequentia autem est propositio hypothetica; constituta enim est ex pluribus propositionibus coniunctis per hanc dictionem 'si' uel per hanc dictionem 'ergo' aut aequivalentem.

2.2.2.1 First attempt

After preliminary remarks on the truth and falsity of propositions, Buridan sets out to give a correct definition of the notion of consequence. In the spirit of medieval philosophical practices, he contends that no theoretical investigation can begin without an accurate nominal definition of its subject matter.[171] His first attempt to provide a nominal definition of the notion of consequence runs as follows:

First Definition:

> [The terms] 'antecedent' and 'consequent' are said correlatively; therefore, they need to be described in terms of each other. Many people say of two propositions that one is the antecedent with respect to the other which cannot be true while the other is not true, so that every proposition is antecedent with respect to any other proposition which cannot be true without the other being true.[172] (TC, 21, 26–32)

Thus formulated, the definition certainly comes across as very familiar to the modern reader. The validity of a consequence seems to be defined in terms of the truth-values of antecedent and consequent. This can be spelled out in different ways, from truth tables to possible-world semantics.

Note on the use of possible-world semantics. The use of this 20th century logical apparatus to account for an aspect of medieval logic may seem inappropriate and anachronistic at first sight. But possible-world semantics will be taken here with no extra metaphysical assumptions. Moreover, medieval logicians made extensive use of the technical notion of *casus* in order to construe hypothetical and counterfactual situations for the purpose of logical analysis, in a way that is remarkably similar to some uses of possible-world semantics. In the present investigation, possible worlds are simply seen as different situations, or contexts; they are not necessarily complete descriptions of a state of affairs in that not all propositions must receive a truth-value. In sum, the term 'possible world' will be taken here as synonymous of 'context', 'situation' and '*casus*'.

This having been said, let $<\mathbf{W}, \mathbf{V}>$ be a possible-world model, in which **W** is a set of worlds w and **V** a set of truth-values, namely $\{T, F\}$. Propositions are seen as functions from possible worlds to truth-values.[173] Under these assumptions, Buridan's definition can be formulated as follows:

Definition 2.2.2.1 'φ, thus ψ' is a valid consequence iff it is impossible for φ to be true while ψ is false $<=>$ There is no w such that $\varphi(w) = T$ and $\psi(w) = F$.

---

[171] TC, 20/21, 3–6.

[172] Antecedens autem et consequens relatiue dicuntur ad inuicem; ideo per inuicem describi debent. Dicunt ergo multi quod propositionum duarum illa est antecedens ad aliam quam impossibile est esse ueram illa alia non existente uera et illa est consequens ad reliquam quam impossibile est non esse ueram reliqua existente uera, ita quod omnis propositio ad omnem aliam propositionem est antecedens quam impossibile est esse ueram illa alia non existente uera.

[173] But notice that propositions are here the very utterances and inscriptions, and not the abstract entities that are often taken to be their senses.

	w_1	w_2	w_3	w_4	...
φ	T	F	F	T	
ψ	T	F	T	F	
φ, thus ψ	M	M	M	M	
χ	T	F	F	T	
λ	T	T	F	T	
χ, thus λ	B	B	B	B	

Figure 2.2.2.1. Definition 2.2.2.1

The definition can also be represented in an equally familiar truth table-like diagram (Figure 2.2.2.1).

Note that, usually, the term used for a 'good' consequence is not 'true consequence' (only sometimes does Buridan use this phrase). In medieval parlance, the term often used is '*bona consequentia*', whence the letter 'B' standing for a 'good' consequence and 'M' ('*mala consequentia*') for a 'bad' consequence. Clearly, a consequence is valid if all its instances are valid, that is, if there is no world in which the antecedent is true while the consequent is false.

2.2.2.2 Second attempt

But Buridan quickly dismisses this (apparently correct) definition of consequence. Following the usual medieval procedure, he comes up with a counterexample, that is, a consequence that does not comply with the definition just given, but which is intuitively a valid consequence.

> But this description is defective or incomplete, for the following is a valid consequence: 'every man is running; therefore some man is running'; still, it is possible for the first propositions to be true and for the second not to be true, indeed, for the second not to be.[174] (TC, 21, 32–35)

That is, according to Buridan, 'every man runs, therefore some man runs' is a valid consequence,[175] and yet it is not always the case that, whenever the antecedent is true, the consequent is true, namely when the consequent does not exist (is not formed). We here encounter Buridan's commitment to tokens for the first time: a proposition can only be said to be true or false insofar as it exists, that is, if it is formed, but a consequence does not require that both antecedent and consequent be formed in every situation for it to be a valid consequence.

Hence, Buridan amends his definition in order to include the clause of the token's existence.

---

[174] Sed haec descriptio deficit uel est incompleta, quia hic est bona *consequentia* 'omnis homo currit; ergo aliquis homo currit', et tamen possible est primam esse ueram secunda non existente uera, immo secunda non existente.

[175] Assuming the existential import of universal propositions.

Second Definition:

Therefore, some people say that this description needs to be supplemented as follows: that proposition is the antecedent with respect to another proposition which cannot be true while the other is not true, when they are formed at the same time.[176] (TC, 21, 36–38)

In terms of the possible-world model defined above, the existence clause effectively means that the set of truth-values **V** must be complemented with the value 'undefined', represented by '*', for propositions which are not actually formed in a given possible world w. Alternatively, propositions can be seen as partial functions, with no truth-value (a truth-value gap), in the situations in which they are not formed. This may seem awkward at first sight, since, properly speaking, propositions which are not formed in a given situation are not propositions in that situation – they are nothing at all. But since Buridan clearly adopts a local perspective (truth being defined with respect to a situation/possible world), we are led to talk about propositions that do not exist in a certain situation, but do exist in others.[177]

Notation: φ is formed in w_n <=> $\varphi \in [w_n]$[178]

Thus, if $\varphi \notin [w_n]$, then $\varphi(w_n) = *$. Similarly, $\varphi \in [w_n]$ iff $\varphi(w_n) = \mathrm{T}$ or $\psi(w_n) = \mathrm{F}$. We can now incorporate the clause of the token's existence to the definition of consequence:

Definition 2.2.2.2: 'φ, thus ψ' is a valid consequence iff it is impossible for φ to be true while ψ is false, when both are formed simultaneously <=> 'φ, thus ψ' is a valid consequence iff there is a w_n such that $\varphi \in [w_n]$ and $\psi \in [w_n]$, and there is no w_n such that $\varphi(w_n) = \mathrm{T}$ and $\psi(w_n) = \mathrm{F}$.

Notice that, under this definition, what invalidates a consequence is not a situation in which either antecedent or consequent is not formed (where there is no instance of the consequence), since the validity of a consequence does not depend solely on the contingent existence of propositions, but rather a situation in which they are both formed and the antecedent is true while the consequent is false.

Again, we can represent the definition on a truth-table (Figure 2.2.2.2).

Note that, if one of the two tokens (antecedent or consequent) fails to be formed in a certain world, then an instance of the consequence does not exist in that world, but the consequence is not invalidated. In other words, a consequence is valid if all its instances are not invalid (i.e., are either valid or non-existent).

---

[176] Et ideo aliqui dicunt dictam descriptionem debere suppleri sic: illa propositio est antecedens ad aliam propositionem quam impossibile est esse ueram illa alia non existente uera illis simul formatis.

[177] In the same manner as we often talk about individuals that do not exist in a given possible world, in most versions of quantified modal logic.

[178] This notation simply reproduces the familiar notion of an object belonging (or not) to the domain of a given possible world. Here, the individuals we are dealing with are the very linguistic entities that I am referring to as 'propositions'.

	w_1	w_2	w_3	w_4	w_5	w_6	…	w_n
φ	T	F	*	F	*	T		T
ψ	T	T	T	F	*	F		*
φ, thus ψ	M	M	*	M	*	M		*
χ	T	F	*	F	*	T		F
λ	T	T	*	*	T	T		F
χ, thus λ	B	B	*	*	*	B		*

Figure 2.2.2.2. Definition 2.2.2.2

2.2.2.3 *Third and final attempt*

But again, this definition does not endure the scrutiny of counterexamples. Buridan presents two propositions that satisfy the conditions above and which, therefore, according to this definition, would be in a relation of consequence with each other, but which intuitively simply do not form a valid consequence.

> I still claim that this description is not valid, for the following is not a valid consequence: 'no proposition is negative, therefore no donkey is running', but on the basis of the given description one should accept it as valid; therefore etc.[179] (TC, 21/22, 38–42)

The putative consequence

(+)No proposition is negative, therefore no donkey is running.

complies with the definition above because 'No proposition is negative' is false every time it is formed (since its very existence falsifies it). Hence, if it is impossible for the antecedent to be true, it is *a fortiori* impossible for the antecedent to be true while the consequent is false. At first sight, this seems to be a case of *ex impossibili quodlibet*, a principle that Buridan is prepared to accept.[180] But the reason why this is not a valid consequence is of a different nature.

A straightforward argument shows that the contrapositive of (+) is not a valid consequence: 'Some donkey is running, therefore some proposition is negative' simply does not qualify as a valid consequence, according to the criterion we have so far, since it is possible for the antecedent to be true (for some donkey to be running) while the consequent is false (for example, if the only existing propositions are the antecedent and consequent, that is, both are affirmative propositions). (For a detailed reconstruction of the argument, see Klima 2004a). But according to the principle of contraposition, a consequence and its contraposition are equivalent; hence, if (+) is a valid consequence, so should its contraposition be. If one is not prepared to give

[179] Sed adhuc dico quod haec descriptio non est bona, quia hic non est bona consequentia 'nulla propositio est negatiua, ergo nullus asinus currit', et tamen secundum dictam descriptionem oporteret eam concedere esse bonam; ergo etc.
[180] TC, 31, 10–12.

up the principle of contraposition – and that is Buridan's case (see Buridan 2001, 952)– then one has to revise the definition of consequence as stated above.
The treatise on consequences is rather brief on this particular issue; for clarifications, one has to turn to his treatise on *sophismata*[181] – namely, the first two sophisms of the eighth chapter, the very passages that inspired (Prior 1969). Buridan's five conclusions to the first sophism all point in the same direction: consequence should not be defined in terms of the truth-values of antecedent and consequent. More specifically, the modal value of propositions cannot be defined in terms of their truth-values, and since Buridan is in search of a modal definition of consequence, the impossibility of defining modalities in terms of truth affects the definition of consequence.

2.2.3 Modalities

So, we need an account of the modal value of a proposition that does not rely on its truth-value in different situations (truth is not 'reliable').

The third conclusion is that some proposition is possible that cannot be true.[182] (Buridan 2001, 954)

In Prior's terms, some propositions are possible but not possibly-true, such as 'No proposition is negative': the situation it describes is a possible one (God may annihilate all negative propositions), but once it is formed, it falsifies itself. Similarly, 'Some proposition is affirmative' is necessarily-true but not necessary.

2.2.3.1 '...holds in ...' and '... is true in ...'

But how are we to account for the modal value of propositions? We must distinguish the situation **in which** a proposition is/is not the case from the situation **of which** it is/is not the case.[183] In other words, we must distinguish the context of formation from the context of evaluation of propositions – the very core of modern two-dimensional semantics. That this idea is present in Buridan's framework is attested, for example, by the following passage, in particular by the phrase '[what a proposition] would signify if it were propounded', where the subjunctive mode clearly expresses the view that the situation in which a proposition is or is not formed is possibly distinct from the situation of its evaluation.

[…] the rule, thus stated, is not true by virtue of the expression, namely, [the rule] that any proposition entails that it is true; indeed, this is not valid: 'A man is; therefore "A man is" is true',[184] for a man could exist even if no proposition existed, and also because it is possible that things could be as the proposition 'A man is' signifies or would signify if it were propounded, whereas they would not be as ' "A man is" is

---

[181] Cf. (Buridan 2004).
[182] Tertia conclusio est quod aliqua propositio est possibilis, quae non potest esse vera. (Buridan 2004, 142 (8–9))
[183] Prior's idea was to relativize the notion of truth to 'truth-**in**-the-sheet-of-paper-*x*', which was distinct from the notion of 'truth-**of**-the-sheet-of-paper-*x*'
[184] Clearly, within Buridan's token-based semantics, Tarski's T-condition does not hold, unless the proviso of the existence of the token is made.

true' would signify. For this would be the case if there were a man, but no proposition existed.[185] (Buridan 2001, 957)

The main idea is thus that a proposition φ may **hold** in a situation w_j, another way of saying that things in w_j are in whatever way is signified by φ. In modern logic, this notion is usually seen as identical to truth in a model; but in a token-based semantics such as Buridan's, the two notions must be distinguished. For Buridan, truth presupposes that the proposition in question be formed in the very situation of which it is said to be true or false, whereas '… holds in …' does not; thus defined, this relation is something like a 'trans-world' relation between propositions and situations.

Notation: things in w_j are in whatever way is signified by $\varphi : w_j \Vdash \varphi$.
things in w_j are other than as signified by φ : $w_j \Vdash\!\!/\ \varphi_i$.

The cases treated by Buridan do not concern the use of indexical expressions; in fact, all propositions at stake here are propositions whose content is stable for all contexts of formation.[186] Therefore, propositions can be seen as functions from ordered pairs of situations $(w_i\, w_j)$ – where w_i is a context of formation and w_j is a context of evaluation – to satisfiability values, without loss of accuracy for the present analysis.

On the basis of this primitive notion of 'holds in', the relation of satisfiability (a relation between a pair of worlds and a satisfiability value, defined by the function associated to a proposition) can be defined. Let <**W, S**> be a possible-world model, where **W** is a set of worlds w and **S** is a set of satisfiability values $\mathbf{S} = \{\mathrm{E}, \mathrm{N}, *\}$ ('E' stands for '*est*' and 'N' stands for '*non est*').

Definition 2.2.3.1 Satisfiability

$$\varphi(w_i, w_j) = \begin{cases} E \text{ iff } w_j \Vdash \varphi \text{ and } \varphi \in [w_i]. \\ N \text{ iff } w_j \Vdash\!\!/\ \varphi \text{ and } \varphi \in [w_i]. \\ * \text{ iff } \varphi \notin [w_i], \text{ for all } w_j. \end{cases}$$

---

[185] … illa regula, sicut ponebatur, non est vera de virtute sermonis, scilicet quod ad quamlibet propositionem sequatur quod ipsa sit vera; immo non sequitur 'homo est; igitur haec est vera "homo est" ', quia homo posset esse, licet nulla propositio esset, et etiam quia ita esse est possibile sicut per istam propositionem 'homo est' significatur, vel significaretur si proponeretur, non existente ita sicut per istam significaretur 'haec est vera "homo est" '. Ita enim esset sicut per primam significatur, si homo esset et nulla propositio esset. (Buridan 2004, 144 (22–28)).

[186] Propositions with unstable contents will be discussed in section 2.3.1.

C.E. / C.F.	w_1	W_2	...
w_1	a_{11}	a_{12}	
w_2	a_{21}	a_{22}	
...			

Figure 2.2.3.2. Contexts of formation and contexts of evaluation

2.2.3.2 *Matrices*

I borrow from Stalnaker the idea of representing propositions on matrices[187] of the following kind: the matrix $\mathbf{A}^2$ representing a proposition φ is a matrix where rows represent the contexts of formation and columns the contexts of evaluation (Figure 2.2.3.2). The elements a_{ij} of the matrix (corresponding to φ being formed in situation w_i and evaluated in situation w_j) belong to the set $\mathbf{S} = \{\mathrm{E}, \mathrm{N}, *\}$.

If a proposition is not formed in a situation w_j, the entire corresponding row must be filled with *. However, the proposition formed in a situation w_i is also evaluated with respect to w_j, and the outcome is element a_{ij} of the matrix.

On the basis of such matrices, Prior's notions of possibly-true, necessarily-true etc. – here termed μ-true modalities – receive a straightforward definition, since they concern the evaluation of propositions in the same situation in which they are formed, that is, the diagonal of the matrix.

Definition 2.2.3.2.1: μ-true modalities

φ is possibly-true <=> For some (w_j, w_j), $\varphi(w_j, w_j) = E$
<=> For some a_{jj}, $a_{jj} = E$.

φ is contingently-true <=> For some (w_j, w_j), $\varphi(w_j, w_j) = N$ & for some (w_i, w_i), $\varphi(w_i, w_i) = E$
<=> For some $a_{jj}, a_{jj} = N$ & for some $a_{ii}, a_{ii} = E$.

φ is necessarily-true <=> For all (w_j, w_j), $\varphi(w_j, w_j) \neq N$ & for some (w_i, w_i), $\varphi(w_i, w_i) = E$
<=> For all $a_{jj}, a_{jj} \neq N$ & for some $a_{ii}, a_{ii} = E$.

φ is impossibly-true <=> For all (w_j, w_j), $\varphi(w_j, w_j) \neq E$ & for some (w_i, w_i), $\varphi(w_i, w_i) = N$
<=> For all $a_{jj}, a_{jj} \neq \mathrm{E}$ & for some $a_{ii}, a_{ii} = N$.

---

[187] But notice that Stalnaker's matrices are representations of a different notion, namely that of propositional concept. 'I have emphasized in a number of places that one should not confuse my use of the two-dimensional apparatus with Kaplan's'. (Stalnaker 1999, 10) It will become clear below that, in Buridan's account, what is at issue is mainly Kaplan's distinction between character and content.

Similarly, the traditional alethic modalities – here named μ-modalities – are defined as:

Definition 2.2.3.2.2: μ-modalities

φ is possible <=> For some (w_i, w_j), $\varphi(w_i, w_j) = \mathrm{E}$
<=> For some a_{ij}, $a_{ij} = \mathrm{E}$

φ is contingent <=> For some (w_i, w_j), $\varphi(w_i, w_j) = \mathrm{N}$ & for some (w_i, w_j), $\varphi(w_i, w_j) = \mathrm{E}$
<=> For some a_{ij}, $a_{ij} = \mathrm{N}$ & for some a_{ij}, $a_{ij} = \mathrm{E}$

φ is necessary <=> For all (w_i, w_j), $\varphi(w_i, w_j) \neq \mathrm{N}$
<=> For all a_{ij}, $a_{ij} \neq \mathrm{N}$

φ is impossible <=> For all (w_i, w_j), $\varphi(w_i, w_j) \neq \mathrm{E}$
<=> For all a_{ij}, $a_{ij} \neq \mathrm{E}$

That is, the μ-true modalities are represented by the diagonal of the matrix, whereas the μ-modalities are represented by the entire matrix.[188] Notice that the modal value (μ-modality) of a proposition is not defined with respect to one particular occurrence of the proposition; rather, its modal value is defined with respect to all its equiform occurrences and the different situations in which they occur.

2.2.3.3 Oppositions

In most systems of modal logic, the contradictory proposition of a necessary proposition is an impossible proposition, whereas the contradictory proposition of a possible proposition is a contingent proposition. But it has been noticed (Spade 1996, 313; Normore 1999, 42) that, in a token-based semantics, if the alethic modalities are defined in terms of truth-values – such as in Ockham's system – the usual oppositions between modalities do not hold. In addition to the difficulties concerning the notion of consequence, the failure of the usual oppositions between modalities is an even more telling sign that, within a token-based semantics, modalities must not be defined in terms of truth-values.

More specifically, the problem is that, while a proposition such as 'Some proposition is affirmative' is considered necessary (according to the 'always true' criterion), its contradictory 'No proposition is affirmative' is by no means impossible, as it should be according to the usual oppositions. In fact, 'No proposition is affirmative' is possible, since the situation in which it is the only proposition formed is possible, and in this situation it is true.

As we now know, within a token-based semantics, what is defined in terms of the truth-values of propositions are the so-called μ-true modalities, and not the desired μ-modalities. As a result, it is not surprising that the usual oppositions fail with

[188] In Stalnaker's analysis, this distinction corresponds to the a 'priori/necessary' distinction (Stalnaker 1999, 83–4).

respect to μ-true modalities, since these are not well-behaved properties – pragmatic elements play too important a role therein.[189]

But according to the definitions of μ-modalities just presented (in terms of all the elements of the matrix representing a proposition) the usual oppositions do hold. In order to prove it, some assumptions concerning the contradictory of a proposition must be made. For convenience, I will represent the contradictory of a proposition φ as $\sim \varphi$ (even though this representation is not entirely accurate, since Buridan's account of negating expressions is not restricted to negation as a contradiction-forming functor[190]).

Definition 2.2.3.3: Negation

$$\sim \varphi(w_i, w_j) = \begin{cases} E \text{ iff } w_j \not\Vdash \varphi \text{ and} \sim \varphi \in [w_i]. \\ N \text{ iff } w_j \Vdash \varphi \text{ and} \sim \varphi \in [w_i]. \\ * \text{ iff } \sim\varphi \notin [w_i], \text{ for all } w_j. \end{cases}$$

For this definition to function, the possible world of evaluation must be a maximal-consistent set of propositions (an assumption that we have not made so far), such that $w_j \not\Vdash \varphi$ guarantees $w_j \Vdash \sim\varphi$. But if w_j is not a complete description of a situation, then it may occur that $w_j \not\Vdash \varphi$ and $w_j \not\Vdash \sim\varphi$. Again, for convenience, let us consider that the possible worlds we are dealing with are complete descriptions at least in the sense that $w_j \not\Vdash \varphi$ guarantees $w_j \Vdash \sim\varphi$.

Theorem 2.2.3.3 The contradictory of a necessary proposition is an impossible proposition.

Proof: Suppose that φ is a necessary proposition. Hence, all the elements a_{ij} of the matrix representing it are either * or E. Given the token-existence clause, there is at least one situation w_i in which the proposition is formed. Hence, there is at least one row whose elements a_{ij}, for any context of evaluation w_j, are different from *. Given the definition of a matrix representing a necessary proposition, all its elements which are not * are E.

Therefore, for all contexts of evaluation w_j, $w_j \Vdash \varphi$. Now, with respect to $\sim\varphi$, two situations are possible:

- For some w_i, $\sim\varphi \notin [w_i]$. Thus, all elements b_{ij} of the matrix representing the proposition $\sim\varphi$, for this specific w_i are * (definition of $\sim\varphi$).
- For some w_i, $\sim\varphi \in [w_i]$. Thus, since for all contexts of evaluation w_j, $w_j \Vdash \varphi$, all elements b_{ij} of the matrix representing the proposition $\sim\varphi$, for this specific w_i are N (definition of $\sim\varphi$).

[189] Therefore, prima facie there do not seem to exist formal relations between the diagonals of matrices representing a necessarily-true proposition and its contradictory, respectively, in the fashion of Stalnaker's one-dimension and two-dimension operators between matrices.

[190] Cf. the distinction between *negatio negans* vs. *negatio infinitans*, in (Buridan 2001, 269–271). *Negatio negans* would correspond to propositional negation.

Situation \ Proposition	Φ
w_1	$\varphi \in [w_1],\ w_1 \Vdash \varphi$
w_2	$\varphi \notin [w_2],\ w_2 \Vdash \varphi$
w_3	$\varphi \in [w_3],\ w_3 \Vdash \varphi$
w_4	$\varphi \notin [w_4],\ w_4 \Vdash \varphi$

Figure 2.2.3.4.1. Propositio est affirmativa

C.F. \ C.E.	w_1	w_2	w_3	w_4
w_1	E	E	E	N
w_2	*	*	*	*
w_3	E	E	E	N
w_4	*	*	*	*

Figure 2.2.3.4.2. Matrix for *Propositio est affirmativa*

Hence, if φ is a necessary proposition, all the elements b_{ij} of the matrix representing $\sim\varphi$ are either * or N. Therefore, $\sim\varphi$ is an impossible proposition. □

Similarly, one can prove that the contradictory proposition of a possible proposition is a contingent proposition. Hence, thus defined, the modalities do satisfy the usual relations of opposition. Therefore, it is clear, as Spade has claimed (Spade 1996, 316–317), that Buridan's definitions of the modalities do not suffer from the same flaws of Ockham's definitions, even though both systems are token-based.

2.2.3.4 Example

φ: ***Propositio est affirmativa.***

For convenience, the different possible situations can be represented on a table, which is not yet the context of formation/context of evaluation matrix (Figure 2.2.3.4.1).

The real matrix is shown in Figure 2.2.3.4.2.

Obviously, for the complete evaluation of the modal value of φ, the enumeration of situations/possible worlds would have to be exhaustive (and since this is not feasible, this is not an effective procedure[191]). But in practice, the description of a few situations should do in most cases, at least for convenience of exposition.[192]

---

[191] That is, according to (Read 1994, 252), one of the reasons why the 'interpretational' view on validity is appealing (although it is mistaken), since the manipulation of these 'possible worlds' seems at first infinitely less feasible than the manipulation of terms.

[192] Cf. (Stalnaker 1999, 80).

In sum, 'Some proposition is affirmative' is necessarily-true, since for all a_{ii}, $a_{ii} \neq$ N, but not necessary; in fact, it is possible and contingent, since $a_{14} =$ N and $a_{34} =$ N, while some other elements of the matrix are E.

2.2.4 *Consequentia*

We are now in a better position to approach the definition of consequence. Here is Buridan's third formulation of it:

Third Definition:

> Therefore, others define [antecedent] differently, [by saying that] that a proposition is antecedent to another which is related to it in such a way that it is impossible for things to be in whatever way the first signifies them to be without their being in whatever way the other signifies them to be, when these propositions are formed at the same time.[193] (TC, 22, 48–51)

φ and ψ must be both formed in at least some world, since the idea of a not-formed (valid) consequence does not seem to make any sense, just as much as the idea of a not-formed (true/false) proposition does not make sense, given Buridan's commitment to tokens. As for the phrase 'what a proposition signifies', soon after Buridan says that it is just a way of speaking (see TC, 22), since, properly speaking, this only holds of propositions in the present tense.[194]

On the basis of the passage just quoted, the equivocal understanding of the notion of *consequentia* seems to remain. On the one hand, Buridan's formulation seems to imply that *consequentia* is a (objective) relation between two propositions, resembling thus the modern notion of consequence. On the other hand, his requirement that the propositions be formed together indicates that the actual assertion of both antecedent and consequent, performed by someone, is required for the relevant *consequentia* to exist – pointing thus in the direction of the notion of inference understood as an act. Accordingly, I will provide two alternative definitions of a valid *consequentia* on the basis of the matrices introduced above, and postpone further discussions.

2.2.4.1 Consequentia *as inference*

Since I follow Normore's contention that Buridan's notion of *consequentia* does not encompass the notion of the (mental) act proper to the notion of inference, it is unclear whether this notion can be accurately attributed to his theory. Therefore, I introduce the notion of **inferability** (potential inference) as the relation between two or more actually formed propositions in one given context, to the effect that, were a person to perform the act of inference between them, it would be correct.

---

[193] Ideo alii aliter diffiniunt, dicentes quod illa propositio est antecedens ad aliam quae sic se habet ad illam quod impossibile est qualitercumque ipsa significat sic ess quin qualitercumque illa alia significant sic sit ipsis simul propositis.

[194] Buridan is interested in the correspondence clauses defined in terms of the supposition of a proposition's terms, which he presents in the first two chapters of TC. What a proposition signifies is things in the world, which is clearly problematic in the case of past, future and possible propositions. To cover these cases, Buridan adds specific clauses.

Thus defined, the notion of inferability receives a straightforward treatment in terms of the matrices representing the propositions involved:

Definition 2.2.4.1: Inferability

ψ is inferable from φ in w_i

<=> φ and ψ are formed simultaneously in some context w_i, and it is impossible for things to be in whatever way φ in w_i signifies them to be, and other than in whatever way ψ in w_i signifies them to be.

<=> If $\mathbf{A}^2$ is the matrix representing φ and $\mathbf{B}^2$ is the matrix representing ψ, then for a given world w_i (where φ and ψ are formed), there is no w_j such that $a_{ij} = \text{E}$ and $b_{ij} = \text{N}$.

2.2.4.2 Consequentia *as consequence*

Alternatively, *consequentia* is also the logical relation between classes of equiform propositions, in different contexts of formation:

Definition 2.2.4.2: Consequence

ψ follows from φ

<=> φ and ψ are formed simultaneously in (at least) some context w_i, and it is impossible for things to be in whatever way φ signifies them to be, and other than in whatever way ψ signifies them to be.

<=> If $\mathbf{A}^2$ is the matrix representing φ and $\mathbf{B}^2$ is the matrix representing ψ, then for all w_i there is no w_j such that $a_{ij} = \text{E}$ and $b_{ij} = \text{N}$.

The main difference between these two definitions is, obviously, that the first one concerns only parts of the matrices – particular rows of each matrix – whereas the second one concerns entire matrices. These definitions outline what I believe to be one of the main differences between the concepts of inference and consequence: inference is a **local** relation between propositions, performed by someone in one given context of formation, whereas consequence is a **global** relation, that is, it holds for all different contexts of formation.

2.2.4.3 Example

φ: ***Nulla propositio est negatiua***.
ψ: ***Nullus asinus currit***.

Description of possible situations (Figure 2.2.4.3.1):

Situation / Proposition	φ	ψ
w_1	$\varphi \in [w_1], w_1 \Vdash\!\!/\ \varphi$	$\psi \in [w_1], w_1 \Vdash\!\!/\ \psi$
w_2	$\varphi \notin [w_2], w_2 \Vdash\!\!/\ \varphi$	$\psi \notin [w_2], w_2 \Vdash\!\!/\ \psi$
w_3	$\varphi \in [w_3], w_3 \Vdash\!\!/\ \varphi$	$\psi \in [w_3], w_3 \Vdash\!\!/\ \psi$
w_4	$\varphi \notin [w_4], w_4 \Vdash\!\!/\ \varphi$	$\psi \notin [w_4], w_4 \Vdash\!\!/\ \psi$
w_5	$\varphi \notin [w_5], w_5 \Vdash\!\!/\ \varphi$	$\psi \in [w_5], w_5 \Vdash\!\!/\ \psi$
w_6	$\varphi \in [w_6], w_6 \Vdash\!\!/\ \varphi$	$\psi \in [w_6], w_6 \Vdash\!\!/\ \psi$

Figure 2.2.4.3.1. Nulla propositio est negatiua/Nullus asinus currit

A^2 (matrix corresponding to φ)

C.E. / C.F.	w_1	w_2	w_3	w_4	w_5	w_6
w_1	N	E	N	E	N	N
w_2	*	*	*	*	*	*
w_3	N	E	N	E	N	N
w_4	*	*	*	*	*	*
w_5	*	*	*	*	*	*
w_6	N	E	N	E	N	N

Figure 2.2.4.3.2. Matrix for *Nulla propositio est negatiua*

B^2 (matrix corresponding to ψ)

C.E. / C.F.	w_1	w_2	w_3	w_4	w_5	w_6
w_1	N	N	E	E	N	E
w_2	*	*	*	*	*	*
w_3	N	N	E	E	N	E
w_4	*	*	*	*	*	*
w_5	N	N	E	E	N	E
w_6	N	N	E	E	N	E

Figure 2.2.4.3.3. Matrix for *Nullus asinus currit*

Thus, there is a a_{ij} such that a_{ij} = E and b_{ij} = N, namely a_{12}, a_{32}, a_{52} and a_{62}, and b_{12}, b_{32}, b_{52} and b_{62}. Therefore, φ does not imply ψ. Intuitively, it means that it is possible for no proposition to be negative while there is a donkey running, such as in situation w_2.

Similarly, this consequence is invalid also according to the local definition of *consequentia* (2.2.4.1), namely in the worlds in which both propositions are formed – w_1, w_3 and w_6. In the other worlds, according to the local definition, the inference is not valid or invalid; it simply does not exist.

Notice that each column is homogeneous with respect to the satisfiability values: this occurs because both propositions have stable contents, which do not vary in function of the context of formation. Cases of consequences which are valid according to the local definition but not according to the global definition will arise only with propositions whose content is not stable (see 2.3.1).

2.2.5 *Consequentia formalis*

There remains one aspect of Buridan's notion of *consequentia* to be discussed, namely his distinction between consequences that hold materially and those that hold formally.

> 'Formal' consequence means that [the consequence] holds for all terms, retaining the form common to all. Or, if you want to express it according to the proper force of discourse, a formal consequence is that which, for every proposition similar in form which might be formed, it would be a good consequence, such as 'what is A is B; thus what is B is A'.[195] (TC 22/23, 5–9)

That is, Buridan advocates what we could call the substitutional notion of formal validity: a consequence is formally valid iff all substitutions of its categorematic (i.e., nonlogical) terms yield a valid consequence. A material consequence is a valid consequence (according to the modal criterion) in which the substitutional condition fails.

Now, this distinction only makes sense if one assumes that there is something in common among different proposition-tokens, so that a given consequence can be said valid in virtue of its logical form (which it obviously shares with all equiform occurrences). It presupposes not only that proposition-tokens form classes of identical propositions (the proposition-types), but also that these classes form classes of classes of propositions having the same logical form (i.e., having only logical terms in common).

Actually, the passage above is puzzling concerning the status of expressions such as 'What is A is B; thus, what is B is A'. The use of schematic letters instead of meaningful terms means that neither the whole consequence nor its parts are propositions properly speaking; should we then call it a 'consequence', or rather a 'consequence-schema'? Buridan's discussion is inconclusive, but this issue is of utmost importance for the present purposes: if consequence-schemata are also consequences, then Buridan's notion is certainly not akin to the modern notion of inference, but rather to the modern notion of consequence – in fact, also covering the modern notion of logical/formal consequence.[196]

---

[195] *Consequentia* 'formalis' uocatur quae in omnibus terminis ualet retenta forma consimili. Vel si uis expresse loqui de ui sermonis, *consequentia* formalis est cui omnis propositio similes in forma quae formaretur esset bona consequentia, ut 'quod est A est B; ergo quod est B est A'.

[196] Another passage in which Buridan attributes the term '*consequentia*' to a consequence-schema: (TC, 31, 30–31).

An expression in which the nonlogical terms are replaced by schematic letters – such as ‘Every A is B’ or ‘Some A is B’ etc. – is usually known as a proposition-schema.[197] Such schemata represent different ‘logical forms’, and shall be represented by Greek capital letters Γ, Δ, Π, Σ etc. A proposition γ belongs to the proposition-schema Σ iff it is a substitutional instance of Σ (where meaningful terms take the place of the schematic letters) – represented by ‘$\gamma < \Sigma$’.

While a proposition-type is represented by a two-dimensional matrix (depicting contexts of formation and of evaluation), a proposition-schema must be represented by a three-dimensional matrix: all its different substitutional instances are represented by the usual two-dimensional matrices, and these are superposed along the third axis of the matrix.

The elements of a three-dimensional matrix $\mathbf{A}^3$ representing a proposition-schema Σ are of the form $a_{\varphi ij}$, where the index ‘φ’ represents the substitutional instance φ of the schema Σ – and thus the corresponding ‘slice’ of the third axis of the matrix – and i and j represent the contexts of formation and evaluation, as usual.

Consider now the situation in which two proposition-schemata Γ and Δ share the same schematic letters, or there is a mapping from the schematic letters of one schema into those of the other – for example, ‘Every A is B’ and ‘Some A is B’. Now consider a substitutional instance of each schema such that the schematic letters are replaced by, respectively, the same meaningful terms – for example, γ:‘Every man is an animal’ and η:‘Some man is an animal’. Let us say that γ and η form a pair of **homogeneous substitution** – HS(γ, η). The notion of formal consequence (whose relata are proposition-schemata) can thus be formulated:

Definition 2.2.5: Formal Consequence

Γ formally follows from Δ <=> If $\mathbf{A}^3$ is the matrix representing Δ and $\mathbf{B}^3$ is the matrix representing Γ, then for all propositions $\gamma < \Delta$ and $\eta < \Gamma$ such that HS(γ, η), for all w_i and for all w_j, for all $a_{\gamma ij}$ and $b_{\eta ij}$, if $a_{\gamma ij} = \mathrm{E}$, then $b_{\eta ij} \neq \mathrm{N}$.

In sum, Γ formally follows from Δ iff for all propositions $\gamma < \Delta$ and $\eta < \Gamma$ such that HS(γ, η), η (materially) follows from γ. Notice the primacy of the logical relation between actually formed propositions over more complex relations (defined in terms of this more fundamental relation, that I have termed inferability), in conformity with his commitment to tokens.

Put in a broader historical perspective, Buridan’s criterion of a formal consequence is in line with Bolzano’s substitutional notion of a valid consequence,[198] and also (to a lesser extent) to Tarski’s interpretational criterion,[199] to use Etchemendy’s terminology in (Etchemendy 1990). Again within this terminology, what is peculiar about

[197] Cf. (Corcoran 2004).
[198] Cf. (Bolzano 1973), in particular §§154–155.
[199] Cf. (Tarski 2002).

Buridanian formal consequences is that they must comply with the representational, modal criterion **as well as** with the interpretational (or substitutional) criterion to be valid formal consequences. In this sense, Buridan seems to come very close to what Shapiro has termed the 'hybrid' or 'conglomeration' notion of (logical) consequence,[200] as will be shown in section 2.3.2 below.

2.3 COMPARISONS

So far, I have tried to remain historically faithful to Buridan's account, although it was often evident that his framework could be expanded to a general analysis of the properties of a token-based semantics. In particular, the kind of interaction between language and context that occurs when the content of an expression is at least partly determined by the expression's context of formation remains to be explored. It is a well-known fact that this phenomenon is particularly acute in the case of what are now known as 'indexical expressions'. The paradigmatic cases are demonstrative pronouns such as 'this' and 'that', personal pronouns such as 'I', 'he' 'she' etc., and expressions designating time – 'now', 'tomorrow' etc. – and place – 'here', 'there' etc. It is evident that these terms' contribution to the meaning of a complex expression in which they are embedded depends on the context. In what follows, I will try to sketch what a 'Buridanian' account of propositions including such terms would be like.

Moreover, I will argue in section 2.3.2 that Buridan's notion of formal consequence is remarkably similar to the view on logical consequence defended by S. Shapiro (1998), and that, for this reason, it can contribute to the ongoing debates about this notion.

2.3.1 Two-dimensional semantics

We have seen that Buridan had to deal with propositions whose very existence interferes with their modal values in order to give a correct definition of the concept of consequence. As is well known, Kaplan identified a similar phenomenon with respect to indexical expressions. Propositions such as 'I am here now' come out true every time they are formed, even though the situation they describe (the speaker being in a certain place at a certain time) is not a necessary one (Kaplan 1989, 508–510). Again, in such cases it is clear that the 'always true' criterion is inadequate to characterize a necessary proposition. In other words, 'the facts that determine what is said need to be distinguished from the facts that determine whether what is said is true' (Stalnaker 1999, 5). In such cases, what is said by a proposition – its content – depends on the context in which it is formed, whereas its truth or falsity depends on the context that the proposition says something about.

---

200 Cf. (Shapiro 1998); (Dutilh Novaes 2005a).

Stalnaker has also pursued this line of investigation in several articles. He is interested in the general topic of the relations between linguistic expressions and their context of use, and identifies two basic directions in which these relations occur:

> First, context influences content, since the expressions used to say something are often context dependent; what they are used to say is a function, not only of the meaning of the expressions, but also of facts about the situations in which they are used. But second, the contents that are expressed influence the context; speech acts affect the situation in which they are performed. (Stalnaker 1999, 4)

Clearly, the phenomena related to indexicals that are Kaplan's main concern belong to the first kind of interaction between expressions and context, that is, the context determining the expression's content. By contrast, the problematic aspects of Buridan's commitment to tokens encountered in the analysis of his notion of consequence belong to the second kind of interaction between expressions and context as identified by Stalnaker. That is, in the foregoing we were mostly dealing with cases in which the very performance of a speech act (the formation of a proposition) influences the context in a non-trivial way. A proposition such as 'Some proposition is affirmative' has the remarkable property of provoking its own truth every time that it is formed, insofar as it modifies the context in such a way that it is automatically verified by that context.

2.3.1.1 Indexicals

First of all, it must be noted that Buridan in particular and medieval logicians in general were not oblivious to the peculiar semantic behavior of indexical terms, especially of demonstrative pronouns. In fact, demonstratives such as 'this' and 'that' play a crucial role in one of the most significant aspects of medieval logic, namely the theories of modes of personal supposition.[201] Moreover, Buridan's view on the tense of verbs is clearly that tense functions as a sort of time-indexing.[202]

And yet, in his analysis of the notion of consequence, Buridan does not discuss the case of indexical expressions. However, since these expressions have been at the core of recent discussions on the two-dimensional semantics of tokens, the examination of such cases is a sensible addition to Buridan's theory.

In the foregoing analysis, the pragmatic aspect being taken into account was the existence or absence of a given token in a given situation, but it was assumed that each proposition would have the same content in all possible contexts of formation. Buridan's examples only deal with propositions whose contents are stable throughout the different contexts of formation. So, a proposition φ receives the same value for all ordered pairs sharing the same second element (the context of evaluation) and in which the context of formation w_i was such that $\varphi \in [w_i]$.

But if φ contains indexical terms, adjustments in the general framework are necessary. The definition of propositions as functions from ordered pairs of contexts into satisfiability values is not sufficient in such cases: it seems inaccurate to say that a

---

[201] Cf. (Buridan 2001, 259/264); section 1.5.2 of this text.
[202] Cf. (Buridan 2001, 941/951).

context of evaluation w_j 'satisfies' a proposition (which I denoted by '$w_j \Vdash \varphi$'). In those cases, the introduction of an extra entity, namely the **content** expressed by the proposition in a given context of formation, is imperative.

So the model has to be expanded in order to include these entities. Thus, let $<\mathbf{W}, \mathbf{S}, \Phi>$ be a possible-world model, where **W** is a set of worlds w, **S** is a set of satisfiability values $\mathbf{S} = \{E, N\}$ and **Φ** a set of contents (these are very much like the modern notion of abstract 'propositions', the meaning of declarative sentences). For convenience, * is included in the set of contents, it represents the 'empty' content. Propositions can be seen as functions from contexts of formation to a **content** 'φ_i' (with respect to the context of formation w_i). Notice that, when treating Buridan's discussion, it was important to stress, for metaphysical reasons, that the propositions in question were actually formed mental, spoken or written propositions; but now that we are considering conceptual modifications of the Buridanian framework, it seems patent that it is conceptually and technically more convenient to take proposition-types as primitives, which are or fail to be formed in a given context. When a proposition-type is actually instantiated in a given context (by means of a proposition-token), it yields a content. Intuitively, it means that the function corresponding to the proposition-type is actually applied to/executed in the context of formation in question, something like $\lambda w_j \varphi w$, yielding a content 'φ_i'. The content is in turn a function from contexts of evaluation to satisfiability values E, N.

In other words, the notion of 'holds in', or 'things are as signified by' must also be modified: it no longer concerns worlds and propositions, but rather worlds and contents.

Notation: things in w_j are according to 'φ_i': $w_j \Vdash$ 'φ_i'.
things in w_j are other than according to 'φ_i': $w_j \Vdash\!\!/$ 'φ_i'.

Propositions and contents are thus defined:

Definition 2.3.1.1: Proposition and content

$$\varphi(w_i) = \begin{cases} `\varphi_i\text{' iff } \varphi \in [w_i], \\ *\text{otherwise} \end{cases}$$

$$`\varphi_i\text{'}(w_j) = \begin{cases} \text{E iff } w_j \Vdash `\varphi_i\text{'}. \\ \text{N iff } w_j \Vdash\!\!/ `\varphi_i\text{'}. \end{cases}$$

Admittedly, this approach may not be acceptable within Buridan's account insofar as contents thus defined – dangerously resembling (platonic) abstract entities – are not acceptable within his ontology. Moreover, for Buridan, the significations of propositions are plainly the actual things named in them.[203] But disregarding these difficulties, these two views on propositions – as functions taking ordered

[203] Cf. (Spade 1996, 175–178).

pairs into satisfiability values and as functions taking contexts of formation into contents – emerge naturally, and are indeed virtually equivalent, as noticed by Stalnaker.[204]

Let us take the proposition 'I am here now' and construct a matrix representing it, for a finite number of situations (Figure 2.3.1.1.1). Consider three people, John, Tom and Paul, two different times t_1, t_2 and two different places p_1, p_2.[205] Now we suppose the following situations:

The matrix representing the proposition in these situations is (Figure 2.3.1.1.2).

Notice that, following the Priorian terminology adopted here, 'I am here now' is a necessarily-true proposition, but not a necessary one, since all the elements in the diagonal of the matrix representing it are other than 'N', but not all elements of the matrix are other than 'N'.

But most importantly, notice that the matrices in section 4, representing propositions with stable content, had homogeneous columns – for a given context of evaluation w_j, for all a_{ij}, $a_{i'j}$ etc, either $a_{ij} = a_{i'j}$ or ($a_{ij} = *$ or $a_{i'j} = *$). Here, however, several elements of the matrix in the same column are distinct (even if both

Situation	Description
w_1	t_1, p_1, J. and T. present, J.: 'I am here now'
w_2	t_2, p_2, J. and P. present, P.: 'I am here now'
w_3	t_1, p_1, J. and T. present, T.: 'I am here now'
w_4	t_2, p_2, T. and P. present

Figure 2.3.1.1.1. 'I am here now'

C.E / C.F.	w_1	w_2	w_3	w_4
w_1	E	N	E	N
w_2	N	E	N	E
w_3	E	N	E	N
w_4	*	*	*	*

Figure 2.3.1.1.2. Matrix for 'I am here now'

[204] 'I will call what a matrix like B represents a PROPOSITIONAL CONCEPT. A propositional concept is a function from possible worlds into propositions, or, equivalently, a function from an ordered pair of possible worlds into a truth-value. Each concrete utterance token can be associated with the propositional concept it determines […]'. (Stalnaker 1999, 81)

[205] Kaplan's apparatus to treat such propositions is much more sophisticated, since it includes a formal account of indexes for time, place and utterer. The present analysis has more modest aims, and perhaps the virtue of simplicity.

are other than '*'), and that reflects the fact that the content of the proposition is not stable.

2.3.1.2 Character vs. Content

Buridan's notion of the signification of a proposition is quite strict: a sentence simply signifies the objects named in it, just as much as simple terms signify the objects that fall under it.[206] Naturally, many objections can be raised against such a view, for example, that it is too limited and therefore that it fails to account for many aspects of a proposition's meaningfulness. Now, a critical examination of Buridan's position is not my goal at present, but I have to emphasize that the framework developed in this section is in fact incompatible with some of Buridan's proclaimed views. However, insofar as Buridan's strict view of signification was mainly motivated by his nominalist convictions, and not by purely logical and semantic considerations, it seems legitimate to investigate how his framework could be adjusted to a more comprehensive view of propositional meaning.

In fact, as can be inferred from the framework as applied so far, one may understand the notion of signification, or meaning, of a proposition either globally, that is, as what all equiform propositions have in common, or locally, that is, as related to one particular use of a proposition in a given situation. Transposing this issue to the modern analysis of indexicals, it is obvious that it corresponds to Kaplan's distinction between **character** and **content** of an expression.

> But there is another sense of meaning in which, absent lexical or syntactical ambiguities, two occurrences of the same word or phrase must mean the same. [...] This sense of meaning – which I call *character* – is what determines the content of an occurrence of a word or phrase in a given context. For indexicals, the rules of language constitute the meaning in the sense of *character*. (Kaplan 1989, 524)

Character is what every equiform occurrence of a proposition has in common with the other elements of its proposition-type class. Under the assumptions made here, the character of a proposition does not have to be seen as the abstract meaning of a proposition-type, but simply as its form, that is, the material disposition of the terms in the proposition. Clearly, the character of a proposition is represented in the entire matrix, since it is what determines its content in each situation, whereas the different contents are relative to a given context of formation, and are represented by each row.

Summing up, character is a global property of classes of equiform propositions – proposition-types – whereas content is a local property of a proposition(-token), relative to a given context of formation. Both properties are rightly said to constitute the meaning of a proposition(-token): the meaning of an expression is, we now know, (at least) two-dimensional.

2.3.1.3 Inference and consequence in a token-based semantics

It is fairly accurate to say that the notions of inference and consequence are typically taken to concern only abstract proposition-types and formulae. It is true that

---

[206] Cf. (Spade 1996, 175–178).

inference, seen as a particular action, implies some degree of particularity, but it is usually assumed that the act of inference performed by someone consists of deriving abstract and hitherto unknown information (the conclusion) from abstract and previously known information (the premises). In this sense, it is assumed that particular occurrences of meaningful propositions are not the relata of an inferential relation.

As for consequence, if it is defined as the relation between two sets of abstract contents, which are independent of any (human) action, then the token perspective is entirely absent. And indeed, since logic (which is the natural terrain for investigations on inference and consequence) is generally concerned with abstract objects, it is not surprising that particular linguistic occurrences are usually of little interest to the logician.

However, one may have philosophical and metaphysical objections to the view that such abstract entities exist, or else one may simply be interested in understanding how, in 'everyday life', people actually practice inferences.[207] In both cases, it seems reasonable to investigate the mechanisms of inference and consequence within a token-based semantics. In the first case, tokens[208] are more palatable to a nominalist's strict ontology; in the second case, it is simply a fact that, in everyday life, we make extensive use of expressions that require the context of formation to be understood, and that modify this very context when produced; now, these are the very expressions we use to express the inferences we draw. Moreover, some extreme uses of language, such as those that lead to paradoxes (the liar paradox, for example), can only be properly understood from a token perspective.

Given the two aspects of the meaning of a token, its character and its content, it is to be expected that different inferential relations would concern each of them. I have argued that Buridan's notion of *consequentia* is equivocal insofar as it covers both the inferential relation between propositions in a given context of formation as well as the relation between classes of equiform propositions. These two notions can be summarized as:

> (**Cto**) a relation between the contents of propositions, relative to a context of formation – taken thus as tokens.[209]
>
> (**Cty**) a relation between the characters of propositions – taken thus as types.

[207] Note that it has been recently argued that adopting the token perspective for operations of inference in the logic of knowledge explains away the notoriously thorny problems of logical omniscience and common knowledge. Cf. (Parikh 2005).

[208] From now on, since my object of analysis is no longer Buridan's theory, I will adopt the modern terminology of types and tokens.

[209] It may seem here that, by adopting contents as the relata of inferences I am in fact betraying the token perspective. However, tokens should not be understood only in their materiality, but rather in their meaningfulness with respect to a given context; in other words, a token is a token only in relation to the content expressed by it in a situation. Therefore, it is equivalent to consider meaningful tokens in a given context or the contents they express in that context as the relata of local inferential deeds. Accordingly, in

Once one has adopted the token perspective, the natural view of the connection between certain tokens would be that inferential relations concern exclusively particular occurrences and the relevant contexts of formation, that is, (**Cto**). That is, from

(a) John will go to the theater tomorrow.

and the fact that one knows that today is November 22nd 2003, one is licensed to infer, and thus assert, the following token:

(b) John will go to the theater on November 23rd 2003.

This inference only holds in this particular context of formation, since it is a relation between the contents expressed by (a) and (b) relative to that context.[210]

But take the following propositions:

(a) John will go to the theater tomorrow.

(c) John will go out tomorrow.

It seems that the inferential relation that exists between them goes beyond particular instances of their occurrences. In fact, every time (a) is stated, anyone would be licensed to draw the conclusion to the effect that (c). Thus, the inferential relation between (a) and (c) seems to be independent of particular contexts of formation, and therefore seems to concern their character rather than their content,[211] corresponding thus to (**Cty**).[212]

Indeed, the two kinds of inferential relations between propositions (**Cto**) and (**Cty**) neatly correspond to the notions of inference and consequence as I have been using them (which confirms my claim that Buridan's notion of *consequentia* is equivocal). Inferences are local deeds, and consequences are global relations between the characters of certain propositions. Notice that a platonic view of propositions is not required: all we need are two (or more) classes of proposition-tokens that share the same character, respectively, in virtue of being equiform.

---

this section the notions of inference and inferability will be said to concern tokens or contents indistinctively, as these two notions are equivalent under these assumptions.

[210] The notion of context is not exclusively pragmatic. In formal logic, the notions of assumption and presupposition can very well be accounted for by the notion of context broadly understood. Indeed, it is said of some inferences that they are valid provided that the assumptions obtain, in which case one would not talk of a consequence that holds, but rather of a valid inference or inference-rule. (This is noticeable, for instance, in constructive type theory.)

[211] Another example: 'I am hungry' implies 'Someone is hungry' in any context of formation.

[212] The inferential relations between (a) and (b) and between (a) and (c) can be turned into formal inferential relations by the addition of extra premises – 'Tomorrow is November 23, 2003', in the first case, and 'To go to the theater is to go out'. The difference between these two cases is also made evident by the fact that, in the first case, the 'missing premise' is a contingent (synthetic) proposition, whereas in the second case it is a necessary (analytic) proposition.

Based on these considerations, let me now reassess the four concepts introduced throughout the text, namely those of inference, inferability, consequence and formal consequence, with respect to the relata involved, their extensions and the dimension of their definitions.

The relata of inference and of inferability are contents.[213] Inference remains essentially an action performed by someone, and therefore the notion of inferability[214] had to be introduced, which is equivalent to the concept of inference except for the exclusion of the actual performance of the inferential act. For the assessment of the validity of an inference and of a relation of inferability, only the rows of each matrix representing a particular context of formation w_i matter – that is, one context of formation and several contexts of evaluation come into play. The modal component is still present, since ψ is inferable from φ in w_i only if in all circumstances (of evaluation) in which 'φ_i' is the case, 'ψ_i' is also the case. But the notions of inference and inferability thus understood are not two-dimensionally modal: they are two-dimensional, since the contexts of formation and evaluation are considered separately, but they are one-dimensionally modal, since the context of formation remains fixed and different possible contexts of evaluation are taken into account.

Consequence, on the other hand, understood as the relation between the characters of the tokens involved, must be two-dimensionally modal, taking into account variation of the context of evaluation as well as of the context of formation. Therefore, the entire two-dimensional matrix must be considered in the definition of a valid consequence.

The notion of formal consequence seems to add yet another dimension, that of the variation of the nonlogical terms. The relata of a formal consequence are propositional schemata, represented in three-dimensional matrices, and these three dimensions must be taken into account in the definition of a valid formal consequence.

On the basis of the definitions, the extensions of each concept can be delineated. Inferability defines the set of potential valid inferences, that is, inferences that would be valid if they were performed, whereas the concept of inference refers to actually performed valid inferences (so the latter is a subset of inferability). Consequence concerns the subset of the extension of potential inferences that is valid in any context of formation (in particular, but not exclusively, those potential inferences which involve propositions with stable content). Formal consequence concerns the subset of the extension of consequence that also satisfies the substitutional criterion. These relations of inclusion are depicted in Figure 2.3.1.3.1.

---

[213] Or, equivalently, tokens understood as meaningful with respect to a context.

[214] In fact, one can say that, if the relation of consequence is primarily defined by the nature of the relation (either as act or as an objectual relation), and not by the relata involved, then the notion of a logical relation between two tokens, excluding the act component of the notion of inference, could also be qualified as consequence. But in the present account, I choose to use the notion of inference for the (local) relation between two or more tokens actually performed by someone, and the notion of inferability for the (local) relation between the tokens (excluding the action), so that the notion of consequence can be applied exclusively to the global relation between equiform occurrences of propositions.

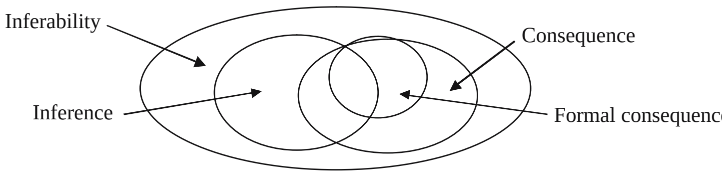


Figure 2.3.1.3.1. Relations of inclusion

	Inference	Inferability	Consequence	Formal Consequence
Relata	Tokens – Contents	Tokens – Contents	Types [215] – Characters	Schemata [216]
An action or a relation?	Action	Relation	Relation	Relation
Dimension of the definition	One-dimensional (local)	One-dimensional (local)	Two-dimensional (global)	Three-dimensional (global)

Figure 2.3.1.3.2. Summary

One may object that validity is only a matter of form, and that only consequences that satisfy the criterion of formality are valid – in sum, that there are no 'materially valid' consequences. Here, as well as in type-based semantics, a plea must be made in favor of the so-called 'representational' view on validity, which shall be discussed in more detail below.

In sum, inference and inferability are not a matter of formality, but a matter of syntax, semantics and pragmatics combined, since these three levels of meaning are required to determine the content of a particular use of a given expression; in addition, the concept of inference involves the essentially pragmatic notion of an action actually performed by someone. Consequence is a matter of syntax and semantics, since it concerns the characters of the propositions involved (including the meanings of non-logical terms). Formal consequence is a matter of syntax and of the meaning of logical constants only. That is, to analyze the notions of inference and consequence one must go beyond mere formal validity, and this is made particularly evident within a token-based semantics.

A summary of these notions is shown in Figure 2.3.1.3.2.

2.3.1.4 Conclusion

a. Buridan's theory features an intriguing combination of semantics and pragmatics. Even though modern semanticists take their investigations in different directions, and only a small part of their theoretical apparatus has been put to use here, this is

[215] That is, classes of tokens formed by strictly equiform tokens.
[216] That is, classes of types formed by types sharing the same logical form.

sufficient evidence to the effect that Buridan should be rightly considered as a pioneer of two-dimensional semantics.

b. Buridan himself says that, in practice, the final definition of consequence is needed in a few cases only; for most cases, the familiar definition in terms of truth-values is perfectly sufficient (see TC, 22). Clearly, those difficult cases are those in which one or both of Stalnaker's forms of mutual influence between context and language occur, since in these cases Buridan's commitment to tokens comes into play.

c. The classical 'paradoxes' of consequence – *ex impossibili sequitur quodlibet* and *necessarium sequitur a quolibet* – are not affected by the present analysis. According to the definition of consequence proposed here, it is still the case that, from an impossible proposition (one whose matrix only has elements other than 'E'), any proposition follows, as much as a necessary proposition follows from any proposition. What the present account seeks to exclude are pseudo-occurrences of these inferential schemata (related to μ-true modalities) (see also Read 2001).

d. Buridan's theory also shows that any account of inference and consequence within a token-based semantics demands amendments to the traditional notions, since it has to include the two kinds of mutual influence between context and linguistic occurrences. I have tried to indicate which amendments would be necessary, so that the present analysis may serve as starting point for further work on the logic of tokens.

2.3.2 The concept of logical consequence

After decades of predominant focus on the notion of logical truth, the philosophical debate on the concept of logical consequence was re-ignited by J. Etchemendy's book *The Concept of Logical Consequence* (1990). His main tenet was that the model-theoretic notion of logical consequence did not capture adequately the corresponding intuitive notion. One of Etchemendy's central claims was that the intuitive notion could be understood essentially from two different perspectives, which he termed 'representational' and 'interpretational' – and that the model-theoretic notion failed to match either.

Some years ago, S. Shapiro (1998) sought to vindicate the model-theoretic notion of logical consequence; one of his arguments was that the dichotomy representational/ interpretational notion of logical consequence was in a certain way infelicitous, since, according to him, a faithful rendering of the intuitive concept would have to have elements of both notions. Clearly, the resolution of the issue as to whether the model-theoretic notion correctly captures the intuitive notion presupposes a minimally adequate characterization of this intuitive notion. Shapiro claimed that Etchemendy hadn't really provided such a characterization,[217] and attempted to formulate one himself. He further claimed that, thus characterized, the intuitive notion was indeed correctly captured by the model-theoretic notion (albeit with some adjustments).[218]

---

[217] Cf. (Shapiro 1998, 143).

[218] 'My claim is that model-theoretic consequence can be made into a good model of this notion [the "hybrid" intuitive notion of logical consequence] and that both the intuitive notion so characterized and its

Here, I do not discuss Shapiro's defense of the model-theoretic notion; rather I examine his contention that the best rendering of the intuitive notion of logical consequence is what he called the 'conglomeration' notion, that is, the hybrid notion that combines both the representational and the interpretational view on consequence. More specifically, I claim that such a hybrid view corresponds precisely to Buridan's notion of formal consequence, and that this fact offers significant historical support to Shapiro's version of the intuitive concept of (logical) consequence.

2.3.2.1 Four notions of logical consequence[219]

Both in Etchemendy's book and in Shapiro's article, four presumably distinct notions of logical consequence are at stake: the (elusive) intuitive, pre-theoretic notion, [220] the model-theoretic notion, the representational notion and the interpretational notion. The goal of most philosophers and logicians interested in this issue (among whom Etchemendy, Shapiro and Tarski[221]) seems to be that of capturing the intuitive, pre-theoretical notion – the real notion, so to say – by means of suitable theoretical constructions (e.g., one of the three other aforementioned notions).

How does one accomplish that? How can one compare different notions of (logical) consequence? The most straightforward way seems to be to focus on their extensions, that is, on the sets of pairs of antecedent and consequent considered to form valid (logical) consequences according to each notion. For this purpose, Etchemendy introduced the very useful concepts of **overgeneration** and **undergeneration**, which can be formulated as follows:

Definition 2.3.2.1: Overgeneration and undergeneration
Consider two notions of logical consequence, **A** and **B**. That a pair <**K**, **X**> (where **K** is a set of sentences and **X** an individual sentence[222]) forms a valid consequence according to **A**, **B** is represented, respectively, by $\mathbf{K} \vDash_A \mathbf{X}$ and $\mathbf{K} \vDash_B \mathbf{X}$; that a

---

mathematical model are useful tools for shedding light on the normative/modal/semantic notion of correct reasoning in natural language, the target of logic'. (Shapiro 1998, 148)

[219] Following the modern terminology, I here use the phrase 'logical consequence' to refer to what hitherto has been referred to as 'consequence' *tout court*.

[220] This variety of notions of consequence was already present in Tarski's seminal 'On the concept of following logically' (Tarski 2002); Tarski referred to the concepts of **logical** consequence, **formal** consequence, **material** consequence and also to what he called in the Polish version of the text the 'everyday concept' of consequence (Stroinska and Hitchcock 2002, 165). Tarski also employed extensively the word 'intuition' connected with the use of the concept of consequence (Stroinska and Hitchcock 2002, 166). Spelling out how these different concepts relate to one another is one of the aims of the present discussion, but for the moment the reader must bear in mind that, even though my main interest is in the concept of **logical** consequence, I will occasionally refer to the concept of consequence *tout court*.

[221] 'The concept of following logically belongs to the category of those concepts whose introduction into the domain of exact formal investigations was not only an act of arbitrary decision on the side of this or that researcher: in making precise the content of this concept, efforts were made to conform to the everyday "pre-existing" way it is used'. (Tarski 2002, 176)

[222] Again, here I switch to the modern terminology and start using the term 'sentence' to refer to what hitherto has been referred to as 'proposition'.

pair $<\mathbf{K}, \mathbf{X}>$ does not form a valid consequence according to **A**, **B** is represented, respectively, by $\mathbf{K} \nvDash_{A} \mathbf{X}$ and $\mathbf{K} \nvDash_{B} \mathbf{X}$.

- **A** overgenerates w.r.t. **B** <==> There is a pair $<\mathbf{K}, \mathbf{X}>$ such that $\mathbf{K} \vDash_{A} \mathbf{X}$ and $\mathbf{K} \nvDash_{B} \mathbf{X}$.
- **A** undergenerates w.r.t. **B** <==> There is a pair $<\mathbf{K}, \mathbf{X}>$ such that $\mathbf{K} \nvDash_{A} \mathbf{X}$ and $\mathbf{K} \Vdash_{B} \mathbf{X}$.

In other words, if a notion **A** either overgenerates or undergenerates w.r.t. a notion **B**, then their extensions do not coincide. When one of the two notions is the desired one, undergeneration is not as threatening as overgeneration; for example, logical systems that are sound but not complete w.r.t a given semantics undergenerate w.r.t this semantics, but are still considered to be, to some extent, reliable. But if a logical system overgenerates w.r.t the semantics in question, then it is unsound, and this is of course reason enough to dismiss the system as utterly flawed – that is, if one is not prepared to revise its semantics, on account of it having some sort of intuitive validity.

2.3.2.1.1 Model-theoretic notion. The birth of the model-theoretic notion of logical consequence, as that of model-theory in general, is usually traced back to the works of Alfred Tarski and Robert Vaught. The historical accuracy of this view is not at issue in the present discussion; for the present purposes, it is sufficient to turn to a precise (and prima facie uncontroversial) formulation thereof, as in the passage below:

Definition 2.3.2.1.1: Model-theoretic notion of logical consequence

- A sentence **X** is a model-theoretic logical consequence of a set of sentences $\mathbf{K}$($\mathbf{K} \vDash_{MT} \mathbf{X}$, for short) just in case every set-theoretic structure which is a model of all the sentences in **K** is also a model of **X**. If an argument $<\mathbf{K}, \mathbf{X}>$ is such that $\mathbf{K} \vDash_{MT} \mathbf{X}$ then any argument $<\mathbf{K}', \mathbf{X}'>$ with the same form will be such that $\mathbf{K}' \vDash_{MT} \mathbf{X}'$, since the model-theoretic notion of logical consequence is intended for languages where any two sentences of the same form have as models exactly the same structures. (Gómez-Torrente 2000, 529)

Notice that it is the notion of equiformity that connects the language in which an argument is expressed to the structures that are models of the sentences of the language, by means of the crucial concept of **form**: sentences of the same form (the linguistic, syntactic level) have the same structures as models (the semantic level). In this sense, both interpretational and representational approaches to logical consequence are to some extent represented in the model-theoretic notion, as will become clear below: the interpretational notion is reflected on the idea of sentences having the same form, whereas the representational notion is reflected on the all-models criterion (a logical consequence is valid iff all models satisfying the premises also satisfy the conclusion).

But this is only part of the story. 'It seems that model theory is not a good model of representational semantics nor is it a good model of interpretational semantics'. (Shapiro 1998, 143). Etchemendy's conclusion had been that the model-theoretic notion came the closest to the interpretational/Tarskian notion of logical consequence,[223] and that the former simply did not capture the representational notion. Shapiro and others have argued that the model-theoretic notion is also not a good model of interpretational semantics: one of the advantages of the model-theoretic notion over the interpretational one is that, in the latter, the domain of discourse is fixed (the entities of the actual world), whereas the former encompasses the idea of different domains, coming thus closer to the (sound) basic intuitions of the representational approach.[224]

In either case, the model-theoretic notion does not seem to capture the intuitive notion of logical consequence: it applies only to language with specific characteristics (mostly artificial, formal languages), and it is dependent on what structures are to be considered as models. One of the arguments offered by Etchemendy against the model-theoretic notion (in this case, of logical truth) is the following: assume that the (finite) universe has exactly *n* entities. A sentence stating that there are exactly *n* entities in the universe will come out as a logical truth (true in all models), even though intuitively the size of the universe is **not** a matter of logical truth (but rather a merely contingent matter).[225] Thus, the model-theoretic notion overgenerates w.r.t. the intuitive notion, and it also undergenerates if one considers the fact that it only captures valid arguments in languages satisfying certain requirements (in particular the requirement of linguistic equiformity corresponding to sameness of models).

2.3.2.1.2 Interpretational notion. The interpretational notion is essentially the Tarskian notion presented in (Tarski 2002), although it has been claimed that Etchemendy's discussion does not entirely do justice to the subtleties of Tarski's argumentation (cf. Gómez-Torrente 1999, Stroinska and Hitchcock 2002). Briefly put, the interpretational notion can be formulated as follows:

Definition 2.3.2.1.2: Interpretational notion of logical consequence

- **X** is an interpretational logical consequence of **K**(**K** $\vDash_{TI}$ **X**, for short[226]) iff for all interpretations of the non-logical terms of **K** and **X**, if **K** is true then **X** is true.

An interpretation of the non-logical terms of **K** and **X** is the assignment of a sequence of objects of the universe to the sequence of these non-logical terms, such that each object is the denotation of one of the non-logical terms in question. A sequence of

[223] Cf. (Etchemendy 1990, 51).
[224] Cf. (Shapiro 1998, 143).
[225] See also (Read 1995, 41)
[226] 'TI' for '**T**arskian**I**nterpretational'.

objects satisfies a sentence or set of sentences if the latter come out true under the given interpretation. If all sequences that satisfy **K** also satisfy **X**, then $\mathbf{K} \vDash_{TI} \mathbf{X}$ (and the converse).

So far, so good. But there are serious problems with the interpretational notion: a sharp distinction between logical and non-logical terms is required (and that is notoriously a difficult task[227]); the interpretational notion fails to capture necessary connections belonging to the level of the meaning of non-logical terms; the domain remains fixed, which means that the modal intuition behind the intuitive notion of logical consequence is at least partially lost. As a result, the interpretational notion undergenerates and overgenerates w.r.t. the intuitive notion.

Take the following example: '*a* is red, consequently *a* is colored'. Intuitively, this is a valid consequence – in fact, many of us would call it an analytic consequence, since there is a connection between the meanings of 'red' and 'colored'. However, under the interpretational notion, when these terms are taken to be non-logical terms, this is not a valid consequence. Hence, the interpretational notion undergenerates w.r.t. the intuitive notion.[228]

Consider another example: 'Andrew Jackson was President, thus Andrew Jackson was male'.[229] (cf. Shapiro 1998, 144). If only 'Andrew Jackson' is taken to be a non-logical term, and different interpretations are assigned to it, then it will turn out that every interpretation that satisfies the premise also satisfies the conclusion. Thus, according to the interpretational notion, this would be a valid consequence. But intuitively, this is obviously not a valid consequence; it is a matter of pure contingency (or in any case of the history of humanity) that, thus far, all presidents of the USA have been men. The problem here is that of the fixed domain: only entities existing or having existed in the actual world are taken into account, and this is not sufficient to establish the logical validity of an argument. Hence, the interpretational notion also overgenerates w.r.t. the intuitive notion.

2.3.2.1.3 Representational notion. The representational notion of logical consequence, or variants thereof, has been in operation for a long time in the history of logic,[230] and can be traced back to Aristotle.[231] It can also be described as the incompatibility notion of consequence, insofar as it holds that the truth of the premises is incompatible with the falsity of the conclusion. In a general formulation, it can be formulated as follows:

---

[227] Indeed, Tarski considered this very issue as perhaps the gravest 'open question' of his analysis (cf. Tarski 2002, 188).

[228] Notice that this argument is also valid according to the representational notion (it is impossible for something to be red without being colored), thus the interpretational notion also undergenerates w.r.t. the representational notion.

[229] It is presupposed, of course that 'President' here means 'President of the USA'; it might have been more polite towards other republics if this presupposition had been made explicit.

[230] Cf. (Martin 1986, 567).

[231] 'A deduction is a discourse in which, certain things being stated, something other than what is stated follows of necessity from their being so'. *Prior Analytics* 24^{b}19–20.

Definition 2.3.2.1.3 Representational notion of logical consequence

- **X** is a representational logical consequence of **K**(**K** $\vDash_{RE}$ **X**, for short) iff it is impossible for **K** to be true and **X** not to be true.

As much as in interpretational semantics, an account of what it means for a sentence **X** to be true must be provided. But the biggest challenge for representational semantics is to give an account of what it means to be **impossible** for **K** to be true and **X** not to be true. The representational notion of logical consequence is rightly seen as essentially modal in nature, and there is a variety of ways to spell out this modal character; currently the most popular one employs the idea of possible worlds. But while interpretational semantics is based on the rather manageable notions of interpretation, sequences of objects and satisfiability, representational semantics is (according to many) inexorably tangled in all kinds of metaphysical webs, owing to the problematic modal notions.

However, here I do not wish to discuss philosophical objections to modal notions and to possible-world semantics.[232] What matters now is how the representational notion of logical consequence fares when confronted to the intuitive notion. At first sight, the representational notion seems a strong candidate for the office of an accurate model of the intuitive notion, since it appears to capture successfully its modal nature. But there are difficulties that must be faced.

Consider the following example: '*a* is a man, consequently *a* is an animal'. Under the representational view, this comes out as a valid consequence, since (assuming a reasonable dose of essentialism) it is impossible for any entity to be a man without being an animal. But is this an intuitively valid **logical** consequence? It seems to differ from the analytic consequence involving the terms 'red' and 'colored' mentioned above, since the connection between the terms 'man' and 'animal' does not seem to pertain to the level of their meanings. It is rather a metaphysical connection: animality pertains to the very nature of manhood. I, in any case, would not be prepared to call it an analytic, or logical, consequence.[233] Thus, if this is not an intuitively valid logical consequence, but it does come out valid according to the representational notion, then the latter seems to overgenerate w.r.t. the intuitive notion.

Another difficulty concerning the representational notion is related to the problem of how many 'possible situations' are to be taken into account, just as in the case of the model-theoretic notion. If the meta-theory is not broad enough and does not

---

[232] Notice though that the apparent manageability of interpretational semantics when contrasted to representational semantics partially explains the appeal of interpretational semantics. cf. (Read 1994, 252) and (Read 1995, 50–51).

[233] Obviously, it can be turned into a logical/formal consequence by the addition of a premise, namely 'Every man is an animal'. The missing premise is a popular approach to the issue of logical and formal validity: a valid argument is one that can be turned into a formally valid one by the addition of a (necessarily true) premise. But this approach is not without difficulties, cf. (Read 1994, 254–259); Buridan, for example, says that the addition of a premise makes the validity of an argument **evident**, but it does not turn an invalid argument into a valid one (cf. TC, 23). So this strategy has at most an epistemological value.

consider **all** possible situations, then an argument may be held valid only because no counterexample has been found among the available situations, while a counterexample (a situation in which the premises are true and the conclusion false) would exist in a meta-theory with a wider range of situations. Again, if the meta-theory is not broad enough, then the representational notion may overgenerate w.r.t. the intuitive notion.

At any rate, it is clear that the representational notion is conceptually and extensionally different from the other two notions considered thus far. Etchemendy seems to be right in insisting on their fundamental differences, and in saying that it is mere contingency if they occasionally intersect.[234]

2.3.2.1.4 Intuitive (pre-theoretic) notion. How far have we got in our quest for the intuitive notion of logical consequence? None of the three candidates seems to be a good model of it, as they all seem to undergenerate or, more importantly, overgenerate (or both) w.r.t. the intuitive notion.

Another aspect that deserves further investigation is the extent to which the requirement of formality – which is present in the model-theoretic and in the interpretational notions, but not in the representational notion – is a part of the intuitive notion of logical consequence.

While some people have very strong intuitions according to which validity is mainly determined by the form of arguments, others may feel that what is really at stake is preservation of truth. In other words, it is very well possible that a single intuitive notion of logical consequence does not exist, and that there are as many 'intuitive notions' as there are people pondering on the matter. A less desperate view on the issue may be that there are perhaps two basic intuitive notions of consequence in operation, one focusing on the formal, logical character of validity and the other on truth preservation.

Clearly, a deeper analysis of the matter is welcome, and that is precisely what seems to be offered by the Buridanian notion of formal consequence.

2.3.2.2 Buridan's hybrid notion of formal consequence

2.3.2.2.1 Material and formal consequence. As already argued, the first formulation of consequence offered by Buridan (which later has to be reformulated to accommodate the behavior of tokens) is very much in the spirit of the representational notion:[235]

> [The terms] 'antecedent' and 'consequent' are said correlatively; therefore, they need to be described in terms of each other. Therefore many people say of two propositions that one is the antecedent with respect to the other which cannot be true while the other is not true, so that every proposition is antecedent with respect to any other proposition which cannot be true without the other being true.[236] (TC, p.21, 26–32)

[234] Cf. (Etchemendy 1990, 63–64).

[235] For the reader's convenience, I reproduce some of Buridan's passages already quoted above.

[236] Antecedens autem et consequens relatiue dicuntur ad inuicem; ideo per inuicem describi debent. Dicunt ergo multi quod propositionum duarum illa est antecedens ad aliam quam impossibile est esse ueram illa alia non existente uera et illa est consequens ad reliquam quam impossibile est non esse ueram reliqua

In practice, Buridan adds a few pages later, this definition holds good in all cases in which the very existence of the tokens in question does not interfere with their own modal values, so for the present purposes this is the definition that matters.

Among the consequences that satisfy the modal/representational criterion, some seem to display an interesting feature, namely the fact that their form appears to be connected to their validity as consequences. These are called formal consequences:

> 'Formal' consequence means that [the consequence] holds for all terms, retaining the form common to all. Or, if you want to express it according to the proper force of discourse, a formal consequence is that which, for every proposition similar in form which might be formed, it would be a good consequence, such as 'what is A is B; thus what is B is A'.[237] (TC 22/23, 5–9)

Notice that the non-logical/logical boundary was less of a problem for medieval logicians such as Buridan because they operated with the distinction between categorematic and syncategorematic terms; the form of a proposition is given by its syncategorematic terms, whereas the categorematic terms define its matter.[238]

A few things are worth noticing about Buridan's definition of formal consequence. First, his criterion is akin to what is usually referred to as the substitutional/variational criterion of validity, to be (re)discovered by Bolzano[239] in the 19th century and refined by Tarski in the 20th century, yielding what Etchemendy calls the interpretational notion of consequence (cf. Etchemendy 1990, chap. 3). But notice that Buridan is not referring to the set of all putative consequences that satisfy the substitutional criterion; rather, he is interested in the subset of the putative consequences that satisfy the modal/representational criterion, which **also** satisfy the substitutional criterion. In particular, what defines a formal consequence is the fact that all its substitutional instances are material consequences, as already discussed above.

2.3.2.2.2 Shapiro and Buridan: the hybrid notion. Based on his critique of the representational/interpretational dichotomy, Shapiro concludes that the intuitions behind both notions are present in the intuitive notion of logical consequence, but that each of them only partially represents the latter. While in the representational account the language remains fixed and the world of evaluation varies, in the interpretational account the reverse occurs. Shapiro demands that both variations – of worlds and of language – occur to qualify a logical consequence. He proposes thus his 'conglomeration' notion of logical consequence, which is indeed a blend of the modal and of the semantic characterizations of logical consequence and, according to Shapiro, a characterization as good as any of the intuitive notion.

---

existente uera, ita quod omnis propositio ad omnem aliam propositionem est antecedens quam impossibile est esse ueram illa alia non existente uera.

[237] Consequentia 'formalis' uocatur quae in omnibus terminis ualet retenta forma consimili. Vel si uis expresse loqui de ui sermonis, consequentia formalis est cui omnis propositio similes in forma quae formaretur esset bona consequentia, ut 'quod est A est B; ergo quod est B est A'.

[238] But Buridan recognized the existence of 'limit cases', for example verbs such as 'to see', 'to want', 'to know', which perform functions relevant to both form and matter. Cf. (Buridan 2001, section 4.3.8.4).

[239] Cf. (Bolzano 1973), in particular §§154–155.

Definition 2.3.2.2.2: Hybrid notion of logical consequence

- Φ is a logical consequence of Γ if φ holds in all possibilities under every interpretation of the non-logical terminology in which Γ holds. (Shapiro 1998, 148)

In other words, the extension of the intuitive notion of logical consequence seems to be the subset of the representationally valid consequences that are also interpretationally valid. This corresponds precisely to Buridan's notion of a formal consequence, as shown in the previous section : a formally valid consequence is a materially valid consequence such that all substitutional instances of its non-logical terms are (materially) valid consequences.

2.3.2.2.3 Extension. In terms of its extension, the intuitive notion of logical consequence simply corresponds to the intersection between the sets of RE-consequences and of TI-consequences, or, in Buridan's framework, to the subset of materially valid consequences that are also formally valid (Figure 2.3.2.2.3).

Purely in terms of extensional adequacy, it is irrelevant whether we consider the representational notion or the interpretational notion as primitive: the subset of the representationally valid consequences that are also interpretationally valid is the same as the subset of the interpretationally valid consequences that are also representationally valid. But we must bear Etchemendy's warning in mind: extensional adequacy is not sufficient, the proper conceptual analysis must also be offered. Etchemendy claims that the model-theoretic notion of logical consequence does capture the extension of the intuitive notion in the case of first-order logic, for example, but that this happens almost by pure chance (cf. Etchemendy 1990, 8), since the underlying conceptual assumptions are not correct. In particular, the same (extensional adequacy) does not happen in the case of second-order logic.

Thus, not only do we seek to capture the extension of the intuitive notion of logical consequence, we also want an adequate conceptual analysis. For this purpose, turning to Buridan's theory seems very fruitful, insofar as, in his account, the representational notion (corresponding to the notion of material consequence) is clearly the primitive notion. There is not even a term for the putative consequences that satisfy the substitutional criterion but not the modal one (see Figure 2.3.2.2.3).

In sum, in both Shapiro's hybrid notion of logical consequence and Buridan's notion of formal consequence, the modal component seems to be the central feature;

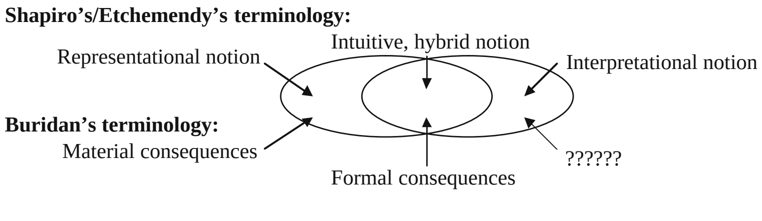


Figure 2.3.2.2.3. Correspondence of terminologies

now, insofar as the intuitive concept of validity seems to be essentially that of truth-preservation, the core of the intuitive concept of logical consequence is indeed the modal intuition. The interpretational/substitutional criterion is applied only once the representational/modal criterion has been applied, to capture the set of valid consequences whose validity is not only a matter of truth-preservation, but also of their form. Hence, the analysis of Buridan's theory seems to yield not only extensional adequacy w.r.t. the intuitive notion, but also the adequate conceptual analysis of the issue.

2.3.2.3 Conclusion and open questions

The conclusion to be drawn from this discussion is that the intuitive notion of logical consequence is most probably what we could call a 'hybrid' notion, as Shapiro claims, and that the analysis of Buridan's theory of *consequentia* is illuminating for the purpose of attaining a better understanding of this hybrid notion (since Shapiro's account thereof is rather brief). But some issues remain to be discussed.

a. Is **formal** consequence equivalent to **logical** consequence? – So far I have been considering the two notions as roughly equivalent; in particular, I claim that Buridan's notion of **formal** consequence is a very good model of the intuitive notion of **logical** consequence. But one may object that there is more to logic than just form – for example, that analytic consequences that are not formal are also in some sense logical. In other words, a deeper discussion of this issue is needed.

b. Consequence and validity are fundamentally modal notions, but **logical** consequence concerns form. That is precisely the role of the technical device of interpretation/substitution: it takes care of the formality requirement of logic (isomorphism of models/variation of interpretation). But in which sense the essential trait of logic is that of being formal is again an issue that deserves a more detailed analysis (cf. MacFarlane 2000; part 4 of the present text).

c. I claim that material consequence must remain the primitive notion (cf. Read 1994). Patterns of logically/formally valid consequences are attempts to model and capture an increasing number of materially valid consequences, but these formal patterns do not offer the very foundations for validity. The issue of logical/non-logical boundary is precisely a sign of that: the very purpose of logic is to model the logical form of propositions – but the models remain incomplete, and are always only partial approximations.

> What must be acknowledged is that belief that every valid argument is valid in virtue of form is a myth, and exclusive concentration on the study of pure forms of argument does a disservice both to logic and to those who can be helped by it. Validity is a question of the impossibility of true[240] premises and false conclusion for whatever reason, and there are arguments which are materially valid and where that reason is not purely logical. (Read 1994, 264)

Whether a given language has sufficient expressive power to capture in formal patterns certain regularities concerning valid consequences – that is, what valid arguments

[240] We now know that if the semantics is token-based, mere truth and falsity are not sufficient to capture the notion of valid consequence, but Read's general point applies to the present analysis *mutatis mutandis*.

displaying a similar general structure have in common – is merely a fact about the language in which the argument is conduced, and not about the validity of the argument itself. Obviously, the recognition of such patterns is one of the aims of logic, given their usefulness for the assessment of validity, but they are not what justifies the validity of a consequence in the first place. This, however, is a general view on the very nature of logic that must be argued for in more detail than what I can offer at present.

2.4 THE BURIDANIAN THEORY OF INFERENTIAL RELATIONS BETWEEN DOUBLY QUANTIFIED PROPOSITIONS

In his *Summula de suppositionibus*[241] and *Sophismata* chap. 3,[242] John Buridan makes several remarks on the logical properties of terms and propositions based on the subdivisions of personal supposition. On the basis of those remarks, in an article entitled 'A theory of immediate inference contained in Buridan's logic' (1993), Elizabeth Karger presents a reconstruction of what could be called the 'Buridanian theory of inferential relations between doubly quantified propositions'.[243] These are propositions such as '*Homo videt omnem asinum*', '*Omnis homo est animal*' etc. The point of view adopted by E. Karger in her reconstruction is primarily syntactical. She analyzes the syntactical features of propositions that force one or another kind of personal supposition upon their terms, thus encountering a complex set of rules. Eventually, four propositional schemata are defined, and the inferential relations between them are established.[244]

Here, I offer a semantic analysis complementary to Karger's results. It will be shown that, on the basis of Buridan's definitions of each mode of personal supposition, the inferential relations just mentioned hold also from a semantic point of view. As Karger does, I will be employing a rather simplified framework, in order to outline the most important features of Buridan's theory. Moreover, I will be employing methods and techniques that might seem anachronistic, insofar as they are typical products of recent model-theoretic logic, but I believe the use of these tools successfully clarifies some aspects of Buridan's theory.

2.4.1 Review of Karger's results

The class of Latin propositions being dealt with (in Karger's terms, 'basic categoricals') is of the form

$$\xi\, a\, \mathbf{R}\, \zeta\, b$$

[241] (Buridan 1998) and (Buridan 2001, Treatise 4).
[242] (Buridan 2004) and (Buridan 2001, 863–875).
[243] The term 'doubly quantified propositions' used for the medieval context is in fact an anachronistic projection, since, for the medieval authors in question, these are simply 'basic categoricals' (more on this in the next section).
[244] In (Ashworth 1978) E.J. Ashworth offers a similar analysis of the post-Buridan, later developments in theories of multiple quantification.

where '*a*' and '*b*' stand for terms, 'ξ' and 'ζ 'for quantifying words (syncategorematic terms) and 'R' for an (extensional) verb,[245] a common case thereof being the copula '*est*'.[246] The terms can be either in the nominative or in the accusative cases. The two quantifying words may or may not be the same in a given proposition. It can also happen that no quantifying word occurs in front of a term, in which case the effect obtained is similar to the effect of an indefinite article, such as the English 'a'. Otherwise, the quantifying words are most often of the kind '*omnis*', '*aliquid*' etc. Signs of negation can occur either preceding the whole proposition, or as a quantifying word ('*nullus*'), or immediately preceding the verb. In the present investigation, however, particular properties of negating signs will not be discussed. Modal and tensed propositions will also not be dealt with at this point, but presumably similar results could be obtained with the correct application of Buridan's notion of ampliation.[247]

A proposition may not always be in the normal form with respect to its word order. Another element of perturbation is the negation, which can be introduced virtually in any place of the proposition, but provoking different effects on the supposition of the terms. Based on Buridan's remarks, Karger proposes rules to 'normalize' any proposition of this kind.

From this very brief sketch, it can be seen that the syntactical analysis of the modes of personal supposition is a complicated matter. This is due most of all to the complexity of the group of expressions one is dealing with, that is, a large corpus of Latin propositions, which do not always follow easily recognizable patterns. Therefore, the rules determining which mode of personal supposition a given term has in a proposition, on the basis of its syntactical features, must encompass this variety of forms and their corresponding semantic behavior.

E. Karger successfully brings these irregularities under control. In order to arrive at a coherent and complete set of rules, she introduces concepts such as the 'mode preserving non-ordinary form' of propositions and deals with the delicate effect of negating signs. Once these propositions have been cast in 'normalized' form (their 'mode preserving non-ordinary form'), the rules defining the modes of personal supposition seem to cover all cases.[248] She then concludes that there are only four

---

245 As noted by Karger (1993, 409), intensional verbs must be excluded.

246 Further on, I will also use variables $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{t}$ etc., ranging over objects, whereas *a* and *b* are used as schematic letters for terms, following the conventions already adopted in the formalization of supposition theory.

247 (Buridan 2001, section 4.6)

248 Karger's treatment of the syntactic rules determining the kind of supposition of a given term sits well with Parson's interpretation of 14th century theories of personal supposition as 'global quantificational effect' (cf. Parsons 1997). According to Parsons, in the 13th century the kind of supposition of a term was determined simply by the sincategorema in question – for example, 'man' in 'not every man is white' would have distributive supposition simply because it is under the scope of a sign of distribution, namely 'every'. By contrast, in the 14th century – that is, the period Karger is dealing with – the modes of personal supposition were no longer seen as defined exclusively by the kinds of syncategorema used, but also by the effect of negating expressions over them. Now, the effect of negating expressions on the modes of supposition of terms is precisely one of the main topics of Karger's paper.

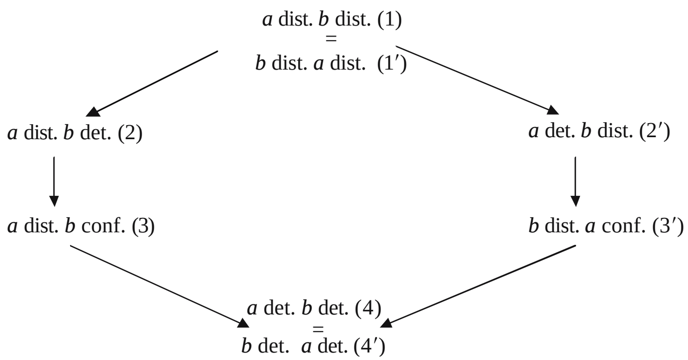


Figure 2.4.1. Inferential relations between interpretational schemata

possible interpretational schemata with respect to the personal supposition of the terms. Let '*a*' stand for the subject and '*b*' for the predicate, 'dist.' for confused and distributive supposition, 'det.' for determinate supposition and 'conf.' for merely confused supposition. The four interpretational schemata are:[249]

– (1) *a* dist. *b* dist.
– (2) *a* dist. *b* det.
– (3) *a* dist. *b* conf.
– (4) *a* det. *b* det.

Since the order of occurrence of '*a*' and '*b*' in the proposition can vary, the same holds if, in the above schemata, '*a*' is replaced by '*b*' and vice-versa. These schemata define classes of 'propositional forms': two propositions have the same propositional form if they share the same structure and only vary with respect to their categoremata and/or verbs. Different propositional forms belong to the same interpretational schema: 'propositional form' is a syntactic notion, whereas 'interpretational schema' is a semantic notion.

On the basis of Buridan's rules, Karger reaches the conclusion that the relations of inference as depicted in Figure 2.4.1 hold between these schemata (the arrow represents the logical implication and the sign of equality the double implication – logical equivalence).

This means that propositions which are identical with respect to their categorematic terms, the verb relating those terms and their polarity (i.e., being affirmative or negative), but varying with respect to the kinds of personal supposition of their terms (presumably as a result of the presence of different syncategorematic terms or of different word order) – thus belonging to different interpretational

[249] We have already encountered this fact in section 1.5.2 above. A remark on notation: the notation to be used in this section is a simplified version of the notation used in section 1.5.2, since here we are interested only in the investigation of the inferential relations between propositions.

schemata – have inferential relations with each other according to the relations depicted in the graph above.[250]

In the final part of her paper, E. Karger argues that this hexagon neatly corresponds to a similar hexagon, drawn on the basis of modern quantification theory, its edges being formulae featuring the two quantifiers (universal and existential) in different orders. This correspondence provides evidence for the correctness of the inferential relations recognized by Buridan, as they seem to mimic inferential relations recognized by modern quantification theory. This is also why I use the term 'doubly-quantified propositions' to refer to the class of basic categoricals here at stake, even though it would be undoubtedly anachronistic and mistaken to simply identify the medieval theory being analyzed here with modern quantification theory.

The validity of such inferential relations is established by Karger on the basis of syntactical elements. However, the definitions of the modes of personal supposition that are found in virtually all medieval treatises on supposition, including Buridan's, provide an alternative way of substantiating these relations. If the semantic analysis of these definitions confirms the syntactical inferential relations just described, then one can say, with the license of anachronism, that this fragment of Buridan's system is 'sound', insofar as the inferential relations syntactically recognized as legitimate are also shown to hold from a semantic point of view. The purpose of the following pages is to offer a 'proof of soundness' for the Buridan/Karger inference rules according to the semantics of personal supposition.

The first step will be to analyze Buridan's formulations of the definitions of the modes of personal supposition, in the hope of extracting sufficient elements from them so as to set down definitions of what it means for a term to have such-and-such personal supposition. Once this is done, it will be possible to establish which models satisfy each propositional schema. Finally, it will be shown that the relations of inference syntactically recognized by Buridan can be proven to hold semantically, in particular that they correspond to relations of containment between classes of models satisfying the interpretational schemata.

2.4.2 Modes of personal supposition[251]

In the secondary literature, there have been numerous studies devoted to the analysis of medieval theories of modes of personal supposition.[252] However, it is reasonable to say that just as many questions were raised as answered by these studies, and that

[250] Example: 'Every man sees every donkey' implies 'A (specific) donkey is seen by every man', which implies 'Every man sees a donkey', which implies 'A man sees a donkey'.

[251] This section follows very closely the material presented in section 1.5.2; the latter concerns Ockham's theory of the modes of personal supposition, but in practice, with respect to this fragment of their logical system, there are few, if any, significant differences between the two authors.

[252] Most of these studies and discussions were devoted to Ockham's theory of supposition, not to Buridan's, but many of the problems raised regarding Ockham's theory have counterparts in Buridan's theory. Some of the studies on supposition theory are: (Swiniarski 1970); (Priest and Read 1977); (Spade 1988); (Willing 1991); (Karger 1984).

these theories remain puzzling most of all with respect to their general purpose. It is particularly unclear what the inferential relations (the 'ascents and descents') between propositions featuring common terms and the corresponding singular propositions featuring demonstratives should account for.[253]

As already argued in section 1.5.2, it seems to me that the definitions of the three modes of personal common supposition can be said to account for the semantic properties of the corresponding 'quantified' terms in a given proposition, that is, what it means, in terms of *supposita* (the entities supposited for), to have such-and-such supposition.[254] Earlier theories of supposition did not rely on ascent and descent for these definitions (in particular Peter of Spain's[255] and William of Sherwood's[256] – but they defined the concepts of distributive mobile/immobile supposition in terms of ascent and descent); in contrast, 14th century theories such as Buridan's rely heavily on these inferential relations to define the kinds of personal supposition. The problem with this development is, as has been argued in the literature, that the definitions so cast simply do not work properly, since there seems to be no logical equivalence between the original proposition and the chains of disjunction and conjunction.[257] Often, a given mode of personal supposition is defined in terms of which ascents and descents are **not** possible, which is obviously not sufficient with respect to definitions *sensu stricto*.

[253] It must be mentioned that the inferential relations of ascent and descent are, according to Buridan himself, merely *de materia* inferences, whereas the relations of inference between basic categoricals that are the object of this study are *de forma* inferences, that is, they hold in virtue of their form.

[254] On p. 416, E. Karger says: 'In *De Suppositionibus,* Buridan provides two sets of such criteria [to identify, for any general term, personally suppositing in a proposition, which one of the three kinds of supposition belongs to it]'. (By these two sets she means a syntactic set and a semantic set). 'If these criteria be regarded as definitions, it follows that Buridan is in effect introducing for each mode two different concepts, the underlying – and unproven – thesis being that they are coextensive. With respect to our present inquiry however, such a duplication of concepts is superfluous: we need no more than one definition of each mode of supposition, or one set of criteria for identifying the mode of a term'. (Her choice, as we already know, will be in favor of the syntactic set).

The idea that both sets of criteria were supposed to do the same job, namely that of determining which kind of (personal) supposition a term has in a proposition, seems awkward (even though it is explicitly expressed by Buridan himself). Why would Buridan, or any other medieval logician, need two sets of criteria to accomplish the same task? Moreover, the information about which kind of supposition a term has in a proposition seems to be only partially sufficient for whichever goal the application of supposition theory was supposed to have. It seems to take us nowhere to know which supposition a term has without knowing what it means to have such-and-such supposition. In this sense, my interpretation of the modes of personal supposition differs from Karger's.

My general view about theories of supposition (as presented in part 1 above) is that they are formal theories for the semantic analysis of propositions, that is, for the analysis of what is asserted by the proposition based on its formal properties. According to this view, the syntactic rules are supposed to determine which kind of supposition a term has, and the semantic rules determine what it means for a term to have such and such supposition. But this point is in fact immaterial for the 'proof of soundness' that I present here.

[255] (Peter of Spain 1972, 82–83).

[256] (William of Sherwood 1995, 136).

[257] Cf. (Spade 1996, 289 *et passim*).

However, even if not good enough for definitions of logical equivalences, the inferential relations of ascent and descent do seem to provide a glimpse at what is asserted by means of a proposition. In fact, the idea that the meaning of a proposition is defined in terms of the inferential relations it entertains with other propositions has been given an elaborate theoretical status in recent developments.[258] Well, it seems plausible that the purpose of these medieval theories was somehow to define the meaning of a given proposition (in particular the 'doubly quantified' ones we are dealing with) in terms of its inferential relations with chains of conjunctions and disjunctions of propositions of the form 'This *a* is *b*'. Indeed, within the inferentialist framework,[259] the fact that the inferential relations between these 'quantified' propositions and the propositions of the form 'This *a* is *b*' are, according to Buridan, materially but not formally valid, only confirms that what is at stake with these inferential relations are the contents of the propositions involved.

Hence, it can be said that the moves of ascent and descent are not so much operational definitions as they are heuristic explanations of meaning. Therefore, in a reconstruction such as this one, we face the task of extracting definitions that are operational from the relations of ascent and descent. The task is made more delicate by the brevity of Buridan's considerations on this particular aspect of his theory. Indeed, in Van der Lecq's edition of *De Suppositionibus*, the semantic considerations on personal supposition cover three pages, whereas the syntactic analysis on which E. Karger has based her study covers seventeen pages.

Still, Buridan's brief remarks seem to supply enough material for the reconstruction of the semantic definitions of the modes of personal common supposition. Some preliminary clarifications: recall that the personal *supposita* of a term are all the entities that it stands for personally, that is, all the entities of which the term in question can be predicated (cf. section 1.5.). The class of these *supposita* is dependent on the time and mode of the verb of the proposition; in the present case, since we are only dealing with non-modal present-tensed propositions and with terms having personal supposition, the *supposita* of a term are all the entities of which it can be predicated in the present, actual world.

2.4.2.1 Determinate supposition

Buridan begins by introducing determinate supposition. He first presents a heavily interpreted version of Peter of Spain's remarks on the topic (from Peter's *Summulae*), and then draws some conclusions himself. Here are the relevant passages:

The supposition of a term is called determinate if it is necessary for the truth of the proposition in which it is posited (or of a proposition similar in form) that it be true for some determinate suppositum. For example,

---

[258] Cf. (Brandom 2000) and (Brandom 1994). According to Brandom (2000, chap. 1), authors such as the young Frege, Sellars, Dummett and Gentzen, among others, are all proponents of this view.

[259] '… according to the inferentialist view of conceptual contents, it is these implicitly recognized material inferential relations that conceptual contents consist in…'. (Brandom 2000, 60)

if 'A man is white' is true, then it has to be true for this man and this white thing, or for that man and that white thing, and so on for each.[260] (Buridan 2001, 262)

Nevertheless, I should say that in determinate supposition the proposition need not be true for one suppositum only; indeed, sometimes it is true for any suppositum. But it is necessary and sufficient that it should be true for one.[261] (Buridan 2001, 263)

On the basis of these semantic definitions, it is, according to Buridan, immediately evident that some specific inferential relations of ascent and descent hold between a proposition with a term in determinate supposition and certain singular propositions (propositions with either a proper name or a demonstrative pronoun as their subjects). Notice that he first presents the semantics of determinate supposition in terms of verification of the proposition, and that the relations of ascent and descent follow from the semantic definition. That is, ascents and descents provide the semantics of the propositions in question only derivatively.[262] They are however a powerful device to grasp the meaning of a proposition, excellent rules of thumb for the determination of the kind of personal supposition that a term has.

So we have to note immediately that there are two conditions for the determinate supposition of some common term. The first one is that from any suppositum of the term it is possible to infer the common term, the other parts in the proposition remaining unchanged. For example, since, in 'A man runs', the term 'man' supposits determinately, it follows that 'Socrates runs; therefore a man runs', 'Plato runs; therefore a man runs', and so on for any singular contained under the term 'man'.[263] (Buridan 2001, 263)

The second condition is that from a common term suppositing in this manner all singulars can be inferred disjunctively, by a disjunctive proposition. For example: 'A man runs; therefore Socrates runs or Plato runs or John runs…' and so on for the rest.[264] (Buridan 2001, 263)

Three aspects seem to be significant for a term in a proposition having determinate supposition (2 and 3 being derived from 1):

(1) A common term having determinate supposition supposits for all its *supposita*, but for the truth of this proposition it is sufficient that it be verified of only one.
(2) From a singular proposition (with a proper name or demonstrative) the proposition with the corresponding common term as its subject in determinate supposition follows.

---

[260] Vocatur autem suppositio determinata alicuius termini, si necesse sit ad veritatem illius vel talis propositionis in forma quod ipsa sit vera pro aliquo determinato supposito. Ut si ista sit vera 'homo est albus', oportet quod sit vera pro isto homine vel pro isto albo, est sic de aliis. (Buridan 1998, 49, 6–10).

[261] Dico tamen quod in suppositione determinata non oportet quod verificetur pro uno solo supposito, immo aliquando est vera pro quolibet; sed hoc requiritur et sufficit quod sit vera pro uno. (Buridan 1998, 49, 19–21)

[262] P.V. Spade seems to defend similar views in (Spade 1996, 293–298).

[263] Unde notandum statim quod duae sunt condiciones suppositionis determinatae alicuius termini communis. Prima est quod ex quolibet supposito illius termini possit inferri terminus communis remanentibus aliis in propositione positis. Verbi gratia, quia in ista 'homo currit' iste terminus 'homo' supponit determinate, ideo sequitur 'Socrates currit; ergo homo currit', 'Plato currit; ergo homo currit', et sic de quolibet alio singulari contento sub 'homine'. (Buridan 1998, 49, 22–50)

[264] Secunda condicio est quod ex termino communi sic supponente possunt inferri omnia disiunctive secundum propositionem disiunctivam. Verbi gratia, sequitur 'homo currit; ergo Socrates currit vel Plato currit vel Johannes currit' et sic de aliis. (Buridan 1998, 50, 5–7)

(3) The **disjunction** of these singular propositions follows from the corresponding proposition whose subject has determinate supposition.

Aspect (1) deserves further analysis. Some medieval authors stated clearly that a term having determinate supposition supposits for one entity; others said that it supposits for all its *supposita*, insofar as any of them could be supposited for. The problem here is that if a term in determinate supposition supposits for all its *supposita*, what is then the difference between determinate supposition and distributive supposition? There would seem to be none. Nuances were then made necessary: some said that a term in determinate supposition supposited for all its *supposita* in a disjunctive way; others said that such terms supposited for all its *supposita*, but only one of them was required for the truth of the proposition.

Let us be faithful to Buridan's text. He clearly follows Peter of Spain in maintaining that a common term having determinate supposition supposits for all its *supposita*,[265] but he adds that it is necessary and sufficient that it be verified of one *suppositum*. One could compare determinate supposition to the existential quantifier in the following way: the variable associated to the existential quantifier ranges over its whole domain, as much as a term having determinate supposition supposits for all its *supposita*, but it is sufficient that one element of the domain satisfy the variable, as much as it is sufficient that one *suppositum* verify the supposition.

As for ascents and descents, the ascent from any relevant singular proposition to the corresponding proposition having the common term in question in determinate supposition and the descent from the latter to the disjunction of relevant singular propositions are valid exactly because of the condition that a proposition having a common term in determinate supposition must be verified for one singular to be true. Given this condition, the relations of inference just mentioned are intuitively correct.

Definition[266] **2.4.2.1:** A term '*a*' has determinate supposition in $\varphi <=>$ It is necessary and sufficient for the truth of φ that it be verified for one *suppositum* of '*a*'.[267]

[265] Cf. (Buridan 2001, 866).

[266] In fact, with respect to definitions 2.4.2.1–2.4.2.3 as well as definitions 2.4.3.1–2.4.3.4 below, it is unclear whether they should be formulated as full-fledged bi-directional definitions. According to my interpretation of supposition theory (cf. Part 1), it is not the semantic properties of a proposition (in particular the relations of ascent and descent) that determine the (personal) supposition of its terms; rather, it is its quasi-syntactic features that determine the supposition of its terms, and this in turn determines its semantic properties. Therefore, in the case of these 'definitions', the left-to-right direction would not only be sufficient for the present purposes, but it would also be a more faithful rendering of my interpretation. However, Buridan himself seems to formulate them as full-fledged definitions, and they are in any case formally correct as they stand (since presumably the left-hand side condition is satisfied whenever the right-side condition is satisfied), even if the right-to-left direction is in practice epistemically inadequate in the supposition framework thus interpreted (since it is not the semantic properties of a proposition that determine the supposition of its terms). By the way, in section 1.5, the formulations of the different kinds of personal supposition are one-direction implications.

[267] Notice that, thus formulated, the definition does not exclude the possibility of the proposition being verified for more than one *suppositum* of '*a*'. What it states is the minimal, sufficient condition for a term to have determinate supposition.

2.4.2.2 Confused and distributive supposition

In contrast, confused and distributive supposition (henceforth, distributive for short) seems to be conceptually simpler. This is how Buridan defines it:

> Confused and distributive supposition is that in accordance with which from a common term any of its supposita can be inferred separately, or all of them at once conjunctively, in terms of a conjunctive proposition. For example, from 'Every man runs' it follows that therefore 'Socrates runs', therefore 'Socrates runs and Plato runs …' and so on for the rest.[268] (Buridan 2001, 264)

> And it is obvious that distributive supposition differs from the other suppositions, for in its case a common term implies any of its singulars separately; whereas the other suppositions do not. Therefore, if the proposition is true, it has to be true for any *suppositum*, which is not required in the other cases of supposition.[269] (Buridan 2001, 264)

It is worth noticing that, for distributive supposition, Buridan provides semantic definitions only in terms of relations of ascent and descent. Making use of the inferentialist idea that the meaning of a proposition is given by its inferential relations with other propositions, it becomes clear that, if any of the relevant singular propositions relative to a given common term, as much as their conjunction, can be inferred from a proposition having as its subject the common term in distributive supposition, this is the case exactly because all its *supposita* must verify the proposition. Hence the following definition:

Definition 2.4.2.2: A term '*a*' has distributive supposition in φ $<=>$ It is necessary and sufficient for the truth of φ that it be verified for all *supposita* of '*a*'.

2.4.2.3 Merely confused supposition

Merely confused supposition is the most delicate kind of personal supposition, conceptually speaking. One of the reasons for this is that this kind of supposition cannot be defined independently of the supposition of the other term in the proposition. In particular, according to the list of propositional schemata established by E. Karger, merely confused supposition only occurs for a term in a proposition when the other term in the proposition has distributive supposition.

> But merely confused supposition is that in accordance with which none of the singulars follows separately while retaining the other parts of the proposition, and neither do the singulars follow disjunctively, in terms

---

[268] Suppositio confusa et distributiva est secundum quam ex termino communi potest inferri quodlibet suppositorum seorsum vel omnia etiam simul copulative secundum propositionem copulativam, ut, 'omnis homo currit; ergo Socrates currit et Plato currit' et sic de aliis. (Buridan 1998, 50, 14–18)

[269] Et manifestum est quod suppositio distributiva differt ab aliis suppositionibus, quia terminus communis secundum eam infert quodlibet suorum singularium seorsum; aliae autem hoc non faciunt. Ideo etiam si propositio sit vera, oportet quod sit vera pro quolibet supposito; quod non requiritur in aliis. (Buridan 1998, 51, 3–6)

of a distjunctive proposition, although perhaps they do follow by a proposition with a disjunct term.[270] (Buridan 2001, 264)

Merely confused supposition, on the other hand, differs from determinate supposition, because in the case of confused supposition the singulars cannot be inferred from the common term by means of a disjunctive proposition, whereas this can correctly be done with determinate supposition. For example, in the proposition 'Every man is an animal', the term 'animal' has merely confused supposition, and the inference 'Every man is an animal; therefore every man is this animal or every man is that animal ...' (and so on for the rest) is not valid, for the antecedent is true and all the consequents are false.[271] (Buridan 2001, 264)

In terms of the relations of ascent and descent, a term with merely confused supposition does not have the inferential relations defining the two other kinds of personal supposition: neither the disjunction nor the conjunction of the corresponding singular propositions follows. What does follow is a proposition whose extreme (subject or predicate, depending on which term has merely confused supposition in the original proposition) is disjunctive.

In '*Omnis homo est animal*', '*animal*' has merely confused supposition, according to the syntactic rules proposed by Buridan and reconstructed by Karger. Buridan says that, from this proposition, '*Omnis homo est hoc animal vel omnis homo est illud animal...*' and so forth, does not follow, but '*Omnis homo est hoc animal vel illud animal...*' does follow (the disjunctive extreme mentioned by Buridan).

The idea is thus that each *suppositum* of the other term in the proposition (which has distributive supposition) is put in a relation with some *suppositum* of the term having merely confused supposition. Ordered pairs of entities being in relation R to one another verify the proposition in question, one of the relata being a *suppositum* of the term having distributive supposition and the other being a *suppositum* of the term having merely confused supposition. There are as many ordered pairs as there are *supposita* for the term having distributive supposition, and the other relatum of each pair is any of the *supposita* of the term having confused supposition.

Definition 2.4.2.3: '*b*' has merely confused supposition in φ <=> '*a*' has distributive supposition in φ, and it is necessary and sufficient for the truth of φ that each *suppositum* of '*a*' be mapped into a *suppositum* of '*b*', while they are not all mapped into the same *suppositum* of '*b*'.[272]

270 Suppositio confusa tantum est secundum quam non sequitur aliquod singularium seorsum retentis aliis in propositione positis, nec sequuntur etiam singularia disiunctive secundum propositionem disiunctiva, licet forte sequantur secundum propositionem de disiuncto extremo. (Buridan 1998, 50, 18–22)

271 Suppositio autem confusa tantum differt a suppositione determinata, quia secundum suppositionem confusam non inferuntur ex termino communi singularia secundum propositionem disiunctivam, quod bene fit secundum suppositionem determinatam. Verbi gratia, in ista propositione 'omnis homo est animal' iste terminus 'animal' supponit confuse tantum. Et non sequitur 'omnis homo est animal; ergo omnis homo est hoc animal vel omnis homo est illud animal' et sic de aliis, quoniam prima est vera et omnes aliae sunt falsae. (Buridan 1998, 51, 7–13)

272 This last clause is necessary in order to distinguish merely confused supposition from determinate supposition.

2.4.3 The models verifying each interpretational schema

Given Buridan's definition of the consequence relation – the consequent follows the antecedent iff it is impossible for the antecedent to be true and the consequent false – we shall be interested in the models that verify each propositional scheme. We shall consider two sets within each model, namely the set of *supposita* of the subject – named 'A' – and the set of *supposita* of the predicate – named 'B'. Moreover, we will consider the sort of mapping established between certain members of A and certain members of B, defined by the relation expressed by the verb of the proposition. So, the copula '*est*' expresses a relation of identity between certain members of A and certain members of B; a verb such as '*videt*' expresses the relation of seeing between certain members of A and certain members of B.

Based on the definitions above, it is possible to define which models satisfy each interpretational schema. In each case, it is presupposed that in the propositions in question its terms are related by the relational term 'R'.

2.4.3.1 Schema (1) – a *dist.* b *dist.*

A proposition φ, whose two terms '*a*' and '*b*' both have distributive supposition, and are related by the relation denoted by 'R', is true in a certain model iff **all** members of A are related (or not, if the proposition is negative) in relation R to **all** members of B. As said, distributive supposition selects all the *supposita* of the term in question to be in relation R. Well, if both '*a*' and '*b*' have this kind of supposition, that means that the relation R maps all members of B into all members of A (and vice-versa). Figure 2.4.3.1 represents the necessary and sufficient condition for φ to be true in a model ℵ (notation: ℵ ⊩ φ) ('*' represents the elements of A and B – the *supposita* of terms '*a*' and '*b*' – and the lines between elements of A and elements of B represent the relation R between them[273]).

In sum, this interpretational schema is true in a model ℵ iff the Cartesian product of elements of A and B occurs in this model. Thus:

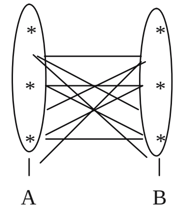


Figure 2.4.3.1. Model satisfying schema (1)

[273] This relation may be affirmative or negative, if the proposition is negative, in which case the proposition asserts that the relation expressed by the verb does not exist between each *suppositum* in question.

Definition 2.4.3.1: In φ its terms '*a*' and '*b*' both have distributive supposition <=> ℵ ⊩ φ iff $x\,\mathrm{R}\,y$, for every element x of A and every element y of B, in model ℵ.[274]

Example: *Omnis homo videt omnem asinum.*

This proposition is true if every man is in a relation of seeing with all donkeys – if every man sees each and every donkey.

2.4.3.2 Schema (2) – a *dist.* b *det.*

A proposition φ, whose term '*a*' has distributive supposition and whose term '*b*' has determinate supposition, and the two terms are related by the relation denoted by 'R', is true in a certain model iff all members of A are related in relation R to one and the same member of B.[275] Distributive supposition 'selects' all elements of A to be in relation R, but determinate supposition 'selects' only one (unspecified) element of B to be in relation R. Thus, all elements of A are mapped into the same element of B (see Figure 2.4.3.2).

Definition 2.4.3.2: In φ its term '*a*' has distributive supposition and its term '*b*' has determinate supposition <=> ℵ ⊩ φ iff $x\,\mathrm{R}\,t$, for every element x of A and one single element t of B, in model ℵ.

Example: *Asinum omnis homo videt.*

According to the regimented use of Latin at Buridan's time,[276] this proposition is true iff every man sees one and the same donkey.

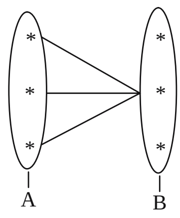


Figure 2.4.3.2. Model satisfying schema (2)

[274] Likewise, if the proposition is negative, then ℵ ⊩ φ iff ∼ $(x\,\mathrm{R}\,y)$, for every element x of A and every element y of B, in model ℵ. The same holds for all of the upcoming definitions.

[275] Notice again that members of A may be related by R to other members of B.

[276] In this case, it regards word order. If the predicate is a general term and precedes the subject with a distributive sign of universality, then the predicate has determinate supposition and the subject has distributive supposition (at least one single member of the domain of the predicate is related to all members of the domain of the subject). This regimentation of the language displays a strong similarity with the idea of **scope** in current quantified logic.

2.4.3.3 Schema (3) – a *dist.* b *conf.*

A proposition φ, whose term '*a*' has distributive supposition and whose term '*b*' has merely confused supposition, and the two terms are related by the relation denoted by 'R', is true in a certain model iff **all** members of A are related in relation R to **some** member of B. That is, there is no element of A which is left 'unattended' (by some element of B), but the elements of B in question are unspecified. Three possibilities can occur, as depicted below.

1 – Surjection. All members of A are mapped into some member of B, but some member of B is related to more than one member of A.

Example: *Omnis homo videt asinum.*

This proposition is true, for example, if each *suppositum* of '*homo*' sees one individual among the *supposita* of '*asinum*', but it is possible that a certain donkey is seen by more than one man, as much as it is possible that another donkey is not seen by any man (Figure 2.4.3.3).

2 – Injection. All members of A are mapped into a member of B, and an element of B is related to at most one member of A, but there are more elements in B than in A, so some members of B are not related to any member of A.

Example: *Omnis homo est animal.*

All men are animals, but not all animals are men.

3 – Bijection. All members of A are mapped into a member of B, and all members of B are related to exactly one member of A.

Example: *Omnis homo est animal rationale.*[277]

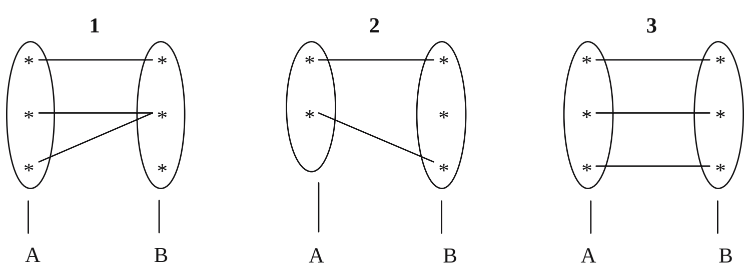


Figure 2.4.3.3. Models satisfying schema (3)

[277] Namely, a bijection is what is established by a correct definition of a term, implying equivalence between the two sides of the definition.

If a proposition φ belongs to this propositional schema, it is not possible to determine from its form alone which of the three cases applies: as far as its form goes, any of these three situations would verify it. Only by means of an analysis of the **content** of the proposition is it possible to know what exactly is being asserted, among these three possibilities. This can be tested for example, in terms of the inferential relations between the proposition in question and similar ones. For example, the converse of 'Every man is a rational animal', 'Every rational animal is a man', is entailed by the former, but 'Every animal is a man' is not entailed by 'Every man is an animal'.

Definition 2.4.3.3: In φ its term '*a*' has distributive supposition and its term '*b*' has merely confused supposition <=> $\aleph \Vdash \varphi$ iff $x \, \mathrm{R} \, y$, for every element *x* of A and some element *y* of B, in model ℵ.

2.4.3.4 Schema (4) – a *det.* b *det.*

A proposition φ, whose two terms '*a*' and '*b*' both have determinate supposition, and are related by the relation denoted by 'R', is true in a certain model iff one member of A is related in relation R to one member of B. As said, determinate supposition corresponds to (at least) one of the *supposita* of the term in question being in relation R. Well, if both '*a*' and '*b*' have this kind of supposition, that means that the relation R maps (at least) one member of A to (at least) one member of B (and vice-versa) (see Figure 2.4.3.4.).

Definition 2.4.3.4: In φ its terms '*a*' and '*b*' both have determinate supposition <=> $\aleph \Vdash \varphi$ iff $x \, \mathrm{R} \, y$, for some element *x* of A and some element *y* of B, in model ℵ.

Example: *Homo est albus.*

This proposition is true iff at least one of the *supposita* of '*homo*' has a relation of identity with one of the *supposita* of '*albus*'.

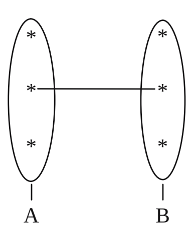


Figure 2.4.3.4. Models satisfying schema (4)

2.4.4 The relations of inference

In sections 2.1–2.3 I have presented a reconstruction of Buridan's theory of *consequentia* where many of its different aspects are taken into account, in particular the effect of his commitment to tokens, the ambiguous status of the medieval concept of *consequentia* – between the modern concept of consequence and the modern concept of inference – and his definition of formal consequence. But I have also remarked (in section 2.3.1.4) that, according to Buridan, in most cases the improved definition of consequence is not required and the modal/incompatibility criterion is sufficient. Moreover, even though Buridan holds that the consequences to be investigated here are formally valid,[278] since the present investigation is essentially semantic, their formal character is now not under scrutiny.

For these reasons, in the simplified framework adopted for the present investigation, such subtle aspects of Buridan's theory of consequence will not come into play. In other words, we shall be content with a definition of the relation of consequence – here understood as roughly synonymous with inference or implication – that disregards the implications of Buridan's token-based semantics. In this sense, the definition will be virtually identical to the standard model-theoretical definition of consequence.

A relation of consequence – an inference – between two propositions φ and ψ (represented by '$\varphi \Rightarrow \psi$') holds iff it is impossible for φ to be true while ψ is false. That is, in all models in which φ is true ψ is also true.

Definition 2.4.4: $\vDash \varphi \Rightarrow \psi \Leftrightarrow$ For every model $\aleph$ such that $\aleph \Vdash \varphi$, $\aleph \Vdash \psi$.

2.4.4.1 Proofs by absurdity

Based on this definition of consequence, in this section I prove that the relations of implication between basic categoricals, established on the syntactical level by E. Karger, also hold semantically.

The proof by absurdity of the truth of a relation of implication proceeds by assuming that this relation does not hold, and from this assumption a contradiction will be derived. To assume that a relation of consequence does not hold (in all models) consists of assuming that there is at least one model $\aleph$ in which $\aleph \Vdash \varphi$ and $\aleph \nVdash \psi$.

2.4.4.1.1 Schema (1) implies schema (2)

Theorem 2.4.4.1.1: If φ and ψ are propositions having the same verb, the same polarity and the same terms '*a*' and '*b*', such that in φ both terms have distributive supposition, whereas in ψ '*a*' has distributive supposition and '*b*' has determinate supposition, then $\vDash \varphi \Rightarrow \psi$.

---

[278] 'Hence these types of supposition are related to each other in such a manner that distributive supposition formally implies determinate supposition and even merely confused supposition, but neither of these latter implies distributive supposition by a formal consequence'. (Buridan 2001, 264)

Proof: Suppose that $\nvDash \varphi => \psi$. Then there is at least one model $\aleph$ such that $\aleph \Vdash \varphi$ and $\aleph \nVdash \psi$.

If $\aleph \Vdash \varphi$ and its terms '*a*' and '*b*' have distributive supposition, then all members of A are mapped into all members of B.

If $\aleph \nVdash \psi$ and its term '*a*' has distributive supposition and '*b*' has determinate supposition, then there is no element of B which is related to all members of A (and that covers the cases in which either A or B or both are empty sets).

But if all members of A are related to all members of B, then there is at least one element of B (in fact, all elements of B) which is related to all elements of A. Hence, we have a contradiction.

So $\vDash \varphi => \psi$. □

2.4.4.1.2 Schema (2) implies schema (3)

Theorem 2.4.4.1.2: If φ and ψ are propositions featuring the same verb, the same polarity and the same terms '*a*' and '*b*', such that in φ '*a*' has distributive supposition and '*b*' has determinate supposition, whereas in ψ '*a*' has distributive supposition and '*b*' has merely confused supposition, then $\vDash \varphi => \psi$.

Proof: Suppose that $\nvDash \varphi => \psi$. Then there is at least one model $\aleph$ such that $\aleph \Vdash \varphi$ and $\aleph \nVdash \psi$.

If $\aleph \Vdash \varphi$ and its term '*a*' has distributive supposition and '*b*' has determinate supposition, then there is at least one element of B which is related to all members of A.

If $\aleph \nVdash \psi$ and its term '*a*' has distributive supposition and '*b*' has merely confused supposition, then there is at least one element of A which is not related to any member of B (and that covers the cases in which either A or B or both are empty sets).

But if all members of A are related to one given member of B, then there is no element of A which is not related to at least one element of B. Hence, we have a contradiction.

So $\vDash \varphi => \psi$. □

2.4.4.1.3 Schema (3) implies schema (4)

Theorem 2.4.4.1.3: If φ and ψ are propositions featuring the same verb, the same polarity and the same terms '*a*' and '*b*', such that in φ '*a*' has distributive supposition and '*b*' has merely confused supposition, whereas in ψ '*a*' and '*b*' have determinate supposition, then $\vDash \varphi => \psi$.

Proof: Suppose that $\nvDash \varphi => \psi$. Then there is at least one model $\aleph$ such that $\aleph \Vdash \varphi$ and $\aleph \nVdash \psi$.

If $\aleph \Vdash \varphi$ and its term '*a*' has distributive supposition and '*b*' has merely confused supposition, then all elements of A are related to some element of B.

If $\aleph \nVdash \psi$ and its terms '*a*' and '*b*' have determinate supposition, then there is not a single element of A which is related to a member of B (and that covers the cases in which either A or B or both are empty sets).

But if all members of A are related to some member of B, then there is at least one element of A (in fact, all elements of A) which is related to an element of B. Hence, we have a contradiction.

So $\vDash \varphi => \psi$. □

Similar reasoning can be applied to prove that schema **(1)** implies schema **(2′)**, that schema **(2′)** implies schema **(3′)** and that schema **(3′)** implies schema **(4).**

Naturally, given the transitive property of the implication, it is also true that **(1)** implies **(3)**, **(3′)** and **(4)**, that **(2)** implies **(4)** and that **(2′)** implies **(4).**

2.4.4.2 Proof by relation of containment

For those who might object the heavy use of the negation in a proof by absurdity (e.g., intuitionists or constructivists), I present alternative semantic proofs of the inferential relations in question in terms of the relations of containment between the classes of models that satisfy each sentential scheme.

We start from the definition of implication.

$$\vDash \varphi => \psi <=> \text{For every model } \aleph \text{ such that } \aleph \Vdash \varphi, \aleph \Vdash \psi.$$

Now, we define $\boldsymbol{UM}_{\varphi}$ <=> The class of models that satisfy φ (its members are all models $\aleph$ such that $\aleph \Vdash \varphi$).

According to the definition of implication, if $\vDash \varphi => \psi$, then for every model $\aleph$ such that $\aleph \Vdash \varphi$, $\aleph \Vdash \psi$. So if $\Vdash \varphi => \psi$ then every model that satisfies φ also satisfies ψ. So $\boldsymbol{UM}_{\varphi}$ is contained in $\boldsymbol{UM}_{\psi}$

$$\boldsymbol{UM}_{\varphi} \subseteq \boldsymbol{UM}_{\psi}.$$

Hence, to prove that $\vDash \varphi => \psi$ it is sufficient to show that $\boldsymbol{UM}_{\varphi} \subseteq \boldsymbol{UM}_{\psi}$.

2.4.4.2.1 Schema (1) implies schema (2)

Theorem 2.4.4.2.1: If φ and ψ are propositions featuring the same verb, the same polarity and the same terms ‘*a*’ and ‘*b*’, such that in φ both terms have distributive supposition, whereas in ψ ‘*a*’ has distributive supposition and ‘*b*’ has determinate supposition, then $\vDash \varphi => \psi$.

Proof: If in φ its terms ‘*a*’ and ‘*b*’ both have distributive supposition and are related by relation ‘R’, then $\aleph \Vdash \varphi$ iff xRy, for every element x of A and every element y of B, in model $\aleph$ (Definition 2.4.3.1).

So in all models that satisfy φ, all elements of A are related to all elements of B. In particular, all elements of A are related to one given element of B.

Now, if in ψ its term ‘*a*’ has distributive supposition and its term ‘*b*’ has determinate supposition, and they are related by relation ‘R’, then $\aleph \Vdash \psi$ iff x R t, for every element x of A and one specific element t of B, in model $\aleph$ (Definition 2.4.3.2).

So all models that satisfy φ also satisfy ψ – $\boldsymbol{UM}_{\varphi} \Vdash \psi$.

But not all models that satisfy ψ also satisfy φ – ***UM***$_ψ$ ⊩/ φ –, since in at least some members of ***UM***$_ψ$ there are elements of B which are not related to all elements of A.

Hence ***UM***$_φ$ ⊂ ***UM***$_ψ$.

So ⊨ φ => ψ □

2.4.4.2.2 Schema (2) implies schema (3)

Theorem 2.4.4.2.2: If φ and ψ are propositions featuring the same verb, the same polarity and the same terms '*a*' and '*b*', such that in φ '*a*' has distributive supposition and '*b*' has determinate supposition, whereas in ψ '*a*' has distributive supposition and '*b*' has merely confused supposition, then ⊨ φ => ψ.

Proof: If in φ its term '*a*' has distributive supposition and its term '*b*' has determinate supposition, and they are related by relation 'R', then ℵ ⊩ φ iff *x* R *t*, for every element *x* of A and one specific element *t* of B, in model ℵ (Definition 2.4.3.2).

So this guarantees that all members of A are related to some element of B in the models that verify φ.

Now, if in ψ its term '*a*' has distributive supposition and its term '*b*' has merely confused supposition, and they are related by relation 'R', then ℵ ⊩ ψ iff *x*R*y*, for every element *x* of A and some element *y* of B, in model ℵ (Definition 2.4.3.3)

So all models that satisfy φ also satisfy ψ – ***UM***$_φ$ ⊩ ψ.

But not all models that satisfy ψ also satisfy φ – ***UM***$_ψ$ ⊩/ φ –, since in at least some members of ***UM***$_ψ$ all elements of A are not related to one single element of B.

Hence ***UM***$_φ$ ⊂ ***UM***$_ψ$.

So ⊨ φ => ψ.

2.4.4.2.3 Schema (3) implies schema (4)

Theorem 2.4.4.2.3: If φ and ψ are propositions featuring the same terms '*a*' and '*b*', such that in φ '*a*' has distributive supposition and '*b*' has merely confused supposition, whereas in ψ '*a*' and '*b*' have determinate supposition, then ⊨ φ => ψ.

Proof: If in φ its term '*a*' has distributive supposition and its term '*b*' has merely confused supposition, and they are related by relation 'R', then ℵ ⊩ φ iff *x*R*y*, for every element *x* of A and some element *y* of B, in model ℵ (Definition 2.4.3.3).

So this guarantees that at least one element of A is related to some element of B in the models that verify φ.

Now, if in φ its terms '*a*' and '*b*' both have determinate supposition and are related by relation 'R', then ℵ ⊩ φ iff *x* R *y*, for some element *x* of A and some element *y* of B, in model ℵ (Definition 2.4.3.4)

So all models that satisfy φ also satisfy ψ – ***UM***$_φ$ ⊩ ψ.

But not all models that satisfy ψ also satisfy φ – ***UM***$_ψ$ ⊩/ φ –, since in at least some members of ***UM***ψ there is some element of A that is not related to any element of B.

Hence ***UM***$_φ$ ⊂ ***UM***ψ.

So ⊨ φ => ψ. □

2.4.5 Concluding remarks

a. Buridan's definitions of each mode of personal supposition, thus reconstructed, validate the inferential relations established syntactically by Elizabeth Karger. Hence, syntax and semantics yield the same results in this fragment of Buridan's system; with the license of anachronism, this fragment can be said to be 'sound'.

b. I have not given a thorough and satisfactory account of the negation in the present analysis. As already remarked, it is to be expected that the general effect of the negation would be to assert the non-existence of the relation R between the relevant *supposita* of the terms, but further investigation is necessary. Probably, thanks to the syntactical analysis, all negative propositions can be reduced to their 'mode preserving non-ordinary form', and again the semantic investigation of the negative fragment of the language could proceed with the same interpretational schemata, *mutatis mutandi*. But again, these are only intuitions that must receive a more rigorous confirmation.

2.5 CONCLUSION

Consequentia is, in many senses, perhaps the medieval logical topic that comes the closest to current investigations in logic and philosophy of logic. While nobody in their sane mind would propose to apply theories of supposition or of obligations as such (although, naturally, lessons can be learned from the analysis of these theories too), the notion of consequence is discussed now as much as it was then, and with remarkable similarities between the two traditions. In this sense, the dialogue between these traditions with respect to this notion is bound to be fruitful.

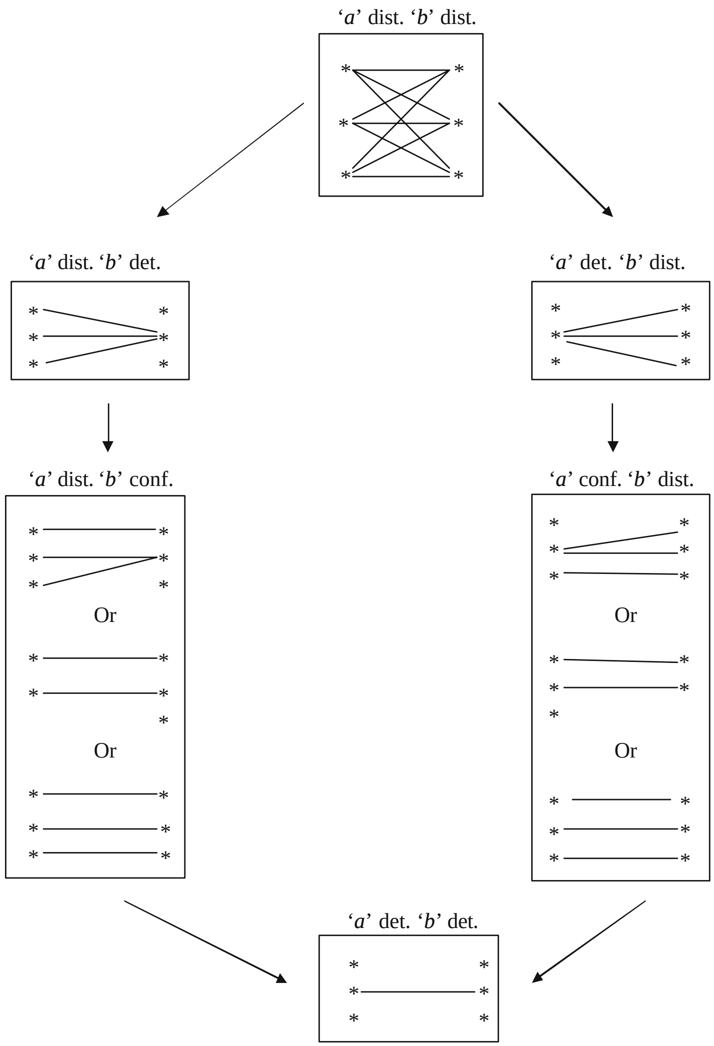
Indeed, as I see it, the historical analysis of Buridan's views greatly benefits from the application of some modern techniques, in particular model-theoretic ones, as certain properties of these medieval theories can thereby be outlined and even proved with rigor. Moreover, we have seen that Buridan was already aware of many of the phenomena related to the interaction between context and language that are currently treated within the framework of modern two-dimensional semantics; for this reason, this latter framework was particularly suitable to reconstruct Buridan's arguments. For the same reason, it is also clear that Buridan should be recognized as one of the first practitioners of two-dimensional semantics.

But not only the historical analysis can benefit from this dialogue: I have attempted to show that Buridan's notion of formal consequence can offer an important contribution to current debates on the notion of logical consequence. I am in fact convinced that the Buridan/Shapiro hybrid notion of (logical/formal) consequence is the best account of this notion proposed so far.

Moreover, his commitment to tokens leads us to reflect on relations of implication and inference also in more mundane contexts, other than strictly logical ones. Interestingly, there is a growing trend of investigations on the logic of knowledge and belief in non-perfect contexts, and it seems to me that the issues that emerge from Buridan's analysis of inference-drawing with tokens may also be relevant for these investigations. In particular, the distinction between the concepts of inference, seen as a local deed relating meaningful tokens and their contents, and of consequence, seen as a global relation relating the characters of the propositions involved, seems especially crucial and yet it has thus far been mostly overlooked. Not that Buridan himself was aware of this distinction, even more so since he seemed to use the same term for both notions; but while reconstructing his theory I was compelled to reflect on these concepts, in a way that might not have happened otherwise.

APPENDIX

A visual rendering of the hexagon of inferential relations



PART 3

OBLIGATIONES AS LOGICAL GAMES*

3.0 INTRODUCTION

Although the medieval form of dialectical disputation known as *obligationes* has attracted the attention of medievalists as well as of philosophers of logic over roughly the last 30 years, much about it still remains mysterious. Here I shall defend the idea that, if *obligationes* are viewed as logical games of consistency maintenance, a great deal of this mystery can be dissolved.

Over the last few years, historical work has greatly contributed to the establishment of reliable manuscripts of various *De Obligationibus* treatises and to the understanding of the intricacies of the theories upheld in these treatises. Within a more theoretical approach, intriguing suggestions have been made as to how these theories should be interpreted from a modern perspective, although no satisfying definitive solution seems to have been found. In this part, I will focus on two of the best-known medieval texts on *obligationes*, namely Water Burley's and Roger Swyneshed's treatises, and on a much less known treatise by Ralph Strode. I present a new interpretation of how to account for the rules of this regimented form of dialectical disputation from the viewpoint of modern logic, and use modern tools to investigate some of the logical properties of obligational disputations. In particular I suggest that *obligationes* be viewed as logical games in the sense of J. Hintikka's game-theoretical semantics (cf. Hintikka and Sandu 1997), P. Lorenzen's dialogue logic (Lorenzen 1961), and, more recently, the trend of games in logic (cf. van Benthem 2001) and recent developments in dialogical logic (cf. Rahman and Keiff 2005, Keiff forthcoming).[279]

---

* This part is largely based on (Dutilh Novaes 2005b), (Dutilh Novaes 2006a) and (Dutilh Novaes 2006b).

[279] Henceforth, I will occasionally refer to *obligationes* as 'the game', but this claim will receive further corroboration. As for a thorough comparison between *obligationes* and these modern versions of games in logic (particularly interesting in the case of dialogical logic), this falls out of the scope of the present investigation and is a topic for future research.

3.1 HISTORY

Although *obligationes* has been a much investigated topic over the last few decades, we are probably still very far from a complete understanding of this kind of disputation. Its popularity during the Middle Ages is attested by the impressive number of treatises on the topic, some of which have been given modern editions, but many of which remain of difficult access (most of them only under the form of manuscripts in libraries across the world). Therefore, a great deal of work still has to be done before the whole picture is uncovered, in particular concerning the edition of manuscripts.[280]

Here, I do not intend to offer an exhaustive historical overview of the topic.[281] However, some considerations must be made in order to clarify the approach adopted here. For the present purposes, probably the most important historical aspect concerning them is that these theories underwent a small revolution in the second half of the 14th century: Walter Burley's theory is the paradigmatic version of the theories known as *antiqua responsio* (section 3.3), whereas the changes introduced by Richard Kilvington and, more importantly, Roger Swyneshed, came to constitute the *nova responsio* (section 3.4). However, it seems that the *antiqua responsio* remained dominant (cf. Spade 1982a, 339), as is attested by the analysis of Ralph Strode's treatise in section 3.5 of this part.

Hence, we shall first be concerned with Burley's theory (i.e., with the *antiqua responsio*). In his treatise, he describes six kinds of *obligationes*: *petitio, sit verum, institutio, positio, depositio* and *dubitatio*. We shall be focusing on *positio* (as this is the mode of *obligatio* to which Burley devotes most of his treatise). The same will be done in the analysis of Swyneshed's and Strode's treatise.

The disputation has two participants, Opponent and Respondent. In the case of *positio*, the game starts with Opponent putting forward a proposition, often called *positum*, which Respondent must accept as true for the sake of the disputation, unless it is contradictory in itself. Opponent then puts forward propositions (the *proposita*), one at a time, which Respondent must either concede, deny or doubt, on the basis of inferential relations with the previously accepted or denied propositions, or, in case there are none (and these are called irrelevant or impertinent propositions[282]), on the basis of the common knowledge shared by those who are present. Respondent loses

---

[280] The development of the obligational genre prior to Burley, for example, is still a matter of controversy. Some (Spade and Stump 1983) have argued that no fully structured treatise on *obligationes* has been written before the end of the 13th century, and in particular that the treatise attributed to Sherwood (Green 1963) could not have been written by him. Others, in particular (Braakhuis 1998), have provided compelling evidence to the effect that full-fledged obligational treatises can be dated to as early as the first half of the 13th century. The history of this genre prior to Burley falls, however, out of the scope of the present investigation. See also (Martin 2001).

[281] For overviews, see (Spade 1982a), (Spade 2003), (Stump 1982b), (Martin 2001) and (Yrjönsuuri 2001a).

[282] Throughout the text, I will use the terms 'relevant' and 'pertinent' as synonymous, as much as 'irrelevant' and 'impertinent'. The terms in Latin are '*pertinens*' and '*impertinens*', but they are often translated as 'relevant' and 'irrelevant', for example in the translation of Burley's treatise.

the game if he concedes a contradictory set of propositions. Opponent loses the game if Respondent is able to maintain consistency during the stipulated period of time. At the end, Opponent and possibly a jury determine whether Respondent responded well.

3.2 OVERVIEW OF THE LITERATURE

3.2.1 Different suggestions

The secondary literature on *obligationes* is vast and ranges over a plethora of different issues related to the theme. Some of these issues concern the authorship of a given treatise (cf. Spade and Stump 1983, on the treatise attributed to Sherwood), the evolution from the *antiqua responsio* to the *nova responsio* (cf. Spade 1982a, Stump 1981, Ashworth 1981), or the editing work on manuscripts (De Rijk 1974). But one of the most puzzling of all these issues is: what can this mode of oral disputation be compared to, in terms of modern logic? What was the general purpose of such disputations? How are we, in the 21st century, to understand them?

Clearly, these issues arise only in the context of modern interpretation of *obligationes*: apparently, this practice was so entrenched in the patterns of medieval intellectual activity that, for the scholars engaged in such disputations, the very question concerning the purpose of this sort of disputation would be superfluous. For them, there was no need to ask for the 'purpose' of *obligationes*;[283] it is we who have to struggle to understand what they were about. But the attempt of interpreting *obligationes* within the current conceptual framework is nonetheless legitimate and relevant. The only caution one must have is not to dismiss the topic too quickly as historical nonsense if we fail to establish the rationale behind it.

At first sight, the view that *obligationes* fulfilled primarily pedagogical tasks – that they were schoolboy's exercises – is very compelling. This seems to have been the view on *obligationes* held by the first analyses of the topic (cf. Weisheipl 1956; Green 1963; Keffer 2001, 52). But more logic-oriented medievalists could not be satisfied with the view that such a sophisticated system of rules had only and merely pedagogical purposes. Another argument against this view is that the most respected logicians of the period, such as Burley and Ockham, dedicated a considerable number of pages to the study and analysis of *obligationes*, indicating that it might have had a more significant theoretical status than mere exercises.

To my knowledge, the first suggestion as to what *obligationes* should be compared to in terms of modern philosophy of logic has been put forward by P.V. Spade. He proposed that *obligationes* be viewed as a framework to explore counterfactual conditionals (cf. Spade 1982b). Several counterarguments to this proposal were advanced by E. Stump (cf. Stump 1985), and Spade retracted his early position in an article wittily entitled 'If *Obligationes* were Counterfactuals' (Spade 1992). (Spade's counterfactual hypothesis is addressed in more details in section 3.2.2 below.)

[283] For some of the medieval 'meta-remarks' on *obligationes*, cf. (Keffer 2001, 52–58).

Alternatively, Stump put forward the view that *obligationes* were primarily intended as a theoretical framework for the analysis of *insolubilia* and *sophismata* (cf. Stump 1982; Stump 1985), in particular self-referential paradoxes raised by reference to either the premises, the participants of the disputation, or an action within the disputation. Indeed, the virtual omnipresence of *insolubilia* and *sophismata* in treatises on *obligationes*, and of obligational vocabulary in treatises on s*ophismata*, is striking; there is no doubt that the two topics were closely related (cf. Martin 2001; Pironet 2001). But it would seem that the solution of paradoxes was not an end in itself of obligational disputations: rather, these were the interesting cases in which one's ability to play the obligational game, or the strength of the obligational system in question, were put to test. The same holds for other medieval logical theories, such as supposition theories: *sophismata* were seen as the ultimate test for a system. In fact, the emphasis on paradoxes only supports the idea that consistency maintenance is the general aim of *obligationes*: nothing is as threatening for consistency than a set of propositions from which both φ and $\neg\varphi$ can be derived; now, the paradoxes discussed in Burley's treatise are typically of this kind (cf. Stump 1985).

C. Martin (cf. Martin 2001, 72) has put forward the suggestion that the general purpose of *obligationes* was the evaluation of cotenability between propositions, in a way similar to the construction of possible worlds by means of Lindenbaum's lemma (with the difference that, in the finite time of an obligational disputation, only a partial valuation of the sentences of the language could be given – see def. 3.3.1.3). His view seems to be the closest, among all commentators, to the idea of consistency maintenance that I defend here.[284]

Another view that emphasizes the inconsistency prevention trait of *obligationes* is Lagerlund and Olsson's suggestion that a theory of *obligationes* be seen as a theory of belief revision (cf. Lagerlund and Olsson 2001). The idea is that a set of beliefs K would be revised by the acceptance of the proposition φ_0 (the *positum*) during the very performance of the disputation. Respondent must include all propositions that follow from φ_0 (that might or might not have been in K at the outset) to his set of beliefs, and exclude all propositions that contradict a proposition that follows from φ_0. As for all the other propositions, he will keep his original beliefs (answering according to their own quality to us, as Burley puts it).

As I see it, there are two main problems with this proposal, one conceptual and the other structural. First of all, belief revision is supposed to concern an agent's actual assent or dissent to a new belief, whereas in an obligational disputation it is merely for the sake of the argument, or of the disputation, that Respondent considers the *positum* as true (without giving it his actual assent). The structural problem is: Respondent cannot 'change his mind' about any proposition that he has previously conceded (he is obligated to them, hence the name *obligationes*), whereas an agent performing a belief revision may conclude that it is better not to change his belief with

[284] Note, though, that others have emphasized the importance of consistency, cf. (Yrjönsuuri 1998) and (Yrjönsuuri 2001a).

respect to a given proposition after all, if it yields too much change in his original set of beliefs.

In other words, the general idea of consistency being the primary aim of an obligational disputation, as maintained by Lagerlund and Olsson, is sound; the problem with their proposal seems to lie in the conceptual background. It is important to understand that, in an obligational disputation, Respondent is committed (obligated) to certain rules (such as the impossibility of going back in his decisions) and has a certain general goal – to 'win' the disputation – that are fundamentally different from those of an agent performing belief revision. In the latter case, the agent's commitments are geared towards reality, to himself and possibly to the accomplishment of a given task; in the case of *obligationes*, Respondent is opposed to Opponent, who is trying to trap him into contradictions (something that could only be compared to some 'Evil Genius' trying to fool an agent into false beliefs, Descartes's *malin génie*). In view of this, it seems more natural to suppose that Respondent and Opponent are actually playing a game, to the rules of which they are committed (obligated). Clearly, these rules do not completely mimic how an agent actually revises his own set of beliefs: at most, they mirror the (implausible and artificial) circumstance of an agent pushing the revision of his set of beliefs by a given proposition as far as he can. Belief revision aims at modeling mental processes that actually take place; *obligationes* are explicitly 'artificial'.

Mindful of the 'artificiality' of *obligationes*, M. Yrjönsuuri has argued that *obligationes* could be seen as thought experiments (cf. Yrjönsuuri 1996), which is an epistemological version of Martin's idea that an obligational disputation mirrors the construction of a state of affairs (or a possible world).[285] Nevertheless, the thought experiment hypothesis seems to me to overlook many important aspects, such as the competing status of Opponent and Respondent with respect to each other, among others.

In sum, underlying most interpretations of *obligationes* is thus the idea of consistency maintenance (although their syntactical aspects must not be disregarded, namely the importance of inferential relations). But nobody seems to have taken very seriously the idea that such a disputation is really like a game, involving two participants, determined by specific and previously agreed-on sets of rules, and of which the aim is to win and beat the other participant. That is, it seems to me that the artificial and regimented character of this practice can be best captured by the assumption that it is a logical game. To my knowledge, only a few scholars have suggested that *obligationes* could be something like games: L. de Rijk's series of articles entitled 'Some 13th-Century Tracts on the Game of *Obligationes*' (de Rijk 1974), following a hint in (Hamblin 1970), J. Ashworth's allusions to 'the game' (cf. Ashworth 1984, 131)

---

[285] Interestingly, van Benthem (cf. van Benthem 2001, 3) observed that a computationally simpler strategy for Respondent to play the game of *obligationes* is to choose one model beforehand, since model checking is computationally simpler (complexity P) than consistency management (complexity NP) – that is, something quite similar to the idea of a thought experiment. (cf. section 3.1)

and, appearing only much later, van Benthem's allusion to *obligationes* as games of consistency maintenance, in his *Logic in Games* (van Benthem 2001, 2–3).

3.2.2 Arguments against the counterfactual hypothesis

Finally, I would like to reassess P.V. Spade's counterfactual hypothesis. It was the first attempt towards a (modern) philosophical interpretation of the medieval theories of *obligationes*, proposed in his 1982 article 'Three theories of *obligationes*: Burley, Kilvington and Swyneshed on counterfactual reasoning'. Since then, this interpretation has been challenged by a number of scholars, but their arguments did not seem to provoke the general agreement to the effect that the counterfactual hypothesis should be considered as untenable. Thus, in what follows I will dare to present once more arguments against the counterfactual hypothesis.

Spade admits himself that the hypothesis works better for Walter Burley's version of *obligationes* than for Swyneshed's version, so I will turn to Burley's theory first. My starting point is, naturally, that Burley's version of *obligationes* is best described as a logical game of consistency maintenance; I will argue that many of the aspects of *obligationes* that Spade claims can be accounted for by the counterfactual hypothesis are better accounted for by the consistency game hypothesis.

Perhaps the best argument against the counterfactual hypothesis is one put forward by Christopher J. Martin (Martin 2001): if counterfactuals at all, what Burley's theory of *obligationes* defines are not **would**-counterfactuals, but rather **might**-counterfactuals. When Respondent has to accept or deny a proposition with respect to the previously accepted or denied propositions, if the proposition is pertinent – that is, if the proposition itself or its contradictory follows from the previously accepted or denied propositions – then we do have a would-counterfactual: since the proposition proposed or its contradictory follows logically from the premises, if all the premises were true, then it would have to be true (or false) as well. But when the proposition is impertinent, then the proposition itself as well as its contradictory **might** be true in a counterfactual situation where all premises are true. Well, if the aim of a theory of counterfactuals is to exclude one of the two possibilities, counterfactual reasoning in general is not meant to explore a situation in which both a proposition and its contradictory could be the case. The indeterminacy of Burley's theory with respect to impertinent propositions does not fit well with the counterfactual hypothesis, and that means that either Burley's theory is a problematic theory of counterfactuals (which Spade is willing to grant), or that it is not a theory of counterfactuals at all.[286]

One of Spade's arguments in favor of the counterfactual hypothesis is that the procedure defined by Burley's theory of *obligationes* progressively constructs the possible world that is as similar as possible to the actual world, except for the (false) *positum* and its consequences. But, in fact, a false *positum* does not always define **one** class of possible worlds or models that is subsequently narrowed by the responses

[286] Similar arguments have been advanced in (Stump 1981).

to impertinent propositions; sometimes, branching – mutually exclusive – classes of models are defined by a false *positum*. Which one of these classes contains the possible world that is most similar to the actual world? The *obligationes* theory falls short of providing decisive procedures for this problem. Moreover, the determination of the possible world as similar as possible to the actual world seems to involve metaphysical discussions that do not belong to the scope of obligational disputations properly speaking. (Both points shall be clarified by the example below.)

Take the following example, from Ralph Strode's treatise on *obligationes* (cf. Ashworth 1993; 1996, 349): the *positum* is 'Every man is running'. It is accepted as a *positum*, since it is possible. Before the game proceeds, two branching classes of possible worlds correspond to the *positum* being accepted, with respect to Respondent (who is a man and is not running): the class of possible worlds in which he is a man and is running, and the class of possible worlds in which he is not running and thus is not a man. Which one of these classes is more similar to the actual world? One may argue that the possible worlds in which Respondent is a man are more similar to the actual world than those in which he is not a man; but this argument does not follow from the rules of *obligationes*, rather it seems to rely on the essentialist contention that the property of being a man is more essential to Respondent than the property of running at time *t*. But the rules of *obligationes* themselves offer no support for this position.

In practice, what happens is that the choice of the class of possible worlds (or models) with which the disputation will go on is made on the basis of purely contingent facts, namely the order in which Opponent will propose propositions.[287] If he first proposes 'You are running', since Respondent has not granted that he is a man, 'You are running' is an irrelevant, false proposition, which thus must be denied. In this case, it is the class of models in which Respondent is not running and is not a man that is 'chosen'. Hence, if Opponent then proposes 'You are a man', Respondent will have to deny it, since its contradictory follows from the previously accepted propositions. However, if Opponent, as a matter of pure contingency (i.e., not prompted by the rules of *obligationes*), proposes 'You are a man' first, then Respondent will accept it as a true, irrelevant proposition, and thus the disputation will go on with the other class of models, in which Respondent is a man and is running.

In sum, the rules of *obligationes* do not provide elements for the determination of **the** possible world that is most similar to the actual world (given the false *positum*), and not even of **the class** of possible worlds that are as similar as possible to the actual world, except for the false *positum*. Given that Opponent can choose the order in which he proposes propositions, the classes of models defined by the disputation are entirely dependent on this order, thus reflecting the great importance of the dynamic character of *obligationes* under Burley's version.

However, if *obligationes* are viewed as games of consistency maintenance, the 'branching phenomenon' just described is entirely unproblematic; if the aim is to

[287] Recall that it is the Burley-style theory being discussed.

keep consistency, then either class of models will do. By means of the order in which Opponent proposes propositions, he forces Respondent to adopt this or that class of models; if Opponent wants to win, it is likely that he will force Respondent to adopt the least straightforward, most counterintuitive class of models, and these models will be as different from the actual world as possible, given the rules of the game. Thus, as for the example above, it seems more likely that Opponent will propose 'You are running' first (which would have to be denied), to force Respondent to admit that he is not a man. The denial of this proposition will probably provoke more revisions in the model corresponding to the actual world than just accepting that he, the Respondent, is a man who is running.

I now turn to the arguments put forward by Spade to support his counterfactual hypothesis.

Argument 1: Spade claims that the rules of *obligationes* provide a clear procedure to determine what would 'happen' if the *positum* were true. Again, as argued, the rules of *obligationes* only define what **might** happen if the *positum* were true, so what is at stake here is primarily the notion of consistency, or 'cotenability' (in C. Martin's terms).

Argument 2: the counterfactual hypothesis 'provides an explanation for the otherwise apparently pointless treatment of "irrelevant" sentences' (p. 12). Spade contends that the role of the rules for irrelevant/impertinent propositions is to generate a possible world as similar as possible to the actual world. As we have seen, the rules for irrelevant propositions are not sufficient to determine **the** possible world most similar to the actual world, given the order-dependence and the so-called 'branching phenomenon'. In fact, according to the game hypothesis, the role of irrelevant propositions is to make the game harder, as I will argue in section 3.3.3, and to exclude 'lazy' strategies. So it seems to me that the game hypothesis gives a better account of the role of irrelevant propositions than the counterfactual hypothesis.

Argument 3: the choice of false propositions as *posita* is meant to mirror counterfactual reasoning, which consists in defining what reality would be like if some particular element of it were different. Again, the game hypothesis also 'provides a rationale' (Spade's terms) to this aspect of *obligationes*: if the *positum* were a true proposition, then – since the actual world would simply be the (consistent) underlying model – it would be sufficient to respond to propositions according to how things actually are in order to keep consistency. In this case the whole activity would be pointless, on account of being too easy. Thus, the counterfactual hypothesis is not the only way to account for the choice of false or impossible *posita*.

Argument 4: 'obligational disputations have many of the characteristic properties of counterfactuals'. (p. 12). Spade sees the relation between a *positum* and any conceded proposition (or the contradictory of any denied proposition) in a given disputation as some sort of 'inference', and he argues that such 'inferences' follow patterns similar to counterfactual inferences, in particular with respect to the failure of strengthening, transitivity and contraposition. But, to see those similarities, one must adopt as a starting point that there are such things as 'obligational inferences'; now, I have already argued that the only restriction on granting a proposition B is that it must be **compossible** (consistent) with a *positum* A. If the relation between a *positum* and the conceded/denied propositions is some kind of inference, then Spade must accept that, in the case of irrelevant propositions, both $A \Box\!\!\rightarrow B$ and $A \Box\!\!\rightarrow {\sim}B$ are valid, which seems very awkward. But most importantly, by assuming that these are inferential relations, Spade is already somehow assuming the point he wants to prove.

Argument 5: 'the counterfactual interpretation of *obligationes* [...] yields a plausible account of the transition from Burley's theory to Kilvington's, and from Kilvington's to Swyneshed's'. (p. 13). Spade argues that Burley's theory is a theory of counterfactuals **with problems**, and these are most of all related to the fact that one might assert that B counterfactually follows from a *positum* A and, in a different disputation, that ${\sim}$B counterfactually follows from the same *positum* A. This is not a surprise; as we saw, in the case of irrelevant propositions, it is possible to accept both B and, in a different disputation, ${\sim}$B from the same *positum* A (*obligationes* being thus at most a case of might-counterfactuals). Hence, the so-called problems of Burley's theory seem to stem rather from Spade's interpretation.

Spade goes on to say that, because Burley's theory had problems, it was to be expected that it had to be revised, and Kilvington's and Swyneshed's theories would be such a revision. I will not offer an extensive analysis of Kilvington's 'theory', since it is disputable whether it was a full-fledged theory of *obligationes* at all (he only made scanty remarks on *obligationes* in his 47th *sophisma*), and since it raises a series of interpretational problems. Swyneshed himself does not motivate the changes he introduces to the theory of *obligationes*, and that is why commentators usually turn to Kilvington's *sophisma* 47 to retrieve what his motivations could have been. It seems that the two main problems of Burley's theory according to Kilvington are: any falsehood compossible with the *positum* can be 'proved' in Burley's theory, and, according to the rules defined by Burley, the same proposition could receive different answers in two different disputations with the same *positum* or in the same disputation, if proposed at different moments (cf. Spade 1982b). But turning to Swyneshed's theory, it seems clear that his main objection to Burley's theory (even though he never says it explicitly) is the extreme sensitivity to the order of the proposed propositions for the responses they should receive, and the fact that, in two disputations with the same *positum*, it could very well happen that, since the responses depend on the responses to previous propositions as well, a given proposition would receive different responses. As will be discussed in section 3.4, Swyneshed's main goal was apparently to abolish the dynamic character of Burley's *obligationes*. Thus, Spade's claim that the revision proposed by Swyneshed was motivated by the fact that Burley's theory was a problematic theory of counterfactuals seems far-fetched. It is clear that Swyneshed thought that Burley's theory was problematic, but nothing seems to indicate that he found it problematic **insofar as** it was a bad theory of counterfactuals.

If these were indeed Swyneshed's motivations to revise the *antiqua responsio*, then his revision of the notion of pertinent/impertinent propositions hit the right spot of

Burley's theory, but none of this seems to have any relation to *obligationes* being a logic of counterfactuals.

Spade's hypothesis becomes even more endangered when one analyzes the outcome of the revision proposed by Swyneshed; if Swyneshed meant his theory to be a better theory of counterfactuals than Burley's, then he failed miserably – his theory is a **worse** theory of counterfactuals than Burley's. By excluding the dynamic elements of Burley's theory, Swyneshed ended up with a theory which is committed to an even stronger form of inconsistency: in Swyneshed's theory, the set formed by the propositions accepted during the disputation and the contradictories of those denied during the same disputation is very likely to be **inconsistent**. Spade is aware of this fact, which leads him to conclude: 'Swyneshed's theory is by no means an attractive account of counterfactuals'. (Spade 1982b, 30). So, since each proposed revision of *obligationes* is worse as a theory of counterfactuals than its predecessor, it seems that the counterfactual hypothesis does not yield a plausible account of the transitions from one *obligationes* theory to the other, as Spade claims.

Nevertheless, in all fairness, it must be said that the game hypothesis does not provide a good account of the transitions between the theories either. From a game-theoretical point of view, Burley's theory is by far the most interesting one. Attempting to suppress the dynamic character of Burley's *obligationes*, Kilvington and Swyneshed seem to have produced a much less interesting theory of *obligationes* with respect to whichever goals these disputations were supposed to accomplish, and this would explain why Burley's theory remained more influential even after alternative theories had been proposed.

3.2.3 Conclusion

In sum, there are problems with each of the modern interpretations previously proposed in the literature. For this reason, the attempt at a new interpretation, departing from different assumptions and in a different framework, seems a welcome enterprise. As said in the introduction, the viewpoint I will be adopting here for the analysis of *obligationes* is that they are to be taken as logical games. As I hope will become clear in the coming pages, this hypothesis has a considerable explanatory power in that it allows us to study with rigor many of the interesting properties of *obligationes*, and to uncover a rationale for many of the aspects of this form of disputation which seem otherwise impenetrable.

3.3 BURLEY'S *OBLIGATIONES*: CONSISTENCY MAINTENANCE

As already said, Burley's theory of *obligationes* was probably the most influential of such theories throughout the 14th century, even after the *nova responsio* was proposed. Hence, it is only natural that I first present the analysis of Burley's theory. In what follows, I present a reconstruction of this theory, based on the assumption that *obligationes* are best viewed as logical games, and making use of tools from the game-theoretic and model-theoretic frameworks.

3.3.1 The rules of the game

3.3.1.1 Preliminary notions

The obligational game is played by two players, opponent (O) and respondent (R). It is defined by the quadruple:

Definition 3.3.1.1.1: The obligational game (Burley)

$$\boldsymbol{Ob} = <\mathbf{K_C}, \boldsymbol{\Phi}, \boldsymbol{\Gamma}, \mathbf{R}(\varphi_n)>$$

$\mathbf{K_C}$ is the common state of knowledge of those present at the disputation complemented by the *casus*. The *casus* was usually a proposition to be assumed as true, often to make it explicit that the *positum* was false. An example from Burley: 'Suppose Socrates is black, and suppose it is posited that Socrates is white' (Burley 1988, 378). That Socrates is black is the *casus*, a proposition that all the participants are to assume to be true, and the *positum* is that he is white (thus in the context a false, contingent proposition).[288]

$\mathbf{K_C}$ is an incomplete model, in the sense that some propositions do not receive a truth-value: for some propositions, it is not known whether they are true or false (e.g., that the pope is sitting or that he is not sitting, if he is not present), although it may be known that they are true-or-false. So, the state of common knowledge is a state of imperfect information: it includes all information that is considered as common sense (that the pope is in Rome, all religious dogmas etc.), plus information circumstantially available, due to the pragmatics of the disputational situation. It is assumed that there is no disagreement between the participants with respect to the truth-value of a proposition according to K_C.

$\boldsymbol{\Phi}$ is an ordered set of propositions, or better put, declarative sentences. However, the Latin term used is *propositio*, so I will, again, be using the term 'proposition' in its Latin acceptation. It is the set of propositions actually put forward by O during an obligation. Each element of Φ is denoted by 'φ_n', where *n* is a natural number, denoting the place of φ_n in the ordering. The order corresponds to the order in which the propositions are put forward by O, starting with φ_0 (the *positum*).

$\boldsymbol{\Gamma}$ is an ordered set of sets of propositions, which are formed by R's responses to the various φ_n. How each Γ_n is formed will be explained below. The ordering is such that Γ_n is contained in Γ_{n+1}.

$\mathbf{R}(\varphi_n)$ is a function from propositions to the values 1, 0, and ?. This function corresponds to the rules R must apply to respond to each proposition φ_n · 1 corresponds to his accepting φ_n, 0 to his denying φ_n and ? to his doubting φ_n.

The **procedural rules** of the game are quite simple: O first puts forward a proposition. If R accepts it (according to $R(\varphi_0)$ defined below), then the game begins. Then O puts forward a further proposition, R responds to it according to $R(\varphi_n)$, and this procedure is repeated until the end of the game.

---

[288] For more on the notion of *casus*, see (Yrjönsuuri 1993) and section 3.5.1.2 below.

The **logical rules** of the game are defined by $R(\varphi_n)$, in the following way:

Definition 3.3.1.1.2: Rules for *positum*

$$R(\varphi_0) = 0 \text{ iff } \varphi_0 \vdash \bot$$

$$R(\varphi_0) = 1 \text{ iff } \varphi_0 \nvdash \bot$$

The rule defining the response that R should give to φ_0 (the *positum*) has interesting consequences for the idea that *obligationes* are games of consistency maintenance. If R is obliged to accept at the beginning a proposition that entails a contradiction – for example, any paradoxical proposition such as liar sentences and the like – then there is no possible winning strategy for R. There is no way that he can maintain the consistency of a set of propositions that, from the outset, contains a contradictory proposition. So the rules of the game stipulate that there always be a winning strategy for R, starting from this restriction upon the *positum* (see more on this in section 3.3.3.1). Burley expresses this clause by saying that it must be in the Respondent's power to satisfy the requirement (of not falling in contradiction).

Another sophisma: I require you to respond badly to 'God exists'. Next, I propose this to you: 'God exists'. [. . .] The solution: The *petitio* should not be admitted, because no *petitio* should be admitted unless it is in the respondent's power to satisfy the requirement.[289] (Burley 1988, 376)

3.3.1.2 Two interpretations of the rules

For φ_n, $n > 0$, some complications arise. It seems that there are two possible interpretations of the obligational game with respect to its logical rules: either the game is entirely determined – that is, its rules leave no scope for Respondent to choose between possible alternatives, and thus such rules should be seen as instructions that he may or may not fail to comply with – or there is room for choice, hence for strategy. I will explore both alternatives and draw some conclusions.

3.3.1.2.1 Deterministic interpretation. If the game is seen as entirely determined, then its rules can be formulated as follows:

Definition 3.3.1.2.1: Deterministic rules for proposita

$$R(\varphi_n) = 1 \text{ iff } \begin{cases} -\ \Gamma_{n-1} \vdash \varphi_n, \text{ or} \\ -\ \Gamma_{n-1} \nvdash \varphi_n,\ \Gamma_{n-1} \nvdash \neg\varphi_n \text{ and } K_C \Vdash \varphi_n \end{cases}$$

$$R(\varphi_n) = 0 \text{ iff } \begin{cases} -\ \Gamma_{n-1} \vdash \neg\varphi_n, \text{ or} \\ -\ \Gamma_{n-1} \nvdash \varphi_n,\ \Gamma_{n-1} \nvdash \neg\varphi_n \text{ and } K_C \Vdash \neg\varphi_n \end{cases}$$

$$R(\varphi_n) = ? \text{ iff } -\ \Gamma_{n-1} \nvdash \varphi_n, \Gamma_{n-1} \nvdash \neg\varphi_n, K_C \nVdash \varphi_n, K_C \nVdash \neg\varphi_n$$

[289] Aliud sophisma. Peto te male respondere ad istam; Deus est. Deinde, propono tibi istam: Deus est. [. . .] Solutio: petitio non est admittenda, quia nulla petitio est admittenda, nisi fuerit in potestate respondentis ut satisfaciat petitioni. Se admissa ista petitione ab aliquot, non potest isti satisfacere. (Green 1963, 43 (20–21) and (29–32)).

On this interpretation, if Respondent fails to recognize inferential relations **and** if he does not respond to a proposition according to its truth-value within common knowledge, then he responds badly and loses the game.

3.3.1.2.2 'Point system'. In another possible interpretation of *obligationes*, the rules for responding to relevant propositions are given much more weight than the rules for responding to irrelevant propositions. That is, if Respondent fails to identify inference relations between a given *propositum* and the previously accepted/denied set of propositions, and accepts a proposition that he should deny, this automatically yields a contradiction and makes him lose the game. But if he accepts an irrelevant proposition that is false according to K_C, this may constitute a bad response and perhaps make him lose 'points', but not lose the game. Under this interpretation, the formulation of the rules of the game would be:

Definition 3.3.1.2.2.1: Non-deterministic rules – relevant propositions

$$\text{If } \Gamma_{n-1} \vdash \varphi_n \text{ then } R(\varphi_n) = 1$$

$$\text{If } \Gamma_{n-1} \vdash \neg\varphi_n \text{ then } R(\varphi_n) = 0$$

As for the irrelevant propositions (i.e., φ_n such that $\Gamma_{n-1} \nvdash \varphi_n$ and $\Gamma_{n-1} \nvdash \neg\varphi_n$), there would be no rigid instructions as to how one should respond to them, but rather something like the advice of remaining as close as possible to K_C. If Respondent decides not to comply with the advice, he does not lose the game but rather 'loses points', which for strategic reasons might be interesting in a given circumstance, or in view of future moves. This interpretation would make the game game-theoretically more interesting, because it introduces the idea of strategic choices made by Respondent, directed by a 'trade-off' principle. In this case, something like an 'attributing points' function $P(\iota_n)$, from responses ι_n to points $[-; +]$ would have to be introduced – responses being an ordered pair whose first element is a proposition φ_n and whose second element is one of the values 1 – accept, 0 – deny or ? – doubt, denoting thus what R actually responds to a given *propositum*, as follows:

Definition 3.3.1.2.2.2: Non-deterministic rules – irrelevant propositions

$$K_C \Vdash \varphi_n \text{ and } \iota_n = [\varphi_n; 1] \Rightarrow P(\iota_n) = +$$

$$K_C \nVdash \varphi_n \text{ and } \iota_n = [\varphi_n; 0] \Rightarrow P(\iota_n) = +$$

$$K_C \Vdash \varphi_n, K_R \nvdash \varphi_n \text{ and } \iota_n = [\varphi_n; ?] \Rightarrow P(\iota_n) = +$$

$$K_C \Vdash \varphi_n \text{ and } \iota_n = [\varphi_n; 0] \text{ or } \iota_n = [\varphi_n; ?] \Rightarrow P(\iota_n) = -$$

$$K_C \nVdash \varphi_n \text{ and } \iota_n = [\varphi_n; 1] \text{ or } \iota_n = [\varphi_n; ?] \Rightarrow P(\iota_n) = -$$

$$K_C \Vdash \varphi_n, K_R \nvdash \varphi_n \text{ and } \iota_n = [\varphi_n; 1] \text{ or } \iota_n = [\varphi_n; 0] \Rightarrow P(\iota_n) = -$$

On this interpretation, the goal of Respondent would be to avoid contradiction **and** to score the highest number of points possible. In this case, it becomes clear that the role fulfilled by K_C and the 'point system' is to exclude some 'lazy' strategies (which will be discussed in the sequel) of consistency maintenance.[290] Moreover, in this case, the choices made by Respondent may be motivated by more complex strategic reasons than the simple rule-following of the first interpretation. Responding unfaithfully with respect to K_C may constitute a bad response (which may mean that R 'loses points' at the final evaluation), but this move may be advantageous in some circumstances, in order to avoid 'traps' being set by O. In sum, responding not according to K_C may make R lose points, but it may mean his best chance to stay in the game.

Admittedly, Burley's text does not provide sufficient elements so that it can be decided which of these two interpretations is more suitable. On the one hand, it would seem that the first interpretation is closer in spirit to Burley's theory, since he does not introduce a point system and does not say explicitly that the rules for relevant propositions have such a different status from the rules for irrelevant propositions. But some remarks in the text seem to introduce the idea of choice and of a 'trade-off'. Burley expresses this idea of a 'trade-off' when he says for example that 'one should choose the lesser of two evils' (Walter Burley, *Obligations*, 376), with respect to a given *sophisma*.

In any case, the second interpretation is, as said above, more interesting from a game-theoretical point of view, and it would seem that it does capture some important elements of how the disputations actually took place. In general, it seems that the propositions put forward during an *obligatio* were fundamentally *sophismata*, that is, difficult propositions, semantic paradoxes etc. In such cases, it might be convenient to deny or doubt a given proposition that one knew to have 'tricky' consequences, even if it is true according to K_C.

3.3.1.3 Stages of the game

Formation of Γ_n. The different sets of propositions accepted by R (i.e., the propositions to which R has committed himself in the game) are formed in the following way:

Definition 3.3.1.3: The sets Γ_n

$$\text{If } R(\varphi_n) = 1, \text{ then } \Gamma_n = \Gamma_{n-1} \cup \{\varphi_n\}$$

$$\text{If } R(\varphi_n) = 0, \text{ then } \Gamma_n = \Gamma_{n-1} \cup \{\neg\varphi_n\}$$

$$\text{If } R(\varphi_n) = ?, \text{ then } \Gamma_n = \Gamma_{n-1}$$

In particular, if $R(\varphi_0) = 1$, then $\Gamma_0 = \{\varphi_0\}$. If $R(\varphi_0) = 0$ or $R(\varphi_0) = ?$, then there is no game.

[290] As mentioned above, Spade has argued (in Spade 1982b) that the otherwise mysterious role of irrelevant propositions is explained by his interpretational hypothesis based on counterfactuals. But here it is shown that it may be much more natural to view the role of irrelevant propositions as a way to make the game 'harder'.

These rules mirror very closely the clauses of Lindenbaum's lemma, the main idea being that propositions are gradually added to a set of propositions (which starts with one single element, the *positum*), while consistency is also maintained. There is a significant difference, though, in that, in the construction of a maximal consistent set according to this lemma, if the set $\Gamma_n = \Gamma_{n-1} \cup \{\varphi_n\}$ formed is inconsistent, then the construction simply continues with Γ_{n-1}, that is, the so far largest consistent set built. In the *obligationes* framework, however, if an inconsistent set is constructed, the procedure comes to a halt.

Outcome. O wins the game if it is recognized that $\Gamma_n \vdash \bot$, that is, if R has conceded a contradictory set of propositions. R wins the game if, after the stipulated time, it is recognized that $\Gamma_n \vdash\!/ \bot$. The clause about the stipulated time concerns the feasibility of the game: a consistent set of propositions can always be extended to a maximal-consistent set of propositions (according to Lindenbaum lemma), where all propositions of the language receive a truth-value. But, in practice, the construction of such sets is not humanly feasible, therefore there must be a limit as to how far the game can go. At the end of the stipulated time, and if Respondent has not yet granted an obvious contradiction, Opponent and possibly a jury judge whether Respondent responded well, that is, if he hasn't granted a contradiction and if he has remained as close as possible to K_C – in the case of the second interpretation of the rules, whether he scored as many points as he could.

3.3.2 Moves and trees

The range of possible moves for R as a response to a given proposition φ_n, and the outcome of each move for the continuation of the game, can be depicted in trees.

A move by O: -·-·-·-·-	R accepts φ_n: ___A___	R doubts φ_n: ___D___
A move by R: ———	R denies φ_n: ___N___	

Let us first examine the case in which R has more 'freedom' as to how he should respond to φ_n. This happens when there are no inferential relations between the set of propositions hitherto constructed and the proposition put forward. Clearly, R is still obliged towards the state of knowledge K_C. However, maintaining consistency (winning or losing the game) is more important an aim than being faithful to K_C. Again, the degree of freedom that Respondent actually has with respect to irrelevant propositions depends on which of the two aforementioned interpretations of the rules of the game one is prepared to adopt (Figure 3.3.2.1).

Note that O's moves are motivated by R's previous responses – O chooses the next proposition depending on how R has played the previous round and therefore on the set Γ_n that has been formed. Thus, it is reasonable to assume that, in each branch, there would be a different φ_{n+1} with which to continue the game, represented by φ'_{n+1} and φ''_{n+1}. For this reason, it does not seem advisable to represent the whole game in such a tree (as done in van Benthem 2001, 3), since the set **Φ** of propositions

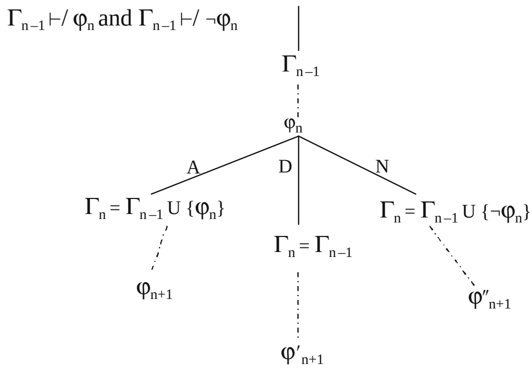


Figure 3.3.2.1. Tree of responses for irrelevant propositions

put forward throughout the game is most likely not predefined by O; he decides what proposition to put forward as the game goes along. A tree structure seems to be adequate to represent only each singular move by R; once there is a branching, each branch has hardly any relation to the other branches any longer (and in fact each branch describes different plays of the game). Therefore, an actual play of the game is best represented by a **deduction**, where the *positum* and the propositions accepted or denied on the basis of K_R would correspond to open assumptions, and the other propositions would be the consequents of the application of inferential rules. All the propositions occurring in a given deduction would form the final set Γ_n of the game. (See section 3.3.3.2 below.)

However, when there is an inferential relation between Γ_{n-1} and φ_n or $\neg\varphi_n$, then an incorrect response may provoke the end of the game and R's defeat (Figure 3.3.2.2).

If $\Gamma_{n-1} \vdash \varphi_n$, the rules of the game stipulate that R should accept φ_n. If he denies it, he loses the game because he forms an inconsistent set of propositions (under the closure of implication – see more on section 3.3.3.4). If he doubts it, he fails to recognize the relation of implication between Γ_{n-1} and φ_n, but does not form an inconsistent set. Therefore, he 'loses points' (under the second interpretation), but the game goes on.

Something analogous happens when $\Gamma_{n-1} \nvdash \varphi_n$ (Figure 3.3.2.3).

Notice that, under the second interpretation of the rules, doubting each proposition is in practice a possible winning strategy, for the starting set of propositions $\Gamma_0 = \{\varphi_0\}$ simply remains consistent. This is perhaps one of the reasons why rules for irrelevant propositions have been created, to prevent such 'lazy' strategies for avoiding inconsistency: R may be able to maintain consistency, but he will 'lose many points' when judged, to the point that this is no longer an advantageous strategy.

3.3.3 Strategies

Now that the general structure of the game has been spelled out, we must discuss some important strategic points concerning the performance of the game.

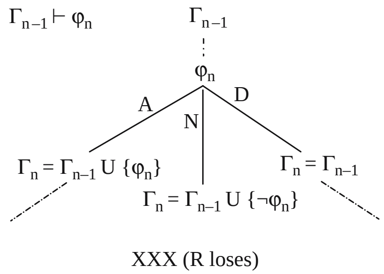


Figure 3.3.2.2. Tree of responses for propositions that follow

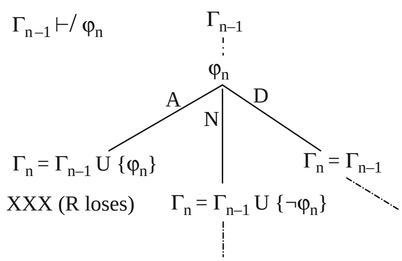


Figure 3.3.2.3. Tree of responses for repugnant propositions

3.3.3.1 Can Respondent always win?

The rules of the obligational game as defined guarantee that there always be a winning strategy for R. This is due to two facts: one is a stipulated rule of the game and the other is a general logical fact, to be discussed shortly.[291] Of course, the possibility of doubting all propositions is also available to R. However, if he wants not only to maintain consistency but also to score points by responding well, then there is always another winning strategy available (admittedly harder to implement, but therefore rewarded with a better score).

The relevant rule of the game is: a paradoxical *positum* should not be accepted. As stated by Burley himself, the point of this clause is exactly to guarantee that R stands a chance to win. Therefore, R always starts out with a consistent set of propositions.

Now, it is a general principle of logic (and the backbone of Lindenbaum's lemma) that any consistent set of propositions can always be consistently expanded with at least one of the two propositions φ_n and $\neg\varphi_n$ (cf. van Benthem 2001, 3).[292]

[291] This fact has already been noticed by J. Ashworth: 'a certain kind of consistency was guaranteed for any correctly-handled disputation'. (Ashworth 1981, 177)

[292] *Proof*: Assume that Γ is consistent. Assume that $\Gamma \cup \{\varphi\}$ is inconsistent. Thus $\Gamma \Vdash \neg\varphi$ (1). Moreover, assume that $\Gamma \cup \{\neg\varphi\}$ is inconsistent. Thus $\Gamma \Vdash \varphi$ (2). From (1) and (2), it follows that $\Gamma \Vdash \varphi \,\&\, \neg\varphi$, that

R starts with a consistent set of propositions (the set composed of the *positum*); so at each move, there is in theory at least one 'correct' way of answering, i.e. either accepting or denying φ_n, which maintains the set of accepted and denied propositions consistent.

The very fact that paradoxical situations – in which it seems that R should both accept and deny a given proposition (a challenge to the principle above) – can be 'resolved', as Burley does in his *De Obligationibus*, is precisely his way of showing that there is always at least one correct answer, given the logical framework of *obligationes*. Burley's strategy is usually either to show that there is only an apparent paradox (so that either φ_n or $\neg\varphi_n$ is in fact to be accepted), or else to identify when the set of accepted propositions had become unnoticeably inconsistent, forcing the situation in which $\Gamma_{n-1} \vdash (\varphi_n \;\&\; \neg\varphi_n)$. Thus, since the obligational framework is particularly suitable for maintaining consistency, it is no wonder that this framework was so widely used to analyze *insolubilia*, the situation par excellence in which consistency is threatened.

J. van Benthem has suggested that a computationally simpler strategy for the game is to proceed by 'model checking' instead of consistency maintenance (cf. van Benthem 2001, 3). The point is that one can choose an arbitrary model – a valuation of all propositions – **beforehand**, and then respond to each *propositum* according to this valuation. This strategy is in theory applicable, but in practice it does not seem to offer much help to Respondent, for the following reasons.

First of all, he would have to be sure that the valuation with which he chooses to play the game is itself consistent. This is an easy task in the case of a simple language such as propositional logic (where the truth-value of each atomic proposition does not 'interfere' in the truth value of other atomic propositions, and the truth values of molecular propositions are defined recursively),[293] but not in the complex and unregimented language (medieval Latin) in which the game was played.

Second, Respondent is obliged to assign the value true to the *positum* in his valuation, but he doesn't know what the *positum* is going to be before the game begins. So, in practice, he has to define the valuation in question *after* the *positum* has been posited, that is, when the game has already begun. Otherwise, a previously defined model has only a 50% chance of complying with the *positum*.

Third, one must not forget that Respondent is also strongly committed to K_C, so his valuation should in fact be as close as possible to K_C. In this sense, the belief revision insight is elucidating: Respondent wants to build a valuation that is as close to K_C as possible, except for the *positum* (and in fact usually the *positum* is a possible but false proposition according to K_C), and which is at the same time consistent. In other words, Respondent must execute a revision of K_C by some φ_0, namely a φ_0 such that $K_C \Vdash\!\!/ \; \varphi_0$. Moreover, K_C is not itself a complete valuation. So the valuation

---

is, that Γ is inconsistent, which contradicts the original assumption. The principle to be proven follows by contraposition.

[293] The same holds, *mutatis mutandi*, for first-order predicate logic, which is still a much simpler language than the language in which the game was played, full of semantic ambiguities and equivocations.

K_C' resulting of the revision of K_C by φ_0 is likely to be much more complex than any valuation that Respondent might have been able to define previously.

Again, the strategy of choosing one model beforehand and playing according to it is the kind of 'lazy'[294] strategy that the obligation towards K_C is supposed to avoid. The strategy of doubting every proposition except for the *positum* is the 'laziest' variation thereof; it corresponds to the following valuation:

$$R(\varphi_0) = 1$$

$$R(\varphi_n) = ?, \quad \text{for } n > 0$$

In sum, although the idea of model checking is at first sight attractive for playing *obligationes*, other traits of the game, in particular the commitment towards K_C, guarantee that it remains hard to play. And indeed, even though there is always a winning strategy available to Respondent, as explained, in practice the game remains demanding. We shall now see why.

3.3.3.2 Why does Respondent not always win?

But why does the game remain hard? If it were easy, then it would not fulfill its pedagogical or theoretical purposes. It seems that *obligationes* remain a difficult kind of disputation for the Respondent for two basic reasons: Opponent makes use of the intricacies (e.g., the phenomena of synonymy and equivocation) of the language being used in the game to set up 'traps' for Respondent; and the game is essentially dynamic (this last aspect will be explored in the next section).

To have a glimpse of the kind of trap Opponent may set up, take a look at the hypothetical disputation represented in Figure 3.3.3.2 (it is not an example taken from Burley's text, but it is very much in the spirit of the examples he proposes himself, including the terms used).

The deduction[295] representing this game would be (φ_0, $\neg\varphi_1$ and φ_2 are open assumptions and φ_3 and φ_4 are the conclusions of inferential operations):

$$\dfrac{\dfrac{\varphi_0 \quad \neg\varphi_1}{\varphi_3} \qquad \varphi_2}{\varphi_4}$$

Where did it go wrong? Why was Respondent forced both to accept and to deny φ_4 in the last round? Could he have avoided the trap? A closer inspection of the propositions shows that φ_2 is not irrelevant (i.e., with no inferential connections to the previously accepted or denied propositions), as it might seem. Actually, from φ_0 and $\neg\varphi_1$,$\neg\varphi_2$ follows: by negating the first disjunct, Respondent has already (logically) committed himself to the second disjunct, that he is the Pope. So if he is not in Rome and he is

294 'Lazy' meaning, of course, computationally simple.

295 Here there is an interesting similarity with modern logical games: the proof of a proposition G is seen as a winning strategy to play game G. (cf. van Benthem 2001, 74)

Proposition	Calculation	Verdict	Outcome
φ_0: You are in Rome or you are the Pope.	Possible	Conceded	$\Gamma_0 = \{\varphi_0\}$
φ_1: You are in Rome.	$\Gamma_0 \nvdash \neg\varphi_1$ $\Gamma_0 \nvdash \neg\varphi_1$ $K_c \Vdash\!\!/ \neg\varphi_1$	Denied	$\Gamma_1 = \{\varphi_0, \neg\varphi_1\}$
φ_2: The pope is in Rome.	$\Gamma_1 \nvdash \varphi_2$ $\Gamma_1 \nvdash \neg\varphi_2$ $K_c \Vdash\!\!/ \neg\varphi_2$	Conceded	$\Gamma_2 = \{\varphi_0, \neg\varphi_1, \varphi_2\}$
φ_3: You are the pope.	$\Gamma_2 \vdash \varphi_3$ (from φ_0 and $\neg\varphi_2$)	Conceded	$\Gamma_3 = \{\varphi_0, \neg\varphi_1, \varphi_2, \varphi_3\}$
φ_4: You are in Rome.	$\Gamma_3 \vdash \varphi_4$ (from φ_2 and φ_3)	Conceded	$\Gamma_4 = \{\varphi_0, \neg\varphi_1, \varphi_2, \varphi_3, \varphi_4\}$

Figure 3.3.3.2. A hypothetical disputation

But $\varphi_1 = \varphi_4$. So $\Gamma_4 \vdash \perp$

the Pope, then the Pope cannot be in Rome. So Respondent should deny φ_2, instead of accepting it as a proposition irrelevant and true according to K_C, even though he has not explicitly granted φ_3 yet.

But how does one capture the relation of inference between φ_0, $\neg\varphi_1$, and $\neg\varphi_2$? Some non-trivial logical concepts are involved: the unique denotation of ‘The Pope’ and ‘You’, the transitivity of terms, and the tacit commitment to ‘You are the Pope’, insofar as it is the second disjunct of an accepted disjunction, whose first disjunct has been denied. To identify this inferential relation, Respondent’s inferential system must account for all these logical concepts. We shall discuss the issue of which inferential system is in question in section 3.3.4.

In sum, given the complex structure of the language in use, simplifying devices such as a complete valuation defined beforehand are not available to Respondent. He must remain alert in order not to be caught up in the webs of the language being used, and to be able to spot all inferential relations, even the well-hidden ones.

3.3.3.3 The game is dynamic

Another source of difficulty in this game is its dynamic character. This is related to the inclusion of irrelevant propositions, accepted or denied according to K_C, in the set of propositions that will be used to respond to each *propositum* still to come. In this section the dynamic aspects of the game will be explored.

First of all, let me show that, if it were not for the inclusion of irrelevant propositions, that is, if only inferential relations prompted acceptance or denial of a proposition, then the game would not be dynamic.

Theorem 3.3.3.3.1
If

$$
\begin{aligned}
R'(\varphi_n) &= 1 \text{ iff} \quad - \Gamma_{n-1} \vdash \varphi_n \\
R'(\varphi_n) &= 0 \text{ iff} \quad - \Gamma_{n-1} \vdash \neg\varphi_n \\
R'(\varphi_n) &= ? \text{ iff} \quad - \Gamma_{n-1} \nvdash \varphi_n, \Gamma_{n-1} \nvdash \neg\varphi_n
\end{aligned}
$$

then the game defined by $R'(\varphi)$ is not dynamic.

Lemma. If $\varphi_0 \vdash \varphi_1$; $\varphi_0, \varphi_1 \vdash \varphi_2$; $\varphi_0, \varphi_1, \varphi_2 \vdash \varphi_3$ [...] $\varphi_0 \ldots \varphi_{n-1} \vdash \varphi_n$ then $\varphi_0 \vdash \varphi_n$

Proof: I prove by induction that, if $\varphi_0, \ldots, \varphi_k \vdash \varphi_{k+1}$, then $\varphi_0 + \varphi_{k+1}$.
Basic Case (n =1): $\varphi_0 \vdash \varphi_1$ (trivial)
Inductive step (n = k):

(Given)	$\varphi_0, \ldots, \varphi_k \vdash \varphi_{k+1}$
(By Induction Hypothesis)	$\varphi_0 \vdash \varphi_i$ for all $i \leq k$
(By Cut)	$\varphi_0 \vdash \varphi_{k+1}$ □

This means that, in this modified version of the game, in order to find out whether some φ_n must be accepted or denied, Respondent has to verify the existence of inferential relations with respect to only one proposition, the *positum*.[296] Moreover, if Respondent does not have to take any of the previously accepted or denied propositions into account (except for the *positum* φ_0) then the order in which the propositions are put forward is irrelevant.

Just as much as the game defined by $R'(\varphi_n)$ is **not** dynamic, the game defined by $R(\varphi_n)$ **is** dynamic. Burley himself attracts the reader's attention to this point:

> Next are rules that do not constitute the practice of this art, but are just useful. One rule concerning the obligational art is this: *One must pay special attention to the order*.[297] (Burley 1988, 385)

This means that, during a disputation, it may occur that (1) $\varphi_0, \varphi_1 \vdash \varphi_2$ but $\varphi_0, \varphi_2 \nvdash \varphi_1$, or else that (2) $\varphi_0, \varphi_1 \vdash \varphi_2$ but $\varphi_0 \nvdash \varphi_2$. (1) is related to the obvious asymmetric character of implication/inference,[298] and (2) to the dynamic nature of

[296] In fact, this is what happens in the *nova responsio*. According to this mode of *obligationes*, 'a relevant sentence was one that either followed from or was inconsistent with the posited sentence alone' (Spade 1982a, 336), as will be discussed in section 3.4.
[297] Sequitur de regulis quae non sunt de esse istius artis, sed solum utiles. Una regula de arte obligatoria est ista: ordo est maxime attendendus. Et ratio huius regulae est ista: quod uno loco est concedendum, alio loco non est concedendum. (Green 1963, 52(1–5))
[298] $(P \rightarrow Q)$ is obviously not equivalent to $(Q \rightarrow P)$.

the game, what I shall call the 'expansion of the informational base Γ_n'. This can be best seen if we take a look at what happens in terms of models during and obligational disputation. For that, here are some definitions:

Definition 3.2.3.2.1: Γ_n = Informational base, that is, a set of propositions.

Definition 3.2.3.2.2: $\boldsymbol{UM}_\mathbf{n}$ = The class of models that satisfy informational base Γ_n.

Definition 3.2.3.2.3: $\boldsymbol{UM}\varphi_n$ = The class of models that satisfy φ_n.

Definition 3.2.3.2.4: $\boldsymbol{UM_n} \Vdash \Gamma_n$ iff $\boldsymbol{UM_n} \Vdash$ P for all P in Γ_n.

A model that satisfies a set of propositions satisfies each of them (i.e., they are all true in this model). It is clear that, if $\Gamma_k = \{\varphi_n\} \cup \{\varphi_m\}$, then $\boldsymbol{UM_k} = \boldsymbol{UM}\varphi_n \cap \boldsymbol{UM}\varphi_m$. So, the set of models that satisfy Γ_k is the intersection of all the models that satisfy each of the elements of Γ_k. Similarly, if $\Gamma_{n+1} = \Gamma_n \cup \{\varphi_{n+1}\}$, then $\boldsymbol{UM_{n+1}} = \boldsymbol{UM_n} \cap \boldsymbol{UM}\varphi_{n+1}$.

Theorem 3.3.3.3.2: If $\Gamma_n \vdash \varphi_{n+1}$ and φ_{n+1} is accepted, then $\boldsymbol{UM_n} = \boldsymbol{UM_{n+1}}$. Assume that, at a given state of the game, $\Gamma_n \vdash \varphi_{n+1}$. According to $R(\varphi_n)$, φ_{n+1} must be accepted, forming $\Gamma_{n+1} = \Gamma_n \cup \{\varphi_{n+1}\}$. Now take $\boldsymbol{UM_n}$, that is, all the models that satisfy Γ_n. According to the model-theoretic definition of implication[299] (i.e., P $\vdash$ Q iff Q is true in all models where P is true, i.e., if $\boldsymbol{UM_P} \Vdash$ Q), if $\boldsymbol{UM_n} \Vdash \Gamma_n$ and $\Gamma_n \vdash \varphi_{n+1}$, then $\boldsymbol{UM_n} \Vdash \varphi_{n+1}$. Since $\Gamma_{n+1} = \Gamma_n \cup \{\varphi_{n+1}\}$, $\boldsymbol{UM_n} \Vdash \Gamma_n$ and $\boldsymbol{UM_n} \Vdash \varphi_{n+1}$, then $\boldsymbol{UM_n} \Vdash \Gamma_{n+1}$. It is defined that $\boldsymbol{UM_{n+1}} \Vdash \Gamma_{n+1}$, so $\boldsymbol{UM_n} = \boldsymbol{UM_{n+1}}$.

Thus, all the models that satisfy Γ_n also satisfy Γ_{n+1}. □

Theorem 3.3.3.3.3: If $\Gamma_n \nvdash \varphi_{n+1}$ and φ_{n+1} is accepted, then $\boldsymbol{UM_{n+1}} \subset \boldsymbol{UM_n}$. Assume that, at a given state of the game, $\Gamma_n \nvdash \varphi_{n+1}$ and $K_C \Vdash \varphi_{n+1}$. According to $R(\varphi_n)$, φ_{n+1} must be accepted, forming $\Gamma_{n+1} = \Gamma_n \cup \{\varphi_{n+1}\}$. $\boldsymbol{UM}_{n+1}$ is the intersection of $\boldsymbol{UM_n}$ and $\boldsymbol{UM}\varphi_{n+1}(\boldsymbol{UM_{n+1}} = \boldsymbol{UM_n} \cap \boldsymbol{UM}\varphi_{n+1})$. But because $\Gamma_n \nvdash \varphi_{n+1}$, $\boldsymbol{UM_n} \nVdash \varphi_{n+1}$. So not all models that satisfy Γ_n also satisfy φ_{n+1}. Since $\Gamma_{n+1} = \Gamma_n \cup \{\varphi_{n+1}\}$, not all models that satisfy Γ_n also satisfy Γ_{n+1}. Thus $\boldsymbol{UM_{n+1}} \neq \boldsymbol{UM_n}$. But Γ_n is contained in Γ_{n+1}, so all models that satisfy Γ_{n+1} also satisfy Γ_n – $\boldsymbol{UM_{n+1}} \Vdash \Gamma_n$. So $\boldsymbol{UM_{n+1}} \subset \boldsymbol{UM_n}$. □

Thus, all the models that satisfy Γ_{n+1} are contained in the set of models that satisfy Γ_n.

Summing up; in an obligational game, $\boldsymbol{UM_{n+1}} \subseteq \boldsymbol{UM_n}$. If $\Gamma_n \Vdash \varphi_{n+1}$, $\Gamma_n \Vdash \neg\varphi_{n+1}$ or $R(\varphi_{n+1})$ = ?, then $\boldsymbol{UM_n} = \boldsymbol{UM_{n+1}}$, otherwise $\boldsymbol{UM_{n+1}} \subseteq \boldsymbol{UM_n}$. That is, the larger the informational base, the fewer models will satisfy it, and greater the

[299] Cf. section 2.3.2.1.1 of the present work.

constraints on the choice between $\neg\varphi_n$ and φ_n will be (a model-theoretic way to see why a larger base implies that more propositions will have inferential relations with Γ_n). Clearly, the base is expanded (and therefore the range of models that satisfy it reduced) only by inclusion of 'irrelevant' propositions. Similarly, if, for all φ_n, $\varphi_0 \vdash \varphi_n$ (the game played according to $R'(\varphi_n)$), then there is no expansion of the informational base Γ_n, and no reduction of the models (valuations) satisfying the game.

It is also interesting to see that these conclusions echo the idea that the obligational game mirrors the construction of maximal-consistent sets of propositions, or possible worlds. From the model perspective, what happens is that, by responding to the *proposita*, Respondent gradually narrows the range of models that satisfy Γ_n, and if this procedure continues so as to cover all propositions (an ideal situation, obviously not to be reproduced on a human scale), then at the end there will be only one model left. If one adopts the first interpretation of the logical rules, where K_C plays a more prominent role, then the model that is built by a play of the game is the model that most resembles K_C, except for φ_0 and its consequents.

A similar result is established if one applies the logic of questions (Stokhof and Groenendijk 1997) to *obligationes*. Applying this framework to *obligationes* is a quite natural step, as each *propositum* φ_n is obviously a question of the yes/no kind (but Respondent may also 'pass' by doubting it). According to Groenendijk and Stokhof, questions of the yes/no kind establish a partition of the logical space: possible worlds in which φ_n is true, and possible worlds in which φ_n is not true (Figure 3.3.3.3.1).

Suppose that Respondent accepts φ_n. When the game goes on with φ_{n+1}, a new partition is made, but not of the entire logical space. Rather, the game goes on with the part of the logical space that was chosen by Respondent, which means that the choices made by Respondent are cumulative (Figure 3.3.3.3.2).

The dynamic character of *obligationes*, corresponding to the role of irrelevant *proposita*, is perhaps one of the most interesting features of this kind of disputation.[300] It is also one of the main reasons why *obligationes* remain a hard game to play. Moreover, the fact that Respondent is committed to his set of beliefs represents an extra constraint upon his responses with respect to irrelevant propositions.

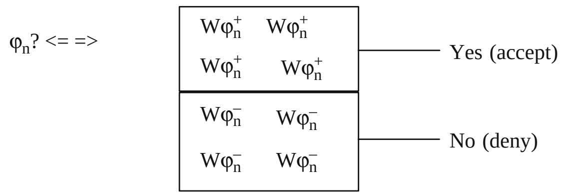


Figure 3.3.3.3.1. Obligationes as partitions of possible worlds I

[300] In this sense, the move to the *nova responsio* really seems to make the game a lot less interesting, as we shall see below.

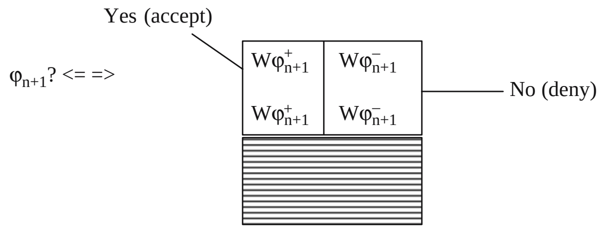


Figure 3.3.3.3.2. Obligationes as partitions of possible worlds II

3.3.4 Problems

The account of *obligationes* based on the idea of logical games presented here may have shed light on some aspects of this still rather puzzling body of medieval logical theories, but some other aspects require further clarification within this account.

The role of Opponent. In other versions of logical games, the Opponent is often characterized as 'Falsifier' (playing against Verifier), as 'Nature' (playing against Me), or as 'Spoiler'. In *obligationes*, what does Opponent try to accomplish? One could describe him as something like 'Tester': he seeks to test Respondent's abilities to maintain consistency, including his recognition of inferential relations. In this sense, he is motivated not necessarily by winning, but rather by measuring the skills of Respondent. So, it is reasonable to imagine that, when Respondent is inexperienced, Opponent may not set up his most difficult traps, with which he knows Respondent will not be able to cope; he may instead adopt an easier line of 'attack', in order to test what Respondent can actually do. In this case, it may really be something like a school exam, which should be imposed on the student according to his own stage of education. But one can also think of circumstances (i.e., a disputation between two students) in which Opponent's purpose is to win.

In this sense, there seem to be no clear instructions as to how Opponent should play the game. The only instruction is that he should put forward propositions to which Respondent must respond. However, it is certainly the case that his choice of φ_n is motivated by Respondent's previous responses. But how he actually conducts his choices of propositions (his moves) is difficult if not impossible to model in the present framework.

Another important task of Opponent (possibly to be performed together with other masters) is to judge whether Respondent has responded well. First of all, Opponent must himself be able to evaluate whether Respondent has maintained consistency; in certain ways, this task is even harder than that of Respondent. Respondent may simply perform model checking, when responding to the *proposita*, by means of comparing them to a model (valuation) previously defined by him. But Opponent must perform consistency management, a task of higher complexity (cf. van Benthem 2001, 3). In this sense, it seems reasonable to imagine that Opponent was typically a more experienced logician, for he had a harder task to accomplish, even though he was not the one under attack.

How should $\Gamma \vdash \varphi$ be defined? This is perhaps the most serious problem still to be dealt with within the account of *obligationes* proposed here. In a model-theoretic version of implication – P implies Q iff, in every model where P is true, Q is true – the relation of implication and inference with consistency is evident. From this point of view, the syntactic (inferential relations) and semantic (consistency maintenance) aspects of the obligational game would seem perfectly integrated. However, as briefly mentioned before, there were different versions of the medieval notion of *consequentia* in operation, just as much as there is a range of different notions of consequence in operation in modern logic (cf. Sundholm 2002a). But clearly the game would be played differently if the participants subscribed to one given inferential system instead of another – for example, in modern times one might opt to play the game based on intuitionistic patterns of inference, instead of the classical ones. Therefore, the performance of the game presupposes a pre-agreement with respect to the notion of inference to be adopted.

Moreover, for the game to be played successfully, one needs a suitable inferential system, which would indeed account for all the relations between any propositions P and Q such that Q is true in all circumstances in which P is true. The example given in section 3.3.3.2 indicates that it is not always a straightforward matter to define inferential rules that would capture, on the syntactic level, relations between propositions that exist on the semantic level, but which are not immediately evident. Explicit rules of inference had then, as they have now, the role of enabling the easy recognition of inferential patters, so that one does not need to resort to the often more complicated semantic level of truth-preservation. But such rules typically do not cover all possible cases in a non-axiomatizable language, and therefore a syntactic system of rules of inference could only resolve part of Respondent's problem.

Falling in contradiction. Finally, it must be clearly defined what it means for Respondent to fall in contradiction. In the example in section 3.3.3.2, when Respondent granted φ_0, $\neg\varphi_1$, φ_2, he had formally already fallen in contradiction, but in practice he had not yet **granted** a contradiction (which would happen only when the contradiction was made explicit). It seems that rhetorical elements also come into play: Opponent must not only make Respondent fall in contradiction but he must also force him to admit his contradiction. This may happen very much in the way in which Socrates, in the Platonic dialogues, shows to his 'respondents' that their position is flawed and that they are ignorant with respect to the topic being discussed. There are psychological, rhetorical aspects involved that seem to go beyond the simple model-theoretical notion of inconsistency. One can say that making Respondent **fall** in contradiction is a different victory criterion from making Respondent **grant** a contradiction.

3.3.5 Conclusion

This analysis of Burley's theory shows that the logical properties and structure of *obligationes* can be explored in a fruitful way with the hypothesis that they are logical games of consistency maintenance. In the next sections, the same hypothesis will be applied to Swyneshed's and Strode's theories, so as to ensure continuity and uniformity in order to compare each theory, even though the game perspective over *obligationes* is admittedly at its best with respect to Burley's theory.

3.4 SWYNESHED'S *OBLIGATIONES*: INFERENCE RECOGNITION

As a natural continuation to the analysis of Burley's theory, I now intend to test the game hypothesis on another important theory of *obligationes*, namely Roger Swyneshed's theory. Burley wrote his *obligationes* treatise roughly at the beginning of the 14th century, and enjoyed unanimous popularity for a certain time. Roger Swyneshed's treatise on *obligationes* seems to have been written some 30 years later, and is clearly a reaction to Burley's theory, proposing many revisions to the rules governing obligational disputations. These two theories, Burley's and Swyneshed's, were described by Robert Fland respectively as *antiqua responsio* and *nova responsio*.[301] Therefore, any serious attempt to understand the obligational genre should approach at least these two theories.

The most striking contrast between the two theories is that, in the case of Swyneshed, the application of the rules simply does not safeguard consistency. Indeed, the present reconstruction shows that Swyneshed's *obligationes* are by no means games of **consistency maintenance**. They can at most be considered to be games of **inference recognition**, but even this claim must be qualified: Swyneshed's goal seemed to be to exclude all dynamic features[302] of Burley's theory, but thereby he ended up excluding its most interesting game-theoretical aspects as well. If Swyneshed's *obligationes* are games at all, they are of a less interesting kind.

3.4.1 Reconstruction

The reconstruction of Swyneshed's theory proposed here follows roughly the same lines as the reconstruction of Burley's theory, to facilitate the comparison. Moreover, as much as in my discussion of Burley, I will focus on Swyneshed's treatment of *positio*, thus disregarding *impositio* and *depositio* (the other kinds of obligational disputations considered by Swyneshed in his treatise).

3.4.1.1 Central notions

In Swyneshed's version, an obligation corresponds to the following quadruple:

Definition 3.4.1.1: the obligational game (Swyneshed)

$$\boldsymbol{Ob} = \langle \boldsymbol{\Sigma}, \boldsymbol{\Phi}, \mathbf{I}, \mathbf{R}\ (\varphi_n) \rangle$$

---

[301] Cf. (Spade 1982b, 3).

[302] In the sense that only the *positum* and the state of common knowledge, and not the previous moves, had to be taken into account at each move. By contrast, I will use the term 'static' to characterize Swyneshed's version of *obligationes* to stress the fact that, in the latter, previous moves and their order do not influence how Respondent should reply to a given *propositum*.

Σ is an ordered set of states of knowledge S_n. This is the first significant difference with respect to Burley's theory. In the latter, all irrelevant propositions were supposed to be answered to according to the static state of common knowledge K_C.[303] Changes in things during the time of the disputation were not supposed to affect the response to (irrelevant) propositions, all the more since, once proposed and accepted or denied, these were included in the 'informational base' of the disputation. So, in Burley's theory, if, at a certain point, 'You are seated' is proposed to Respondent, and Respondent is indeed seated, he should accept the proposition. Subsequently, if Respondent stands up, and Opponent proposes 'You are not seated', Respondent should deny it, because it contradicts the set of previously accepted/denied propositions, and this logical relation has priority over reality.

In Swyneshed's theory, since irrelevant accepted or denied propositions are not included in the informational base of the disputation, as we shall see, the state of knowledge is not required to be static. So the response to irrelevant propositions, according to Swyneshed's theory, should take into account the changes in things during the time of the disputation; therefore, what we have is a series of states of knowledge S_n, ordered according to their index *n*, which is a natural number and corresponds to the stage of the disputation in which the state of common knowledge must come into play.[304]

Φ is an ordered set of propositions φ_n. These are the propositions proposed during the disputation; their index *n* is a natural number and corresponds to the place they occupy in the order in which the propositions are proposed. (No difference here with respect to Burley's theory.)

I is an ordered set of responses $\iota_n = [\varphi_n; \gamma]$. Responses are ordered pairs of propositions and one of the replies 1, 0 or ?, corresponding to Respondent's response to proposition φ_n. Note that the index of the response need not be the same as the index of the proposition, in case the same proposition is proposed twice, in different moments of the disputation.

In my reconstruction of Burley's theory, responses were not primitive constituents of the game, and were introduced only to account for the 'point system' of the second, non-deterministic interpretation of Burley's theory (cf. section 3.3.1.2.3). But to express some of the interesting properties of Swyneshed's theory, the notion of responses is crucial.

R(φ) is a function from propositions to the values 1, 0, and ?. This function corresponds to the rules that Respondent must apply to respond to each proposition φ_n. 1 corresponds to his accepting φ_n, 0 to his denying φ_n and ? to his doubting φ_n. This definition is identical to the definition of **R**(φ) in the reconstruction of Burley's

303 Cf. (Ashworth 1996, 352).

304 But why use states of knowledge, and not simply states of affairs? Because, both in Burley's and Swyneshed's theories, proposed propositions whose truth-value is unknown to the participants of the disputation – for example, 'The Pope is sitting now' – should be accordingly doubted. We are dealing here with imperfect states of information.

theory, but the **function** corresponding to the rules of Swyneshed's theory is different from the function of Burley's theory, since the rules are different.

3.4.1.2 Rules of the game

The **procedural rules** of the game are quite simple, and identical to the procedural rules in Burley's theory. Opponent first puts forward a proposition. If Respondent accepts it (according to $\mathbf{R}(\varphi_0)$ defined below), the game begins.[305] Then Opponent puts forward a further proposition, Respondent responds to it according to $\mathbf{R}(\varphi_n)$, and this procedure is repeated until the end of the game.[306]

By contrast, the logical rules of Swyneshed's game are quite different from Burley's.

3.4.1.2.1 Positum/Obligatum. Swyneshed's analysis of the requirements for a proposition to be accepted as *obligatum* (i.e., the first proposition proposed, named *positum* in the specific case of *positio*) is less extensive than Burley's. Since an inconsistent *positum* gives no chance of success for Respondent, Burley clearly says that the *positum* mustn't be inconsistent. Swyneshed does not follow the same line of argumentation; rather, he requires that a proposition be contingent to be a *positum*, which he phrases in the following way:

> It must be known that every proposition which, out of the time of the obligation, must receive different answers because of changes in things are here to be obliged [accepted as *obligatum*], and no other.[307]

That is, Swyneshed requires of a proposition that it be sometimes known to be true, sometimes known to be false, and sometimes not known to be true and not known to be false, for it to be accepted as an *obligatum* – a situation which would prompt

---

[305] After the positing of these rules it must be seen with which signs, with which propositions and in which manner in this species of obligation the obligation occurs, and when and for how long the Respondent will be obliged. For the first <point> it must be known that by means of these signs 'I posit' or 'it is posited', 'I suppose', or 'it is supposed', the obligation happens.

Positis regulis videndum est quibus signis et per quas propositiones et qualiter in hac specie obligationis contingit obligare et quando et per quantum tempus respondens erit obligatus. Pro primo est sciendum quod mediantibus istis signis 'pono' vel 'ponitur', 'suppono' vel 'supponitur' contingit hic obligare. In Spade 1977, (§72). All quotations from Swyneshed's treatise will refer to this edition, and the reference (§xx) is to the paragraph in which they appear. The translation of Swyneshed's passages are my own, in collaboration with Marije Martijn and Mariska Leunissen.

[306] With respect to the fourth and fifth <points> taken together, it must be known that, once one of these signs is uttered with an *obligatum*, if the Respondent assents, he is obliged, and for some time, until the phrase 'Let the time of obligation be over' is uttered, the Respondent will continue to be obliged. And once the phrase 'Let the time of obligation be over' is uttered, the Respondent is no longer obliged.

Pro quarto et quinto conjunctim est sciendum quod dicto aliquo illorum signorum cum obligato, si respondens assentiat statim obligatur, et aliquo tempore quousque dicatur illa oratio 'Cedat tempus obligationis' continue erit respondens obligatus. Et dicta illa oratione 'Cedat tempus obligationis' non est respondens obligatus amplius. (§75)

[307] *... est sciendum quod omnis propositio ad quam extra tempus obligationis per mutationem ex parte rei est varianda responsio et nulla alia est hic obliganda.* (§73). I am indebted to Prof. E.J. Ashworth for having spotted this passage for me.

a variation in the answers it would receive, were it to be proposed out of the time of an obligation. This excludes impossible propositions – always false – and necessary propositions – always true – and that is a necessary requirement in view of the *ex impossibili sequitur quodlibet* rule: if Swyneshed's rules of *obligationes* are indeed meant to test Respondent's abilities to recognize inferential relations, an impossible *obligatum* would make the game trivial (any proposition would follow).[308] Moreover, from a necessary proposition only necessary propositions follow, so if the *obligatum* is a necessary proposition, then the game becomes that of recognizing necessary propositions, that is, a deviation from its original purpose.

So the rule for accepting the *positum* could be formulated as:[309]

Definition 3.4.1.2.1.1: Rules for *positum*

> $R(\varphi_0) = 0$ iff, for all moments *n* and *m*, and for one reply γ, $\iota_n = [\varphi_0; \gamma]$ and $\iota_m = [\varphi_0; \gamma]$.
> $R(\varphi_0) = 1$ iff, for some moments *n* and *m*, and for two replies γ and $\varkappa$, $\gamma \neq \varkappa$, $\iota_n = [\varphi_0; \gamma]$ and $\iota_m = [\varphi_0; \varkappa]$

Moreover, Swyneshed also gives instructions as to how to respond to the *positum* if it is **posited again** during the disputation (§§ 62–64). A *positum* which is re-proposed must be accepted, except in the cases of a *positum* which is inconsistent with the very act of positing, admitting and responding in an obligational context. The paradigmatic example is 'Nothing is posited to you': it should be accepted as a *positum*, according to the rules above, but if it is again proposed during the same disputation, it should be responded to as if it were an irrelevant proposition. In this case, it would be denied, even though it had been accepted as *positum*.

In effect, from the start, the set of all propositions (not only those put forward during the disputation, which constitute **Φ**) is divided in two sub-sets, namely the set of propositions that are **pertinent** with respect to the *positum* φ_0[310] – denoted $\Delta_{\varphi 0}$ – and

---

[308] Notice that Swyneshed's reason for excluding impossible propositions is different from Burley's – trivialization of the game *versus* absence of a winning strategy for Respondent. (Keffer 2001) has also remarked that impossible (and true) *posita* have a *Trivialisierungseffekt* on both kinds of responses, but for different reasons (pp. 158–164). Notice also that Swyneshed applies the obligational framework to non-obligational situations to define a contingent proposition.

[309] Like Swyneshed, I am making use of the obligational conceptual framework to describe a situation **out** of the time of a disputation.

[310] A proposition is either pertinent or impertinent to the *obligatum*. And of the ones that are pertinent to the *obligatum*, they either follows from or are repugnant to the *obligatum*.

Propositionum alia est pertinens obligato, alia est impertinens obligato. Et pertinentium obligato alia est sequens ex obligato, alia repugnans obligato. (§4)

A pertinent proposition is a proposition that is not obliged [that is not the *obligatum*], which, in whichever way it signifies, must be accepted or denied in virtue of the *obligatum*.

Propositio pertinens est propositio non obligata quae, qualitercumque significet, propter obligatum est concedenda vel neganda. (§7) the set of those that are **impertinent** with respect to the *positum* φ_0[311] – denoted $\Pi_{\varphi 0}$. The sets are defined as follows:

Definition 3.4.1.2.1.2: Pertinent and impertinent propositions

$$\Delta_{\varphi 0} = \{\varphi_n \in \Delta_{\varphi 0} : \varphi_0 \vdash \varphi_n \text{ or } \varphi_0 \vdash \neg\varphi_n\}$$

$$\Pi_{\varphi 0} = \{\varphi_n \in \Pi_{\varphi 0} : \varphi_0 \nvdash \varphi_n \text{ and } \varphi_0 \nvdash \neg\varphi_n\}$$

Assuming that any proposition implies itself, the *positum* φ_0 belongs to $\Delta_{\varphi 0}$.[312] E. Stump (Stump 1981, 167) mentions the possibility of allowing for a second *positum* at any given moment of the disputation. In this case, obviously the two sets defined above must be revised, and the set of pertinent propositions is defined by the conjunction of the two (or more) *posita*.

3.4.1.2.2 Proposita. The rules for responding to the proposed propositions other than the *positum* are better formulated in two steps, first for the pertinent, then for the impertinent propositions, as this division is in fact the decisive aspect of the game in Swyneshed's version.

Definition 3.4.1.2.2: Rules for *proposita*
Pertinent propositions ($\varphi_n \neq \varphi_0$, $\varphi_n \in \Delta_{\varphi_0}$)[313]

$$\mathbf{R}(\varphi_n) = \mathbf{1} \text{ if } \varphi_0^{314} \vdash \varphi_n$$

$$\mathbf{R}(\varphi_n) = \mathbf{0} \text{ if } \varphi_0 \vdash \neg\varphi_n$$

311 An impertinent proposition is a proposition that is not obliged, which should not be conceded or denied in virtue of the *obligatum*.

Propositio impertinens est propositio non obligata, et propter obligatum nec est concedenda nec neganda. (§8)

312 Except for the *posita* that are (pragmatically) repugnant to the act of positing (cf. §64); according to Swyneshed, those should be answered to as if they were impertinent, thus belonging to $\Pi_{\varphi 0}$.

313 The response to pertinent propositions, and not the response to impertinent propositions, must vary in virtue of the *obligatum*. This is clear. For only the response to a proposition that either follows [from] or is repugnant [to the *obligatum*] must vary in virtue of the *obligatum*.

Ad propositionem pertinentem et non ad impertinentem propter obligatum est responsio varianda. Hoc patet. Nam propter obligatum non est varianda responsio nisi ad sequens vel ad repugnans. (§24)

Second rule: every proposition that follows from the *positum*, without obligation pertinent to it, and which is not repugnant to the positing during the time of the obligation must be conceded.

Secunda regula: Omne sequens ex posito sine obligatione ad hoc pertinente non repugnans positioni in tempore obligationis est concedendum. (§67)

Third rule: every proposition that is repugnant to the *positum*, without obligation pertinent to it, and which is not repugnant to the positing during the time of the obligation must be denied.

Tertia regula: Omne repugnans posito sine obligatione ad hoc pertinente non repugnans positioni in tempore positionis est negandum. (§68)

314 Clearly, if the introduction of extra *posita* occurs, then this definition holds for the set of *posita*, instead of for the first *positum* only.

Impertinent propositions ($\varphi_n \,\varepsilon\, \Pi_{\varphi_0}$) [315]

$$\mathbf{R}(\varphi_n) = 1 \text{ if } S_n \Vdash \varphi_n$$
$$\mathbf{R}(\varphi_n) = 0 \text{ if } S_n \Vdash \neg\varphi_n$$
$$\mathbf{R}(\varphi_n) = ? \text{ iff } S_n \not\Vdash \varphi_n \text{ and } S_n \not\Vdash \neg\varphi_n$$

As is clear, the most fundamental disagreement between Burley's and Swyneshed's theories concerns the definition of pertinent/impertinent propositions.[316] For Burley, a pertinent proposition is one that follows from (or whose contradictory follows from) the conjunction of all previously granted propositions and the contradictories of all previously denied propositions. If the intuition that Swyneshed wanted to suppress all dynamic aspects of Burley's *obligationes* is correct, then he certainly hit the bull's eye by modifying the definition of pertinent/relevant propositions. The changes in R(φ) simply follow from this modification.

3.4.1.2.3 Outcome. The game ends when Opponent says '*Cedat tempus obligationis*'. From Swyneshed's text, it seems that Opponent can say it at any time; he will say it when Respondent has made a bad move, and thus has lost the game,[317] but he may say it when he is satisfied with the performance of Respondent, who until then has not made any bad move, and therefore has 'won' the game. However, Swyneshed's remarks on this are rather brief (more on a criterion of loss below).

---

[315] The response to an impertinent [proposition] must not vary in virtue of the *obligatum*. Therefore, if an obligation is not pertinent to the latter, it follows that to such impertinent proposition the response must not vary in virtue of the *obligatum* nor in virtue of the obligation. Therefore, to such impertinent [proposition] one must respond during the time of the disputation as [one would] outside the time of the disputation. Now, outside the disputation, any of these would have to be accepted by anyone immediately aware that it signifies just as things are. Therefore, [the same holds for] during the time of the disputation.

Ad impertinens non est varianda responsio propter obligatum. Ergo, si obligatio non sit pertinens ad illud, sequitur quod ad tale impertinens non est responsio varianda propter obligatum nec propter obligationem. Ergo, ad tale impertinens sic est respondendum infra sicut extra. Sed extra quaelibet talis foret concedenda a quocumque sciente principaliter sibi significare sicut est. Igitur, et infra. (§26)

Fourth rule: an impertinent [proposition] without obligation pertinent to it must be responded to as if for that which is primarily grasped.

Quarta regula: Ad impertinens sine obligatione ad hoc pertinente velut per illud quod principaliter concipitur respondendum est. (§69)

[316] This fact has been acknowledged by virtually all studies on medieval *obligationes*, including (Stump 1981), (Ashworth 1981), (Ashworth 1993), (Spade 1982b), (Keffer 2001) etc., so I claim no novelty here.

[317] If you accept, the time of the obligation is over. You have both accepted and denied [the same proposition] during the time of the disputation. Therefore, you have responded badly, and because there is not a change in things. If you deny, let the time of the obligation be over. [The proposition] follows from the *positum*. Therefore, it must be accepted.

Si conceditur, cedat tempus obligationis. Idem concessisti et negasti infra tempus obligationis. Igitur, male respondisti eo quod non est mutatio facta ex parte rei. Si negatur, cedat tempus obligationis. Illa sequitur ex posito. Igitur, concedenda. (§98)

3.4.2 Characteristics of Swyneshed's game

On the basis of this reconstruction, some of the relevant aspects of Swyneshed's version of the obligational game can be explored.

3.4.2.1 The game is fully determined

Only one answer to each proposition is correct at a given point. In this aspect, Swyneshed's *obligationes* resemble Burley's *obligationes* under the first interpretation (section 3.3.1.2.1 above), that is, the interpretation according to which Respondent has no space for maneuver and must answer according to R(φ). That this is the case is seen from the fact that R(φ) really is a function, assigning exactly one value to each argument of its domain (the class of propositions). Swyneshed's rules divide the class of propositions in two sets and in five sub-sets: pertinent propositions – 1. repugnant to or 2. following from the *positum* – and impertinent propositions – 3. which are known to be true; 4. which are known to be false; 5. which are not known to be true and are not known to be false. These five subsets exhaust the class of propositions, and for each of them there is a defined correct answer. The same occurs in Burley's theory under the first interpretation, with the difference that these five subsets are relative to each moment of the disputation and to the informational base built at each point.

However, as we have seen, Burley's theory seems to give rise to an alternative interpretation, in which Respondent has some choice. If maintaining consistency is the ultimate goal of Burley's game, then irrelevant propositions can be either accepted or denied, and there is always the possibility of using the option *dubito*. I have expressed this flexibility in terms of a 'point system' in the previous section. In contrast, in Swyneshed's treatise there is no mention at all to flexibility of choice. At each stage of the disputation, Respondent's moves are totally determined by the rules of the game.

3.4.2.2 The game is not dynamic

The game played according to the *antiqua responsio* is, as aforementioned, dynamic in that the player must take into account all previous moves of the game in their corresponding order (cf. Keffer 2001, 179). By contrast, the game played according to the *nova responsio* is 'static': the response to a proposition is entirely independent of the order in which it occurs during the disputation, as it is entirely independent of all previous moves except for the first one, relative to the *positum*. As said above, this seemed indeed to be one of the main goals of Swyneshed's revision of Burley's theory, which he accomplished by modifying the notion of pertinent/impertinent proposition. In effect, for any proposition φ_n, at any round *n* of the disputation, the reply to φ_n is always the same γ, where γ is either 1, 0 or ?:

$$\iota_n = [\varphi_n; \gamma]$$

In particular, given two rounds *n* and *m* of the game, we must have

$$\iota_n = [\varphi_n; \gamma] \text{ and } \iota_m = [\varphi_n; \gamma]$$

Indeed, the great difference with respect to Burley's theory is that, in Swyneshed's version, the game is totally determined once the *positum* has been posited, from the start, and not only at each move. All Respondent has to do is to correctly determine the two sets of pertinent and impertinent propositions from the outset. Opponent can do nothing to interfere in Respondent's winning strategy, as it simply consists of assessing correctly the presence or absence of relations of inference between the *positum* and the proposed propositions. In this sense, Swyneshed's *obligationes* are much less of a game than Burley's; in Burley's version of the game, the moves of each participant were decisive for the choice of subsequent moves by the other participant. This does not occur in Swyneshed's version. In a way, it is as if it was a game with **one** participant (Respondent), similar to a game in which, once the *positum* has been established, Respondent simply draws cards with arbitrary propositions on them and answers to these propositions according to the rules of the game (like solitaire or similar card games).

Indeed, in section 3.3.3.3, I have proved (by applications of the cut-rule) that a version of Burley's game in which impertinent propositions are all doubted, or receive no answer (resembling thus Swyneshed's version), is equivalent to a game in which only the *positum* determines pertinence (the aim was to show that the dynamic aspect of Burley's game was due to the role of impertinent propositions).

Once more, the fact that the game is totally determined from the moment the *positum* is posited means that the order of presentation of the *proposita* does not matter, and that Opponent cannot do much to make the game harder for Respondent. Moreover, it also means that, during a disputation, only one response is the right one for a given proposition, independent of when it is proposed. In Burley's game, it can happen that a proposition is first doubted (as impertinent and unknown) and then accepted or denied (it has become pertinent in the meantime, given the expansion of the informational base), that is, it is possible that, for two rounds *n* and *m*, for two replies γ and $\varkappa$,

$$\iota_n = [\varphi_n; \gamma] \text{ and } \iota_m = [\varphi_n; \varkappa]$$

where $\gamma \neq \varkappa$. This cannot occur in Swyneshed's game.

There is one exception to this rule: impertinent propositions whose truth-values change during the course of the disputation. Swyneshed says that these propositions should be answered to according to the state of knowledge of that moment, and therefore the response depends on the moment in which they are proposed – but not on the moment **within the disputation** in which they are proposed (their relative position with respect to other propositions). Similarly, if such propositions are proposed twice during the same disputation, they may receive different answers, as a consequence of a change in things.

Thus, let φ_n be an impertinent proposition with respect to a *positum* φ_0 ($\varphi_n \, \varepsilon \, \Pi_{\varphi_0}$). Let φ_n be proposed at round *n* of the disputation. Suppose that $S_n \Vdash \varphi_n$. According to R(φ), the response to φ_n at round *n* should be $\iota_n = [\varphi_n; 1]$. Then suppose that φ_n is again proposed at round *m* of the disputation, and that $S_m \Vdash \neg\varphi_n$. According to R(φ), the response to φ_n at round *m* should be $\iota_m = [\varphi_n; 0]$. But since such changes are caused by changes in things, Swyneshed does not see this as a problem to his theory, and he seems to think that such changes should be taken into account in the rules of *obligationes*.

Interestingly, this is in a way a more daring position than Burley's concerning different responses to the same proposition. In Burley's theory, a proposition can be doubted and then subsequently denied or accepted, but it can never be so that, for two rounds *n* and *m*,

$$\iota_n = [\varphi_n; 1] \text{ and } \iota_m = [\varphi_n; 0] \text{ or}$$
$$\iota_n = [\varphi_n; 0] \text{ and } \iota_m = [\varphi_n; 1]$$

The difference is of course that, for Burley, any proposition can potentially receive different responses at different rounds of a disputation, whereas for Swyneshed this can only occur with propositions that are impertinent with respect to the *positum* and whose truth-values according to reality change during the time of the disputation.

Thus, one criterion of loss for Respondent is if he gives two different answers to the same proposition (which is not an impertinent proposition whose truth-value changes during the disputation); since the game is determined and not dynamic, each proposition only has one correct response at **any** time during the disputation, so if a proposition receives two different responses at different times, one of these responses is necessarily incorrect, and therefore Respondent has responded badly in at least one of the two moves. In Burley's game, on the other hand, this is not a criterion of loss, even though Respondent cannot deny and then grant the same proposition (or vice-versa); but he can first correctly doubt and then deny or accept the same proposition.

3.4.2.3 Two disputations with the same positum *will prompt the same answers, except for variations in things*

This is perhaps the main motivation for the changes introduced by Swynehed to the obligational game. In many passages, he emphasizes that the response to impertinent propositions must vary only in virtue of changes in things, and not in virtue of other previously accepted/denied propositions. Indeed, the crucial element of a winning strategy for Swyneshed's game is the accurate definition of the two sets of propositions relative to a *positum* (the set of pertinent propositions and the set of impertinent ones). This can be done once the *positum* has been posed, and thus before any other proposition is proposed. Once the two sets are formed, the application of the rules of the game should follow in a straightforward way.

So, if the game is defined once the *positum* is posited, then any two disputations with the same *positum* have **the same winning strategy**, that is, the establishment of the same two sets of pertinent and impertinent propositions.

Since the propositions proposed by Opponent may vary, two disputations with the same *positum* will not necessarily be identical. But any given proposition proposed in both disputations will belong to the same set of propositions – either pertinent or impertinent – in both cases. Moreover, if a proposition is pertinent and is proposed in two different disputations with the same *positum*, it should obviously receive the same response in both cases, as the logical relation of following from or being repugnant to the *positum* is independent of other contextual elements of the disputation.

In the case of impertinent propositions, as is to be expected, some of them may receive different responses in two disputations that do not take place simultaneously and which have the same *positum*, but that is caused by a change that occurred in things, and not by elements of the disputation itself. This situation is analogous to a proposition being proposed twice during the same disputation and receiving different responses because of a change in things during the time of the disputation.

Again, the dissimilarity with Burley's theory is striking. In Burley's version of the game, the *positum* was merely one of the propositions constituting the set according to which a proposed proposition was to be evaluated as pertinent or impertinent (the others being the previously accepted/denied propositions). So in two disputations having in common only the *positum*, a given proposition proposed in each of them was most likely bound to receive different responses.

3.4.2.4 Responses do not follow the usual properties of the connectives

One of the most discussed aspects of the *nova responsio*, not only among medieval authors but also among modern commentators, is the non-observance of the usual behavior of some sentential connectives, in particular conjunction and disjunction. This is a corollary of the basic rules of the *nova responsio*, as the proofs below show, and it was thought to be one of its distinctive traits (cf. Stump 1981, 139). (Keffer 2001 offers similar proofs in his reconstruction, pp. 176–178.)

Theorem 3.4.2.4.1: 'One need not grant a conjunction in virtue of having granted all its conjuncts'. (Stump 1981, 138)

Proof: there is at least one situation in which two propositions must be granted, but not their conjunction.[318]

Suppose that φ_0 satisfies the condition to qualify as a *positum* (i.e., it is contingent). Then pose φ_0 as *positum*. The correct response is:

$$\iota_0 = [\varphi_0; 1]$$

[318] A conjunctive proposition does not have to be accepted in virtue of the concession of its parts, nor does any part of a disjunctive proposition have to be accepted in virtue of the concession of the disjunction. The first part of the conclusion is proved as follows: Let *a* be a conjunction made out of the false *obligatum* and of an impertinent [proposition] primarily signifying things as they are. Let *b* be the *obligatum*. Thereupon the parts are accepted, [but] the whole conjunction is impertinent to the *obligatum* and is known to signify primarily things in another way. Therefore, [the conjunction] must be denied.

Propter concessionem partium copulativae non est copulativa concedenda nec propter concessionem disjunctivae est aliqua pars ejus concedenda. Prima pars conclusionis probatur sic: Sit a una copulativa facta ex obligato falso et impertinente significante principaliter sicut est. Sit b illud obligatum. Tunc concessis istis partibus tota copulativa est impertinens obligato scita principaliter significare aliter quam est. Igitur, neganda. (§32)

Suppose then that φ_n ε $\Pi_{\varphi 0}$ and that $S_n \Vdash \varphi_n$; φ_n is proposed; the correct response is:

$$\iota_n = [\varphi_n; 1]$$

Then propose φ_m: φ_0 & φ_n. Clearly, φ_m ε $\Pi_{\varphi 0}$, since one of the conjuncts does not follow from the *positum*. Moreover, we have that $S_m \Vdash \varphi_n$ but $S_m \Vdash \neg\varphi_0$ (no change has occurred in things with respect to φ_0 and φ_n). So $S_m \Vdash \neg(\varphi_0 \,\&\, \varphi_n)$, that is, $S_m \Vdash \neg\varphi_m$. Therefore, the correct response to φ_m is:

$$\iota_m = [\varphi_m; 0]$$

Hence, two propositions have been accepted, but their conjunction has been denied. □

Theorem 3.4.2.4.2: 'One need not grant any disjunct of a disjunction in virtue of having granted that disjunction'. (Stump 1981, 138)

Proof: there is at least one situation in which a disjunction must be granted, but neither of the disjuncts must be granted.[319]

Suppose that φ_0: $\varphi_0{}'$ *v* $\varphi_0{}''$ satisfies the condition to qualify as a *positum* (i.e., it is contingent). Then pose φ_0 as *positum*. The correct response is:

$$\iota_0 = [\varphi_0; 1]$$

Let $\varphi_0{}'$ be proposed at *n*, and suppose that $S_n \Vdash \neg\varphi_0{}'$. Clearly, $\varphi_0{}'$ ε $\Pi_{\varphi 0}$, so the correct response is:

$$\iota_n = [\varphi_0{}'; 0]$$

Then let $\varphi_0{}''$ be proposed at *m*, and suppose that $S_m \Vdash \neg\varphi_0{}''$. Clearly, $\varphi_0{}''$ ε $\Pi_{\varphi 0}$, so the correct response is:

$$\iota_m = [\varphi_0{}''; 0]$$

Hence, the disjunction of two propositions has been accepted, but both disjuncts have been denied. □

Apparently, these two corollaries have struck some of Swyneshed's contemporaries as very odd, and were for them sufficient reason to reject the *nova responsio* as a

[319] The second part of the conclusion is proved as follows: let *c* be the opposite of such a conjunction, while *b* is the *obligatum*, just as before. It is argued: *c* is a disjunction, as it is the opposite of a denied conjunction. Therefore, it must be conceded. And that both parts must be denied is patent. For one part is the opposite of the *obligatum b*. Therefore, it must be denied. And the other part is an impertinent proposition immediately signifying other than what it is. Therefore, it must be denied.

Secunda pars conclusionis probatur sic: Sit c oppositum talis copulativae, b existente obligato sicut prius. Et arguitur sic: c disjunctiva est oppositum copulativae negatae. Igitur, illa est concedenda. Et quod utraque pars sit neganda patet. Nam una pars est opposita b obligato. Igitur, illa est neganda. Et alia est impertinens significans principaliter aliter quam est. Igitur, est neganda. (§32) whole. However, in careful inspection, it is only in appearance that two of the most fundamental laws of logic – the truth-conditions of conjunction and disjunction – are being challenged. As M. Yrjönsuuri suggested (Yrjönsuuri 1993, 317), it is as if the bookkeeping of a Swyneshed-style obligational disputation featured two columns of responses, one for pertinent propositions and one for impertinent propositions. **Within each column**, the laws for conjunction and disjunction are in effect observed. So, if in one of the columns two propositions have been correctly granted, then their conjunction will also be granted (disregarding changes in things); similarly, if in one of the columns a disjunction has been correctly granted, then at least one of the disjuncts will have to be granted too.[320] This fact only emphasizes the idea that the crucial aspect of playing a Swyneshed-style game of *obligationes* is the correct division between pertinent and impertinent propositions.

The apparent conflict arises in cases in which a conjunction or disjunction is formed by propositions taken from both columns.[321] To analyze such cases, one must distinguish pertinent propositions that follow from the *positum* from those pertinent propositions that are repugnant to the *positum*, since obviously this makes all the difference for the inferential relations between a conjunction or disjunction and the *positum* (and therefore for the status of a disjunction/conjunction as pertinent or impertinent). The same does not hold of impertinent propositions: it is irrelevant for the purposes of pertinence whether the (impertinent) parts of a disjunction/conjunction have been accepted or denied. In practice, we are dealing with three 'truth-values', which can be denoted by 'P+' (the propositions that follow from the *positum*), 'P−' (the propositions that are repugnant to the *positum*) and 'I' (the impertinent propositions). The truth-values that conjunction and disjunction should receive, according to each pair of conjunct/disjunct, are depicted in the table in Figure 3.4.2.4 (the table's content can be easily verified[322]).

In the case of conjunction, if both conjuncts belong to the pertinent set of propositions with respect to the *positum*, then the conjunction is a pertinent proposition;

φ_n	P+	P+	P+	P–	P–	I
φ_m	P+	P–	I	P–	I	I
$\varphi_n \& \varphi_m$	P+	P–	I	P–	P–	I
$\varphi_n \vee \varphi_m$	P+	P+	P+	P–	I	I

Figure 3.4.2.4. Conjunctions and disjunctions

320 For simplicity, I am disregarding impertinent propositions whose truth-value may change during the disputation.

321 I owe this remark to an anonymous referee.

322 If $\varphi_0 \vdash \varphi_n$, $\varphi_0 \vdash \varphi_m$ (by hypothesis) and $\varphi_n, \varphi_m \vdash \varphi_n \& \varphi_m$ (def. of conjunction), then $\varphi_0 \vdash \varphi_n \& \varphi_m$ (application of the cut rule); if $\varphi_0 \vdash \varphi_n$, $\varphi_0 \vdash \neg\varphi_m$ (by hypothesis) and $\varphi_n, \neg\varphi_m \vdash \neg(\varphi_n \& \varphi_m)$ (def. of conjunction), then $\varphi_0 \vdash \neg(\varphi_n \& \varphi_m)$ (application of the cut rule); if $\varphi_0 \vdash \varphi_n$ (by hypothesis) and $\varphi_n \vdash \varphi_n \vee \varphi_m$ (def. of disjunction), then $\varphi_0 \vdash \varphi_n \vee \varphi_m$ (application of the cut rule); if $\varphi_0 \vdash \varphi_n$, $\varphi_0 \nvdash \varphi_m$ (by hypothesis – φ_m is impertinent) and $\varphi_n, \varphi_m \vdash \varphi_n \& \varphi_m$ (def. of conjunction), then $\varphi_0 \nvdash \varphi_n \& \varphi_m$ etc.

similarly, if the conjuncts belong to the impertinent set of propositions with respect to the *positum*, then the conjunction should be judged as an impertinent proposition. But if one conjunct belongs to the set of pertinent propositions and the other to the set of impertinent propositions, then one must distinguish the case of a conjunct being a repugnant proposition from the case of a conjunct being a proposition that follows from the *positum*, as in the table.

As for disjunction, the disjunction itself may belong to the set of pertinent propositions (if it is the *positum*) and yet both disjuncts shall belong to the set of impertinent propositions (and be both denied). On the other hand, if either one of the disjuncts follows from the *positum*, then the disjunction follows from the *positum* too (and hence it is a pertinent proposition). If both disjuncts belong to the set of impertinent propositions, then the disjunction will be an impertinent proposition too, and therefore should be judged as such (unless it is the *positum*).

3.4.2.5 The set of accepted/denied propositions can be inconsistent

Perhaps the most surprising feature of Swyneshed's *obligationes* is the little importance attributed to consistency maintenance. That is, if one takes the set of all propositions granted and the contradictories of all propositions denied during a disputation, this set is very likely to be **inconsistent**, and this feature struck many medieval authors as very odd (cf. Keffer 2001, 164–166). There are two main sources of inconsistency in Swyneshed's game: the most obvious one is the case of impertinent propositions which receive two different responses in different times of the disputation (in particular if they are first denied and then accepted or vice-versa), in virtue of changes that occurred in things during the time of the disputation. The second source of inconsistency for this set is the behavior of conjunctions and disjunctions explained above.[323]

But again, the bookkeeping metaphor implies that this corollary is not as awkward as it seems. Since the set of propositions that follow from a proposition is always consistent, the column for pertinent propositions will always be consistent, – for each contradictory pair of propositions (A, $\neg A$), a given proposition B implies either one of them ($B \rightarrow A$), or the other ($B \rightarrow \neg A$), or none, but never both contradictory

---

[323] Solution: the conclusion to the effect that three repugnant [propositions] must be accepted, and four and so forth, must be accepted. And similarly, a pair of contradictories must be accepted in virtue of a change in things, for example if 'You are in Rome' is posited, and afterwards 'You are seated' is proposed, it must be accepted. If afterwards, during the time of the obligation you stand up, and 'You are not seated' is proposed to you, it must be accepted. And thus contradictory [propositions] must be accepted during the time of the obligation. And this is true as long as at least no contradictory repugnant to the *positum* is accepted during the time of the obligation.

Solutio: concedenda est conclusio quod tria repugnantia sunt concedenda et quattuor et sic deinceps. Et duo contradictoria similiter sunt concedenda propter mutationem rei ut si illa 'Tu es Romae' poneretur, et postea proponatur 'Tu sedes' illa est concedenda. Si postea durante tempore obligationis tu stares, et tibi proponeretur illa 'Tu non sedes', illa foret concedenda. Et sic contradictoria infra tempus obligationis forent concedenda. Et hoc est verum dum tamen nullum contradictorium repugnans posito concedatur infra tempus obligationis. (§101) propositions. Therefore, it will never be the case that a *positum* forces the granting of a proposition A and of its contradictory $\neg A$.

In the case of *posita* that are pragmatically paradoxical (such as ‘Nothing is posited to you’), and therefore can be denied if they are re-proposed, they are to be treated as impertinent propositions when re-proposed, and therefore would be written down in the impertinent column. Hence, such cases do not introduce inconsistency in the pertinent column. By contrast, the column for impertinent propositions can very well be inconsistent, in the case of impertinent propositions whose truth-value changes during the disputation and which are in fact proposed twice (and receive different responses).

These considerations indicate thus that Swyneshed has no interest whatsoever in the set formed by all granted and denied propositions during a disputation, and that he is perfectly willing to accept its inconsistency. For Burley, on the contrary, the ultimate goal of the *obligationes* game is to keep this very set consistent. So the differences between the two versions of the game do not only regard the rules governing them, but seemingly the very motivations for playing the game.

3.4.3 What is Respondent’s task then?

Clearly, in Swyneshed’s version of *obligationes*, since the set of propositions formed by the performance of a disputation is most likely **not** consistent if the game is played by a correct application of the rules, Respondent’s main task is clearly **not** that of consistency maintenance. Moreover, the hypothesis that *obligationes* are a logic of counterfactuals is, as we have seen, even less felicitous in the case of Swyneshed’s *obligationes* than in the case of Burley’s.

An analysis of Swyneshed’s rules for *obligationes* shows that the most crucial aspect of a winning strategy for Respondent is the correct definition of the two sets of relevant and irrelevant propositions with respect to the *positum*; if this is done correctly, and in particular if Respondent identifies accurately which propositions follow from and which are repugnant to the *positum*, then the rest of the game poses no further challenge, since Respondent is simply expected to answer to irrelevant propositions as he would if he were not playing the obligational game at that moment.

Thus, Swyneshed’s version of *obligationes* seems to be primarily a test for the **recognition of inferential relations** between propositions – namely, the *positum* and each proposed proposition. In Burley’s version, inference recognition also plays a role, as the failure of spotting a relation of inference can yield inconsistency. But in Burley’s game, especially under the second, non-deterministic interpretation, there seem to be other strategic elements involved in consistency management, for example, the possible denial of a proposition with notorious ‘tricky’ consequences. In Swyneshed’s game, no strategy seems to be necessary as long as Respondent possesses the skills of inference recognition; he cannot play the game ‘better’ or ‘worse’ in virtue of the presence or absence of strategic decisions.

In sum, inference recognition is related to consistency maintenance, but in fact these seem to be two different tasks.[324] True enough, one method for performing consistency maintenance would consist of applying inference recognition between the set of propositions which have hitherto received a valuation in the model being built and the proposition being proposed; if there is an inferential relation between this set of propositions and the proposition newly proposed (or its contradictory), then adding this proposition (if it follows from) or its contradictory (if it is repugnant to) to the set of propositions being formed will obviously guarantee consistency (this fact follows naturally from the model-theoretic definition of implication, cf. (Keffer 2001, 149) and section 3.3.3.3 of this text). If there is no such inferential relation, and if this fact is correctly identified by the player, then, with respect to consistency, the player may add either the proposition proposed or its contradictory to the set of propositions being formed, and in either case consistency will be maintained. That is why, in Burley's version of the game, inference recognition plays a crucial role in the task of performing consistency maintenance (it takes priority over the 'semantic' rules concerning common knowledge), but it is not the only element that the player may and should take into account to perform this task. Moreover, in Swyneshed's game, only inferences with single premises (the *positum*) are at stake, and therefore the latter is clearly easier to play than Burley's game, where inferences with multiple premises are at stake.

E.J. Ashworth suggests that the difference between the *antiqua responsio* and the *nova responsio* rests on two different conceptions of inference: the *antiqua responsio* adopted a 'semantic' notion of inference, based on the idea of consistency, while the *nova responsio* adopted a 'syntactical' notion of inference, based on logical relations between propositions (cf. Ashworth 1996, 359). She gives compelling textual evidence to substantiate her claim. As said, in the *antiqua responsio*, inference recognition does play a central role, but if in this framework inference is understood as essentially a semantic notion, as argued by E.J. Ashworth, then there is no doubt that *obligationes* played according to the *antiqua responsio* are essentially a semantic game. By contrast, the *nova responsio* seems to define an essentially syntactical game (although an exclusively syntactical, comprehensive and effective definition of inference and consequence is arguably not to be found in any medieval treatise). If this is indeed the case, then it is even more evident that the two games, Burley's and Swyneshed's, involve the performance of different cognitive tasks.

Again, Burley's game is harder to play than Swyneshed's. In Burley's game, what is tested is the ability to recognize inferential relations between propositions and **sets** of propositions, and that would be done on the spot (during the process). The role of irrelevant propositions is to make it a dynamic game, whereas Swyneshed's game is static.

---

[324] Recent research in neurosciences seems to indicate that, in effect, these two tasks correspond to different brain activities, in particular that different areas of the brain are activated when each of the tasks is performed. Cf. Parsons and Oserson, 'New Evidence for Distinct Right and Left Brain Systems for Deductive versus Probabilistic Reasoning' (*Cerebral Cortex* 2001). Probabilistic reasoning is not exactly what is at stake in Burley's *obligationes*, so the parallel should not be taken too far, but it does seem to add an interesting element to the argumentation. I owe this reference to Prof. Wilfrid Hodges.

3.4.4 Conclusion

With the changes he introduced to the obligational rules, Swyneshed created a game that is not necessarily more problematic that Burley's (the so-called 'inconsistencies' are not real inconsistencies), but simply one that is perhaps more tedious and less effective for testing Respondent's abilities. His motivations for introducing the changes seemed to be unrelated to the game aspect of *obligationes*; rather, they concerned some 'odd' properties of Burley's game – especially the fact that any false proposition could be 'proved' from a false *positum*, and that two disputations with the same *positum* did not necessarily force Respondent to give the same replies to the same *proposita*.

However, if *obligationes* were indeed meant to test and train logical abilities, as is very likely they were, it is no wonder that Burley's version remained dominant, since it is more effective for these purposes. Of course, *obligationes* had also become a logical genre in themselves, whose framework gave rise to logical investigations that went beyond the scope of schoolboy's exercise. In this sense, Swyneshed's *nova responsio* certainly offered new elements for further investigation. In particular, authors who, after him, wrote *obligationes* treatises basically subscribing to Burley's *antiqua responsio* had a more fine-grained perception of the properties of the game than Burley had had, given that the challenge posed by the *nova responsio* prompted further reflection. Indeed, in the next section we shall see how Strode reacts to defend the *antiqua responsio* from Swyneshed's attacks.

The relation between *obligationes* and *sophismata* remains to be considered.[325] In section 3.3.3 I have argued that one of the reasons why the *obligationes* game is hard to play, even though there is always a winning strategy for Respondent, is that Opponent makes use of the intricacies of the language to set up traps for Respondent, involving for example self-reference, ambiguity etc. – that is, typically the phenomena of language that give rise to sophisms and fallacies. Indeed, there is a vast range of studies dedicated to the relation between *sophismata* and *obligationes*. However, there is no unanimous agreement as to the very nature of this relation. I have briefly mentioned in section 3.2.1 that it seems reasonable to suppose that *obligationes* were not primarily meant to solve *sophismata*, but rather that *sophismata* were typically ways of testing whether a given obligational system of rules was sound. One wonders whether the actual performance of obligational disputations (assuming that they did take place) also involved *sophismata* to such a large extent, or whether this was the case only in the treatises, given their theoretical character. *Sophismata* would then have a kind of meta-role with respect to *obligationes*.

In any case, it seems to me that the framework adopted here is not adequate to treat *sophismata* within obligational contexts; these *sophismata* are essentially semantic puzzles, related to the specific meaning of the terms involved, and therefore are (in general) resistant to symbolic formalization; here, I have used schematic letters as placeholders for propositions, and thus their specific meanings were disregarded. In

325 This has been done thoroughly by (Keffer 2001, chap. 5).

fact, it seems that the different *sophismata* can only be treated one by one, whereas the present analysis aimed at giving a general account of the rules of *obligationes*.

3.5 STRODE'S *OBLIGATIONES*: THE RETURN OF CONSISTENCY MAINTENANCE

We have seen how Swyneshed intended his *obligationes* theory to mend the 'flaws' of Burley's theory. However, Burley's set of rules for obligational disputations appears to have resisted Swyneshed's attacks and to have remained more influential. It has been argued in the secondary literature that this is due, to a great extent, to the fact that Burley-style *obligationes* are indeed more efficient for the putative purpose of testing the logical abilities of Respondent. Moreover, the rules introduced by Swyneshed seemed to generate even more serious oddities in the obligational framework than the alleged oddities generated by Burley's rules. So a few decades after Swyneshed composed his treatise, a couple of treatises on *obligationes* were written which essentially subscribed to the *antiqua responsio*, and whose authors openly criticized the *nova responsio*. However, they were not a mere repetition of Burley's treatise; the challenge posed by the *nova responsio* had prompted a more fine-grained understanding of many of the logical properties of this form of disputation.

Therefore, the study of these later texts can greatly contribute to the understanding of the logical structures underlying the rules of *obligationes*, as well as to the understanding of each of the two most influential styles of obligational disputation. In the secondary literature, more and more work is being done on these later treatises, but the main focus of research has been, so far, still on Burley's and Swyneshed's treatises.

In what follows, I analyze Ralph Strode's treatise on obligations. I have used a hitherto unpublished edition of the text (based on 14 manuscripts) made by Prof. E.J. Ashworth; and I am enormously grateful to her for having shared it with me.[326] The importance of Strode (active in the second half of the 14th century, in England) as a logician has been increasingly recognized over the last years, but the publication of most of his writings is still to take place.[327]

The only other study dedicated to Strode's treatise is (not surprisingly) (Ashworth 1993); her main topic in that study is the analysis of the notions of consistency and inconsistency that emerge from his treatise. In the present study, I also deal lengthily with these notions, as they are indeed crucial in the treatise, but I attempt to complement her analysis by outlining some of its other interesting aspects as well.

In the first part of this section, I give a brief description of Strode's text, which is all the more necessary given the fact that it is not (yet) available to the average reader; I also offer a reconstruction of the rules proposed by Strode, following the

---

[326] It is a 43-page typed text. I will refer to this edition as *Strode Obl.*, and will refer to the pages of the text in its current, unpublished form.

[327] To my knowledge, a major project for publishing Strode's texts is now being coordinated by Prof. A. Maierú; but, as always with such ambitious projects, its completion is a matter of many years.

style of reconstruction used in my analysis of Burley's and Swyneshed's rules – that is, essentially based on the idea that *obligationes* can be viewed as logical games.[328] In the second part, I address Strode's explicit arguments *contra* Swyneshed. In the third part, I discuss Strode's epistemic and pragmatic approach to *obligationes*. Indeed, he participated in the general 'epistemic turn' in logic that took place in the second half of the 14th century in England (cf. Normore 1993, Boh 2001), and this is clearly seen in his treatise on *obligationes*, as I attempt to show.

3.5.1 The essentials of Strode's treatise

3.5.1.1 Description of the text

Strode's text has a very clear structure. Chapter I is the *prefatio*. In the introduction (chap. II), Strode puts forward four remarks, five suppositions and four conclusions. Chapter III questions three opinions concerning *obligationes*. Chapters IV–VIII discuss objections to each of the five suppositions (one supposition per chapter); chapters IX–XII discuss objections to each of the four conclusions. Finally, chapters XIII–XVI deal with forms of *obligationes* other than *positio*.

In sum, Strode's rules of *obligationes* are in fact all presented in the introduction, which is thus by far the most important chapter. Chapters IV–XII argue for the soundness of the rules proposed in the typical medieval way of examining objections and *sophismata* that could threaten these rules, but which are dealt with and explained away.

As in the analysis of Burley's and Swyneshed's theories of *obligationes*, I only deal with *positio*; therefore, the four final chapters of Strode's text are not looked into in the present discussion, and the reconstruction below only concerns *positio*.

3.5.1.2 Remarks, suppositions and conclusions

As just said, the rules governing Strode-style *obligationes* are all presented in the introduction of his treatise, in the form of four remarks, five suppositions and four conclusions, which are described below.

Remarks

1. Therefore, first should be described certain terms with a view to our subject.[329]

The first remark only gives a general description of the obligational genre, in particular of the role of the terms – '*pono*', '*depono*' etc. – defining the different forms of *obligationes*. Some of the procedural rules of the game are also introduced in this passage, and they are identical to Burley's and Swyneshed's procedural rules.

---

[328] Concerning Strode's treatise, the logical game metaphor seems less helpful than with the two other treatises analyzed here, so I will not insist on it as much as I have done before. I will though occasionally resort to the metaphor, if it appears to be illuminating for a particular point.

[329] Primo ergo describendi sunt quidam termini ad propositum. (*Strode Obl.* 2) The translation of Strodes passages are my own in collaboration with E.P. Bos.

2. Secondly, it must be remarked that the time of the obligation lasts from the moment that the casus is admitted until the opponent states: ‘let the time of the obligation be over’, or something like it, or until he passes over to dispute about another matter, or until he totally stops disputing.[330]

This remark introduces an important aspect of the procedural rules of the game, namely when it begins and when it ends. Notice that Strode uses the term ‘*casus*’, and this deserves a commentary. In Burley’s treatise, *casus* and *positum* were distinct notions. The *casus* defined the hypothetical situation that was to be considered as true for the sake of the disputation; often, the *casus* to be accepted was simply things as they really were (‘*sit rei veritas*’), but occasionally a *casus* diverging from the actual state of things was to be accepted.[331] The *positum*, however, was not to be taken as true, and Respondent should only be interested in the possible inferential relations between the *positum* and the subsequently proposed propositions.

In later authors, the two notions were often interchangeably used, and the *casus* was often simply the proposition prompting the disputation. Yrjönsuuri conjectures that these authors may have followed Ockham’s suggestion that following the *casus* (in contexts other than obligations, for instance with respect to *sophismata*) amounts to the same as following the *positum*, methodologically speaking (cf. Yrjönsuuri 1993, 310). Indeed, in Strode’s case, it seems at times that he distinguishes the two notions, but more frequently they appear to be taken as equivalent, for example, when he says

But what remains, after the sign of position has been put, is called the *casus* or the *positum*, for instance ‘you are running’ remains in that position, viz. I put to you ‘you are running’.[332]

In practice, Strode appears to use the term ‘*casus*’ for the very situation of positing the *positum*, including thus the content of the *positum* as well as the fact that it has been posited – but the *positum* itself does not have to be considered as true.

3. Thirdly, it must be noted that a proposition is said to be pertinent when it follows from or is repugnant to the thing or things that has or have been conceded. Now, a proposition is said to be impertinent when it neither follows nor is repugnant [to what has been conceded].[333]

This remark is crucial. As widely acknowledged in the secondary literature, and as discussed in (Dutilh Novaes 2006a), the core of Swyneshed’s revision of the *antiqua responsio* was his redefinition of pertinent/impertinent propositions. In the *antiqua responsio* a pertinent proposition is one that follows from (*sequens*) or is repugnant to (its contradictory follows from – *repugnans*) the *positum* and/or the granted propositions and the contradictories of the denied propositions. According to

[330] Secundo notandum quod tempus obligationis durat ab instanti quo casus admittitur donec dicat opponens : cedat tempus obligationis, uel aliquod tale, uel se transferat ad disputandum in alia materia, uel penitus dimittat disputare. (*Strode Obl.*, 2)
[331] Cf. (Yrjönsuuri 1993, 304).
[332] Sed quod remanet deposito signo positionis dicitur casus et positum, ut ista: ‘tu curris’ in ista positione, scilicet: pono tibi istam: ‘tu curris’. (*Strode Obl.*, 2)
[333] Tertio notandum quod propositio pertinens dicitur que sequitur uel repugnat concesso uel concessis. Sed impertinens dicitur que nec sequitur nec repugnat. (*Strode Obl.*, 2)

Swyneshed, a pertinent proposition is one that follows from or is repugnant to (its contradictory follows from) the *positum* only; the subsequently proposed propositions are not taken into account. As a consequence, I have argued, Swyneshed excludes the dynamic aspect of the *antiqua responsio*: in the *nova responsio*, the order in which propositions are proposed does not matter, and Respondent does not have to take his previous moves into account to make a good move – in fact he may as well 'forget' the irrelevant propositions proposed as much as his responses to them (cf. Yrjönsuuri 1993, 317).

Strode returns to Burley's definition of pertinent/impertinent propositions, and this fact alone means that the dynamic character is again an important element of Strode-style *obligationes*.

4. Fourthly, it must be noted that something [a *propositum*] that should be denied as regards our subject is called that which deserves to be denied; a *propositum* that must be conceded is one that deserves to be conceded, and a *propositum* that must be doubted is one that deserves to be doubted. Now, when a proposition must be conceded, denied or doubted will be made clear in the coming suppositions and rules.[334]

This remark may seem rather redundant and uninformative, but it does stress the normative character of the obligational rules. The rules properly speaking are presented subsequently.

Suppositions

1. First: that every [proposition that is] possible, that is known to be possible, and non-repugnant with respect to some *positum*, or what has been put forward and admitted, must be admitted by you, but only when it is put forward. And on account of such a possible [proposition], which is posited and admitted, something impossible is not to be conceded. You must indeed deny an impossible proposition whenever it is put forward as a *positum* or as a *propositum*.[335]

The rule for the admittance of a *positum* is that it be possible. Moreover, an impossible proposition should never be accepted as the *positum*, and since it never follows from a possible proposition, if proposed as a *propositum*, an impossible proposition is never pertinent to the *positum*. Therefore, an impossible proposition should always be denied.

2. Second supposition: that everything known to be a *positum* and correctly admitted during the time of the position, which is proposited under its appropriate form, must be conceded, and also whatever follows from it; and whatever is repugnant to it must be denied.[336]

---

[334] Quarto notandum quod negandum in proposito dicitur quod est dignum negari, et concedendum quod est dignum concedi, et dubitandum quod est dignum dubitare. Sed quando propositio debet concedi, negari, uel dubitari patebit in suppositionibus et regulis inferius ponendis. (*Strode Obl.*, 3)

[335] Primo: quod omne possibile scitum esse possibile non repugnans alicui posito uel admisso cum ponitur est a te admittendum, et solum tale cum ponitur. Et propter tale possibile positum et admissum non sit aliquod impossibile concedendum. Debet enim negari propositio impossibilis quandocumque proponitur uel ponitur. (*Strode Obl.*, 3)

[336] Seconda suppositio: quod omne scitum esse positum et bene admissum in tempore positionis, sub debita forma sua propositum, est concedendum, et quodlibet sequens ex illo; et quodlibet repugnans illi est negandum. (*Strode Obl.*, 3)

That is, what follows from or is repugnant to every proposition proposed and rightly granted during the disputation must be granted/denied, and not only what follows from or is repugnant to the *positum*.[337] Notice also the '*sub debita forma*' clause, which relates to a proposition first put forward as *positum* and then again as *propositum*: it should be granted provided that it is in the right form. This clause is meant to avoid pragmatic inconsistencies of the kind that will be discussed below.

3. The third supposition is that to every impertinent proposition one must answer according to its quality, that is if the proposition is known to be true it must be conceded, if it is known to be false it must be denied, and if it is doubtful it must be doubted.[338]

This is the general rule for impertinent propositions, present in Burley's treatise as well as in Swyneshed's. In itself, it is an incomplete rule, since it depends on the proper definition of an impertinent proposition (remark 3).

4. The fourth supposition is that out of the time of the obligation the truth of things must be disclosed, because, when someone concedes something false or denies something true when not obliged, he responds badly.[339]

The intention of this remark seems to be twofold: on the one hand, it may indicate the fact that matters of truth, including the evaluation of Respondent's performance during the disputation, were to be discussed only once the disputation was over (Ashworth 1993, 366); on the other hand it stresses the fact that, during an obligational disputation, it is not truth that is at stake, as in other forms of disputation, but rather logical notions such as following from, being repugnant to and being consistent with. Partisans of the *nova responsio* had objected that, in the *antiqua responsio*, any falsehood could be proved; here, Strode seems to be defending the view that, given the very nature of *obligationes*, this is not at all problematic.

5. The fifth supposition is that all responses within the time of the obligation must be bent back to the same instant, that is, responses should be given continuously with respect to the instant when the casus is posited.[340]

This is a modification – an improvement? – with respect to Burley's original theory. Burley does say that all responses must be reduced to one instant, namely one in which the *positum* can be true, but he does not judge it is necessary to determine which specific instant that would be (cf. Yrjönsuuri 1993, 308). Perhaps this indetermination gave

---

[337] Further on Strode states this rule in a more general form (in his discussion of the first conclusion): 'Super quam regulam fundatur talis regula: omne sequens ex posito et bene admisso cum bene concesso uel consessis uel opposito bene negati uel opposites bene negatorum cum proponitur est concedendum, et si quid talibus repugnans, illud est negandum'. (*Strode Obl.*, 3)

[338] Tertia suppositio est quod ad omnem propositionem impertinentem respondendum est secundum sui qualitatem, i. si sit scita esse uera est concedenda, si scita sit esse falsa est neganda, si dubia dubitanda. (*Strode Obl.*, 3)

[339] Quarta suppositio est quod extra tempus obligationis rei ueritas est fatenda, quia cum quis concedit falsum uel negat uerum non obligatus male respondet. (*Strode Obl.*, 3)

[340] Quinta suppositio est quod omnes responsiones infra tempus obligationis retorquende sunt ad idem instans, i. responsiones sunt dande continue pro eodem instanti quo casus est positus. (*Strode Obl.*, 3)

way to Swyneshed's introduction of several instants according to which impertinent propositions should be judged (namely, the instants in which each is respectively proposed). But this stipulation gave rise to even more inconsistency within Swyneshed-style *obligationes* (see section 3.5.2.2 below), so Strode felt compelled to determine explicitly which instant was to be taken into account for the response to impertinent propositions – namely the instant in which the disputation begins.

Conclusions

1. First conclusion: that this does not follow: if you concede something false which is known to you to be false, or if you deny something true which is known to you to be true, or if you concede something doubtful which is known to you to be doubtful, then you respond badly.[341]

This conclusion is related to the fourth supposition (the truth is not what is at stake during an obligational disputation), and it follows from the first and second suppositions. Strode remarks that the *casus*/*positum* is usually false, and that the response to *proposita* should be according to the second supposition, and thus not according to their truth-value in the case of pertinent propositions.

2. Second conclusion: that it is not unsuitable that a Respondent sometimes concedes that he responds badly, or denies that he responds well.[342]

This conclusion concerns the pragmatic character of Strode's approach to *obligationes*, which is discussed in detail in 3.5.3 below. What we could call 'performative contradictions' are not considered to be real contradictions by Strode; for him, just as much as a contingent false proposition may be admitted at some point of the disputation, 'you are responding badly' is a contingent (hopefully false!) proposition, thus it can be granted without generating inconsistency. A wholly different situation is when, outside of the time of the disputation, one grants that he responded badly during the disputation, since, according to the fourth supposition, outside of the time of the disputation, the truth must prevail.

3. Third conclusion: that when a *positum* that is possible is repugnant to the act of positing, it should be admitted, and when [the same] is proposed it should be conceded, but the act of positing if it is proposed should be denied.[343]

This conclusion also concerns *posita* that generate performative contradictions (see 3.5.3.3 below and Ashworth 1984). If such a *positum* is put forward (e.g., 'Nothing is posited to you'), it should be accepted, and if it is re-proposed as a *propositum* it should be granted (presumably as following from the *positum*, according to the

[341] Prima conclusio: quod non sequitur : tu concedis falsum scitum a te esse falsum, uel negas uerum scitum a te esse uerum, uel concedis dubium scitum a te esse dubium, ergo male respondes. (*Strode Obl.*, 3)
[342] Secunda conclusio: quod non est inconueniens quandoque respondentem concedere se male respondere, uel negare se bene respondere. (*Strode Obl.*, 3)
[343] Tertia conclusio: quod cum positum possibile repugnat positioni, debet admitti, et cum proponitur concedi, sed positio proposita debet negari [. . .].(*Strode Obl.*, 3)

validity of the principle that every proposition implies itself). But if the situation describing the *positio* is put forward as a *propositum*, as in ' "Nothing is posited to you" is posited to you', then it should be denied (although it is true), as it is repugnant to the *positum*.

4. Fourth and last conclusion is that, when responding in an obligation, the order must be especially paid attention to.[344]

This conclusion is a maxim to be born in mind rather than a rule of the disputation properly speaking. By stating it explicitly, Strode simply stresses what I have described as the dynamic character of the obligational disputation according to the *antiqua responsio*; when following its rules, Respondent should always take into account his previous moves. The very same proposition proposed at different times of two disputations otherwise identical is very likely to receive different responses. The example mentioned by Strode is the *positum* 'Every man runs' and the responses to be given to 'You are running' and 'You are a man', depending on their relative order of being proposed, which I have discussed at length in (Dutilh Novaes 2006b).

Notice though that, once a proposition has been accepted or denied, its relative order with respect to the other *proposita* no longer matters.[345]

3.5.1.3 Reconstruction

A Strode-style obligational disputation is defined by the following quadruple:

Definition 3.5.1.3.1: the obligational game (Strode)

$$\boldsymbol{Ob} = <\mathbf{S_0}, \mathbf{\Phi}, \mathbf{\Gamma}, \mathbf{R}\ (\varphi_n)>$$

$\mathbf{S_0}$ is the state of common knowledge of those present at the disputation **at the moment that the *positum* is posited** (supposition 5). It is an incomplete model, in the sense that some propositions do not receive a truth-value. Concerning S_0, Strode differs radically from Swyneshed. Swyneshed wants impertinent propositions to be judged according to the state of common knowledge of the very moment in which the proposition is proposed; so in Swyneshed's *obligationes* the first element of the quadruple is an ordered set of successive states of common knowledge.

Here, if the *casus* happens to be distinct from the *positum*, then $\mathbf{S_0}$ should follow the *casus* and be complemented by the actual state of common knowledge at the moment in which the disputation begins, as explained above.

As for the other three elements of the quadruple, they are virtually identical to those of the quadruple defining Burley's game. For the reader's convenience, I comment

[344] Quarta conclusio et ultima est talis quod in respondendo per obligatoria sit orde maxime attendendus. (*Strode Obl.*, 3)

[345] The order of the premises is irrelevant for the existence of inferential relations, that is: if A, B $\Rightarrow$ C is valid, then B, A $\Rightarrow$ C is just as valid.

on these notions again: $\mathbf{\Phi}$ is an ordered set of propositions, it is the set of propositions actually put forward by Opponent during an obligation. Each element of $\mathbf{\Phi}$ is denoted by 'φ_n', where *n* is a natural number, denoting the place of φ_n in the ordering. The order corresponds to the order in which the propositions are put forward by Opponent, starting with φ_0 (the *positum*). $\mathbf{\Gamma}$ is an ordered set of sets of propositions, which are formed by Respondent's responses to the various φ_n. How each Γ_n is formed will be explained below. The ordering is such that Γ_n is contained in Γ_{n+1}. $\mathbf{R}(\varphi)$ is a function from propositions to the values 1, 0, and ?. This function corresponds to the rules Respondent must apply to respond to each proposition φ_n. 1 corresponds to his accepting φ_n, 0 to his denying φ_n and ? to his doubting φ_n. The logical rules of the game are defined by $R(\varphi)$.

Positum: Strode holds that any possible *positum* must be accepted, even if it is pragmatically inconsistent with the act of positing, admitting, or more generally participating in the obligational disputation. The nature of this pragmatic inconsistency will be discussed in section 3.5.3.3 below, but for now we must understand what it means for a *positum* to be **possible**.[346]

In my reconstruction of Burley's *obligationes* I used a syntactic criterion to define a possible proposition as a proposition from which absurdity cannot be derived. I have also stressed the importance of the *positum* not being an inconsistent proposition; if this was the case, Respondent would have no chance of keeping consistency, and thus of winning the game.

The same appears to be true of Strode, but a semantic characterization of a possible proposition also seems more than welcome, given the very terms used by Strode in supposition 1. A proposition φ_n is judged to be possible according to S_n, the state of common knowledge at the moment it is proposed; in the case of the *positum* φ_0, this would be S_0. Using the diamond '$\Diamond$' to represent possibility and '$\Vdash$' to represent the relation of 'being true in' between a proposition and a state of knowledge, we have:

Definition 3.5.1.3.2: Rules for the *positum*

$$R(\varphi_0) = 0 \text{ iff } \quad S_0 \not\Vdash \Diamond\varphi_0$$

$$R(\varphi_0) = 1 \text{ iff } \quad S_0 \Vdash \Diamond\varphi_0$$

As it stands, this definition is not very informative, since we still do not know what it means for a proposition to be possible according to a given state of knowledge. Let us thus introduce a relation of accessibility R between different states of knowledge S_n and S_m, corresponding to the notion of conceivability: $S_n \mathrel{R} S_m$ iff S_m is conceivable as a state of knowledge (i.e., does not contain absurdities) in the state of knowledge S_n.

[346] The issue as to which propositions were to be considered 'possible', in view of pragmatic paradoxes, was widely discussed in several obligational treatises. Cf. (Ashworth 1984).

This does not mean that every proposition that holds in S_n also holds in S_m, but only that S_m does not contain propositions that are not conceivably true according to S_n, and it also does not contain contradictions – that is, $S_m \Vdash \varphi_n$ and $S_m \Vdash \neg\varphi_n$, for some proposition φ_n. We can thus define a possible proposition in the following familiar way:

Definition 3.5.1.3.3: Possible proposition

$$S_n \Vdash \Diamond\varphi_n \Longleftrightarrow \text{There is some } S_m \text{ such that } S_n \mathbf{R} S_m \text{ and } S_m \Vdash \varphi_n$$

The definition thus stated also takes care of *posita* that are only performative contradictions, and which thus should be accepted according to Strode: 'you are asleep' is not true in S_0, since Respondent is participating in the disputation, but a state of knowledge in which 'you are asleep' is true is easily conceivable in S_0. Similarly, propositions that are falsified by their own existence, for example 'No proposition is negative' (which were dubbed impossibly-true), also come out possible according to this criterion – 'No proposition is negative' describes a state of affairs conceivable in S_0, namely the state of affairs where only affirmative propositions are formed, and thus this very proposition is not formed. By contrast, a logical contradiction such as 'You are in Rome and you are not in Rome' is not conceivably true in any of the states of knowledge conceivable in S_0.

The rules for responding to *proposita* are virtually identical to Burley's rules, including the all-important definition of pertinent/impertinent propositions; the only difference is that S_0 is referred to in the responses to impertinent propositions, and not some undetermined instant, as in Burley's obligations.

Definition 3.5.1.3.4: Rules for *proposita*

$$R(\varphi_n) = 1 \text{ iff } \begin{cases} -\ \Gamma_{n-1} \vdash \varphi_n, \text{ or} \\ -\ \Gamma_{n-1} \nvdash \varphi_n,\ \Gamma_{n-1} \nvdash \neg\varphi_n \text{ and } S_0 \Vdash \varphi_n \end{cases}$$

$$R(\varphi_n) = 0 \text{ iff } \begin{cases} -\ \Gamma_{n-1} \vdash \neg\varphi_n, \text{ or} \\ -\ \Gamma_{n-1} \nvdash \varphi_n,\ \Gamma_{n-1} \nvdash \neg\varphi_n \text{ and } S_0 \Vdash \neg\varphi_n \end{cases}$$

$$R(\varphi_n) = ? \text{ iff } -\ \Gamma_{n-1} \nvdash \varphi_n, \Gamma_{n-1} \nvdash \neg\varphi_n, S_0 \nVdash \varphi_n, S_0 \nVdash \neg\varphi_n$$

I here present the rules of the game viewed as a determined game, that is, where there is only one correct move at each stage. However, it appears that Strode's treatise offers even more compelling evidence than Burley's treatise to a non-deterministic interpretation of *obligationes* (cf. section 3.3.1.2.2 above), corresponding to some degree of freedom for strategic playing concerning impertinent propositions.

Formation of Γ_n. The different sets of propositions accepted by Respondent (i.e., the propositions to which he has committed himself in the game) are formed in the following way:

Definition 3.5.1.3.5: Formation of Γ_n

$$\text{If } R(\varphi_n) = 1, \text{ then } \Gamma_n = \Gamma_{n-1} \cup \{\varphi_n\}$$
$$\text{If } R(\varphi_n) = 0, \text{ then } \Gamma_n = \Gamma_{n-1} \cup \{\neg\varphi_n\}$$
$$\text{If } R(\varphi_n) = ?, \text{ then } \Gamma_n = \Gamma_{n-1}$$

In particular, if $R(\varphi_0) = 1$, then $\Gamma_0 = \{\varphi_0\}$. If $R(\varphi_0) = 0$, then the disputation does not begin.

3.5.2 *Contra* Swyneshed: consistency maintenance re-established

Throughout the treatise, it is clear that Strode is reacting to Swyneshed, as much as Swyneshed had reacted to the *antiqua responsio* in general and to Burley in particular. Swyneshed is even explicitly named in chapter III, where Strode discusses and rejects three specific opinions. Strode also appears to be reacting to other theories of *obligationes* that questioned the principles of the *antiqua responsio*, such as Kilvington's and that proposed in an anonymous treatise known as the 'Merton treatise'[347] (cf. Ashworth 1993, 375). But his main enemy really seems to be Swyneshed, so in this section I discuss some of Strode's explicit objections against Swyneshed.

3.5.2.1 Swyneshed spotted the wrong problems

The first thing to notice about Ralph Strode's reaction to the theories of his predecessors is that he did not agree with their diagnosis of the problems at issue. (Ashworth 1993, 379)

The two main problems attributed to the *antiqua responsio* by authors such as Swyneshed were the fact that any randomly-chosen falsehood could be proved – for example, if the *positum* is a disjunction of two contingent falsehoods and the first *propositum* is one of the disjuncts, this *propositum* should be denied, and subsequently the other *propositum* must be accepted when proposed – and the fact that, in two disputations having the same *positum*, the same proposition could be accepted in one and denied in the other.

That the first difficulty is not a real problem for Strode can be seen from his fourth supposition, where he stresses that the truth of things is to be made ('*rei veritas est fatenda*') outside of the time of the obligational disputation; therefore, granting a

[347] (Kretzmann and Stump 1985).

falsehood, any falsehood, which follows from the *positum* and the previously granted propositions, or the contradictories of previously denied propositions, is simply a consequence of the rules of this form of disputation.[348] Even though this so-called randomly-chosen falsehood seems at first irrelevant, since it is impertinent to the *positum* (even though it is one of the disjuncts, from a disjunction the individual disjuncts do not follow), it becomes pertinent as the disputation continues.

As for the second difficulty, again Strode does not see what the problem is with a proposition receiving different responses in two disputations with the same *positum*. The response to a proposition is essentially dependent on the relative order within the disputation in which the proposition is proposed, as Strode points out several times (e.g., in the fourth conclusion – see also chapter XII, where possible objections to this conclusion are dealt with). We also encounter here traces of the 'pragmatic' approach to obligational disputations proper to Strode; the response to a given proposition must depend on the actual course of things in a given disputation, and cannot be determined abstractly from the *positum* only.

3.5.2.2 An even worse form of inconsistency?

Strode's main accusation against the *nova responsio* is that the application of its rules yields the worst form of inconsistency within the obligational framework (far worse than the two 'inconsistencies' just discussed), namely that in some occasions Respondent may be forced by the rules of the *nova responsio* to grant the contradictory of the *positum*/*casus*. Strode's arguments are presented in chapter III ('*Contra tres opiniones*'), and are discussed in (Ashworth 1993, 381–383). Ashworth concludes that Strode's argument as she reconstructs it is not sound and that it could be blocked by Swyneshed.

Strode's arguments goes as follows:

φ_0: 'Every man is running' – *positum*, accepted
φ_1: 'You are a man' – irrelevant and true, accepted
φ_2: 'Every man is running and you are a man' – irrelevant and false, denied
φ_3: 'Not every man is running or you are not a man' – irrelevant and true (the contradictory of a correctly denied proposition, 3), accepted
φ_4: 'Not every man is running or you are not a man, but you are a man, therefore not every man is running' – *bona consequentia*, accepted
φ_5: 'Not every man is running' – must be accepted, but is the opposite of the *positum*

According to Strode, φ_5 must be accepted by *modus ponens* on the basis of φ_1, φ_3 and φ_4, and therefore the contradictory of the *positum* must be accepted. Ashworth claims

[348] Et cum proponitur, debite conceditur per secundam [regulam], que fundatur super admissionis officium, ut prius dictum est. Unde patet quod aliquando concedens falsum bene respondet secundum regulas in ista arte positas, et non male. (*Strode Obl.*, 3).

(cf. Ashworth 1993, 383) that the problem with Strode's argument is that Swyneshed would not be obliged to grant the conjunction of φ_1 and φ_3 as impertinent and true (impertinent because both conjuncts are impertinent); rather, he would be entitled to deny this conjunction as pertinent and repugnant to the *positum*.

As I see it, the problem with Strode's arguments does not concern the status of a conjunction formed by two irrelevant (accepted) propositions; rather, what Strode seems to be claiming is that, once one has granted φ_1, φ_3 and φ_4 (regardless of the status of the conjunction of φ_1 and φ_3, or whether it has actually been proposed as a conjunction), one must grant the consequent of the *consequentia* in φ_4, simply because one has granted its antecedent (by *modus ponens*). The fault of the argument is to assume that, in the *nova responsio*, what has been granted as impertinent and true must also function as premise of an inference so that its conclusions can be inferred. But as noted before, impertinent propositions that are granted as true are simply 'forgotten' if one plays the game according to the *nova responsio*, they seem to have no 'assertive force' whatsoever; hence, they cannot be used as premises in an application of *modus ponens*, since Respondent is not committed to their truth. They would of course have this 'assertive force' in a game played according to the *antiqua responsio*, and that is perhaps the source of Strode's miscast argument.

In this case, it is true that Respondent does grant a full-blown inconsistent set of propositions ('Every man is running', 'You are a man' and 'You are not running'), since it is granted that every man is running as the *positum*, but of all men who are not running, 'you are a man' and 'you are not running' must be accepted as true and impertinent.[349] However, contrary to what Strode claims, Respondent will still not be forced to grant the contradictory of the *positum* ('Not every man is running'), since he is not committed to the truth of 'You are a man' and 'You are not running', even though he granted them. Moreover, as already noticed, Swyneshed is happy to accept that inconsistent sets of propositions are granted. But Strode's attribution of this form of inconsistency (i.e., granting the contradictory of the *positum*) to Swyneshed's theory is unfounded, since Respondent is not obliged to accept the contradictory of the *positum* on the basis of granted but irrelevant propositions.

If the *positum* is an existential proposition, there can arise a form of what is now called an ω-inconsistency with respect to the *positum*, based on the concept of ω-consistency introduced by Gödel in his famous incompleteness theorem (cf. Gödel 1986). A theory is ω-inconsistent if both of the following hold in it, for some predicate 'F':

There is an x such that $F(x)$, but

$\neg F(x_1), \neg F(x_2), \neg F(x_3)$ etc., for all x_n.

[349] Provided that they are not proposed simultaneously, as a conjunction, in which case this conjunction would be repugnant to the *positum* and thus should be denied. So the actual logical form of this argument is:

$\forall x\, (F(x) \Rightarrow G(x))$ is granted, but for some x_n such that $F(x_n)$ is granted, $G(x_n)$ is denied.

That is, if the *positum* is an existential proposition that is accepted because it is possible but in fact, at that moment, no individual satisfies the predicate expressed in it (e.g., 'There is an antichrist'), then of each individual ('He is the antichrist') it will be denied that (s)he satisfies the predicate, generating ω-inconsistency with respect to the *positum* properly speaking. Since all propositions of the form $F(x_n)$ that are denied as irrelevant and false do not keep any kind of assertoric force, even if every single individual of the (finite) salient domain is covered, at the last individual Respondent is still not obliged, on account of the *positum*, to grant that it has the property F. By contrast, in the *antiqua responsio*, if Respondent denies the property F to all individuals but the last individual x_n, in the last case he must grant $F(x_n)$, since from $\exists x\ F(x)$ and $\neg F(x_1), \ldots, \neg F(x_{n-1})$ (all of which propositions to whose truth Respondent is committed), $F(x_n)$ follows. In the *nova responsio*, however, Respondent is not committed to the truth of $\neg F(x_1), \ldots, \neg F(x_{n-1})$, even though he granted them (or, equivalently, denied their affirmative forms), so he may also grant an ω-inconsistent set of propositions.

These facts are also related to the notoriously awkward behavior of disjunctions and conjunctions within the *nova responsio*, insofar as a universal proposition can be seen as equivalent to an infinite conjunction, whereas an existential proposition ('particular proposition', in medieval terms) can be seen as equivalent to an infinite disjunction. Now, it is well known that for Swyneshed a disjunction can be accepted as *positum* while every disjunct must be denied as false and irrelevant; but here the situation seems even more awkward, since even if the *positum* is a conjunction (or, equivalently, a universal proposition), some parts of the conjuncts must be denied as false and irrelevant. Notice, however, that, given the *positum* 'Every man is running', if 'You are a man' and 'You are running' are proposed simultaneously, as a conjunction, then the conjunction must be granted, since it follows from the *positum*. What need not be granted is 'You are running' after 'You are a man' having been granted (as true and irrelevant), but this is not surprising, as it only confirms the non-dynamic nature of the *nova responsio* game.

In sum, a Respondent following the *nova respontio* may form ω-inconsistent (if the *positum* is an existential proposition) or outright inconsistent (if the *positum* is a universal proposition) sets of granted propositions (something that Swyneshed gladly accepts), but Strode's argument fails to show that Respondent is also obliged to grant the contradictory of the *positum* on the basis of the *nova respontio* rules.

3.5.2.3 The core of the matter: definition of pertinent/impertinent propositions

As already noted, the core of Strode's strategy to 'correct' the *nova responsio* and exclude the inconsistencies it generated in obligational disputations is to return to Burley's definition of pertinent/impertinent propositions – just as much as the core of Sywneshed's strategy to 'correct' the *antiqua responsio* had been to reformulate this distinction and let it depend only on the *positum*.

In section 3.3.3.1, I have proved that, if one follows the rules proposed by Burley (and thus by Strode), in particular with respect to the definition of a pertinent/impertinent proposition, then Respondent can always maintain the consistency (and even the ω-consistency[350]) of the set of propositions formed during the disputation by the propositions he grants and the contradictories of the propositions he denies. In other words, there is always a winning strategy for Respondent, assuming that the goal of the game is consistency maintenance.

Hence, the same holds here; the strange features of the *nova responsio* were all related to different forms of inconsistency being produced by the application of its rules. Since the fundamental change introduced by Swyneshed was the new definition of the notion of pertinent/impertinent propositions, all Strode had to do to re-establish consistency was to return to the old distinction, which differentiates a pertinent from an impertinent proposition on the basis of all previously granted/denied propositions in a given disputation (cf. his third remark).

> To this it must be replied that 'following from' [pertinent] in the foregoing is not only understood as following the positio, but as following from the *positum* together with the thing or things that have been correctly conceded, as said in the first conclusion.[351]

3.5.2.4 Avoiding time-related inconsistency

Besides introducing a new notion of pertinent/impertinent propositions, another controversial move by Swyneshed was the stipulation that changes in things during the time of the disputation should be taken into account in Respondent's responses to impertinent propositions. The result of this stipulation was that the set composed of granted/denied impertinent propositions could be inconsistent not only on the level of molecular propositions (conjunctions and disjunctions), but even on the atomic level. If at the beginning of the disputation 'You are sitting' was proposed to Respondent and he was in fact sitting at that moment, then he would have to grant this proposition as impertinent and true. But if soon after 'You are not sitting' was proposed, and indeed in the meantime Respondent had stood up, then he would have to grant this proposition as irrelevant and true, since the first proposition had been 'forgotten' in the meantime as it had not been added to the informational base of the disputation (composed only of the first *positum* and possibly of other *posita* subsequently posed[352]).

Strode, of course, is not willing to accept this form of inconsistency. The mere redefinition of the notion of pertinent propositions is in fact sufficient to exclude this

---

[350] Presumably, in a Burley-style obligation, if a *positum* is an existential, contingently false proposition, and each of its individual instances is gradually proposed and denied as impertinent and false, then the very last individual instance proposed would have to be granted as *sequens*, even though it is false (just as when a disjunction is the *positum*: if all other disjuncts have been proposed and denied, the last disjunct must be granted as *sequens*).

[351] Respondetur igitur quod non solum capitur 'sequens' in locis predictis pro sequente ex posito, sed pro sequente ex posito cum bene concesso uel consessis, ut dicebatur in prima conclusione. (*Strode Obl.*, 5)

[352] Swyneshed accepted the possibility of a second or third *positum* being posed during the disputation, so that in practice the actual *positum* became the conjunction of the two accepted *posita* (cf. section 3.4.1.2.1). By comparison, in the *antiqua responsio* it is as though every granted or denied proposition (pertinent or impertinent, *positum* or *propositum*) became a new *positum*.

form of inconsistency, as logical relations between propositions have priority over the actual state of things. So, in the example above, when 'You are not sitting' is proposed, according to Strode's definition of a pertinent proposition, it is not impertinent (as it is for Swyneshed), but rather pertinent – in fact, repugnant, since it contradicts a previously granted proposition. This is why, in Burley's theory, no specific moment had to be referred to in the responses to impertinent propositions; the notion of pertinent propositions alone was sufficient to exclude this form of inconsistency.

But Strode goes further. He is more aware of this issue than Burley was, having been exposed to the *nova responsio*, so he stipulates that one specific instant, namely the moment when the disputation begins, should be considered for the responses to impertinent propositions (fifth supposition).

One possible objection to this stipulation is discussed by Strode in the chapter dedicated to the fifth supposition (chapter VIII). Say the *positum* is 'you are running or the king is sitting'. Then 'the king is sitting' is proposed; Respondent must doubt it, since at moment S_0 when the disputation began, he had no idea whether the king was sitting or not. Then 'you are running' is proposed; he must deny it, since at S_0 he was not running. Then 'the king is sitting' is proposed again, and must be granted. Now, if all responses refer to the same instant, how can Respondent first doubt whether the king is sitting and then accept it?

The same objection could have been raised against Burley, but since Burley does not explicitly say that all responses refer to the same specific instant, this situation appears to be less awkward within a Burley-style obligational disputation. Strode, however, must explain how it is possible that at the same moment Respondent grants and doubts the same thing. In his discussion, he proposes many ways around this difficulty, but an accurate and straightforward reply would simply be to say that, in its first occurrence, 'the king is sitting' was impertinent, and therefore indeed judged according to S_0, but that in its second occurrence it had become a pertinent proposition, and therefore the response to it simply had no bearing on S_0 whatsoever. Thus, S_0 is not thereby proved to be an inconsistent state of knowledge.

3.5.2.5 Conjunctions and disjunctions

Finally, what was considered by many as the most embarrassing characteristic of the *nova responsio*, namely the behavior of conjunctions and disjunctions, was also discussed by Strode (in particular in his analysis of the first of the three opinions he objects to in chapter III). Obviously, Strode is not happy with the inconsistencies related to conjunctions and disjunctions that appear in a Swyneshed-style obligation, so in chapter III he proposes ways to avoid this phenomenon.

Let us recapitulate. According to Swyneshed, it is entirely possible that both conjuncts are granted when proposed as individual propositions, but that their conjunction is denied when proposed as a conjunction – for example, in the case of the conjunction of a false *positum* and a true impertinent proposition already granted (the conjunction is then a false impertinent proposition, and therefore should be denied). Similarly, it is possible that a disjunction is granted and that subsequently both disjuncts are denied (if they are both false, impertinent propositions) (cf. section 3.4.2.4).

How do conjunctions and disjunctions behave in Burley's obligational framework? Such inconsistencies do not occur, but it is worth examining why it is so, which I shall do by means of some examples.

Disputation 1. Suppose that φ_0, a false but possible *positum*, is granted. Then propose φ_1, a true impertinent proposition; it is granted. Then propose φ_0 & φ_1; the question is then whether $\varphi_0, \varphi_1 \vdash \varphi_0 \& \varphi_1$ is a valid consequence. Obviously, it is, so φ_0 & φ_1 must be granted as a pertinent proposition, which follows from the *positum* together with what has been granted.[353] So the consistent set $\Gamma_2 = \{\varphi_0, \varphi_1, \varphi_0 \& \varphi_1\}$ is formed.

Disputation 2. Now suppose that, in a different disputation, the same propositions are proposed, but in a different order. We start with the false *positum* φ_0, and then the conjunction φ_0 & φ_1 is proposed (φ_1 alone has not been proposed yet). At this stage, the conjunction is clearly impertinent; moreover, it is false, since one of the conjuncts (namely φ_0) is false. So it is denied. Then φ_1 is proposed, and the question is: is it a pertinent proposition? Yes it is, namely a repugnant proposition, because the consequence $\varphi_0, \neg(\varphi_0 \& \varphi_1) \vdash \neg\varphi_1$ is valid. So φ_1 must be denied, and the consistent set $\Gamma_{2'} = \{\varphi_0, \neg(\varphi_0 \& \varphi_1), \neg\varphi_1\}$ is formed.

As for a disjunction, the situation is even simpler. **Disputation 3**. If the *positum* is a disjunction and one of the disjuncts is denied in the next round, then whenever the other disjunct is proposed, it must be accepted as a pertinent proposition, since $\varphi_0 \vee \varphi_1, \neg\varphi_0 \vdash \varphi_1$ is a valid consequence. (If the disjunct first proposed is accepted, then the second disjunct remains an impertinent proposition, which thus should be judged according to its own quality).

So what guarantees that the behavior of conjunctions and disjunctions in a Burley-style obligation does not generate inconsistencies is the notion of pertinent/impertinent propositions, and the influence of the order in which propositions are proposed.

Interestingly, Strode proposes yet a different way of avoiding this kind of inconsistency. One wonders why he does so, since the Burley rules as they stand are sufficient to avoid the problem. Strode's proposal is essentially to consider the *positum* as true (*Strode Obl.*, 6; cf. Ashworth 1993, 381). If one does so, Disputation 2 above has a different outcome, even if played according to the *antiqua responsio* canon: if one considers the *positum* φ_0 as true, then the conjunction φ_0 & φ_1, proposed just after the *positum*, becomes an impertinent **but true** proposition, and therefore must be granted. Incidentally, the same response would be given by someone playing the game according to the *nova responsio* canon (the difference between the two canons would appear in the next round: according to the *antiqua responsio*, φ_1 would be a pertinent proposition, while according to the *nova responsio* it would be an impertinent proposition, since it does not follow from the *positum* alone).

---

[353] By comparison, for Swyneshed the conjunction in this case is an impertinent proposition because the consequence $\varphi_0 \Rightarrow \varphi_0 \& \varphi_1$ is not valid.

Notice though that stipulating that the *positum* is true does not seem to solve the inconsistency created in Disputation 3 if played according to the *nova responsio*: if a disjunction is taken to be true, then one of its disjuncts is taken to be true, but Respondent still does not dispose of sufficient information to determine which disjunct it should be, and therefore must still deny both disjuncts if they are proposed and happen to be false propositions.

Why does Strode propose that the *positum* be viewed as true to avoid inconsistencies related to conjunctions? As argued, the usual Burley rules already guarantee that there be no such inconsistencies. One possibility is that Strode is proposing an amendment **to the *nova responsio***, so that such inconsistencies would not occur even if one played the game according to the *nova responsio* canon. Strode was probably aware of the fact that the *antiqua responsio* as it was avoided the generation of such inconsistencies, but the puzzle remains as to whether he thought that considering the *positum* as true was a necessary amendment to the *antiqua responsio*, or only to the *nova responsio*. It is also possible that Strode was not defending this position as his own, but only proposing it as a possible way of handling the *positum*.[354]

3.5.2.6 Conclusion

Strode's critique of Swyneshed sheds new light in the debate involving *antiqua responsio* and *nova responsio*. He does not simply return to Burley's position; he does so by arguing for its superiority and by outlining the oddities of Swyneshed's position. For this reason, he offers an extremely valuable contribution to our grasp of the historical as well as conceptual intricacies of the obligational genre.

3.5.3 Focus on epistemic/pragmatic elements of the disputation

As already mentioned, it has been argued by many scholars that, in the second half of the 14th century, especially in England, an epistemic turn in logic took place.[355] This is to be seen in particular in theories of consequences,[356] and one of its main signs is the recast definition of 'formal consequence'. While in the first half of the 14th century the notion of formal consequence was usually defined in terms of Topical, semantic, logical or substitutional criteria, in the decades following this period many

---

[354] This position was explicitly held by the anonymous Merton author in (Kretzmann and Stump 1985), (cf. Ashworth 1993, 375).

[355] But it has also been argued that an epistemic notion of consequence is also to be found in Boethius and Abelard (cf. Martin 2001), and in some 13th century logicians, such as Kilwardby (cf. Ashworth 2002). So, arguably, the English logicians of the second half of the 14th century were not introducing an entirely novel way of approaching logic, but rather returning to an old approach, which had lost its predominance during the Burley-Ockham-Buridan period.

[356] 'The fourth and perhaps the most important phase [in medieval epistemic logic] arose within the theory of consequences; Philosophers came to see that not only the most general rules of prepositional logic and alethic modalities, but also those involving epistemic, obligational and other modalities need to be recognized as the most basic principles of reasoning in various realms. Strode's *Consequences*, written probably in 1360s, is certainly the best representative of this stage'. (Boh 2000, 129/30)

authors began to formulate it in epistemic terms: a formal consequence is such that the consequent is understood in the antecedent, so that whoever understands the antecedent (as true) will understand the consequent (as true). In other words, a formal consequence is thus defined with respect to the knowledge and understanding of a hypothetical agent, and not anymore purely on the basis of features of the consequence itself. Strode's own definition of a formal consequence runs very much along these lines:

> A consequence is called sound by form when, if the way in which facts are adequately signified by the antecedent is understood, the way in which they are adequately signified by the consequent is also understood; for instance, if anyone understands that you are a man, he will understand also that you are an animal. (Seaton 1973, quoted in Boh 2001, 156/7)

The same focus on the knowing agent is noticeable in Strode's treatise on *obligationes*, as I will try to show in this section. This is all the more interesting insofar as in the other treatises on *obligationes* thus far mentioned, Burley's and Swyneshed's, this epistemic approach is less noticeable. Simultaneously, Strode's focus on the agent also outlines what we could call the pragmatic nature of obligational disputations (and of logic in general) as an actual activity, taking place in space and time, and the issues that must be dealt with as a result of this approach.[357] These include the issue of the actual logical knowledge of Respondent, the essentially verbal nature of these disputations, and how to handle propositions referring to the very circumstances of the disputation as well as to the very moves made within it.

3.5.3.1 Epistemic clauses

Both in Burley's and Swyneshed's treatises, epistemic elements were present in the rules of how to respond to impertinent propositions:[358] these propositions should be responded to on the basis of Respondent's actual knowledge at the moment of the disputation. If he knew an impertinent proposition to be true, it should be granted; if he knew it to be false, it should be denied, and if he did not know whether it was true or whether it was false, then it should be doubted. But the epistemic clause was not extended to the evaluation of the *positum*, nor to the assessment of inferential relations between the *positum* and the *proposita*.

By contrast, in Strode's treatise, phrases such as '*falsum scitum a te esse falsum*', '*verum scitum a te esse verum*' (cf. first conclusion), '*possibile scitum esse possibile*' (cf. first supposition), '*scitum esse positum*', (cf. second supposition) '*scitum esse uerum non repugnans*', '*scitum esse falsum non sequens*' (cf. second conclusion), are abundant. Indeed, such epistemic clauses were characteristic of Oxford

---

[357] A similar pragmatic approach is noticeable in, for instance, Buridan's staunch commitment to proposition-tokens as the bearers of truth-value, and in the amendments made necessary to his logical system as a result of this commitment. (cf. Dutilh Novaes 2005c)

[358] Burley and Swyneshed had the same rule on how to respond to impertinent propositions, but, as I have stressed many times, they disagreed on which propositions were impertinent.

logic.[359] At first sight, these seem harmless and in fact conceptually interesting additions to the obligational framework, related to Strode's general epistemic penchant and to the focus on Respondent as an actual, non-omniscient agent. But, under a given interpretation, this epistemic approach might also be seen as problematic, yielding a confusion between the normative and the descriptive characters of the obligational rules.

It has been stressed by many scholars (cf. Yrjönsuuri 1993, 302) that the rules of *obligationes* are essentially **normative**; commentators have expressed this normative character in different ways, such as pointing out their connection with deontic problems (Knuuttila and Yrjönsuuri 1988), stressing the importance of the notion of 'correctness' (cf. Keffer 2001, 123–127, 147–150), or viewing *obligationes* as rule-governed games, as I do here. But Strode's epistemic clauses appear to introduce a **descriptive** approach to those rules, and this creates difficulties.

Consider the following formulations of some of the obligational rules, in the spirit of the phrases containing the term '*scitum*' quoted above:

> (R1) Respondent grants the *positum* iff he knows it to be a possible proposition.
> (R2) Respondent grants a *propositum* iff he knows it to be a true, not *repugnans* proposition.
> (R3) Respondent denies a *propositum* iff he knows it to be a false, not *sequens* proposition.

According to this formulation of R1, if Respondent denies a given *positum* that is in fact possible, but that he does not know to be possible, he is properly speaking **not** infringing R1. Similarly, suppose that a *propositum* is in fact repugnant, besides being true, but that Respondent fails to identify this logical relation between this *propositum* and the *positum* (together with other previously granted/denied propositions, according to the *antiqua responsio*), and that he therefore grants this proposition; again, properly speaking, he is not infringing R2, given the failure of the epistemic clause. The same holds for R3.

Hence, if Respondent's logical knowledge is deficient, given the epistemic clause, he is strictly speaking not responding badly. But the point of obligational disputations (or at least one of them) is precisely to **test** Respondent's logical abilities and knowledge; he will respond well iff he knows his logic well. A situation in which Respondent makes logical mistakes and can still be said to have responded well in some sense (perhaps 'to the best of his abilities') is against the whole spirit of the game. So the epistemic clauses, albeit realistic, jeopardize the normative character of *obligationes*.

Indeed, one way to make sense of the obligational rules thus formulated is to view them as descriptive rules; in reality, Respondent will only accept a *positum*

---

[359] I owe this information to Prof. E.J. Ashworth. See also (Ashworth 1985).

if he knows it to be possible, a *propositum* as following from or repugnant to the *positum* only if he knows it to be such etc. But the purpose of obligational treatises is presumably that of spelling out rules of how to respond **correctly** (how to win the game) and not that of serving as a description of how such disputations actually take place.

Another way to interpret such clauses is to view them as related to the explicit formulation of the propositions being put forward – as opposed to their 'mental' formulation. As will be discussed below, Strode was against the view that some propositions were implicitly granted during a disputation; similarly, if the *positum* was 'I posit to you the proposition that I am thinking about', Respondent could neither accept nor deny such a *positum*, since he didn't know the (unuttered) proposition to be possible or impossible. But if this is so, a perhaps more intuitive formulation of this clause would have been something like '*falsum scitum a te*', instead of '*falsum scitum a te esse falsum*', stressing thus that the proposition in question must be known to Respondent, but not necessarily known to be such-and-such.

Thus, one wonders to what extent Strode was only reproducing some of the standard Oxford formulations of these rules, or whether there really was an epistemic-descriptive component in his conception of the obligational rules.

3.5.3.2 Only explicitly proposed propositions belong to the informational base

Although Strode's main opponent in this treatise seems to be Swyneshed, in chapter III he also criticizes a particular view on *obligationes* not defended by the latter. In fact, this view was defended by the anonymous author of the so-called Merton treatise (cf. Kretzmann and Stump 1985; Ashworth 1993, 375–379).

The specific view in question was that some propositions, even if not proposed, are implicitly granted during an obligational disputation, and therefore should be taken into account by Respondent when granting or denying the *proposita*. The typical example illustrating this view goes as follows: the *positum* is 'Every man is running'. The first *propositum* is then 'You are running': according to both Burley and Swyneshed, Respondent must deny this *propositum* as impertinent and false (Respondent is in fact not running at that moment). But according to the Merton author,[360] Respondent is tacitly committed to the truth of the proposition 'You are a man', even though it has not been explicitly proposed in the disputation, and therefore Respondent must **grant** 'You are running' as following from the *positum* together with the tacitly granted proposition 'You are a man' (cf. *Strode Obl.*, 6).

Strode is radically opposed to this view. He starts his counter-argumentation by noticing that 'granting' can be understood in two ways, verbally and mentally, and that according to some (the Merton author, for instance), it is not more reasonable to grant a proposition that follows from the *positum* together with a verbally granted

[360] (Kretzmann and Stump 1985), pp. 246 (Latin text), 255 (translation), 266/7 (comments).

proposition than it is to grant a proposition that follows from the *positum* together with a mentally granted proposition.[361] Obviously, this position puts at risk many of the crucial aspects of Strode-style *obligationes*, such as the importance of the order in which propositions are proposed, and the fundamentally pragmatic character of a disputation, seen as an actual deed performed by those involved in it. One could add the objection that, since those mentally granted propositions are only accessible to Respondent, it becomes virtually impossible to judge objectively whether he has responded well or not at the end of the disputation, as the total set of denied/granted propositions is not publicly available.

Strode's argument against this view, if I understand it correctly, runs roughly as follows: if the *positum* is 'Every man is running', to which mental propositions is Respondent actually committed? To the proposition that he is a man or to the proposition that he is not running? They are both true propositions, but if they are both granted (mentally or otherwise) together with the *positum*, then clearly an inconsistent set of propositions has been granted.

Thus, if mentally granted propositions are included in the informational base of the disputation (the set of propositions on the basis of which pertinence is judged), then the rules of the game no longer determine the correct response(s) to a proposition. Given the *positum* 'Every man is running', if the first *propositum* is 'You are running', one can either deny it, as it is an impertinent and false proposition, or else grant it as following from the *positum* and the mentally granted proposition 'You are a man'. Similarly, if 'You are a man' is the first *propositum*, then it can either be granted as impertinent and true, or denied as repugnant to the *positum* together with the mentally granted proposition 'You are not running'.[362]

Strode also notices that granting mental propositions infringes the fourth conclusion, namely that the order of granting and denying is crucial. While verbally granting and denying takes place in time, and therefore the order of responses is well-established, mentally granting and denying subverts the order principle, since presumably all mentally granted or denied propositions are granted or denied at the same time. So the order principle can no longer be applied.[363]

Strode's critique of this position indicates that he is well aware of the essentially pragmatic (as opposed to ideal) character of obligational disputations. Such disputations are dependent on elements of the actual situation in which they take place, such as the actual state of knowledge of those present, the propositions that are actually – that is, verbally – proposed, and the order in which they are proposed. Here, the game metaphor comes in handy, as it stresses the role of the participants

---

361 Non enim eis uidetur esse maior ratio quare debet aliqua propositio concedi que sequitur ex posito cum scito esse uero uel bene concesso vocaliter, quam quando sequitur ex posito cum scito esse uero uel bene concesso mentaliter. (*Strode Obl.*, 6)

362 I have argued that, with respect to impertinent propositions, Respondent seems to have some freedom in how to respond to them. But here what is at stake is indeterminacy even of whether a proposition is pertinent or impertinent.

363 [. . .] videtur quod ista positio transponit uerum ordinem respondendi. (*Strode Obl.*, 6).

playing the game and the specific moves made by each of them, in a certain order. A 'mental move' is simply not a move in the game – intersubjectivity is a fundamental trait of any game. In obligational disputations, all moves are and must be verbal, explicit moves.

3.5.3.3 Self-referential posita

Propositions referring to the situation of the very disputation in which they are proposed are discussed in virtually all *obligationes* treatises of the 14th century (and, for that matter, in 12th and 13th century treatises as well). This fact has led to the hypothesis that *obligationes* were essentially a framework to deal with *sophismata* and self-referential, Liar-like paradoxes (cf. Stump 1982b). Even if one does not agree with this hypothesis (as e.g., Yrjönsuuri 2000, 216), one must still provide an explanation for the abundant presence of *sophismata* in obligational treatises. I have argued (in sections 3.2.1 and 3.4.4) that these seem to be limit cases, which serve to test the soundness of the obligational theory being proposed; if the theory can deal with such difficult cases without becoming incoherent (e.g., without yielding conflicting instructions on how Respondent should reply), then it is a sound theory.

In any case, the discussion of self-referential propositions in Strode's treatise is not a particular trait of this text. What is perhaps novel in Strode's treatise (with respect to his predecessors such as Swyneshed and Burley) is that he formulates specific rules on how to deal with them (for instance, in the first supposition and in the second and third conclusions), in such a way that these *sophismata* are no longer discussed as limit cases. Rather, they seem to be at the core of the theory.

But instead of viewing this fact as an indication that *obligationes* were really meant to solve *sophismata*, I tend to believe that, in Strode's case, it is a sign of the general epistemic/pragmatic focus of his treatise.

Worth considering is Strode's discussion of *posita* that lead to what can be called performative contradictions: *posita* that contradict the very act of positing by Opponent, the *positio* ('Nothing is posited to you'), or *posita* that contradict the very act of accepting the *positum* by Respondent, the *admissio* ('You are sleeping now') (cf. Strode's third conclusion). Even though these *posita* seem to create contradictions (of the performative, pragmatic kind), Strode states very clearly that they should be accepted as *posita*. He remarks that denying these as *posita* on the basis of the first supposition (according to which impossible *posita* should not be accepted) is nothing but an escape for the miserable, who do not know the force of this art.[364]

But if a proposition describing the paradoxical *positio* – that is, the very act of positing the given *positum* – is proposed as a *propositum*, then it should be denied, since (although true) it is repugnant to the *positum*. That is, suppose that the *positum* is 'Nothing is posited to you'. It should be granted, as it is not impossible. Then, if ' "Nothing is posited to you" is posited to you' is proposed, it should not be granted,

[364] Ex quibus patet quod tales casus negare non est nisi fuga miserorum, nescientium istius artis uigorem. (*Strode Obl.*, 3).

even though it is true in the disputational situation, because it is repugnant to the *positum* (it is an instantiation of 'Something is posited to you'). Therefore, since inferential relations take priority over truth in the case of pertinent propositions, it should be denied.

Also worth noting is Strode's second conclusion: Respondent can grant that he is responding badly and yet not lose the game (notice that it follows from the fourth supposition and the first conclusion: if denying a true proposition does not necessarily mean responding badly, Respondent can grant that he is responding badly even though he is in fact responding well). Consider the following disputation:

φ_0: 'Every man is responding badly'. Granted, possible
φ_1: 'You are a man'. Granted, impertinent and true.
φ_2: 'You are responding badly'. Granted, follows from φ_0 and φ_1.

What is paradoxical about this situation is that the very act of granting 'You are responding badly' corresponds to Respondent's responding **well**, whereas denying it would correspond to his responding badly.[365] But again Strode does not consider these performative contradictions to be vicious contradictions, to be excluded from the obligational framework.

How can we make sense of these paradoxical situations? I propose here to consider the distinction between the context of utterance of a proposition and its context of evaluation, familiar from two-dimensional semantics, and already applied in part 2 of the present text. It is as if all granted or denied pertinent propositions (including the *positum*) were being evaluated with respect to a different context, not that of the disputation. Impertinent propositions, on the other hand, are evaluated with respect to the very context of the disputation. So Respondent can grant that nothing is posited to him, that he is dead, that he is responding badly etc., because it is as though these referred to a different situation, not to the very disputation.

Here, the thought-experiment hypothesis[366] is also helpful: in the case of pertinent propositions, one should reply to them as if one were creating a counterfactual situation, a thought-experiment, in which the *positum* is true. These pragmatic/performative paradoxes are thus resolved if, with respect to pertinent propositions, one supposes that Respondent is not referring to the very situation of the disputation – in which he is uttering these performative paradoxes – but rather to a different situation. From this point of view, such performative contradictions are not very different from plain contingent false propositions.[367] Indeed, in the example above, '"Nothing is posited to you" is posited to you' is false in

[365] Something like uttering 'I am not speaking now'.

[366] Some scholars, in particular Yrjönsuuri (2000, 219) have contended that an adequate modern interpretation of the obligational framework is to view it as a technique to build thought-experiments.

[367] A similar conclusion was reached with the application of this distinction to Buridan's 'No proposition is negative': it is impossibly-true, since it is false every time it is uttered, but it is not impossible, since the situation it describes is not impossible – it is in fact a possible proposition.

the situation in which that nothing is posited to you is the case, following the *positum*.[368]

But it remains crucial to identify correctly whether a *propositum* is pertinent or impertinent; if it is impertinent, it should be responded to according to the very context of utterance (and in this case there is no distinction between context of utterance and context of evaluation), but if it is pertinent, it should be responded to according to the relevant inferential relations. These responses progressively create a counterfactual situation that is, as it were, the context of evaluation of pertinent propositions.[369]

An issue that I have not yet addressed is whether Strode would be willing to accept all kinds of pragmatic and performative inconsistencies, or only those discussed so far. E.J. Ashworth argues that he might be seen as ruling out *posita* of the kind 'You are in Rome and no conjunction is posited to you' (Ashworth 1993, 366). I am not sure whether such a *positum* differs so much from 'Nothing is posited to you' or similar *posita*. On the basis of the distinction between context of utterance and context of evaluation, Respondent could then simply consider a situation in which he is in Rome (participating in a disputation or not), and either nothing is posited to him, or a *positum* is posited but it is not a conjunction. But if explicit reference is made to the very disputation taking place – 'You are in Rome and no conjunction is posited to you in this very disputation' – then it seems that we would have a real paradox, similar to Liar-like paradoxes with explicit self-reference ('This proposition is false').

3.5.3.4 Some rules that do not hold

In his influential treatise on consequences (Seaton 1973), Strode presents rules of consequence specifically valid within the realm of *obligationes*. These basically pertain to the issue of whether the consequent or antecedent of a *consequentia* known to hold should be granted/denied/doubted once the antecedent or consequent has been granted/denied/doubted (see (Boh 2001, 162/3) for a formal reconstruction).

There are however a few obligational 'consequences' that might be expected to hold, given the intuitive character of their counterparts in contexts other than *obligationes*, but which fail to hold within Strode's obligational framework. Interestingly, the counterexamples to these invalid schemata are often related to the self-referential phenomena and performative contradictions just discussed.

---

[368] The gist of this distinction can also be found in the following remark by Ashworth: 'As he [Strode] subsequently pointed out, there is, after all, nothing to prevent the Respondent from following through on the logical consequences of what he has admitted. He has simply to deny that he himself is speaking, or debating, or granting, or denying or engaging in any action which is incompatible with his being non-existent, or merely asleep. This approach depends on a careful distinction between uttering and making a statement about the conditions of utterance. It is all right for the Respondent to reply, but he must never grant that he is replying'. (Ashworth 1993, 367).

[369] It is worth noting though that Respondent does not start out with a given model to serve as the context of evaluation for pertinent propositions; rather, this model is created as the disputation progresses. Cf. section 3.3.3.1 and (Yrjönsuuri 2000, 220), against 'semantic interpretations' of obligations.

Another source of counterexamples are propositions that refer to the very moves being made in the game, for example, as in '*p* ought to be granted' (where *p* is any proposition); in such cases, the meta-evaluation of a given move takes place within the very disputation, and can prima facie be iterated as many times as one wishes. It is as if there were various layers of disputation within the very same disputation – the disputation itself, talking about the disputation, talking about talking about the disputation etc. These different layers are particularly confusing when they are 'tossed' together in the same proposition, for example, in the case of a disjunction or conjunction whose members refer to different layers – '*p* or *p* ought to be granted'. Going back and forth the different layers may create what Ashworth has termed 'obligational inconsistencies' (see below), but nevertheless several authors, including Strode, seemed to be prepared to deny the validity of schemata that would otherwise prevent obligational inconsistencies from arising.

In what follows, I shall represent some of these schemata by means of basic propositional logic and some modal-like operators. So 'T[*p*]' corresponds to the statement that proposition *p* is true; 'G[*p*]' to the statement that proposition *p* is granted, and 'OG[*p*]' to the statement that proposition *p* ought to be granted (square brackets represent the nominalization of *p*). Iteration of these operators is possible. Within the obligational context, G[*p*] is in fact equivalent to *p*, insofar as stating a proposition is, in this context, the same as granting it.

The first schema that is usually expected to hold in other disputational or argumentative contexts is the one that associates the duty of granting a proposition with its truth. Indeed, as shown by Yrjönsuuri (2000, 207), originally – in the Aristotelian theory of disputations as well as in earlier medieval treatises on the topic – Respondent's main commitment is with truth, or at least with his knowledge of what is true (Respondent is not expected to know all truths as true). The schema can be represented as:

(Sc1) OG[*p*] <==> T[*p*]

In other words, if *p* is true it ought to be granted, and if it ought to be granted, it is true. But (as shown in Yrjönsuuri 2000, 209), the significant turn introduced by the obligational treatises is precisely that Respondent's duty towards truth becomes overridden by his commitment to the recognition of certain inferential relations. Not only is Respondent committed to accepting false *posita* in the context of *obligationes*, but he is also committed to granting all propositions that follow from the *positum*, regardless of their truth-value. Indeed, Strode is very much aware of this subversion of the original purposes of disputations, and states clearly in his first conclusion that one does not necessarily respond badly when one grants a false proposition or denies a true one.

Another schema that does not seem to hold in a Strode-style obligational disputation is the one associating the granting of *p* with the granting that *p* is true.

(Sc2) G[*p*] <==> OG[T[*p*]]

Prima facie, this appears to be a variation of the Tarskian T-schema and, in effect, within most medieval logical systems, the T-schema would not hold. This is related to the medieval view that tokens are the truth-value bearers: 'p iff T[p]' does not hold because, for a proposition to be true, it must exist (it must be formed); so this schema holds only under the proviso of the token's existence (cf. Ashworth 1993, 368).

But here we seem to encounter a different situation, insofar as the reason why (Sc2) does not hold within Strode's (and other authors', such as Burley's) obligational framework seems to be of a different nature. Recall the distinction between context of utterance and context of evaluation; suppose that p is the *positum* or a *propositum* having been accepted as pertinent and following from the *positum*, but that p is actually false. Because it is a pertinent proposition, its context of evaluation is not that of the disputation, which is its context of utterance. But T[p] is an irrelevant proposition, and therefore should be evaluated according to the context of utterance; now, in this context, p is false, thus T[p] is also false, and therefore should be denied.[370]

We obtain thus the awkward situation of having granted p and having denied T[p]. Discussing Burley's contention that (Sc2) does not hold, Yrjönsuuri (2000, 220) takes this to be a sign that semantic interpretations of *obligationes* (as a framework for the analysis of counterfactuals, thought-experiments or belief revision) are not adequate, as it seems absurd that, in the same situation, p holds and T[p] does not hold. But perhaps the problem is with the view that only one situation (or one class of situations) is at stake during an obligational disputation; what may be required is the distinction between the fictional situation progressively being constructed, and the actual situation of the disputation taking place[371] – or similarly, 'between uttering and any metalinguistic assessment of the speech act involved or of the utterance itself' (Ashworth 1993, 368).

This being said, I am extremely sympathetic to Yrjönsuuri's suggestion that 'obligational disputations aim at constructing consistent sets of sentences' (Yrjönsuuri 2000, 221), taking place thus essentially on a linguistic rather than semantic level.[372] While it may seem absurd that, in a sensible description of a possible state of affairs w, p holds and T[p] does not hold, a set containing both p and $\neg$T[p] is not necessarily inconsistent (in particular if the T-schema does not hold).

The last two schemata that I would like to discuss, which do not seem to hold in Strode's theory of *obligationes*, have also been discussed in (Knuuttila and Yrjönsuuri 1988, 197–199) with respect to *sophismata* proposed in Burley's and Sherwood's (?) treatises. They both concern the evaluation of moves of the disputation within the very

---

[370] Admittedly, at some point Strode seems to be defending the thesis that the *positum*, once accepted, should be considered as true (even if actually false), as discussed in section 3.5.2.5. But nothing is said about false *proposita* that are granted in virtue of being *sequens*, and, presumably, these are not to be considered as true.

[371] In section 3.2.1 I have argued (following a suggestion by Christopher J. Martin) that playing an obligational game is equivalent to the process of constructing a maximal consistent set of propositions, following Lindenbaüm's lemma, which would correspond to a possible world, if possible worlds are defined as maximal consistent sets of propositions.

[372] This point is also related to the essentially verbal nature of obligational disputations discussed in section 3.5.3.2.

disputation, by means of normative terms such as ‘*concedendum*’ and ‘*negandum*’ occurring in propositions proposed (as *posita* or as *proposita*). Properly speaking, such ‘deontic operators’ can be iterated ad infinitum, creating the confusing effect of different layers of discourse within the same proposition.

Consider the following schema:

(Sc3) OG[p] <==> G[p]

An obligational disputation where this schema holds is clearly a fully determined game, in the sense that whatever is correctly granted ought to be granted. In other words, in such disputations Respondent has no space for maneuver, and there is at each turn only one move that will avoid him losing the game. As already mentioned, there are elements in Burley’s treatise suggesting that the deterministic interpretation is not the only one possible, and that some space for maneuver and strategic playing seems to be left with respect to impertinent propositions (since in such cases granting it as well as denying it allow Respondent to keep consistency).

In some passages, Strode seems to be defending precisely this non-deterministic view of the game: Respondent would have the duty to grant pertinent propositions that are *sequens* and to deny pertinent propositions that are *repugnans*, but as for impertinent propositions no such duty would apply. In chapter XI, he presents the following *sophisma*, a clear counterexample to (Sc3):[373]

φ_0: ‘ “Nothing is posited to you” is posited to you’. Granted, possible.
φ_1: ‘Something is posited to you’. Accepted, impertinent and true.
φ_2: ‘ “Something is posited to you” ought to be granted’. Denied, ‘Something is posited to you’ does not follow from φ_0.

That is, the reason he gives for denying φ_2 is that ‘Something is posited to you’ is impertinent; so even though Respondent has granted φ_1, he was under no obligation to do so. (Sc3) is also related to the normative/descriptive dichotomy discussed in section 3.5.3.1 (also discussed in (Knuuttila and Yrjönsuuri 1988, 197)). If this schema is supposed to be descriptive, then obviously it fails, since not everything that ought to be granted is in fact granted – namely, when Respondent actually responds badly. But in Strode’s theory, (Sc3) seems not to hold even under a normative reading, insofar as not everything that is granted ought to be granted (namely, impertinent propositions that are granted, under the non-deterministic interpretation). By contrast, under the normative reading, the other direction (from left to right) of the schema obviously holds: if Respondent does not grant what ought to be granted (a pertinent *sequens* proposition), then he responds badly.

Moreover, in special cases (e.g., when the *positum* is a disjunction featuring the term *concedendum* in one of the disjuncts: ‘*p* or *p* ought to be granted’ – cf. Knuuttila

---

[373] Respondeo admittendo casum [‘nihil est tibi positum’ sit tibi posita et a te bene admissa]; et cum proponitur: ‘aliquid est tibi positum’, concedo, et nego quod illa sit a me concedenda, quia dico quod non sequitur ex mihi posito et bene admisso. (*Strode Obl.*, 27)

and Yrjönsuuri 1988, 198), it can occur that a proposition *p* ought be granted (as *sequens*), but that '*p* ought to be granted' must be denied as repugnant, even though it is true (it describes a true fact about the disputation). In sum, in such cases *p* ought to be granted, but '*p* ought to be granted' does not have to be granted – in fact, since Respondent's commitment towards inferential relations has priority over his commitment to truth, it ought to be denied. Thus, the schema (Sc4), where the deontic operator 'OG' is iterated, was often rejected.[374]

$$\text{(Sc4) } OG[p] \Longleftrightarrow OG[OG[p]]$$

In a recent and yet unpublished article,[375] E.J. Ashworth argues that there were basically two trends among the authors of obligational tratises concerning (Sc4), namely those who rejected it as a valid principle governing the disputation, and those who accepted it, presumably under the argument that denying this principle would generate obligational inconsistencies – for example, if Respondent denies *p* but grants that *p* ought to be granted. Possibly, according to the first group of authors, such 'inconsistencies' would not be very different from other pragmatic inconsistencies, such as 'Nothing is posited to you', which many authors, among whom Strode, were glad to accept as possible *posita*. So one may conjecture that Strode would tend to deny the validity of (Sc4), but in his text I have not found conclusive elements as to which trend he belonged on this particular issue.

The question is of course, once these different layers of discourse about the disputation are introduced, whether the very rules of obligation force Respondent to grant inconsistencies. I have proved in section 3.3.3.1 that, provided that only the lowest level, the object-level, of discourse within the disputation comes in play, Respondent can always maintain consistency. But with the different layers of discourse this may not be possible, in any case if obligational 'inconsistencies' are viewed as real inconsistencies. If they are viewed as mere pragmatic and performative inconsistencies, then there is no real problem. But if they are more than pragmatic and performative inconsistencies, then they require a dedicated discussion, which remains for now a subject for future research.

3.5.4 Conclusion

One of the most interesting aspects of the study of the obligational literature is the fact that the medieval authors really seemed to be conversing with one another, yielding a lively debate. Indeed, an author such as Swyneshed was clearly reacting to Burley, whereas someone like Strode is clearly reacting to Swyneshed, not to mention the other, as-of-now less known participants of this debate. The result was that

---

374 'Neither "You ought to grant *P* if and only if you ought to grant "P ought to be granted"" nor "*P* if and only if *P* is true" were accepted as principles governing an obligational disputation'. (Ashworth 1993, 368)

375 (Ashworth 2003).

the discussions presented in the *obligationes* treatises were progressively sharper and deeper.

I have attempted to show that Strode's main contribution to this debate revolves mostly around his criticism of the *nova responsio* and his epistemic and pragmatic version of the *antiqua responsio*. The rules of the game presented by Strode, properly speaking, differ very little from Burley's rules, but the conceptual analysis accompanying the presentation of his theory seems to have a different flavor, which makes it particularly interesting.

3.6 CONCLUSION

I hope to have shown that the game vantage point is extremely fruitful for a better understanding of the mechanisms and the rationale underlying *obligationes*. To my mind, it offers a more appealing account of the obligational theories and practices as a whole, and it also clarifies some of their aspects that are otherwise rather obscure or difficult to account for, such as the role of impertinent propositions and the essentially dynamic nature of *obligationes*.

Admittedly, the game hypothesis works better for some versions of *obligationes* than for others – Swyneshed's theory of *obligationes* viewed as a game turns out to be a particularly tedious kind of game. But it is thus perhaps for this reason that the *antiqua responsio* remained more influential and resisted the challenge posed by Swyneshed's *nova responsio*.

One may object that the game interpretation of *obligationes* is dangerously anachronistic; but it is undeniable that it successfully captures the rule-governed and goal-oriented character of *obligationes*. As for consistency maintenance, I am convinced that this really is the key idea behind *obligationes*, or at any rate in their *antiqua responsio* version. This idea is broad enough to reconcile most if not all the other proposed interpretations of *obligationes*, discussed in section 3.2: after all, in theories of counterfactual conditionals, in analysis of *sophismata* and *insolubiles*, for the idea of compossibility, in belief revision etc., the imperative of consistency maintenance is overall dominant.

In terms of modern logical games, there seems to be nothing quite equivalent to the obligational game. In some respects, it makes one think of tableau methods; in others, of model construction games (cf. van Benthem, chap. 3).[376] In any case, a more thorough comparison with modern logical games seems to constitute a promising topic for future work.

---

[376] In J. van Benthem's terms, *obligationes* are an 'in-between genre'.

PART 4

THE PHILOSOPHY OF FORMALIZATION

4.0 INTRODUCTION

While formalizing the fragments of medieval logic presented here, I have followed certain methodological guidelines which were not made explicit; however, my project would clearly remain incomplete without an analysis of methodological issues. Moreover, having carried out these formalizations, I was naturally led to reflect on the philosophical aspects of formalizing. Therefore, in this chapter I discuss formalization as such, and its methodological and philosophical implications.

It goes without saying that the philosophy of formalization is an important topic, given the widespread use of formalization in philosophy as well as in other fields. However, surprisingly little seems to have been written on the topic recently;[377] to find analyses of what formalization is and why it is undertaken, one has to turn to the literature of the early days of formalization. The seminal work on this topic is undoubtedly Frege's *Begriffsschrift*; moreover, some mid-20th century logicians have also felt the need to reflect on what it means to formalize – in particular in the 1950s, arguably a period of evaluation of the astonishing progress in metamathematics and logic of the previous decades. Among these analyses, (Kleene 1951), (Church 1956), (Wang 1955), (Curry 1957), (Curry and Feys 1958) stand out.

Nowadays, formalization is usually carried out without further reflection on what precisely is being done or why. If this kind of analysis is undertaken at all, the focus is on the advantages and disadvantages of formalizing – what one could call the pragmatic aspects of formalizing. To my mind, this is an almost inexplicable lacuna, and therefore I deem it essential to offer some considerations on the philosophical import of what I have been doing so far; moreover, under the light of the broader use of formalizations since the 1950s, it is time to take stock and perhaps complement the aforementioned analyses. Thus, I intend to go beyond the pragmatics of formalizing (although this aspect of the matter will be addressed as well); I seek to understand what it means to formalize, what precisely is involved in a formalization – the 'metaphysics' of formalization, so to speak.

---

[377] (Hansson 2000) is one of the few recent texts on the topic.

As a first approximation, formalizations can be seen as *models* – and thus are essentially simplifications. That is to say: typically, some features of what is being modeled are chosen as the cornerstones of the model, while others are ignored. In effect, a good model is always a tradeoff between resemblance to what is being described and simplicity, insofar as the two main goals of modeling, faithfulness to its object and manageability, are often conflicting. Therefore, every model is bound to be a simplification, an interpretation of its subject-matter, and this holds in particular of formalizations. Typically, some features of the object of formalization are considered to be more fundamental than others, while other features are simply altogether ignored.

> We have to make allowance for the fact that a formal theory may be an idealization, rather than a mere transcription of experience. Thus if certain physical theories were formalized as formal systems, they would doubtless contain statements not translatable into statements which can be tested experimentally. (Curry and Feys 1958, 22)

Moreover, formalizations are models of a particular kind, namely models constructed with a specific sort of tools – formal, logical tools. This is far from being a satisfactory definition, since formal tools are also used for modeling in other fields such as physics, but one would not call these formal models 'formalizations'. In practice, formalizations in philosophy are almost always applications of the apparatus of existing familiar logical systems – mainly first-order predicate logic, propositional logic, or else some of the many extensions of these systems – to philosophical issues and theories (cf. Hansson 2000, 170).

Hence, to explain formalization as the application of formal tools is not sufficient. What exactly goes on when formalizations of this kind are undertaken? The starting point of my analysis is the idea that what is usually meant by 'formalization' in philosophy (and in other fields, for that matter) is a conflation of three related but distinct notions: axiomatization, symbolization, and conceptual translation into a given formal theory.[378] *Axiomatization* concerns the **internal articulation** of what is being formalized (a theory, an argument), and consists in spelling out how different statements are related to one another. *Symbolization* concerns the **language** in which the theory or argument is expressed, and usually consists of a 'transcription' of vernacular pieces of language into artificially created languages. *Conceptual translation* into a formal theory concerns the **conceptual makeup** of the object of formalization, and usually consists of (besides the use of a given symbolism) a mapping between the key concepts of the latter into the key concepts of the formal theory in question.

This chapter starts with a preliminary overview of some crucial notions, namely those of objects of formalization, the relation between the notions of formal and of formalization, and the notion of 'formal' properly speaking. Subsequently I turn to the

---

[378] Note that Kleene (1951, 59–68) proposes a very similar view on formalization. More specifically, he stresses the importance of the two different tasks of axiomatization and symbolization.

three aspects that I claim are involved in formalizing – axiomatization, symbolization and conceptual translation. I discuss some foundational issues concerning each of them and pragmatic aspects of their use, in particular in philosophy.

4.1 PRELIMINARY NOTIONS

4.1.1 Objects of formalization

A straightforward search in any database listing works in philosophy, with the term 'formalization' as input, will yield entries from a remarkable variety of topics, which indicates that there is a great diversity of candidates for formalization. Indeed, it seems that almost any philosophical theme, or in fact any topic at all, is amenable to formalization. Hence, this does not seem to be a very promising starting point in our efforts to attain a better understanding of what it means to formalize; formalizations are not defined by the nature of their objects. Nevertheless, we must have a precise picture of what can be formalized, so let us take a closer look at these putative objects of formalization.

First of all, it is important to notice the distinction between, on the one hand, the different fields of knowledge that are prone to formalizations, and, on the other hand, the specific kinds of fragments of knowledge that can be formalized. In other words, while there can be formalizations in biology, physics, philosophy, sociology etc. (this will be discussed below), the individual object of formalization can range from a very simple argument, or even a sentence, to a highly complex theory. For instance, Russell's analysis of the logical form of sentences such as 'The king of France is bald' (Russell 1905) is a good example of formalization on the sentential level.[379] Taking a step further, one may wish to formalize an argument by spelling out the logical structure that licenses the move from premises to conclusion, which can be done by decomposing the argument in its constituent elements. The same can be done with (mathematical or other) proofs; the formalization of a proof (especially in mathematics) may be carried out in order to 'make explicit all the implicit steps involved and to write down the expanded result once and for all' (Wang 1955, 229).[380] Finally, whole theories can be formalized, as exemplified by the different axiomatizations of set theory. In sum, the objects of formalization can be of different kinds, ranging from single sentences to full-blown theories.

Moreover, these objects of formalization can belong to virtually any field of knowledge. A mathematical theory can be formalized, such as set theory,[381] arithmetic[382]

[379] As we shall see, formalization on the sentential level typically amounts to the establishment of the sentence's logical form, which is then often expressed by means of schemata; cf. (Corcoran 2004).
[380] An example of the formalization of a proof/theorem: Kikuchi, M. and Tanaka, K., 'On formalization of model-theoretic proofs of Gödel's theorems'. *Notre Dame Journal of Formal Logic* **35**, 1994.
[381] Cf. (Zermelo 1908).
[382] Cf. Boxho, F., 'Sur la formalisation de l'arithmetique élémentaire'. *Logique et Analyse* **30**, 1987.

or geometry;[383] the same holds of physics,[384] biology,[385] the social sciences,[386] language studies[387] etc. It can of course happen that the first formulation of a theory already displays the regimented form that is supposed to be the outcome of a process of formalization. In such cases the formalization occurs concomitantly with the formation of the theory itself, so to speak – this is particularly frequent in the domain of the formal sciences, especially in logic and mathematics.

It may be argued that not only theoretical artifacts such as arguments, proofs and theories are possible objects of formalization, but that pre-theoretical phenomena such as uses of language, human behavior or even physical events can be formalized as well, in the same way as modeling occurs in physics. But here I assume that a formal model describing such phenomena is itself already theory-laden, insofar as one must always chose some elements of the phenomenon to be formalized and discard others when construing the formalization. Therefore, any modeling, any formalization, is already a theoretical construction, an interpretation. If the theory is directly proposed in a regimented form, then the formalization is embedded in the very formulation of the theory, as just said. But I maintain that no formalization is 'neutral' or pre-theoretical, insofar as the very process of formalizing requires theoretical choices in the first place.

In this chapter, I will be making a rather loose use of the term 'theory', as referring to virtually any kind of theoretical construction, which can be the object of formalization. Obviously, several complications are involved with this notion, as it seems to be one of those words for which no unified meaning can be given (it seems at best a case of family resemblance). But this particular issue will be disregarded; I will rely on a intuitive understanding of the concept of theory, as I am confident that this will have no undesirable consequences for the present investigation.

Naturally, for the present purposes we are mainly interested in formalizations in philosophy. Now, there are two main sorts of formalization in philosophy: (1) the formal treatment of some fundamental philosophical notions, such as the concepts of necessity and possibility (modal logic), of knowledge (epistemic logic), ethical concepts (deontic logic, game-theoretical notions), linguistic phenomena (formal semantics) etc. – this group of theories is usually referred to as 'philosophical logic',[388] and involves, at least potentially, the use of logical tools for the analysis of philosophical issues; (2) the formalization of philosophical theories that were originally formulated in an informal way, in ordinary language.[389] In particular, the second group comprises formalizations of theories given in times when specific

---

[383] Cf. Von Plato, J., 'Formalization of Hibert's Geometry of incidence and Parallelism'. *Synthese* **110**, 1997.
[384] Cf. Day, M., 'An axiomatic approach to first law termodynamics'. *Journal of Philosophical Logic* **6**, 1977.
[385] Cf. Niven, B., 'Formalization of the basic concepts of animal ecology'. *Erkenntnis* **17**, 1982.
[386] Cf. Taylor, C., 'Formal theory in social science'. *Inquiry* **23**, 1980.
[387] Cf. (Ranta 1994).
[388] For a discussion of the implications of formalizations in philosophical logic, cf. (Hansson 2000) and (Jacquette 1994).
[389] Cf. Krysztofiak, W., 'Noemata and their formalization'. *Synthese* **105**, 1995.

formal tools were not known. Indeed, even though formalizations can be carried out on recent philosophical theories as well, in practice formalizations are often used as an instrument for the study of the history of philosophy and, in particular, of the history of logic.

Applications of formalization to past philosophical and logical theories are indeed done with a certain regularity, but at the same time such formalizations end up falling between two different fields of research, to wit historical studies vs. logic proper, while in some sense not belonging to either. Historians of philosophy often see formalizations as undue anachronism (cf. Hansson 2000, 170), and logicians tend to consider historical studies irrelevant for their work – after all, given the assumption of progress that is shared by many, current logic can only be more advanced than past logic. Nevertheless, interesting formalizations of ancient[390] and medieval[391] logic have been carried out that shed a different light on their objects of analysis. Hence, even though they may not be unanimously accepted as legitimate, formalizations in the study of the history of philosophy and logic are by now a well-established and respected form of investigation, albeit with a small number of practitioners.

4.1.2 Formal vs. formalized

At first sight, one might think that there isn't really any major difficulty in defining the concept of formalization. True enough, to formalize means, literally, to make formal, so it is obvious that what it means to formalize depends on what it means to be formal – and defining the concept of formal is not an easy task. However, given a fitting definition of formal, formalization could simply be defined as the process through which the (non-formal?) object of formalization is rendered formal. Hence, a formalization would involve the object to be formalized, an agent, the action, and the outcome of the process, as depicted below. The outcome is presumably the counterpart of the object being formalized which differs as little as possible from it, except for the feature of being formal. To be exact, I here use the term 'formalization' to refer to the very deed of formalizing, while it is common practice to use this term also when referring to the result of the deed; however, when ambiguity is a danger, I shall use terms such as 'outcome of a formalization' to refer to the formalized object. These notions are represented in Figure 4.1.2.

Thus set out, formalization seems to be a rather uncomplicated process. But this very characterization raises several issues, which indicate that this might not be a straightforward matter after all.

---

[390] Thom, P., 'Aristotle's Syllogistic', *Notre Dame Journal of Formal Logic* **20**, 1979; Martin, J.N. 'Proclus and Neo-Platonic Syllogistic', *Journal of Philosophical Logic* **30**(3), 2001; Martin, J.N., 'A tense logic for Boethius', *History and Philosophy of Logic* **10**, 1989; Corcoran, J., *Ancient Logic and its Modern Interpretations*, Reidel, Dordrecht, 1974; Corcoran, J. 'Completeness of an Ancient Logic', *Journal of Symbolic Logic* **37**, 1972.

[391] (Priest and Read 1977), (Spade 1978), (Bird, 1961), (Klima 1993b).

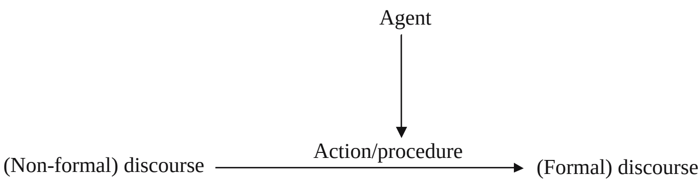


Figure 4.1.2. The act of formalizing

(1) IS THE OBJECT OF FORMALIZATION **MADE FORMAL** BY THIS TRANSFORMATION? OR IS THIS OBJECT ORIGINALLY FORMAL IN SOME SENSE, AND ITS FORMAL CHARACTER IS BROUGHT OUT OR **MADE EXPLICIT** BY THE TRANSFORMATION?

In other words, there seem to be two competing conceptions of formalizing, one according to which to formalize is to turn an informal object (theory, argument) into a formal one (by adding 'missing' premises, for example, cf. (Wang 1955, 227), (Hansson 2000, 166)), and another according to which to formalize is to make explicit the (implicit) formal character of the object of formalization.

In the first case, if formalization transforms its non-formal object in such a way that it becomes formal, one may wonder whether this is a legitimate transformation. If to be formal or non-formal is an essential feature of the object in question, this transformation is really more of a violent mutation (similar to what happens to the main character of the film 'The Fly', whose DNA is mixed with that of a fly, or to Gregor Samsa in Kafka's *Metamorphosis*). Suppose that an argument is not formal; if one turns it into a formal argument, by adding premises, for example, is it still the same argument, or even a legitimate counterpart thereof? There seems to be a good case for the claim that the product of such a transformation is in fact a transfiguration of the original object.

If, however, to formalize consists solely of making explicit the implicit but already present formal character of a given object, then a formalization becomes essentially the epistemic process of revealing something that is already there.[392] In other words, formalization does not radically change the nature of the object undergoing a transformation, but only brings out its formal character **to us**. Of course, if a formalization is an epistemic process, issues concerning subjectivity emerge: is formality a subjective concept, differently attributed to the same object by different people? This may seem a hard pill to swallow, in particular for those who rely on formality as a criterion for logicality.

Another view on this issue consists in separating the concept of formality from the concept of (logical) validity. That is, an argument or theory may be considered valid according to criteria other than formality. In this case to formalize would correspond to turning an informally valid argument/theory into a formally valid argument/theory. This is for example John Buridan's view on formal consequences (cf. Hubien 1976; section 2.2.5 of the present work): according to him, the validity of an already

[392] Along the lines of Brandom's notion of logic as 'making it explicit'. Cf. (Brandom 1994).

materially valid consequence is made evident to us by the transformation of this consequence into a formally valid one (usually by the addition of a necessary proposition as a premise). The problem is then to present a satisfactory criterion of validity that does not rely on the concept of formal; Buridan's own criterion is essentially modal, stressing the incompatibility of things being as the antecedent says they are with things being other than what the consequent says they are.

Alternatively, according to the notion of the formal as abstraction from content (see section 4.1.3 below), formalization can also be seen as the process by means of which one abstracts from content, in order to obtain the general patterns of the class of objects being formalized, beyond their individual specificities. Take a linguistic phenomenon as, say, questions, and investigations on the logic of questions (cf. Stokhof and Groenendijk 1997): the purpose of such an investigation is to identify patterns common to all occurrences of the phenomenon, beyond the specific meaning of each question. Thus seen, formalizations really correspond to generalizations by means of abstraction of content, and have a normative as well as a descriptive character (descriptive at first, by abstraction, and normative insofar as the patterns identified become the canon for the reproduction of the phenomenon). In this sense, a formalization would not consist in turning something informal into something formal, but rather in separating what is formal from what is not in the object of formalization.

That is to say: if the property of being formal is an attribute of objects, the transformation process of formalization may be illegitimate; but if it is seen as an epistemic attribute, lack of objectivity may arise. One way out of this dilemma is to view formality as having an epistemic import, but to distinguish it from validity and to view the latter as an essential (and thus objective) attribute (which, of course, should still receive a suitable definition). Another seemingly reasonable position is to view formalization as the process of separating form from matter – in this sense, everything is, to some extent, formal, insofar as everything is the combination of form and matter. At any rate, it is clear that the characterization of formalization as a process of making something formal raises difficult questions, and therefore that it is not as straightforward as it may seem.

(2) ARE THERE DEGREES OF FORMALITY?

The concept of formal can also be understood as a definite concept, with sharp boundaries, or as a relative attribute, that is, formalization-candidates (theories, arguments) are not formal as such, but rather more or less formal than other objects. If it is seen as a concept without sharp boundaries, then to formalize would mean 'to render it **more** formal', whereby the degree of formality would be increased. If a formalization is performed, the outcome of this process will be more formal than the original object of formalization, but this very outcome can in turn be the object of yet another formalization, and so forth. This is indeed what often seems to be the case in mathematics, as described in the passage below:

> Consider, for example, an oral sketch of a newly discovered proof, an abstract designed to communicate just the basic idea of the proof, an article presenting the proof to people working on related problems, a textbook formulation of the same, and a presentation of it after the manner of *Principia Mathematica*. The

proof gets more and more thoroughly formalized as we go from an earlier version to a later. [...] Each step of it should be easier to follow since it involves no jumps. (Wang 1955, 227)

One could take the formalization even further, and, from the presentation of the proof 'after the manner of *Principia Mathematica*', construct an even more detailed presentation of it using lambda-calculus, for example. At each version, more and more steps of the proof are made explicit, and therefore one could say that the proof becomes increasingly formal.

This is not, however, the picture that one usually associates with the concept of formal, which is typically understood as an absolute concept. At the same time, the idea that formality is a matter of degrees, and thus that a formalization increases the degree of formality of its object, has undeniable appeal. It seems to correspond to what formalizations actually bring about in most cases, and it also explains away the difficulties mentioned in the previous section, related to the essential vs. epistemic views on the notion of the formal. Indeed, if there are degrees of formality, raising the degree of formality of an object is not more problematic than, say, heating up water and thus augmenting its temperature (if it is an essential attribute), or else coming closer to a given object and thus being able to see it better (if it is an epistemic attribute).

Hence, the claim according to which to formalize means to make **more** formal is attractive not only insofar as it avoids philosophical complications, but also because it seems to correspond to what actually takes place in typical uses of formalizations. Wang (1955, 230) mentions that different levels of formality are also peculiar to different fields: the logician may think that the mathematician is much too careless in his practice, while the mathematician may find the foundational worries of logicians nothing but 'learned hair-splitting'. But mathematicians often abhor the sloppy use of the mathematical framework by physicists, who in turn reproach engineers. Hence, in the light of actual practice, that there are degrees of formality seems to be almost a truism.

(3) Is it possible to use the concept of formalization to define the concept of formal?

In section 4.1.3 below, we will see that to produce an apt and unified definition of what it means to be formal is far from being an unproblematic endeavor. At the same time, the crucial importance of such a definition cannot be overestimated.

Given this, one idea springs to mind: is it possible to reverse the roles of *explanandum* and *explanans*, and to take formalization as the primitive notion, on the basis of which the concept of formal can be given? This idea may seem counterintuitive at first sight, as the concept of formal appears to be more primitive than that of formalization. But if we have clear criteria of what is to count as a successful formalization – that is, the defining characteristics of the result of a successful formalization – then we may use the formalizing procedure almost as an empirical test for the degree of formality of a given piece of discourse.

More exactly, if we have an object (theory, argument) of which it is unclear whether it can be said to be formal, we may apply what are considered to be legitimate procedures of formalization to this object, and then evaluate the outcome. If it possesses the characteristics defining the result of a successful formalization, then we may say that the original object of formalization was already formal (in some sense). This holds in particular if the process of formalization is seen as augmenting the degree of formality of the object, since in this case an object of formalization must be in some sense formal, to start with, for the formalization to be successful. Similarly, if the process of formalization consists of 'making it explicit', according to Brandom's motto, then again a formalized object may make explicit the formal character that its non-formal version implicitly possessed. In sum, a formalization may make it possible to recognize the implicit formal character of the original object.

In practice, formalizations of past logical systems are often made with this purpose, among others, namely that of showing in what sense these past theories are logical also according to the current concept of what is to count as logical. This, for sure, is one of the goals of the present investigation.

(4) Is formalization a reversible process?

If formalization is the procedure by means of which an object is made more formal, and which can be iterated, one is naturally led to wonder whether, besides allowing for iteration, it is also reversible. In other words, what would be the converse of a formalization, and why would anyone want to perform it?

As noted by Ranta (1994, 7), computer scientists often make use of (formally defined, algorithmic) procedures that transform strings of formal notation into strings with the same meaning, but in more readable form. Such procedures are called **sugaring**. Since the languages used in computer science must be sufficiently detailed in order for the computer to perform the relevant calculations (by means of algorithms), these languages are often not very 'human-friendly', in the sense that a human agent cannot readily grasp the meaning of a given string. What the computer does is to manipulate 'meaningless' strings, so the role of these languages is to allow for computation; they are not intended to convey meaning in the first place. Hence, it seems natural that, from the highly regimented strings that are the outputs of computer calculations, one may want to pass to strings having the same meaning, where, for example, the symbols that seem superfluous to the human agent (but not to the computer) are deleted, and the order of the remaining symbols is changed so as to make it more easily interpreted by a human agent. Ranta (1994, 7) gives the example of the arithmetical term

$$_*(+(2, _*(3, 5)), 4)$$

which is sugared in the term

$$(2 + 3_*5)_*4.$$

That is, some parsing symbols (',' '(' and ')') have been deleted, and the order of the remaining symbols has been changed. The second term is clearly more readable than the first, but they both refer to the same calculation. In other words, sugaring often serves the epistemic purpose of clarification, insofar as full-blown formalized expressions often do a worse job at conveying meaning than their simpler versions.

Indeed, one of the main objections to formalization is that it frequently serves to make more obscure and complicated what was originally simple. At best, it is a

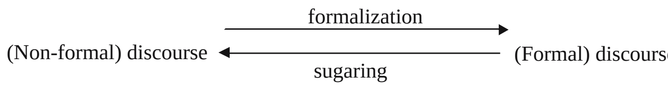


Figure 4.1.3. Formalization vs. sugaring

convoluted way of achieving a goal (presumably, that of conceptual analysis) that could be achieved with simpler means – in Wang's witty simile, it is like taking the airplane to visit a friend in the same city: one may want to do that, for the pleasure of flying the airplane, but this is certainly not the simplest way to achieve the goal of visiting the friend (Wang 1955, 233). While it is often the case that a formalization is made with the purpose of facilitating the communication of a proof or theory (cf. Wang 1955, 227), it may also happen that it only becomes more inaccessible and mysterious (e.g. to those not used to dealing with formal languages).

Hence, the converse of formalization would be the procedure of taking formalized strings into expressions having the same meaning but being less regimented (and more readily understandable), in particular into expressions of ordinary language. Computer scientists usually refer to such procedures as 'sugaring', which seems a fitting name for the general activity of 'undoing' a formalization, or else of rendering more intuitive strings that are written in a specific formal language (Figure 4.1.3).

But ultimately, whether a given formulation of an argument or theory is easily understandable or not remains essentially a subjective issue, as well as a matter of degrees. In the same way that a computer can only 'understand' expressions in their pre-sugaring form, a trained mathematician or logician may be much happier with the symbolic, formalized formulation of certain concepts than with their ordinary language counterpart, for reasons of rigor, brevity etc., whereas a student of logic may require a sugared formulation (in ordinary language or other) of a logical theory in order to develop the right intuitions. So the reasons for performing a sugaring seem essentially pragmatic, and one wonders whether the same holds of the reasons for performing a formalization. This possibility (among others) will be discussed below, when we turn to the three distinct notions associated with that of formalization. But before that, we must turn to the very notion of the formal.

4.1.3 The notion of the formal

An investigation of the concept of formalization obviously presupposes an examination, albeit brief, of the notion of the formal. For this reason, in this section I present a concise overview of some of the possible ways of understanding the concept of formal.[393] In practice, it is surprisingly difficult to provide a unified definition of formal.

[393] (MacFarlane 2000) is a valuable source of information on the different notions of formality. However, his goal is to find a definition of formality to serve as a criterion for logicality, whereas I am interested in the notion of the formal insofar as it (partly) determines the content of the concept of formalization. The central notion of the formal for MacFarlane is formality as generality/topic neutrality. I deliberately

Therefore, the list of proposals below is only meant to help us understand, in the next sections, how the different facets of formalization (axiomatization, symbolization and conceptual translation) are related to specific conceptions of 'formal'.

Two senses of formal seem particularly relevant for the present investigation: formal as regimentation (as in strict application of rules), and formal in the sense that form is opposed to matter. Variations of these two senses are also to be found in the literature, and some of them are worth examining in the present context.

(1a) The formal as regimentation. In everyday life, the most intuitive notion of the formal seems to be the one related to the notions of rules and regimentation. Typically, 'formal' is defined as accordance with established (albeit unwritten) rules, conventions and regulations: a person is formal if she follows the rules of etiquette; formal education is education that has been acquired according to the standard conventions and guidelines; a formal procedure is a procedure carried out strictly according to the appropriate rules etc.

At first, it may seem that this mundane sense of formal is not relevant for the present investigation, nor to logic in general. But isn't what distinguishes formal from so-called 'informal' logic the fact that, in formal logic, the rules governing inferential steps are made **explicit**, and that they must be strictly observed? It is also in this sense that a formal language is formal: a formal language is a set of strings recursively generated from a finite alphabet, by means of the application of rigidly defined rules, while in an 'informal' language the rules for the generation of sentences from words are presumably less rigid, with numerous exceptions. Moreover, a speaker of 'ordinary languages' is able to make sentences even if she is not aware of the 'rules' being applied, that is, if she hasn't learned grammar as such; in the case of informal languages, the rules of grammar are an attempt (usually only partially successful) to capture and model certain patterns tacitly used by the speakers.[394]

To equate formality with regimentation and accordance with rules may not seem very informative at first sight, as it concerns too many uses of 'formal' that are irrelevant for the present investigation. But, clearly, the role of explicitly stated rules is crucial for any formalization: to formalize an argument or theory is typically to make explicit the assumptions and rules used, so in this sense to formalize is to make explicit which rules are being complied with. Curry's comparison between a postulate system and a formal system is illustrative of this point:

> [The notion of a formal system] is quite a different notion from that of a postulate system, as naively conceived a half century ago. In the older conception a mathematical theory consisted of a set of postulates and their logical consequences. The trouble with this idea is that no one knows exactly what a logical

---

leave out the notion of the formal defined by the appeal to the 'general laws of thought' (MacFarlane's 1-formality). Important though it is, it is only tangentially related to the issues surrounding formalization, and therefore a discussion thereof falls out of the scope of the present investigation.

[394] Of course, some have gone as far as saying that these rules for sentence formation are in fact crucial for the mastery of an informal language, and that such rules must be in some sense innate. Compare for example, Montague's claim that English is a formal language (cf. Montague 1970).

consequence is […]. In the modern conception this vague and subjective notion is replaced by the objective one of derivability *according to explicitly stated rules*. (Curry 1957, 1) (emphasis added)

Hence, it is clear that the sense of formal as accordance with (explicit) rules is also essential for logic in general and for formalizations in particular, as it is at the core of the concepts of formal language and formal system.

(1b) Algorithmic notion of the formal. A theoretical counterpart (in particular in logic and mathematics) of the notion of the formal as regimentation, as strict rule-following, is the idea of algorithmic, mechanical reasoning. According to this view, a theory is formal, or is formulated in a formal way, if inference-making within the theory can be carried out by a machine (e.g., a computer) by the blind application of the relevant rules. Gödel says of the rules of inference in his system that they are:

… purely formal, that is, refer only to the outward structure of the formulas, not to their meanings, so that they could be applied by someone who knew nothing about mathematics, or by a machine. (Gödel 1995, 45)

What is important is that no subjective intuition or extra knowledge be required for proving theorems inside a formal theory (this is also one of the motors behind Frege's formulation of his *Begriffsschrift*, cf. (Frege 1879) and section 4.3.2.1 below), which must thus be entirely self-contained; the theorems should simply follow from the mechanical, blind application of its principles (clearly, this process ought to yield the same results independently of the agent carrying it out). For this purpose, artificially created languages, displaying a high degree of regimentation (but not necessary meaningless), tend to be more suitable; indeed, the vision of a calculating, artificial language dates at least as far back as Leibniz[395] and has been an important source of inspiration for the development of mathematical logic.

This may appear to be too strict a notion of the formal, as if formal reasoning was only a matter of algorithms and computation,[396] but at the same time it is very appealing insofar as it seems to give a smooth account of issues concerning subject-independence, effectibility and tractability of knowledge. It may not be a necessary, but it does seem to be a sufficient condition for formality.

[395] '[Everything that is true about the things that are expressible in the language obtained by assigning characters to the primitive notions] can be demonstrated by a calculus alone, or by merely manipulating the characters according to a certain form, without any effort of the imagination or the mind, in a word just as it is done in arithmetic and algebra'. (Leibniz, VE 195)

'Therefore, if someone is sure to have encompassed most of the more customary primitive notions by the Alphabet, it is therefore certain that most of the truths one may need can be proved by a mere calculus'. (Leibniz, VE 196). Both passages quoted in Maat 1999, 260/1.

[396] Indeed, it has been argued that the idea of mechanical computability cannot account for all the aspects of mathematical practices, and if mathematics is viewed as the formal science par excellence, one can easily infer that there must be more to formality than mechanical reasoning. In fact, it all seems to indicate that 'the class of mental procedures is not exhausted by mechanical ones' (Sieg 1994, 71), and, thus, that we are more than computers performing calculations even when we perform strictly formal reasoning.

(2a) The formal as structure and abstraction from content. Stemming from the old opposition form vs. matter – traditionally viewed as the two constitutive aspects of all entities, following the Aristotelian account of substances[397] – is the view that formality corresponds to abstraction from matter. In particular, it amounts to the opposition structure vs. content, when applied to entities such as theories and arguments:

> Traditionally, (formal) logic is concerned with the analysis of sentences or of propositions and of proof with attention to the *form* in abstraction from the *matter*. This distinction between form and matter is not easy to make precise immediately […]. (Church 1956, 1)

That is, of itself, this notion of the formal still does not tell us exactly what it means to be formal – what exactly are the formal features of an object, as opposed to its matter – since the distinction between form and matter is not easily formulated. Nonetheless, many of the (purported) different notions of formal, implicitly and explicitly adopted in philosophical discussions, are in fact variations of the form vs. matter opposition.

Logic, for example, can be said to be formal insofar it concerns essentially the (inferential) relations between its objects (propositions), and not their matter, that is, the content of each of them (their meanings or their truth-values). In other words, logic is concerned with inferential structures.[398] It is in this sense that axiomatization – the organization of the different statements of a theory in a defined structure – is such an important formalizing tool. Even when it comes to the sentential level, logic is not concerned with specific meanings and contents, but only with the logical forms of sentences, and primarily insofar as these logical forms determine the relations of inference between sentences. Similarly, according to the view known as structuralism in mathematics, mathematics is formal because it only deals with unspecified abstract structures, and not with the contents of specific operations such as addition and multiplication.[399]

(2b) 'Absence of meaning' notion of the formal. Connected to the idea of formality as abstraction from matter[400] is the view that, in formal theories such as logical and

---

397 'Originally, indeed, matter and form are introduced as twins: substances are in a sense composite entities, their component 'parts' being matter and form. And originally, matter and form are simply stuff and shape: a bronze sphere – Aristotle's standard example – is an item composed of a certain stuff, namely bronze, and a certain shape, namely sphericality'. (Barnes 1995, 97)

398 True enough, the view according to which logic **is** about the truth-values of propositions is just as influential. But on this matter I side up with those who focus on inferential relations as the main object of logic.

399 'What really distinguishes this new view of pure mathematics from the more traditional view is not so much the emphasis on relations between elements, as opposed to the elements themselves […] as the demand that these relations be capable of being made explicit without any appeal to spatial or temporal intuition. Pure mathematics, as conceived by contemporary "structuralists" in the philosophy of mathematics, concerns only *pure* or "freestanding" structures […]'. (MacFarlane 2000, 12)

400 This notion is also closely related to the algorithmic notion of the formal (1b), as can be inferred from the following passage by Kleene: '[…] it should be possible to perform the deductions treating the technical terms as words in themselves without meaning. For to say that they have meanings necessary to the deduction of the theorems, other than what they derive from the axioms which govern them, amounts

mathematical theories, symbols must be treated as mere objects, and not as meaningful expressions, as a way of achieving abstraction from content – that is, their form in its most extreme, typographical aspect. In other words, according to this concept of formal, 'logic can be treated purely syntactically, without reference to the meanings of expressions' (MacFarlane 2000, 31). This view is typical of the metamathematical turn that occurred roughly in the 1920s and 1930s, whose main motor was Hilbert's school in Göttingen (cf. (Hilbert 1927), (Awodey and Reck 2002, 21), and (Tarski 2002, 176) on the 'initial success' of this approach). Prominent logicians such as Carnap, Tarski and Gödel were proponents of this view on formality:

A theory, a rule, a definition, or the like is to be called formal when no reference is made in it either to the meaning of the symbols (e.g., the words) or to the sense of the expressions (e.g., the sentences), but simply and solely to the kinds and orders of the symbols from which the expressions are constructed. (Carnap 1937, 1)

… in constructing a deductive theory, we disregard the meaning of the axioms and take into account only their form. It is for this reason that people when referring to those phenomena speak about the purely FORMAL CHARACTER of deductive sciences and of all reasonings within these sciences. (Tarski 1959, 128)

To Hilbert is due […] the emphasis that strict formalization of a theory involves the total abstraction from the meaning, the result being called a *formal system* or *formalism*[401] (or sometimes a *formal theory* or *formal mathematics*). (Kleene 1971/51, 61/2)

No one needs to be reminded of the important results obtained within the meta-logical tradition – which is still very influential – attesting to its fruitfulness; nevertheless, this approach has also been criticized for jettisoning the meaningfulness of these logical languages. Thus seen, these 'languages' are strictly speaking, no longer languages at all (cf. (van Heijenoort 1967), (Sundholm 2001), (Sundholm 2003)).

Be that as it may, this notion of the formal is certainly one of the main motivations behind formalization. If it is assumed that formal theories must deal with meaningless symbols and with the relations between them, then it seems very convenient, or even imperative, to dispose of meaningful 'ordinary' words in favor of artificial symbols. The extent to which this is the case will be discussed in the section focusing on symbolization.

(2c) Variational[402] notion of the formal. Applied to sentences, the form vs. matter dichotomy results in the notion of (logical) 'form' of sentences. Indeed, the much-discussed medieval distinction between categorematic and syncategorematic terms is a precursor of this notion: categorematic terms were those that had signification taken

to saying that not all of their properties which matter for the deduction have been expressed by axioms'. (Kleene 1951, 59/60). In practice, viewing the terms of a theory as meaningless may be seen as a way of testing whether purely mechanical inference-drawing within the theory is possible.

401 Notice, however, that the term 'formalism' is not due to Hilbert: 'He *never* used this word to describe his position but always spoke of "axiomatics", and, especially, "metamathematics" and "proof theory" The mis-reading of his philosophy as treating mathematics as mere marks on paper dates from the 1920s onwards and is still thriving happily […]. The term "formalism" was put onto Hilbert by L.E.J. Brower from 1927; it was probably meant as a criticism and was meant to be so when they had a famous row in the late 1920s'. (Grattan-Guinness 2000, 162)

402 I borrow the term 'variational' from Bolzano, cf. (Sebestik 1992, chap. 3). Etchemendy (1990) uses the term 'substitutional' for roughly the same notion, and MacFarlane (2000) uses the term 'schematic'.

in isolation, while syncategorematic terms were terms such as ‘every’, ‘some’, ‘not’, ‘is’, which only contributed to the meaning of sentences. The latter terms constituted the form of a sentence, while the categorematic terms its matter (what it was about). Based on this distinction, it was common practice – in fact since Aristotle, in the *Prior* and *Posterior Analytics*[403] – to represent the form of a sentence by replacing its categorematic terms by schematic letters, yielding schemata (cf. Corcoran 2004). Aristotle does not explicitly relate his use of schematic letters (to represent the different syllogistic patterns) to the notion of the formal, but it is clear that what makes a specific syllogism valid or invalid is its being a variational instance of valid/invalid syllogistic patterns.[404]

In the 14th century, Buridan defined the notion of formal consequence (as opposed to material consequence) precisely in terms of the variation criterion; a formally valid consequence is a materially valid consequence of which all substitutions of its categorematic terms by other terms of the same kind yield materially valid consequences.

> ‘Formal’ consequence means that [the consequence] holds for all terms, retaining *the form common to all*. Or, if you want to express it according to the proper force of discourse, a formal consequence is that which, for every proposition *similar in form* which might be formed, it would be a good consequence, such as ‘what is A is B; thus what is B is A’.[405] (TC 22/23, 5–9) (my emphasis)

Therefore, he is perhaps the first to have explicitly associated the notion of the formal to that of variation, producing the variational notion of the formal. The same basic idea is found in Bolzano’s logic of variation (cf. Sebestik 1992, chap. 3) and in Tarski’s notion of logical consequence.[406]

But the variational notion of the formal has a few drawbacks; one of them is the requirement of a sharp distinction between those parts of sentences to be substituted and those that constitute their ‘form’ and thus are not to undergo variation – the notoriously thorny problem of the boundary between logical and nonlogical terms (cf. Tarski 2002, 188). Moreover, as noted by Tarski, the notion of logical (formal) consequence defined in variational terms is only a necessary but not a sufficient condition for the validity (i.e., formal validity) of consequences, since, for it to hold as a sufficient condition as well, one must presuppose that the language under consideration contains names for all objects (Tarski 2002, 185).[407] This is one of Tarski’s main motivations for the introduction of his object-oriented variational notion of the formal, which corresponds to the notion of the formal as indifference to particular objects discussed below.

---

403 Aristotle makes use of schematic letters throughout these texts, for example (*Prior Analytics*) 68^{a}, 59^{b}; (*Posterior Analytics*) 86^{b}, 87^{a}, 80^{a}, among many others. See also (Corcoran 2003).
404 MacFarlane notices that, while he is the inventor of logic as we know it **and** of hylomorphism, Aristotle did not invent logical hylomorphism. (MacFarlane 2000, Appendix A).
405 Consequentia ‘formalis’ uocatur quae in omnibus terminis ualet retenta forma consimili. Vel si uis expresse loqui de ui sermonis, consequentia formalis est cui omnis propositio similes in forma quae formaretur esset bona consequentia, ut ‘quod est A est B; ergo quod est B est A’.
406 In fact, with Tarski we have the passage from the schematic notion of the formal to the notion of the formal as indifference to particular objects. Cf. (Etchemendy 1990, chap. 3), (Tarski 2002), (Simons 1987).
407 See also (Read 1995, 41) for shortcomings of the variational notion.

Nevertheless, the variational notion remains an intuitive and tractable approach to formality; it is also an important motivation for formalization (in its symbolization facet), since the use of schematic letters is a very convenient manner of expressing which parts of a sentence are to undergo variation. As a matter of fact, it is clear that this notion of the formal is at the heart of the first efforts towards formalization in the history of logic, that is, the use of schematic letters.

(2d) The formal as indifference to particular objects. The notion of the formal based on the form vs. matter dichotomy can also receive a precise formulation with a switch of perspective from the linguistic level to the objectual level. On this level, formality does not concern the forms of expressions of a language, but rather the forms of the very objects referred to by the language of a theory. Again according to the original Aristotelian view, the form of an object corresponds to its essence, and thus to its species, which it shares with all other objects pertaining to the same species. To be exact, while matter can be understood as content and meaning (on the level of linguistic entities), it can also be understood as what individuates objects sharing the same form, the materiality that gives each individual its numerical identity; if form and matter are thus understood, formality amounts to indifference to the identities of objects numerically distinct.

Even though the distinction between form and matter of **objects** is almost as old as philosophy itself, this distinction does not seem to have been used in connection with the concept of formal in logical contexts before[408] Tarski's use of the notions of satisfaction and model to define the concept of logical (formal) consequence.[409]

An arbitrary sequence of objects which satisfies each sentential function of the class L′ we shall call a model of the class. (Tarski 2002, 186)

Note the term 'arbitrary'; indeed, with respect to satisfaction and models, all that matters is whether a given object or a given sequence of objects satisfies a sentential function; other peculiarities of such objects are irrelevant. In virtue of this, within a class of objects that satisfy a given sentential function, permutation is allowed with respect to this sentential function – presumably in virtue of the fact that these objects all share a common form. In other words, formality is also defined as indifference to specific objects, or as concern with unspecified objects:

[…] a formal system …is defined as a body of theorems generated by objective rules[410] and *concerning unspecified objects*. (Curry and Feys 1958, 12) (emphasis added)

---

[408] But hints of it are already to be found in Frege's *Begriffsschrift*, contrary to what MacFarlane claims (see next note): 'The most reliable way of carrying out a proof, obviously, is to follow pure logic, a way that, *disregarding the particular characteristics of objects*, depends solely on those laws upon which all knowledge rests'. (Frege 1879, 5) (emphasis added). Notice also the appeal to the notion of 'laws of thought' as what is distinctive about logic.

[409] MacFarlane (2000, 243) argues that this notion of formality 'is of little importance historically in the demarcation of logic'. In particular, he claims that Kant did take logic to be formal in the sense of indifference to specific objects, but in this sense logic was no different from arithmetic and algebra. By contrast, Frege simply did not take logic to be formal in this sense.

[410] That is formality as regimentation.

Currently, the notion of the formal as indifference to particular objects is very influential, in particular in the shape of **permutation invariance**. 'Permutation invariance can be regarded as a precise technical gloss on the idea of indifference to the particular identities of objects'. (MacFarlane 2000, 59). Indeed, permutation invariance is now often thought to be **the** distinctive trait of formality, accompanying the model-theoretic turn in logic of the last five or six decades. But this view has its limitations as well, as shown by Etchemendy's (1990) critique of the model-theoretic notion of logical consequence, and by MacFarlane's contention that many hidden and untractable assumptions are in play in the use of permutation invariance as a defining criterion for formality and logicality (MacFarlane 2000, 6.8). In other words, on the philosophical level, the extent to which the notion of permutation invariance is really the key for the correct definition of formal is still an ongoing debate; but in practice, it is one of the most distinctive traits of actual practices in the formal sciences, especially in logic.

Hereby I end my list of different notions of formal. I do not wish to present any of these alternative notions as the only correct notion; rather, each notion plays a motivating role for the specific practices of formalization that will be discussed below, and therefore must all be borne in mind. My next step now is to investigate how these practices actually serve the purpose of making their objects (more) 'formal', following these notions.

4.2 AXIOMATIZATION: STRUCTURING[411]

The axiomatic method is a thoroughly explored topic in the history and philosophy of mathematics – and deservedly so, since axiomatization is among the central notions of current mathematics and of the formal sciences in general. Indeed, its introduction is a turning point in the recent history of these sciences. True enough, understood in a broad sense, the axiomatic method goes as far back as Euclid's *Elements*, when it essentially consisted of a method to organize the concepts and statements of a given theory, or in a given discipline, in such a way that, from more primitive concepts and statements, other statements could be deduced, enhancing clarity and certainty.

But the end of the 19th century witnessed the emergence of what is now called formal axiomatics, with Dedekind's and Peano's axiomatizations of the natural numbers and elementary arithmetics, and Hilbert's axiomatization of geometry. The purpose of formal axiomatics went beyond that of 'plain' axiomatics:

> In formal axiomatics the purpose is not primarily to increase certainty, nor is it merely to clarify and organize the concepts and theorems of a mathematical discipline in a systematic way. Rather, an additional aim is to treat the objects of mathematical investigation more abstractly, and then to characterize them *completely* […]. (Awodey and Reck 2002, 5) (my emphasis)

In other words, as it is now understood, the axiomatic method does provide a way to organize the concepts and statements of a theory or discipline (mathematical or other), but, more importantly, it yields a *complete* characterization of the latter. Thus

---

[411] This section is largely inspired by (Awodey and Reck 2002).

seen, it is natural that the appeal of this method should go beyond its application in mathematics, and that, in view of its great achievements in mathematics, attempts were made to transpose this method to other disciplines, to philosophy in particular.

In effect, when formalizing a theory, one typically starts with its basic assumptions, its axioms. Thus, an analysis of the notion of formalization would remain incomplete without an examination of the concept of axiomatization. Therefore, in this section, I present a concise analysis of the notion of axiomatization; my aim is to investigate the extent to which to axiomatize is to render (more) formal.

4.2.1 Axioms and rules of transformation

An axiom is generally thought to have two distinctive characteristics: (i) it is a proposition more fundamental than others, on the basis of whose truth the truth of other propositions can be known; (ii) it is a proposition whose truth is accepted without proof. Each of these aspects deserves further analysis.

(i) Axioms are the pillars on the basis of which knowledge is construed. This can be understood in absolute terms, with respect to knowledge in general, or in relative terms, with respect to a specific theory or field. In the first case, axioms would correspond to the basic principles guiding all intellectual activity, and can receive an ontological or epistemological foundation – if viewed ontologically, axioms should describe the essential principles organizing the being of things; if viewed epistemologically, they would correspond to the essential principles organizing our knowledge.

It is not, however, the absolute sense of the concept of axiom that is relevant for the present investigation. Here, we are interested in axioms of specific theories or domains of knowledge (and from now on I will always use the term ‘axiom’ understood as relative to a specific theory, unless otherwise noted). Dedekind gives an apt description of what axioms are, in the context of a specific domain, the theory of natural numbers:

> What are the mutually independent fundamental properties of the sequence N, that is, those properties that are not derivable from one another but from which all others follow? (Dedekind 1890, 99/100)

When these fundamental properties are established and formulated in propositions, the latter will be the axioms of the theory. Thus seen, axioms are equivalent to postulates or assumptions, that is, to principles underlying all other claims within the theory, but for which there is no proof (as shall be discussed below). There should be as many axioms as necessary to allow for the derivation of all statements valid within the theory (hence, all implicit assumptions must be expressed in the form of axioms[412]), but as few as possible so as to avoid redundancy, in such a way that no axiom can be derived from another – since, by definition, an axiom is a proposition not derivable from another proposition.[413]

---

[412] Frege notes the need to ‘prevent anything intuitive from penetrating here [in his reduction of the concept of ordering to that of logical consequence] unnoticed’. (Frege 1879, 5)

[413] ‘Fundamentally, we strive to arrive at an axiom system which does not contain a single superfluous statement, that is, a statement which can be derived from the remaining axioms and which, therefore, might be counted among the theorems of the theory under construction’. (Tarski 1959, 131)

(ii) Axioms are accepted without proof. Indeed, proofs of axioms are not only superfluous; they are in fact impossible.[414] Since, by hypothesis, axioms are the most primitive elements of a theory, there is nothing else on which their proofs could be based.

Now, one of the most important questions concerning axioms is then: what motivates their acceptance, if not proofs of their truth? In other words, what is the cognitive status of axioms? Two main views on the status of axioms suggest themselves, which one could call principled and pragmatic views.

Principled/unitarian views. Axioms describe essential properties, either of objects or of thought. According to such views, in each case, it is not any set of axioms that would fulfill a foundational role, but only those specific, uniquely defined axioms describing these essential properties. Principled views on axioms are usually of two kinds:

– Realist views, according to which axioms should describe the fundamental properties of the **objects** under consideration, the ur-truths about them (cf. Detlefsen 1992, 2). According to these views, if one is concerned with foundations for knowledge in general, axioms should correspond to the basic principles of organization of being as such, that is, the most fundamental ontological principles (in the sense of Aristotle's *Metaphysics*, for example). If relative to a given theory or domain of knowledge, axioms should describe the essential properties of the objects which the theory is about; on this approach, the axioms of set-theory, for example, should describe the very properties of sets, understood as self-sufficient entities.[415]
– Idealist views, according to which axioms should describe the fundamental properties of our understanding.[416] According to these views, axioms (taken in the absolute sense) ought to describe the laws of thought as such, as in Kantian

[414] Of course, there are cases in which axioms are derived, such as when neo-logicists attempt to derive the Peano axioms. But in this case the point is precisely that the Peano axioms are no longer axioms properly speaking, that is, the most primitive statements of a theory, precisely because the underlying assumption is that mathematics can be derived from logic, and thus that the most primitive mathematical statements – its axioms – are not mathematical statements (such as the Peano 'axioms'), but rather logical statements.

[415] For instance, Whitehead and Russell defended a realist interpretation of the axioms of *Principia Mathematica*: 'The axioms were intended to be believed, or at least to be accepted as plausible hypotheses concerning the world'. (Kleene 1951, 45). For this reason, the axiom of reducibility was for them clearly not 'the sort of axiom with which we can rest content' (in the introduction to the second edition (1925) – quoted in Kleene 1951, 45), since to postulate the existence of properties instead of construing them seemed to be an abuse of the purpose of axioms. But with this axiom included in their system, the right results seemed to follow, so it was accepted as an axiom in spite of the malaise felt by the realists Russell and Whitehead.

[416] '… [This position] has a less metaphysical, more purely epistemic foundational thrust. It too searches for ultimate truths, but not in the sense of final metaphysical grounds of truth. Rather, it seeks to find premises that are epistemically ultimate – truths, that is, that are (at least in certain respects) epistemically unsurpassed'. (Detlefsen 1992, 2)

idealism, or (in the relative sense) the basic principles of our thinking within a particular theory or domain of knowledge, such as with intuitionism in mathematics (cf. (Brower 1923), on the principle of excluded middle).

In both cases, the acceptance of an axiom is usually attributed to its being evident – its truth is evident, indeed self-evident; it cannot rely on anything else, as there is nothing more primitive and fundamental than what is described by an axiom. Moreover, its falsity is impossible, inconceivable. Descartes' *cogito* was intended as a paradigmatic case of a self-evident axiom, resisting all doubt, even in its most radical forms.

For a long time, the axioms of Euclidean geometry were considered as the paradigmatic case of self-evidence, corresponding to the only correct description of the real properties of space. In other words, the unique status of Euclidean geometry seemed to reflect that axioms must describe essential properties of reality and of thought, and that there is only one correct description thereof (and, thus, only one correct system of axioms). Indeed, what I call the principled view on axioms was unanimous well into the nineteenth-century, roughly for as long as Euclidean geometry was considered to be the paradigm of certainty. As has been described by (Nagel 1939) (see also (Torretti 2003)), the emergence of alternative geometries was one of the main reasons for the appearance of a different conception of the status of axioms and the axiomatic method. This revolution in geometry led to a different view on axioms, which will be discussed next.

Anti-essentialist/pluralist views. What distinguishes what I call pluralist views on axioms from the principled/unitary views is not so much the attribution (or not) of truth to them; rather, the fundamental difference seems to be that, following the principled views, there is one unique set of propositions describing the fundamental properties of the domain in question – presumably, on account of these fundamental properties being themselves uniquely determined. By contrast, a pluralist interpretation of axioms consists in accepting the possibility of alternative, equally legitimate sets of axioms for the very same theory or field of knowledge. Indeed, if there are no such things as essential, fundamental properties of objects or of reasoning, axioms can be taken merely as a starting point in the construction of knowledge; accordingly, there would be scope for choice among alternative sets of axioms. In a nutshell: the pluralist rejects the foundational view that there are uniquely determined truths on which all other truths are based.[417]

This pluralist approach to axioms does not imply that axioms are not true (or false, for that matter); a realist/pluralist may uphold that the axioms she chooses to build

[417] 'More recently, these traditional conceptions have given way to less foundationally oriented conceptions of mathematical justification. Axioms are no longer taken as giving the metaphysical grounds for theorems. Nor are they taken to be unsurpassably, or even unsurpassedly, evident. Rather, they are seen only as being evident enough (i) to make a search for further evidence seem unnecessary, or perhaps (ii) to remove sufficiently much of whatever rational doubt or indecision it seems possible and/or desirable to remove from the theorems, or maybe just (iii) to be attractively simple and economical, and evident enough to serve the purposes at hand, etc'. (Detlefsen 1992, 2)

his theory on are true indeed, but she will say that they are not more primitive than other truths within the theory, namely the theorems deduced from the chosen axioms. At any rate, their truth is not what justifies their status as axioms. For example, the statements of a theory can be related to one another in a circular way, in which case the theorist may simply choose certain statements to be considered as his starting point, following criteria such as simplicity and elegance. But, in this case, many of these alternative (sets of) statements could be taken as axioms: since, by hypothesis, the deductive structure of this particular theory is circular, from several alternative sets of its statements all other statements could be deduced.

In practice, though, the pluralist view on axioms is often associated to the idea that the truth of axioms is a non-issue altogether, as much as the truth of the theory as a whole. For example, according to an important trend partaking of this approach, all that can be proved is the consistency or inconsistency of the theory, and thus the main desideratum for a set of axioms is that it does not contain contradiction; hence, any consistent set of statements also satisfying other requirements, such as completeness, is a legitimate set of axioms for the theory. In other words, the purpose of axioms is not to describe the fundamental properties of objects, in fact it is not to describe (real) properties of objects at all, but only to serve as a starting point for the study of the inferential structure of the theory.

Hilbert's work is representative of this approach to axioms. His axiomatization of geometry (1899) had already introduced the possibility of 'tempering' with axioms:

> Hilbert availed himself of this feature [isomorphism] of axiomatic theories for studying the independence of some axioms from the rest. To prove it he proposed actual instances (models) of the structure determined by all axioms but one, plus the negation of the omitted one. (Torretti 2003)

In other words, Hilbert saw the interest in studying an axiomatic system for its own sake, and not only insofar as a given system is 'true', that is, describes real properties of objects.

Two decades later, after the new impulse given by the publication of *Principia* (cf. Zach 2003b), Hilbert returned to foundational issues and began to work on what is now known as 'Hilbert's program'. The emphasis was laid on proving the consistency of axiomatic systems; for this, axioms and statements were to be taken only with respect to their forms – strings of signs being the very objects of mathematical theory – disregarding their descriptive contents and thus their alleged truth or falsity (cf. Hilbert 1927, among others).[418] Notice that this is indeed quite a long way from Dedekind's quest for the most fundamental properties of natural numbers, especially insofar as no claim of exclusivity is made concerning any specific axiomatic system.

The pluralist approach to axioms (and to theories in general) is also often associated with instrumentalist views of knowledge: the aim of theories is not to describe how things actually are, but rather to be used as instruments for prediction and

[418] 'And in mathematics, in particular, what we consider is the concrete signs themselves, whose shape […] is immediately clear and recognizable'. (Hilbert 1927, 465)

explanation.[419] But again, this connection is not of necessity: pluralist views on axioms can be either instrumentalist or realist.

Naturally, in line with these pluralist views, the acceptance of axioms is not founded on their (self-)evidence; the criteria guiding the determination of a suitable set of axioms for a theory are typically 'pragmatic' criteria such as efficiency of deduction, simplicity and completeness. Factors such as the intuitiveness of an axiom may also be taken into account, but are typically overruled in favor of simply 'getting the right results'.[420] Most importantly, it is very unlikely for one specific set of axioms to establish itself as the only suitable set of axioms for a given theory or field; usually, there is scope for choice, depending on the (often conflicting) desiderata one wishes to comply with (see for example, the different axiomatizations of set theory).

What these considerations add to the present analysis is the idea that, when an axiomatization is undertaken, different criteria may guide the determination of the axioms of a theory. At one extreme, one may seek the axioms truly describing the most fundamental principles of the theory in question; at the other extreme, axioms can be seen as a mere starting point, expected to be no more than merely convenient and to allow for the 'right' results.[421]

Rules of Transformation. Hence, in first instance, to axiomatize is to determine a finite number of principles describing the fundamental[422] concepts of a body of knowledge. So far, we have discussed how these axioms are chosen. But this is still only half of the story; determining the suitable axioms is only a partial axiomatization of a theory. To use a game metaphor, it is like defining how a game must start without defining the subsequent moves of the game – obviously, more is needed to play the game. Indeed, just as important as the axioms are the rules of transformation/inference of a theory, permitting the deduction of theorems from the axioms established and thus paving the passage from axioms to theorems.

> [The rules of inference] describe the kind of transformations to which statements of this theory may be subjected in order to derive other statements from them; each definition has to be laid down in accordance with the rules of definition, and each proof must be COMPLETE, that is, it must consist in a successive application of rules of proof to sentences previously recognized as true. (Tarski 1959, 133)

As noted by Awodey and Reck (2002, 19), in the early days of formal axiomatics, a suitable (formal) account of the rules of transformation effecting the passage from axioms to theorems was still not available. There was awareness of the ins and outs involved in determining the axioms of a theory, but less so of how to treat

---

419 But notice that Hilbert's position goes in many respects beyond pure instrumentalism. (Cf. Zach 2003b)
420 See footnote on Russell and Whitehead above.
421 'It is important to realize the fact that we have a large degree of freedom in the selection of the primitive terms and axioms; it would be quite erroneous to believe that certain expressions cannot be defined in any possible way, or that certain statements can, on principle, not be proved'. (Tarski 1959, 130)
422 'Fundamental' here need not be understood in a heavy, metaphysical sense.

formally the inferential steps connecting the different statements of a theory.[423] Naturally, logicians and philosophers of different times had had a sharp interest in issues surrounding deductive inference (as far back as Aristotle and the Stoics, including medieval logicians such as Buridan, and 19th century logicians such as Bolzano and Boole). But the rules of deductive inference had not yet received the same kind of regimented treatment as that given to axioms by mathematicians such as Dedekind, Peano and Hilbert.

Even though the first formalized treatment of deductive inference fully meriting this title is probably Frege's *Begriffsschrift*, its impact was not immediate. In fact, it was only with *Principia Mathematica* that the practitioners of axiomatizations such as Hilbert and Carnap became convinced of the need for a formal approach to the rules of inference of theories as much as for axioms (cf. Awodey and Reck 2002). With *Principia*, rules such as *Modus Ponens*, disjunctive syllogism etc., which had been known as sound for many centuries, received a uniform and formalized treatment.

Beyond these historical details, it is clear that a full axiomatization of a theory requires not only that its axioms be determined, but also that the legitimate moves from the axioms to the truths of the theory be specified. The question is now: are there alternative sets of rules of inference, as much as there may be alternative sets of axioms for the same theory? Logical universalists maintain that there is only one logic, and thus one correct system of inference, insofar as logic is topic-neutral and permeates all valid reasoning regardless of what it is about. By contrast, logical pluralists accept the possibility of more than one legitimate system of inference,[424] but in two different ways. (1) On the one hand, some may hold that, as much as there are different sets of axioms according to different fields and theories, there are different systems of inference proper to each of the latter, but that to each field or theory there corresponds only one sound system of inference. (2) On the other hand, others may hold that for one and the same theory there is more than one suitable system of rules of inference, and that the choice of a deductive system may be guided by the same pragmatic reasons that guide the choice of axioms.

What is relevant here is that the axiomatization of a theory or field of knowledge involves not only the choice of suitable axioms, but also the choice of suitable primitive rules of inference. Again, even if one maintains that there is only one correct system of inference for each domain, there remains the task of actually establishing, for each case, which system is the right one. The opposition classical vs. intuitionistic logic with respect to mathematics is a good example of conflicting opinions concerning the deductive system to be used in the axiomatization of a specific field. Moreover, within a more radical version of logical pluralism, the choice of deductive system (among the various *prima facie* equally legitimate options) for a given axiomatization will not be guided by principled arguments, but rather by pragmatic considerations.

It must also be noted that the boundary between axioms and rules of inference can be rather vague. Certain rules of inference can be expressed under the form of

[423] 'From a contemporary point of view the main ingredient missing in the works considered so far is a precise and purely formal notion of deductive consequence'. (Awodey and Reck 2002, 19)
[424] Cf. (Restall 2002).

conditional propositions, which can be taken as axioms. For example, the rule:

$$\frac{A \,\&\, B}{A}$$

can be roughly expressed by the proposition $A \ \& \ B \ => \ A$, which can be taken as an axiom in case the rule above is to be taken as valid. Still, a theory with only axioms and no rules of inference cannot take off the ground, so to say, because axioms do not encompass the **instruction** to actually perform the passage from premise to conclusion. Therefore, a theory may be formulated with many axioms, but it must always have a minimal deductive system; a typical candidate is a deductive system composed exclusively of *modus ponens* (which amounts to the instruction of, having $A \ => \ B$ and A, effectively moving to B).[425] Indeed, a theory with axioms but no rules of inference is like a recipe for a dish never executed for lack of initiative.

Similarly, a theory may have a complex deductive system, but it must also always feature some kind of axiom system, minimal as it may be,[426] otherwise there is no starting point for the performance of inference-making within the theory (unless axioms are considered as zero-premise rules of inference). A theory with rules of inference but no axioms is like a recipe for a dish never executed for lack of ingredients.

Another point worth mentioning is that it has become common practice, when axiomatizing, to formulate axiom-schemata rather than axioms, since the use of axiom-schemata (where the non-specific vocabulary is represented by schematic letters or variables[427]) renders the problematic rule of substitution superfluous.[428] But for the present purposes, the use of axiom-schemata rather than axioms properly speaking does not interfere with the fundamental aspects of axiomatizing as discussed here (the use of schematic letters and variables will be discussed in more detail in the section on symbolization).

Be that as it may, underlying any axiomatization project is the assumption that knowledge has an essentially deductive nature: it is by inferring new knowledge from previously known (and possibly more fundamental) truths that knowledge in general is expanded. To axiomatize would then consist in (re)-organizing a theory in its due structure, its proper conceptual shape, which may have gotten lost at some point during its construction. Following this assumption, to axiomatize is to provide a theory with its most intuitive and correct formulation, that is, its proper deductive structure.

---

[425] Examples of theories with several axioms and with a minimal deductive system are the modal logics S4 and S5, as formulated in (Hugues and Cresswell 1968), which have only *modus ponens* and the rule of Necessitation as deductive rules.

[426] An example of a theory with a rich deductive system that is parsimonious with axioms is Gentzen's sequent calculus, where the only axioms (all derived from the same axiom-schema) are sequents of the form $A => A$.

[427] Concerning the difference between schematic letters and variables, see (Corcoran 2004).

[428] 'In early 20th century formalizations of logic, it was common to use a substitution rule and a finite set of axioms instead of schemata. Church (1956, 158) credits von Neumann with "the device of using axiom schemata", which rendered the (notoriously difficult to state) substitution rule unnecessary'. (Corcoran 2004)

However, objectors to these claims may hold that knowledge is **not** essentially deductive, in any case not in such a straightforward way, articulating more fundamental truths to other truths in a smooth and systematic manner – that is, they may question the Aristotelian foundationalist picture of knowledge, as presented in the *Posterior Analytics*. For one thing, circularity may occur, as already mentioned, in which case the establishment of 'more fundamental' truths within the theory is really quite arbitrary. Moreover, the net of relations between different statements of a theory may resemble an intricate tangle rather than a neat web. Even more radically, one may argue that knowledge is not exclusively deductive, and that some propositions (that are not axioms) are just 'known' without having been deduced from others.[429] Still, the advocate of axiomatizations may reply that, even if not always the case, every respectable theory ought to have a clear deductive structure underlying it; this seems to hold at least in fields such as logic and philosophy, which are the ones we are concerned with now.

4.2.2 Why axiomatize?

Even if one grants that an axiomatization is supposed to give a theory its correct deductive formulation, one may still wonder: what exactly can be accomplished once this formulation is established? If theories can be used with success in their non-axiomatized forms, which may even be more intuitive than their regimented and rather artificial axiomatized counterparts, what are the benefits of axiomatizing? As I see it, there seem to be two main kinds of advantages in axiomatizing a theory, besides the obvious one of giving it an organized and structured formulation: to achieve completeness and to allow for a meta-perspective.

4.2.2.1 Completeness

From the outset, the quest for completeness was one of the main motivations (if not the chief one) behind the development of formal axiomatics at the end of the 19th century. Indeed, completeness is what differentiates Euclidean informal axiomatics from the more recent, formal approach to the axiomatic method. But for this claim to be informative at all, a suitable definition (or definitions) of completeness must be given;[430] an axiomatized theory is said to be 'complete' only with respect to (implicit or explicit) specific goals.[431] Some of these goals may be:

– that all the true propositions of the theory be deduceable from its foundations (deductive completeness);

[429] Cf. some forms of reliabilism, namely the view that true knowledge is attained by reliable cognitive process, but not necessarily by means of inferences (cf. R. Brandom 1998, 'Insights and blindspots of reliabilism', *Monist* 81(3), 371–392).

[430] Awodey and Reck (2002) list 6 independent definitions of completeness.

[431] 'In general, notions of completeness arise in contexts where axiomatizations are being undertaken with specific goals in mind. To say that an axiomatization is complete is, then, to say that the axiomatizers have achieved their goal, in particular that no further addition of "new axioms" is called for'. (Awodey and Reck 2002, 5)

– that all presuppositions involved in inference-making within the theory be made explicit;
– that for every proposition of the language in which the theory is expressed, either itself or its contradictory be a theorem of the theory.[432]

These formulations of the notion of completeness, as well as others not mentioned here (e.g., those mentioned in (Awodey and Reck 2002)) typically feature some sort of higher-order universal quantification (in the cases above, over assumptions and propositions). In effect, in its most intuitive sense, the term 'complete' is related to the notion of totality, which is very naturally expressed by universal quantification.

But how is higher-order quantification over a theory possible? Most (if not all) non-trivial theories are not finite, in that all the propositions held to be true by the theory cannot be listed in a finite enumeration. Typically, a theory consists of a general framework applicable to infinite particular cases (of the suitable kind), and each possible (not merely actual) instance of application is part of the theory. How can one achieve any kind of completeness with respect to 'infinite entities' such as theories?

What is required is a finite, and thus manageable, way of describing and characterizing infinite theories. In other words, one needs a finite set of instructions allowing for the generation of infinitely many entities – that is, the truths of a theory. Formal axiomatics provides precisely this finite method for generating and describing infinity. If the axioms and rules of inference are sound, the successive application of the rules to the axioms can inductively generate the (potentially infinitely many) theorems of the theory. In a fully axiomatized theory, its theorems form an inductive class.

In effect, the fact that the first serious attempts at (formal) axiomatizations concerned the natural numbers, the case par excellence of an enumerable but infinite structure, is very illuminating. The most instinctive method for describing a structure such as that of natural numbers is, obviously, the plain enumeration of all its members (in the correct order), and their respective properties; but, just as obviously, complete enumeration can thereby never be achieved, let alone 'completeness of proofs', as mentioned in the passage by Dedekind quoted above. But the same structure of natural numbers is described (but not completely, as we know from Gödel) by very succinct sets of axioms, for example, the Dedekind-Peano axioms (a set of four axioms – cf. Peano 1889). And if the axioms and rules of inference of a given axiomatization (of natural numbers or other) are sound, then all the logical consequences of the axioms with respect to the relevant rules of inference shall be truths about its objects, insofar as they are theorems of the theory.

The quest for finitism by means of the axiomatic method has reached its peak with Hilbert's finitary project:

> This methodological standpoint consists in a restriction of mathematical thought to those objects which are 'intuitively present as immediate experience prior to all thought', and to those operations and methods

[432] Cf. (Tarski 1959, 135).

of reasoning about such objects which do not require the introduction of abstract concepts, in particular, without appeal to completed infinite totalities. (Zach 2003b)

Some of the facets of this project are the emphasis on **signs** as the objects of mathematical reflection par excellence (on account of their finitism), and finitist methods such as epsilon calculus (Avigad and Zach 2002, Zach 2003a).

True enough, while some mathematical theories can be finitely axiomatized in a first-order language, others cannot, for example, set-theory and number theory. In the latter cases, the axioms of these theories must be formulated as axiom-schemata, or in a second-order language, since higher-order predicates are required (cf. (Corcoran 2004), (Shapiro 1991), (Awodey and Reck 2002, 21)). Nevertheless, insofar as completeness amounts to the idea of finitely many axioms describing potentially infinite structures, regardless of the order of their terms, such axiomatizations do reach some sort of completeness, thus understood.

Of course, to support the sweeping claim that axiomatization ensures completeness, one would have to show, for each specific notion of completeness, in which way a given axiomatization paves the way from finite sets of axioms to universal quantification over (presumably infinitely many) higher-order entities. Such an inquiry belongs to the meta-level of investigation, to be discussed in the next section. Moreover, some axiomatizations are not intended to ensure completeness at all (at least not according to some relevant notion of completeness);[433] in other cases, completeness (under specific formulations) simply cannot be achieved by axiomatization, in virtue of the very nature of the systems or theories in question, as Gödel's theorems have shown.

Hence, axiomatization should not simply be equated to completeness, even in the case of successful axiomatizations: the goal of completeness is, in some cases, superfluous, and in others, impossible. Nevertheless, as the history of formal axiomatics shows, the quest for completeness is really one of the main motors behind its development, and in the majority of cases the goal of completeness, in one or many of its variants, is indeed attained by means of axiomatization.

4.2.2.2 Meta-perspective

Granting that the main goal of early axiomatics was to attain completeness (of several sorts), it is not difficult to conceive that the next issue faced by the practitioners of early axiomatics had to be: do the axiomatizations construed by us actually attain completeness? In other words, mathematical rigor demanded that the attribution of completeness to given axiomatic systems go beyond mere hunch.

However, the need for such a proof was not evident at first. According to Awodey and Reck (2002, 19) Dedekind, Hilbert, Huntington and Veblen all seemed to be under

[433] 'Of course, the axiomatic method has also been applied very successfully in cases where such "completeness" of the axioms is not required, or even desirable, for example, in the case of groups or topological spaces. In such cases it is not a matter of characterizing one particular mathematical structure, but of studying various different, non-isomorphic, systems all satisfying certain general constraints'. (Awodey and Reck 2002, 5)

the impression at some point that completeness – in particular the equivalence between deductive and semantic consequence – simply followed from the appropriate axiomatic systems. Awodey and Reck give an elucidating description of the state of affairs at the time:

> By 1908 we have axiomatizations for several main areas of then-contemporary mathematics: the theories of the natural numbers, the real numbers, and Euclidean and projective geometry. In each case 'completeness' is stated as an explicit goal, a criterion of adequacy for the axiomatization. What 'completeness' means, more or less explicitly, is primarily categoricity, secondarily semantic completeness (in various equivalent forms), and in some cases even relative completeness or logical completeness. Also, semantic completeness is repeatedly recognized to be a direct consequence of categoricity, *although no proof of that is ever given*; and sometimes the two notions are conflated, or apparently treated as equivalent. (Awodey and Reck 2002, 19) (my emphasis)

It was only in the 1920s and 1930s that the need for metatheoretic analyses of the axiomatic systems was fully understood, so as to establish beyond doubt (i.e., by proof) that certain goals were attained. Hilbert and his school were again a driving force behind these new developments, whose culmination came, as is well known, with Gödel's and Tarski's works. Hilbert's program consisted precisely in applying the very mathematical tools and techniques to metamathematical issues such as the consistency and completeness (cf. Zach 2003b).

> With this new way of providing a foundation for mathematics, which we may appropriately call a proof theory, I pursue a significant goal, for I should like to eliminate once and for all the questions regarding the foundations of mathematics, in the form in which they are now posed, by turning every mathematical proposition into a formula that can be concretely exhibited and strictly derived, thus recasting mathematical definitions and inferences in such a way that they are unshakable and yet provide an adequate picture of the whole science. (Hilbert 1927, 464)

Indeed, the two main properties of axiomatic systems that require investigation were completeness, under various forms, and consistency.

> The consistency and the completeness of the axiom system of a deductive theory now give us a guarantee that every problem of the kind mentioned [i.e., propositions] can actually be decided within the theory, and moreover decided in one way only; the consistency excludes the possibility that any problem may be decided in two ways, that is, both affirmatively and negatively, and the completeness assures us that it can be decided in at least one way. (Tarski 1959, 136)

The case of completeness has been discussed in the previous sections. As for consistency, Russell's discovery of the inconsistency of Frege's axiomatization of logic had shown that the consistency of an axiomatic system should not be taken for granted, even in the case of apparently sound axiomatizations. Moreover, paradoxes of the same family threatened Cantor's naive set theory, in such a way that it became evident that an axiomatization of set-theory had to discard at least one of its (seemingly) intuitive axioms so as to avoid antinomies and inconsistency (cf. von Neumann 1925). As a matter of fact, proving the consistency of axiomatizations of mathematical theories became the goal par excellence of a great deal of the foundational work in mathematics of the 1920s: according to Hilbert and his followers, in order to establish foundations for mathematics, it was sufficient to show that mathematical theories are consistent, as truth (understood as correspondence to something extrinsic to the theory) was a non-issue within mathematics.[434]

Wherever the axiomatic method is used it is incumbent upon us to prove the consistency of the axioms. In geometry and the physical theories this proof is successfully carried out by means of a reduction to the consistency of the arithmetic axioms. This method obviously fails in the case of arithmetic itself. By making this important final step possible through the method of ideal elements,[435] our proof theory forms the necessary keystone of the axiomatic system. (Hilbert 1927, 472)

Aside from these important historical facts, what matters for the present analysis is that a meta-perspective over a theory (mathematical or other) is indeed very much facilitated by an axiomatization of it. Naturally, one can always make informal metatheoretical remarks about a theory even if it is not formulated in an axiomatized form; but if one wishes to attain the degree of certainty and rigor conferred by proofs, a mathematical regimentation of the theory becomes virtually indispensable.

As already noted, all the theorems of an axiomatized theory can be generated inductively by means of the application of the rules of inference to the axioms. In fact, the same procedure is applied to generate formal languages and formal systems:

[F]irst, certain initial elements are specified; second, certain procedures for constructing new elements from given elements are described; and third, it is understood that *all* the elements of the class are obtained from the initial elements by iteration of these procedures. Such a set of specifications is called an inductive definition, and a class so defined is called an inductive class. (Curry and Feys 1958, 18) (emphasis added)

So, in order to make a general claim concerning the whole theory (i.e., each one of its theorems) or the whole (formal) system or language, it is not necessary to proceed by tedious enumeration of each theorem or well-formed formula, since the theory thus defined (by an axiomatization) is an inductive class; instead, it is sufficient to show that property X holds of all axioms A of the theory, and of all iterative applications of the rules of inference. That is, let a theory T be defined as $T = \{A, R\}$, where A is a set of axioms and R a set of rules of inference. Let a property X be ascribed to the theory, that is, to each theorem t of the theory (axioms are to be viewed as zero-premise theorems). This is proved inductively if the following holds:

– For every $a \,\varepsilon\, A$, a is X.
– For every $r \,\varepsilon\, R$ and for every t such that t is X, $r(t)$ is X.

Obviously, this proof-schema tends to become a great deal more complex in actual cases, but its simplicity still attests to the power of formal axiomatics for metatheoretical analyses.

The case of consistency is slightly different, as it is a property to be attributed to sets of propositions, and not to individual propositions, as in the above. Still, to

434 'For Hilbert, it is not about truth (of the axioms and theorems), but only about consistency. "Classical mathematics shall be formulated as a formal axiomatic theory, and this theory shall be proved to be consistent, that is, free from contradiction" '. (Kleene 1951, 53)

435 'Ideal elements' are precisely 'formulas that – just like the numerals of contentual number theory – in themselves mean nothing but are merely things that are governed by our rules and must be regarded as the *ideal objects* of the theory'. (Hilbert 1927, 470)

prove that a theory is consistent without axiomatization would amount to proving that the considerably large set of all its theorems is consistent, a major endeavor indeed. If however a theory is formulated in axiomatized form, to prove its consistency it is sufficient to prove that its axioms are consistent with one another (i.e., that no contradiction can be derived from them – which is still a considerable endeavor, but surely simpler). Again, let a theory T be defined as $T = \{A, R\}$, where A is a set of axioms and R a set of rules of inference. Assume that no contradiction can be derived (by applications of $r \varepsilon R$ – henceforth R-derived) from A, but that a contradiction can be R-derived from a subset T' of T: $\sim(A =>_R \bot)$ and $T' =>_R \bot$. But $A =>_R T'$, and if $T' =>_R \bot$, then by transitivity $A =>_R \bot$, which falsifies the initial assumption. Hence, if the axioms of a theory are consistent, then the set of theorems of the theory is consistent; that is, to prove the consistency of a theory it is sufficient to prove the consistency of its axioms.

Besides allowing for the study of properties of theories (of which completeness and consistency are only two examples, albeit very significant ones), the axiomatic method also allows for the comparison between different theories, of between different axiomatizations of a given theory (such as set-theory). Consider two theories $T_1 = \{A_1, R_1\}$ and $T_2 = \{A_2, R_2\}$ (which may or may not be expressed in the same language; in the latter case, the appropriate translation is required). If $A_1 =>_{R_1} A_2$ and if for all t, t', if $t =>_{R_1} t'$ then $t =>_{R_2} t'$, then every theorem of T_2 is a theorem of T_1. In such cases, T_1 is said to be (strictly) stronger than T_2, since it contains T_2. If the converse holds, then the two theories are equivalent in their deductive power. In other words, in order to compare two theories it is sufficient to compare their axioms and rules of inference, and that also represents an enormous economy of efforts if compared to the task of checking, for each theorem of T_1, if it is also a theorem of T_2.

In short, the development of formal axiomatics led quite naturally to the development of the metatheoretical perspective, insofar as it was necessary to prove that the goals motivating the execution of axiomatizations were actually attained. As it turned out, formal axiomatics also allowed for the study of many more properties of theories, so it seems fair to say that the meta-perspective eventually went beyond proofs of adequacy of given axiomatizations.[436] Indeed, formal axiomatics is an extremely valuable tool for metatheoretical analysis, to the point that one wonders whether metatheory is possible at all without axiomatics. Still, some have voiced doubts concerning the necessity of the relation between metatheory and axiomatics.

> If we had never used logistic systems at all, the many interesting results about logistic systems (such as those of Skolem, Herbrand, and Gödel) would, of course, never have been expressed in the specific form in which they are now being expressed. But it is not certain that essentially the same results might not have been attained, though in other contexts and as the results about other things. (Wang 1955, 226)

But even if this relation is not one of necessity, it cannot be stressed enough, from a historical (i.e., concerning the actual development of the metatheoretical project) as

[436] But: 'Some might say that it actually never got all the way beyond, as Hilbert's program simply collapsed with Godel's incompleteness results'. (Zach 2003b)

well as conceptual perspective, how tight the connection between axiomatization and metatheory turned out to be.

4.2.3 In what sense to axiomatize is to formalize

The aim of the considerations on axiomatization just presented is mainly that of showing how axiomatizations thus described actually serve the purpose of rendering the object of axiomatization (more) formal – that is, in what sense to axiomatize is to formalize. For this purpose, I return to the different senses of formal distinguished in section 4.1.3 above; more specifically, formality as accordance to rules, formality as structure, formality as absence of meaning and algorithmic formality are particularly relevant for the present discussion.

(1a) Axiomatization and the formal as accordance to rules. It is immediately obvious that axiomatization enhances formality understood as accordance to rules. An axiomatization specifies the rules being followed in the inferential steps within a theory, that is, how new information is deduced from previously known information by means of the application of the theory in question. The axioms in the axiomatization make explicit what is 'taken for granted' within the theory, while the rules of inference provide the canon of what is to be considered a legitimate move according to it. Hence, it is obvious that to axiomatize is tantamount to regimenting and to ensuring accordance with explicitly stated rules – that is, to rendering an object (more) formal according to this sense of the formal.

Notice also that the establishment of formal languages and formal systems can also be seen as an 'axiomatization', insofar as the same recursive procedures are used to generate the theorems of a theory from its axioms and rules of inference, the elements and theorems of a formal system from its definitions, and the sentences of a language from its alphabet and generative rules.

(1b) Axiomatization and algorithmic formality. As discussed above, according to this view on formality, formal reasoning is defined by the possibility of blind, mechanical application of the valid rules of inference in order to obtain the desired conclusions. For this to be possible, it is required that no implicit intuition (extrinsic to the system wherein reasoning is being conducted) be used, and thus that every premise and rule used be explicitly stated. This said, it is easy to see the relation between this notion of the formal and axiomatization: the basic idea underlying axiomatizations is that every single element necessary for obtaining certain conclusions (the theorems of a theory) must be made absolutely explicit, under the form of either axiom or rule of inference. Hence, the goal of algorithmic reasoning within a theory seems to be virtually unattainable without a thorough axiomatization of it, that is, without the establishment of the correct deductive structure underlying the theory. In sum, if formality is understood as algorithmic inference-making, then axiomatization (understood as the explicit presentation of every premise and rule of inference used) seems to be a necessary and perhaps even sufficient condition for formality.

(2a) Axiomatization and formality as structure. I have emphasized that the axiomatization of a theory has as one of its main goals the organization of its claims in their appropriate deductive structure. An axiomatization spells out the inferential relations between the different propositions composing a theory, by means of the stipulation of its fundamental principles (its axioms) and of the rules leading from these principles to the theorems of the theory. Now, it seems very natural to view this deductive structure of a theory as its form; therefore, from this standpoint, what an axiomatization does is to make explicit the implicit form of a specific object – namely, a theory – as discussed in section 4.1.2 above. In other words, in this sense it is clear that axiomatizations do make theories (more) formal.

(2b) Axiomatization and formality as absence of meaning. Concerning the notion of the formal as absence of meaning, it seems to me that the relation between this notion and that of axiomatization is merely tangential. True enough, it must be noted that one of the cornerstones of Hilbert's program, which is one of the most radical efforts towards the axiomatization of mathematics, is to consider the very signs in which the axiomatic systems are expressed as its objects, and thus to disregard their meaningful dimension. Indeed, axiomatization facilitates the meta-perspective over theories, as argued above, which in turn typically consists of considering the formulae of a theory as objects in themselves, and not as meaningful expressions about (other) objects. In other words, it is difficult to conceive how the project of considering signs as the very objects of logical and mathematical investigation could be carried out in the absence of thorough axiomatization; but successful axiomatization does not presuppose disregarding the meanings of the expressions involved, as is attested by numerous significant axiomatizations carried out in strictly meaningful languages.

As for the other two notions of formal, formality as variation (2c) and formality as indifference to particular objects (2d), they do not seem to be particularly relevant with respect to the concept of axiomatization. The former does not seem to have any interesting connection with this concept, beyond the fact that axiom-schemata are often used instead of 'regular' axioms, as already noted. Similarly, the same use of axiom-schemata may allow for permutation invariance, which is the technical counterpart of indifference to particular objects; but in both cases the issues at stake actually pertain to the question of symbolization, to be discussed subsequently. It may also be added that, in the case of axiomatizations of particular theories or fields, there is at most a restricted indifference to particular objects, that is, bearing only on the objects which the theory is about. On the other hand, a recursively defined formal system (which is, as argued, a form of axiomatization) typically has as one of its main characteristics the fact that it concerns unspecified objects; thus, in practice, axiomatization and indifference to particular objects often go hand in hand.

4.2.4 Conclusion

It can only be concluded that the act of axiomatizing must be considered as one of the main components of the act of formalizing. Given the appropriate objects (theories), one may go as far as saying that axiomatizing is a necessary condition for formalizing, insofar as, in order to make a theory (more) formal, the establishment of its form, that is, its deductive structure, is a crucial requirement.

A side remark: although one usually does not use the term ‘axiomatization’ for the regimentation of a specific argument, what has been said here concerning theories holds, *mutatis mutandi*, of arguments. Their deductive structure may need to be spelled out, as much as it may be required to outline the implicit premises being used. In this sense, we may be licensed to talk about the ‘axiomatization’ of objects other than theories, although properly speaking the term only applies to theories and specific domains of knowledge.

As far as the formalizations presented in the previous chapters are concerned, it seems unwarranted to speak of full-blown axiomatizations. The fragments of medieval logic presented here are not meant to be efficient theorem-provers, nor is axiomatic reasoning strictly speaking a real concern in those cases. I have, however, undertaken a certain organization of the different statements of each theory, trying to determine which statements were to be taken as primitive (its axioms, so to say), and which statements were derived from the primitive statements. More often, the latter have taken the form of definitions, such as the definitions of the different kinds of supposition, or the definitions of material and formal consequence. As for rules of transformation, since the theories analyzed here are not inferential systems properly speaking, such rules have occupied a modest position. Nevertheless, what I have termed ‘rules of supposition’ are in fact, in some sense, rules of transformation insofar as they define the moves allowed within this theory; the same holds of the logical rules defining the game of *obligationes*.

In sum, the formalizations presented here are not axiomatizations *strictu senso*, that is, full-fledged deductive structures, with clearly defined axioms and rules of inference. But they are axiomatizations *latu senso*, insofar as I have attempted to outline and organize the deductive relations between the different statements of each theory.

4.3 SYMBOLIZATION

Frege’s *Begriffsschrift*, arguably the founding text of modern logic, is innovative in many senses (cf. van Heijenoort 1967); but, for the present purposes, the most important innovation introduced in that text is the idea of an artificial language, created with the explicit purpose of being used for reasoning within science, which disposes of ordinary words in favor of symbols: an ideography. Of course, he was not the first to flirt with the possibility of expressing knowledge with a tailor-made system of symbols and notations (Leibniz and Boole,[437] among others,[438] had made

[437] Concerning Boole, see for example, (Peckhaus 2004) and (Corcoran 2003).

[438] The project of an artificial language was quite fashionable in the 17th century (cf. Maat 1999), but Ramon Lull, in the 13th century, is usually viewed as the pioneer for the idea of mechanical, calculating manipulation of language, with his *figura universalis*.

attempts in this direction); but he was arguably the first to bring this project to a fruitful completion.

The impact of his ideography on subsequent logic and philosophy cannot be overestimated; since Frege, logic is synonymous with symbolic logic, that is, with logic expressed primarily by means of especially designed symbols, instead of ordinary words. (It is true, though, that in the first instance Frege's influence took place essentially in an indirect way, through Whitehead and Russell's *Principia*.) Since Frege, it is almost as if logic ought to be symbolic in order to be logic at all.[439] Clearly, this assumption is one of the main motivations behind formalization projects: it is often thought that, for particular theories to deserve the title of formal/logical, they must be translatable into a language other than that of ordinary words. Now, the assumption of the equivalence between symbolization and formality, which underlines most of 20th century logic, is at best questionable and unsubstantiated, and at worst plainly false (cf. Dipert 1995). Nevertheless, an examination of the concept of formalization clearly requires an analysis of its symbolic aspect.

But Frege is not the only founding father of logic present in the coming pages; besides Frege's, Peirce's ideas will also be of crucial importance. Peirce's work in logic is arguably just as seminal and groundbreaking as Frege's, albeit perhaps less well known; however, for the present analysis, his work in semiotics will be even more relevant, in particular his conception of sign and his famous triad symbol-icon-index.

As will become clear, issues concerning symbolization are more complex than they might seem at first sight. To begin with, it is not entirely obvious how to draw a line distinguishing so-called 'ordinary words' from so-called 'artificial symbols'; in view of this, determining what exactly is accomplished by means of symbolizations is also not a straightforward matter. In what follows, these issues shall be addressed, in particular against the background of some notions borrowed from Peirce's semiotics.

As in the previous section on axiomatization, to conclude this section I shall examine in what ways to symbolize is to formalize, to make it (more) formal.

4.3.1 Words vs. symbols

At first sight, the distinction between ordinary words and artificial symbols may appear to be a very straightforward matter. Ordinary words are simply those we use to talk to other people, to write letters, the words we read in newspapers – in short, the language of life [*Sprache des Lebens* for Frege (1879, 6)]; by contrast, artificial symbols are those expressly created to convey a specific message, notably but not exclusively in different fields of science such as logic and mathematics. However, if we look for definitions distinguishing the former from the latter in a univocal, mutually exclusive way, we soon realize that the boundaries between them are not as sharp as might be expected; in what follows, I will argue that, conceptually, it turns

[439] Cf. (Dipert 1995); (Shin 2002, 17/18).

out to be particularly difficult to characterize what ordinary words are, as opposed to what artificial symbols are.

4.3.1.1 Languages: natural vs. conventional vs. artificial

Nowadays, one often encounters the expression 'natural language' with reference to the languages we speak and write in 'normal' life (783 hits for a search with the phrase 'natural language' in the *Philosopher's Index*); in particular, so-called 'semantics of natural language' is a popular field of research.[440] Not only within philosophy is natural language an important topic; computer scientists, for example, are also extremely interested in natural language processing, which is not at all surprising considering the huge field of applications for it.

Clearly, the term 'natural' here is opposed to the artificiality of 'arbitrarily' created languages, that is, languages created for a specific purpose, such as programming languages or logical languages, among others. In fact, I wonder when the phrase 'natural language' was first used in this acceptation; in Frege's *Begrifsschrift*, the term used to describe the language that we normally speak and write is 'ordinary language'. Now, the term 'natural' may seem harmless, but it does bring along a series of connotations, some of which are at odds with most of the influential views in recent philosophy of language, as I intend to show.

But first let us focus on the concept, rather than on the nomenclature. Clearly, the emphasis on a language that is **ordinary** presupposes an opposition with other languages that are **not** ordinary; indeed, Frege's main motivation for the creation of an ideography is the inadequacy of ordinary language for expressing scientific concepts. So from the outset the creation of a special language such as Frege's *Begriffsschrift*[441] gives rise to another concept, that of an ordinary language. Naturally, if there is no other kind of language, there is no need for an adjective to characterize the languages we speak and write; hence it is only by contrast with so-called non-ordinary languages that it makes sense to speak of ordinary, natural languages.

Interstingly, at some point, the language of life became 'natural language' in the philosophical jargon. Within the tradition stemming from the *Begriffsschrift*, the critique of the imperfections and limitations of 'natural' languages became a leitmotiv (and notice that the mathematical-logical study of these languages – now known as 'natural language semantics' – is a relatively recent phenomenon, which arguably only really got started with Montague's work). For science in general, and for logic in particular, there was almost unanimous agreement that artificial, especially designed

---

[440] See for example (de Zwart 1998).

[441] But remember that Frege's was not the first artificial language, as the idea of a *lingua characteristica* dates at least as far back as the 17th century. Moreover, it would be erroneous to think that prior to the introduction of fully artificial languages such as *Begrifsschrift* all there was were written and spoken languages: in mathematics, for example, the use of specific notation was widespread at least since the 17th century. But these specific notations were not encompassing enough to be considered as full-fledged languages.

languages were required to obtain the rigor and expressive power necessary (cf. the introduction of (Church 1956)).

Clearly, in this context, the concept of 'natural' is associated to imperfection and limitation. But most importantly, 'natural' here is opposed to artificial and man-made; but now comes the problem: aren't the languages we speak in ordinary life also man-made? In what sense are these languages products of nature, while so-called 'artificial' languages are not? More importantly, many of the most influential views in the philosophy of language of the last decades insist on the **conventionality** of ordinary languages, precisely in a sense that is opposed to naturalness. According to a tradition initiated by Wittgenstein's *Philosophical Investigations* (Wittgenstein 1953), what characterizes our uses of language (if anything at all, as the later Wittgenstein was notoriously wary of such generalizations) is the idea of rule-following; now, there is nothing transcendental or 'natural' in these rules, as they are conventionally determined and accepted by the participants of the different language-games. If they are only a matter of convention, in what sense is ordinary language 'natural', and in what sense does it differ from the 'artificial' languages of logic? The latter also consist essentially of rule-following, namely the rules determining how each symbol should be interpreted and combined with other symbols (such as in Part I of Frege 1879).

In other words, there does not seem to exist a principled distinction between so-called 'natural languages' (which are in truth essentially man-made and conventionally defined) and so-called 'artificial languages'. One may argue that a principled distinction is not necessary, that one simply 'knows' which languages are natural (in the sense of ordinary) and which ones are not. To this we may reply that perhaps even more threatening than the lack of a principled distinction is the lack of sharp boundaries between 'natural' and 'artificial' languages; rather, this would be a case of a fluid continuum going from one end to the other.

> The contrast between natural and artificial languages suggests a sharp distinction. Russian is natural, while Esperanto is artificial. But is the language of the biologists or that of the philosophers natural or artificial? […] So far as the development of human scientific activities is concerned, the creation of the language of the classical mechanics or of the axiomatic set theory was rather natural. (Wang 1955, 236)

Indeed, on the one hand, as also argued in (Hansson 2000, 163), every theoretical language, even if essentially consisting of 'ordinary' words, is artificial and regimented in that it often deviates from the 'ordinary meaning' of its terms; on the other hand, science and knowledge themselves are quite natural human phenomena, not very different from our use of language.

Moreover, what is considered to be natural or artificial also varies through time. For example, at first the term 'traumatized' (in the psychological sense) was a highly technical/artificial term, coined within the framework of Psychoanalysis, but now it is widely used in a variety of situations and by all kinds of people, not only specialists; that is, it has now been entirely incorporated to our ordinary vocabulary.

> To introduce an artificial language is to make a revolution. Unless there are compelling natural needs, the resistance will be strong and the proposal will fail. On the other hand, when an artificial language meets existing urgent problems, it will soon get generally accepted and be no longer considered artificial. (Wang 1955, 237)

In sum, the destiny of a so-called artificial language is either to disappear or to become 'natural'. This fact only suggests that there isn't much 'naturalness' in 'natural languages', and that what is considered natural today was, in all likeliness, considered 'artificial' at some point.

It has been suggested recently (cf. Stokhof 2005) that the disparity between the languages we speak and write and the artificial languages used in logic, mathematics and computer science, for example, can be compared to the difference between our hands and the tools, such as a hammer, that we use to perform certain tasks. While both are instruments, our hands are constituent parts of our bodies, whereas a hammer is obviously not. Illuminating though it is, this metaphor has the shortcoming of presupposing a sharp distinction between these two kinds of instruments, ordinary vs. artificial languages, just as much as there is a sharp distinction between the organic parts of our body and extrinsic objects.[442] Now, we have just seen that the former distinction is far from straightforward.

That the signs of the languages we speak and write are not natural was already a tenet of medieval philosophers such as Ockham and Buridan. They both insist on the intrinsically conventional character of these languages, forerunning Wittgenstein (cf. (Ockham, *Summa I*, chap. 1), (Buridan 2001, section 4.3.2)). But, according to them, there is a language which is indeed natural, and this language is **mental** language (cf. Panaccio 1999). Mental language is natural because it is a product of the causal process by means of which mental concepts are formed in the intellect when the latter comes across various entities; in other words, these entities cause the existence of mental terms. Most importantly, the relation between mental terms and what they signify (the objects) is entirely out of the scope of the intellect's will; the causal process takes place regardless of the intellect's deliberation. Therefore, there is nothing conventional or arbitrary in mental language, as both conventionality and arbitrariness depend on one's capacity to choose and deliberate.

The medieval dichotomy between natural (mental) vs. conventional (spoken and written) languages suggests that the current use of the term 'natural' to refer to ordinary languages is in fact unwarranted. Indeed, there is no principled distinction between what are now called natural and artificial languages, at least not in the way that there is a clear distinction between the medieval concepts of mental and conventional languages. In the first case (the modern dichotomy natural vs. artificial), both languages are essentially conventional and man-made, a product of the participants' will and deliberation; in the second case (the medieval dichotomy natural vs. conventional), there exists a fundamental dissimilarity concerning the process by means of which mental language is created – based on the causal relations between things and concepts – and the non-causal, non-natural, conventional process of establishment of the meaning of spoken (and, derivatively, written) words.

---

[442] Naturally, the existence of prostheses and the science-fiction vision of half-organic, half-robotic beings may blur this purported sharp distinction, but these are rather far-fetched examples that do not stand on the same level as the absence of a clear-cut distinction between 'natural' and 'artificial' languages.

If so-called natural languages are not natural, and if artificial languages are not (entirely and perpetually) artificial, what to say of the distinction between ordinary words and the (artificial) symbols introduced as a special notation in the course of a formalization? Can we speak of a fundamental difference between words and symbols? Aren't words symbols too? In order to clarify these issues, I now turn to some concepts borrowed from Peirce's semiotics.

4.3.1.2 What is a symbol?

Semiotics has **signs** as its main subject-matter, in the widest possible acceptation of this term. For this reason, semiotics is a good vantage point for the present investigation; since we seek to understand what distinguishes ordinary words from the symbols and notations used in formalizations, we mustn't assume that there is such a distinction to start with. Rather, we must begin with as neutral a framework as possible, and semiotics, with its emphasis on what is essential to all signs, seems to fit this profile rather well. Here I will focus on a few of the basic principles of Peirce's semiotics. What is particularly interesting in Peirce's work, with respect to the present investigation, are the connections he established between his semiotic work and his logical work; in fact, his semiotics is at the basis of his philosophy of logic (cf. Shin 2002, chap. 2).

From the semiotic point of view, almost everything is or can be a sign: smoke is a sign of fire, the red traffic light is a sign of the imperative to stop, footprints on the sand are a sign of walking people, the word 'Socrates' is a sign of Socrates etc. A sign is whatever points in the direction of something other than itself, with respect to an observer, thus making the observer think of this object.[443] Peirce distinguishes three basic elements defining a sign: its materiality (its form), which he calls the *representamen*; the idea (also a sign) provoked in somebody's mind, which he calls the *interpretant*; the object pointed at by the sign, which he calls (in a rare moment of jargon simplicity) the *object*.

> A sign … [in the form of a *representamen*] is something which stands to somebody for something in some respect or capacity. It addresses somebody, that is, creates in the mind of that person an equivalent sign, or perhaps a more developed sign. That sign which it creates I call the *interpretant* of the first sign. The sign stands for something, its *object*. (Peirce 1931–1958, 2.228, quoted in Chandler 1995 chap. 2)

Notice in particular that, according to semioticians, the sign is distinct from what is also sometimes called the 'sign vehicle', which Peirce refers to as *representamen*, that is, the material aspect of a sign (although sometimes the term 'sign' is also used equivocally to refer to the sign vehicle). In the case of the red traffic light, it is not only the equipment, with its bulbs and fuses, that constitutes the sign; the equipment is only the sign vehicle. It only becomes a sign upon the bestowal of meaning to

---

[443] Notice that Peirce's notion of sign is in many senses similar to some medieval views on signs, in particular John Duns Scotus' (cf. Boler 1963) and John of St. Thomas' (cf. Beuchot and Deely 1995). Indeed, a resemblance between Peirce's conception of sign and that of Ockham's, discussed in section 1.2.1 above, can be immediately perceived.

the sign vehicle, namely that the red light signalizes that one should stop. For those who are not familiar with this convention, a traffic light is no sign at all, it is just a piece of equipment emitting colorful lights. In the case of (written) words, if they are considered merely as blueprints, then they are not signs, properly speaking, but only sign vehicles (this fact has repercussions for the conception of logic and formal as pure manipulation of blueprints, which shall be discussed later).

It is clear that almost everything can be a sign, and, more dangerously, everything can be a sign of everything else, by means of a succession of such relations – a chain that has been termed 'unlimited semiosis' (cf. Chandler 1995, chap. 2). If it stayed at that, Peirce's semiotics would not be very rich in explanatory power; in truth, it might near triviality. But, obviously, this is not the entire story; in fact, one of its main strengths lies in Peirce's impressive effort towards a taxonomy of signs. He offers multiple taxonomies of signs, which are intended to be neither mutually exclusive nor exhaustive, but rather to shed light on different aspects of signs (cf. Chandler 1995, chap. 2).[444] Interestingly, they are almost always tripartite distinctions. In particular, one of his tripartite distinctions is especially influential, indeed the one he considered himself to be ' "the most fundamental" division of signs (Peirce 1931–1958, 2.275)' (cf. Chandler 1995, chap. 2): the distinction between symbols, icons and indexes. Precisely this distinction is of interest for the present analysis.

Before I explain these three notions, it is important to notice that these are not three kinds of signs: rather, they are three kinds of modes of relation between signs and objects. In other words, the same sign can be in some respects a symbol, in others an icon, and yet in others an index.

Symbols. The relation between a sign which is a symbol and an object is based on conventions, on rules conventionally determined and accepted. A symbol is only a sign to those who are familiar with these conventions. Therefore, there is a great amount of arbitrariness in the relation between a symbol and its object. Examples of signs that are symbols: words in general, alphabets, morse code, traffic lights etc.

Icons. The relation between a sign which is an icon and an object is based on resemblance or imitation. Naturally, a sign resembles an object in some specific sense, as the image is not meant to be a duplicate, an exact copy of the object. Examples of signs that are icons: portraits, diagrams, onomatopoeia, sound effects meant to imitate certain sounds etc.

Indexes. The relation between a sign which is an index and an object is based on a direct connection between sign and object, usually of a physical or causal nature. Examples of signs that are indexes: natural signs (smoke, thunder), medical symptoms, recordings (photography, film), demonstrative pronouns.

[444] '[…] Charles Peirce was a compulsive taxonomist and he offered several logical typologies (Peirce 1931–1958, 1.291, 2.243)'. (Chandler 1995, chap. 2)

These three modes of relation between signs and objects are presented in decreasing order of conventionality: symbols are entirely conventional, subject only to agreement upon the rules regulating their use to indicate certain objects;[445] icons are to some extent conventional, as resemblance only occurs with respect to certain aspects (e.g., a portrait may resemble somebody's face but not insofar as the latter is three-dimensional and the former is two-dimensional), but one does not have to learn rules in order to recognize resemblance; finally, indexes have a naturally tight connection with their objects, as these objects are typically the very causes of existence of such signs. Again, these three modes are not mutually exclusive and are not intended to be a taxonomy of vehicles of signs, of *representamens*, but rather of signs with meaning, that is, understood as corresponding to the triad sign vehicle-mental sign-object.

Concerning the triad symbol-icon-index, it is important to notice that, according to this distinction, the spoken and written words of ordinary language and the artificial notation of logical languages and formalizations are signs of the same mode, to wit, symbols. Now, the very idea of the distinction between so-called natural and so-called artificial languages discussed in the previous section is based on the assumption that there is some essential difference between these two kinds of signs; however, as we saw, this purported difference is particularly difficult to get hold of. Now, according to Peirce's taxonomy, there simply is no such distinction, since artificial notation and ordinary words actually belong to the same category of signs. It may be argued that this taxonomy simply fails to identify an actually existing essential dissimilarity; but, as shown in the previous section, the most obvious strategies to account for this purported dissimilarity do not succeed in capturing it either. What we have is thus a theoretical framework that says of ordinary words and of special notations that they are essentially the same kinds of signs, against a gut intuition that they are not, but which we haven't been able to formulate in a theoretically satisfactory way. Surely, in this situation, the burden of the proof is now on the partisans of a dissimilarity, who must provide a suitable account of their view.

Hence, at this point we are compelled to accept the view that there isn't a fundamental difference between so-called natural and so-called artificial languages. We do feel, though, that they are different. One possible way to account for this intuition is to argue that, just as much as the distinction symbol-icon-index corresponds to different degrees of conventionality, the distinction between these two kinds of symbols – ordinary words and artificial notations – lies in the degree of conventionality involved. Indeed, while some have insisted on the radical conventionality of written and spoken languages, others have argued that there is much more to them than plain

---

445 'For Peirce, a symbol is "[…] a sign which refers to the object that it denotes by virtue of a law, usually an association of general ideas, which operate to cause the symbol to be interpreted as referring to that object" (Peirce 1931–1958, 2.249). We interpret symbols according to 'a rule' or 'a habitual connection' (ibid., 2.292, 2.297, 1.369). 'The symbol is connected with its object by virtue of the idea of the symbol-using animal, without which no such connection would exist' (ibid., 2.299). It 'is constituted a sign merely or mainly by the fact that it is used and understood as such' (ibid., 2.307). It 'would lose the character which renders it a sign if there were no *interpretant*' (ibid., 2.304). (Chandler 1995, chap. 2)

conventionality, as the meanings of words are a product of complicated historical and social processes. By contrast, symbols introduced for the purpose of formalization, within theoretical languages such as those of logic and mathematics, would be purely conventional, a convention based on the stipulation that a given symbol is to be understood in a certain way, when the logician presents the syntax and semantics of the language (s)he is working with. In this sense, even within the category of symbols, one could define a sub-taxonomy in function of the degree of conventionality of each symbol. Of course, here again we would have the same lack of sharp boundaries between these presumed sub-categories, just as much as in the case of symbols, icons and indexes. Abbreviations,[446] expressions such as 'etc.' for example, would have an intermediary degree of conventionality, between ordinary words and special symbols. Hence, we would still not be able to claim that there is a clear-cut dissimilarity between these kinds of symbols, but we would be in state to claim that they should not be seen as equivalent kinds of symbols.

However, an important distinction seems to be overlooked here, namely that between conventionality and arbitrariness (cf. Chandler 1995, chap. 2). Indeed, taking this distinction into account, the objection mentioned above against the attribution of strict conventionality to ordinary languages – that is, that they are a product of contingent historical and social processes – appears to be slightly beside the point. If a product of such processes,[447] then ordinary languages may not be arbitrary, but they are nonetheless conventional insofar as these processes correspond precisely to the constant negotiation of conventions involving the members of a language community. What seems to be the case thus is that ordinary languages are just as **conventional** as so-called artificial languages, but that the latter appear to be more **arbitrary** than the former. Indeed, it is a widespread opinion concerning the choice of notations for logic that this choice is entirely arbitrary and inessential:

> It is clear that a formal system can be communicated only through a presentation. It is also clear that the particular choice of symbolism does not matter much. So long as we satisfy one indispensable condition – namely that distinct names be assigned to distinct ob[ject]s – we can choose the symbolism in any way we like without affecting anything essential. We can, therefore, regard a formal system as something independent of this choice, and say that two presentations differing only in the choice of symbolism are presentations of the same formal system. (Curry and Feys 1958, 20)

But this view has its shortcomings: as any experienced logician or mathematician knows, the choice of notation for a given logical language/theory must be very well thought through; it cannot be entirely arbitrary. In fact, the less arbitrary it is, the more likely it is to become widely accepted[448] – for reasons that will become clear in the

---

446 Interestingly, the original inspiration for Dalgarno's Universal Writing is precisely his interest in shorthand. (cf. Maat 1999, section 3.2.2)

447 This fact is accounted for by Wittgenstein in the *Investigations* with the concept of 'forms of life'.

448 See the impressive (Cajori 1928), vols. 1 and 2, on the history of mathematical notation. Here is what he says in his introduction: 'In this history it has been an aim to give not only the first appearance of a symbol and its origin (whenever possible), but also to indicate the competition encountered and the spread of the symbol among writers in different countries. […] The rise of certain symbols, their day of

following section, when I argue that logical symbolism must strive to be some sort of depiction of what is being represented. Moreover, even if the choice of individual symbols is to some extent inessential, the same cannot be said of the rules of concatenation for the underlying formal language; there must be at least a tight isomorphism in the rules of two alternative formal languages (and thus in the expressions generated by them) representing the (presumed to be) same formal system. The crucial importance of the rules of concatenation of a formal language will also become evident in the next section. In any case, the view that the choice of symbolism in formalizations is entirely arbitrary seems simply unjustified.

Hence, if ordinary words and special logical notation are both equally conventional but non-arbitrary symbols, we seem again to fall short of a criterion delimitating these two kinds of symbols; in other words, we still haven't succeeded in finding a principled distinction separating them. Of course, we can keep searching, but it may be more reasonable simply to accept that the dissimilarity between them is just not as natural or easy to account for as might be expected at first.

At any rate, if 'symbolization' is the transformation turning entities that are not symbols into entities that are symbols, then, properly speaking, one cannot use this term for the process transforming ordinary words into artificial symbols, since ordinary words already **are** symbols, just as much as the latter. In sum, the processes we usually refer to as symbolizations are, strictly speaking, no symbolizations at all.

4.3.2 Expressivity

Nevertheless, even though the concept of symbolization, understood as the transcription of ordinary words into certain special symbols, is, to say the least, problematic, the very success of this sort of procedure throughout the history of logic and mathematics prevents us from dismissing it as pure nonsense. Clearly, something is achieved when such procedures are applied; something is enhanced in the theories that are transcribed into special symbols. In this section, I investigate what exactly is achieved; I argue that the main accomplishment of symbolizations is to enhance expressivity.

4.3.2.1 Inadequacy of ordinary language

First of all, we might do well to reflect on the motivations behind the introduction of artificial languages by key figures such as Leibniz and Frege. Indeed, turning to recent practitioners of symbolization would not be very helpful; symbolization is so entrenched in current practices of sciences such as logic, mathematics etc., that its justification seems to have become superfluous – as already said, for many logicians, it is as though 'symbolic', 'formal' and 'logical' were all synonymous. It is only

popularity, and their eventual decline constitute in many cases an interesting story. Our endeavor has been to do justice to obsolete and obsolescent notations, as well as to those which have survived and enjoy the favor of mathematicians at the present moment'. (Cajori 1928 vol. 1, 1) by stepping out of the current paradigm that we may get a glimpse of the (original) reasons for symbolizing.

Of course, the introduction of a new theoretical tool is typically motivated by the inadequacy of the tool hitherto used for the task(s) in question; now, inadequacy can only be attributed to a tool with respect to a specific goal, which one seeks to attain. In the case of symbolization, the main reason for introducing artificial languages was obviously the inadequacy of the ordinary languages thus far used to conduct and express reasoning and knowledge; but as the goals sought may be different, the introduction of artificial notation may come to remedy different kinds of inadequacies. Indeed, the two main lines of complaints against ordinary language are its inadequacy as a proof-machinery and as a means of expression (which seemingly gave rise to the 'logic as language vs. logic as calculus' dichotomy); in this section, I will argue that symbolization is particularly adequate to deal with the expressivity inadequacy. Naturally, a suitably defined language may greatly facilitate the mechanical task of conducting proofs, but it is in the very nature of the latter to be ultimately indifferent to the symbols being used (provided that the machinery is aptly designed); by contrast, the expressivity of a language is inherently dependent on the symbols used and on the rules for their concatenation.

At first sight, the most obvious 'inadequacy' of ordinary languages is precisely the fact that there are many of them, and thus that communication is hindered by this diversity. To remedy this situation, a plethora of 'universal languages' have been proposed at different times. Of course, it has often been the case that a given language had such a dominating status over others that it became a *lingua franca*: this was the case notably of Greek in the ancient world, of Latin throughout the Middle Ages up to at least the 17th century, and to some extent of French in the 18th and 19th century and of English nowadays.[449] Interestingly, when acquiring this status, these languages often lose the spontaneity of ordinary spoken and written language and become rather regimented, such as later medieval Latin. But instead of choosing an already existing language to be a *lingua franca*, the idea of manufacturing a new language to occupy this position has seemed attractive to different people, at different times. Such new languages can be basically composed of regular words of ordinary languages, as in the case of Esperanto, or they may be composed of special, tailor-made symbols (obviously, in the case of languages essentially meant for writing).

Indeed, as shown in (Maat 1999), the main idea behind some of the most significant projects of 'philosophical languages' in the 17th century, namely Dalgarno's and Wilkins', which were essentially composed of tailor-made notations, was that of avoiding communication barriers caused by the diversity of languages (cf. Maat 1999, 329). Leibniz's project of a universal language, however, which came about very much against the background of these other projects, was motivated by different reasons.

[449] Naturally, here I am only referring to a very small portion of the world, namely Europe and what is now (wrongly) referred to as the 'Western World'. In other spheres, many other languages played and play the role of *lingua franca*.

Leibniz's dissatisfaction with ordinary language did not bear on its lack of expressive power, or on its inaccuracy, but rather on the difficulty, or even impossibility, of attaining complete, objective knowledge with ordinary language alone. What Leibniz sought to define (clearly in his *Dissertatio de Arte Combinatoria*, but also in later writings) was a unified, algorithmic procedure for the acquisition of new knowledge, so that the same fundamental procedure could be carried out throughout generations, and, thus, complete knowledge could be attained (cf. Maat 1999, 5.2 and 5.3). For the same reasons, the outcome of this process, the knowledge thereby produced, would be entirely objective, as subjective intuitions would be totally excluded from the process. To reach this goal, Leibniz introduced the idea of a calculating language, to be expressed by artificial symbols so as to facilitate the calculus. That is, Leibniz's project of a symbolization was mainly motivated by the idea of knowledge as calculus, rather than by considerations of expressivity.[450]

Frege, by contrast, attributes inadequacy to ordinary languages for different reasons. The following passage, from the preface to *Begriffsschrift*, seems to me to encompass the heart of the matter for Frege:

> My first step was to attempt to reduce the concept of ordering in a sequence to that of *logical* consequence, so as to proceed from there to the concept of number. To prevent anything intuitive from penetrating here unnoticed, I had to bend every effort to keep the chain of inferences free of gaps. In attempting to comply with this requirement in the strictest possible way I found the inadequacy of language to be an obstacle; no matter how unwieldy the expressions I was ready to accept, I was less and less able, as the relations became more and more complex, to attain the precision that my purpose required. This deficiency led me to the idea of the present ideography. Its first purpose, therefore, is to provide us with the most reliable test of the validity of a chain of inferences and to point out every presupposition that tries to sneak in unnoticed. (Frege 1879, 5/6) (my emphasis)

Frege's intent was to formulate the concept of number by means of strictly logical tools. To accomplish that with rigor, it was necessary to 'keep the chain of inferences free of gaps' – a desideratum that we have already encountered when discussing axiomatizations. But ordinary languages have the tendency to let presuppositions 'sneak in unnoticed', that is, these languages tend to prevent us from seeing clearly all the premises involved in our reasoning. In other words, the expressive power of ordinary languages is inadequate for scientific purposes because more is said implicitly than is said explicitly. In order to establish adequacy between what is actually said and what is presupposed and/or said implicitly, the language had to be bent and twisted in such a way that its formulations became overly cumbersome and almost impenetrable; what is worse, in a further level of complexity, not even this gymnastics with language could ensure the precision and rigor sought after by Frege.

That is, ordinary languages have the bad habit of allowing for implicit presuppositions to play a role in proofs, and when one attempts to remedy this deficiency, one

---

[450] For a brief but insightful analysis of Leibniz's *lingua universalis* in connection with the issues discussed here, see (Peckaus 2004). In particular, footnote 5 informs us that *lingua characteristica*, a phrase attributed to Leibniz by Frege, was actually never used by him.

realizes that their expressive power with respect to rigor and precision is far from satisfactory. Hence, it is not so much a matter of **what** ordinary language can talk about (virtually anything), but rather **the manner** in which it talks about scientific objects and domains, that is, with a shortage of rigor and precision. Frege created ideography precisely with the goal of offering accuracy of expression in scientific contexts. It is essentially a project for expressivity, but not in the sense of merely increasing expressive power; the goal is to allow for exactness of expression, even if this involves the reduction of the range of objects that may be referred to. On this matter, Frege's microscope metaphor is particularly illuminating:

> I believe that I can best make the relation of my ideography to ordinary language clear if I compare it to that which the microscope has to the eye. Because of the range of its possible uses and the versatility with which it can adapt to the most diverse circumstances, the eye is far superior to the microscope. Considered as an optical instrument, to be sure, it exhibits many imperfections, which ordinarily remain unnoticed only on account of its intimate connection with our mental life. But, as soon as scientific goals demand great sharpness of resolution, the eye proves to be insufficient. The microscope, on the other hand, is perfectly suited to precisely such goals, but that is why it is useless for all others. (Frege 1879, 6)

Clearly, a microscope is not solely intended to let us to see more, but rather to let us to see more sharply the objects that can be examined with this instrument (even though in practice it also allows us to 'see' entities invisible to the human eye, such as microbes). Indeed, ideography is not suited to all purposes,[451] and it does not represent an increase in expressive power with respect to range (even though, as the microscope allows us to see microbes, it may occasionally allow us to attain new concepts[452]); if it did, it would have been more fitting to compare it to a binocular. That is, the inadequacy identified by Frege in ordinary languages mainly concerns expressivity, or the lack of precision in their formulations when it comes to scientific matters. In order to increase their degree of precision, these formulations become more and more cumbersome. Another flaw of ordinary languages (with respect to scientific contexts) is the existence of ambiguity, in particular the phenomenon of words having different senses and/or standing for different things.

> To every expression belonging to a complete totality of signs, there should certainly correspond a definite sense; but natural languages[453] often do not satisfy this condition. (Frege 1948, 211)

This too is a threat to rigor and precision, and should be excluded from languages to be used for scientific purposes. In a regimented language such as ideography, each symbol has exactly one sense and one referent: there must be no ambiguous terms, no synonymy, and each term must have exactly one referent (i.e., so-called empty terms and multiple denotation mustn't occur).

---

[451] Cf. (Frege 1879, 6).

[452] Take for example the technique of diagonalization, which produces predicates that do not belong to our original realm of concepts, or that are very counterintuitive in the latter.

[453] Clearly, in 1948 this expression was already in use.

The gist of Frege's project for improved expressivity seems to me to be particularly well described by R. Brandom:

> The task of the work [*Begriffsschrift*] is officially an expressive one: not to prove something but to say something. Frege's logical notation is designed for expressing conceptual contents, making explicit the inferential involvements that are implicit in anything that possesses such content. […] Talking about this project, Frege says: 'Right from the start I had in mind the expression of a content… But the content is to be rendered more exactly than is done by verbal language…Speech often only indicates by inessential marks or by imagery what a concept-script should spell out in full'. The concept-script is a formal language for the explicit codification of conceptual contents. […] The explanatory target here avowedly concerns a sort of inference, not a sort of truth, and the sort of inference involved is content-conferring material inferences, not the derivative formal ones. (Brandom 2000, 57/8)

The question is now whether Frege is defending the view that the language of science must be other than ordinary languages as a matter of principle, or as a matter of pragmatics, that is, because the practice of science becomes exceedingly complex with ordinary languages alone, and therefore it is advisable to introduce a suitable regimented language. In other words, is the practice of science impossible without a specially designed language, or is the latter convenient rather than essential? Stokhof (2005) argues that there are (at least) three possible views on the relation between ordinary language and an artificial language created to remedy the inadequacies of the former: the latter may be seen as an extension, as an improvement or as a (principled) reform of the former. In the first case, the vocabulary and syntax of the artificial language is added to the ordinary language but does not replace portions of it, such that the scientist has a wider range of forms of expressions at hand. In the second case, the artificial language is seen as doing a better job at conveying scientific ideas, but the same could in principle also be done with ordinary language, albeit in a less efficient way. Finally, the third view is that ordinary languages are simply not adequate to fulfill expressive tasks in scientific contexts, and therefore a thorough reform of the language to be used for science must take place.

Stokhof argues that, while it may at first appear that Frege is defending the second, pragmatic view, closer inspection of the terms he uses in the relevant passages of *Begriffsschrift* reveals a principled, reformist position on the status of artificial languages. In fact, Stokhof shows that there are tensions in Frege's views on the matter, to an extent that the debate as to which view he actually held cannot be resolved here. It remains unclear (to me) whether Frege saw his ideography mainly as a convenient tool, with which the practices of sciences were very much facilitated, or as an essential feature of science.

By contrast, the instrumentalist-pragmatic view was explicitly defended by Tarski, some fifty years after Frege, as the passage below shows. This passage is also illuminating in that it indicates how powerful the notation of logic and mathematics can be, in terms of precision and brevity, in particular with the use of variables and schematic letters as placeholders.

> From what has been said it does not follow, however, that it would be impossible in principle to formulate the latter [mathematics] without the use of variables. But in practice it would scarcely be feasible to do without them, since even comparatively simple sentences would assume a complicated and obscure form.

As an illustration let us consider the following theorem of arithmetic:

$$\textit{For any numbers } x \textit{ and } y, x^3 - y^3 = (x - y) \cdot (x^2 + xy + y^2)$$

Without the use of variables, this theorem would look as follows:

The difference of the third powers of any two numbers is equal to the product of the difference of these numbers and a sum of three terms, the first of which is the square of the first number, the second the product of the two numbers, and the third the square of the second number. (Tarski 1959, 13)

A mere comparison of the length of the strings involved in each formulation of the same content suggests how powerful notational devices such as variables can be with respect to brevity and accuracy.[454]

Naturally, there may be abusive uses of these notational devices, namely when the transcription of portions of ordinary language into special notation does not correspond to a significant gain in clarity or brevity, in fact much to the contrary: as artificial notational systems are less familiar than ordinary language, the gain in using the latter has to be considerable in order to justify the deviance from familiarity. Otherwise, the use of special notation may result in obscure and pedantic formulations of what might just as well be formulated with ordinary words (cf. (Hansson 2000, 170), (Wang 1955, 233)). But, as argued in section 4.2.2 of this part, whether a given notation is obscure or not may also be a matter of the reader's training, the extent to which (s)he is familiar with this sort of device.

Now, talk of the inadequacy of ordinary language may induce the idea of an *ideal language* for science (or other), a system so well defined that no imperfection of expression would occur and where transparency would prevail. This was indeed Leibniz' view on the ideal character of his (never completed) *lingua universalis*,[455] and perhaps Frege's reformist view as well. But as later developments have shown, and given the fact that, as argued by Stokhof (2005), the view that artificial languages are to be ideal means of communication is based on dubious assumptions, the ideal of a perfect language is precisely that, an ideal that cannot be attained.[456]

In practice, special languages may fare better than ordinary languages in some aspects, with respect to specific goals, but there is always a tradeoff of gains and losses.[457] True enough, when used well, notational devices may enhance clarity and brevity of expression tremendously, but there isn't anything magic about them. They are a tool, and as any other tool are fundamentally dependent on the practitioner's

[454] See also (Wang 1955, 228) on this matter.

[455] Leibniz was, however, aware of the utopian character of this project, and seemed to have taken it more as a pole to direct his researches than as a goal to be attained in the short run (cf. Peckhaus 2004).

[456] 'The task of constructing a comprehensive ideal language is in many ways similar to that of finding a mechanical procedure to decide answers to all problems of mathematics. They are equally impossible'. (Wang 1955, 236)

[457] Differently from their contemporary Leibniz, Dalgano and Wilkins were very much aware of the necessity to compromise: 'Finally, whereas both Dalgarno and Wilkins were aware that compromises between conflicting goals were necessary to achieve a practicable language. Leibniz unconditionally believed that a perfect language was possible'. (Maat 1999, 329)

skills for these goals to be attained. In other words, among the three possible views on artificial languages identified by Stokhof, the second pragmatic view seems to be the least problematic and most plausible one, and is indeed the view adopted in the present study. The inadequacy of ordinary languages for scientific purposes is at most a pragmatic inadequacy, that is, there may be more efficient means to express scientific contents, but these could, in principle, be expressed with ordinary words as well.

4.3.2.2 Displaying/depicting

But what is a good use of notational devices? In this section I explore the principles guiding the good use of special notations: what criteria should these special symbols comply with in order to fulfill their task? For this purpose, I will turn to Peirce's semiotics again, in particular to his idea of an iconic logic, and to Wittgenstein's picture theory.

Here, I defend the view that a notation, a special symbol, must strive to **display** the objects and facts it represents. As previously mentioned, while it may be thought that an artificial notational system is even more arbitrary and conventional than ordinary words, at the same time the fact that artificial notations have lack of familiarity running against them (as opposed to the weight of widespread usage providing legitimacy for ordinary words) means that special notations must be chosen with extra care. In this sense, it is advisable that their degree of conventionality somehow decrease, in order to tighten their link to the objects and facts they represent, and thus to facilitate their assimilation into usage. Now, a natural way of achieving this is the attempt to depict the object or fact being represented.

Clearly, what we have here once more is the old issue of the foundations for the relation between signs (sign vehicles) and that which they are signs of. Following Peirce's taxonomy, there seem to be three basic kinds of such relations: those that exist in virtue of conventions, those that exist in virtue of resemblance, and those that exist in virtue of causation. Naturally, each of these concepts brings with it a number of philosophical intricacies, which shall only be partially discussed here, since what is of interest for us now are just the different kinds of foundations that can be given to the link between sign vehicles and objects.

But notice that the idea of an artificial language depicting its objects is not a 20th century invention: as the passage below shows, this intuition already permeated what is probably the first comprehensive effort of symbolization in philosophy, namely Leibniz's 'logic of invention'.

> For let the first terms, of the combination of which all others consist, be designated by signs; these signs will be a kind of alphabet. It will be convenient for the signs to be as *natural* as possible – for example, for one, a point; for numbers, points; for the relations of one entity to another, lines; for the variation of relations or of terms, kinds of angles in the lines. If these are correctly and ingenuously established, this universal writing will be as easy as it is common, and will be capable of being read *without any dictionary*; at the same time, a fundamental knowledge of all things will be obtained. The whole of such a writing will be made of *geometrical figures*, as it were, and of a kind of *pictures*, just as the ancient Egyptians did, and the Chinese do today. Their pictures, however, are not reduced to a fixed alphabet, that is, to letters, with the result that a tremendous strain on the memory is necessary, which is the contrary of what we propose. (Leibniz 1966, 11, with revisions as quoted in Maat 1999, 243; my emphasis)

It may come as a surprise to see Leibniz attributing to his ideal signs precisely many of the characteristics of Peircean icons: they are ‘a kind of pictures’; they can be understood without the use of dictionaries (since their meaningfulness does not depend on conventionally established rules); they are as natural as possible, and thus not strictly conventional (although they cannot reach the level of naturalness proper to indexes). Notice also that, in his characteristic optimism, Leibniz deemed it possible that not only the relations between objects be pictorially represented, but also that each individual sign be a faithful picture of the object(s) it stands for. Now, in later developments, this latter goal would often be viewed as impossible to attain, and, as we shall see, later iconic theories of meaningfulness such as Wittgenstein’s and Peirce’s will settle for isomorphism of elements rather than individual pictorial representation.

4.3.2.2.1 Wittgenstein on depicting. Let us first consider Wittgenstein. Now, if we apply the distinction symbol-icon-index to the different theories of meaning he proposed at different periods, it is clear that the Wittgenstein of the *Tractatus* (Wittgenstein 1963) saw the relation between facts and signs (which are also facts, as we shall see) as primarily based on resemblance, while the Wittgenstein of the *Investigations* focused on conventionality as the foundation for this relation. Now, in the previous section I have argued that the concept of conventionality does not seem to offer a satisfactory vantage point for the analysis of special notations, as the choice thereof mustn’t be entirely arbitrary if the notation is to be successful. Therefore, it is mainly the Wittgenstein of the *Tractatus* that may help us shed new light on issues pertaining to symbolization.

It is important to notice, though, that, unlike Peirce, Wittgenstein is not so much interested in the relation between incomplex sign vehicles and objects; since the world is the totality of facts, not of things (*Tractatus* 1.1), it is the relation between facts and signs representing them that is at the core of his investigation. Now, a fact is represented by a complex articulation of sign vehicles, each of them being a sign of one of the objects involved in the fact. This complex articulation is a **proposition**, and more important than the individual mapping of sign vehicles (names) into things is the **relative position** of each sign with respect to other signs, which must correspond to the relative position of the object in question with respect to the other objects involved in the fact (and which are also represented in the proposition).

2.13 To the objects correspond in the picture the elements of the picture.

2.131 The elements of the picture stand, in the picture, for the objects.

2.14 The picture consists in the fact that its elements are combined with one another in a definite way.

2.141 The picture is a fact.

2.16 If a fact is to be a picture, it must have something in common with what it depicts.

2.161 There must be something identical in a picture and what it depicts, to enable the one to be a picture of the other at all.

2.17 What a picture must have in common with reality, in order to be able to depict it – correctly or incorrectly – in the way that it does, is its pictorial form.

2.2. A picture has logico-pictorial form in common with what it depicts.

3.1431 The essential nature of the propositional sign becomes very clear when we imagine it made up of spatial objects (such as tables, chairs, books) instead of written signs.

The mutual spatial position of these things then expresses the sense of the proposition.

3.21 The configuration of objects in a situation corresponds to the configuration of simple signs in the propositional sign.

A proposition is thus a picture of the fact it represents/signifies, since the spatial position of its terms relative to one another depicts the position of the objects in the fact. The sense of the proposition is not defined merely by the objects named in it; far more important are the relations between them that the proposition asserts to be the case. Even when taken in isolation, what matters most about names is not so much the objects they name, but the possibilities of articulation with other names, as these mirror the possibilities of articulation of the named objects.

Hence, the core of Wittgenstein's picture theory of meaning is not the claim that each name of an object must be a portrait of this object, but rather that the spatial articulation of names in propositions depict the relations between the objects named by the names in the fact signified by the proposition, that is, that there is isomorphism between picture and what is depicted. Consider the representation of a subway line: the names of the stations are not meant to be pictures of the stations themselves, and the distances between the points representing each station are not meant to correspond exactly to the actual distances between the stations, not even to be a proportionate projection of these distances. What this representation shows is the relative order of the stations in the linear trajectory of the train. A picture must always have something identical to that of which it is a picture (*Tractatus* 2.162), but at the same time the picture (proposition) and the fact it depicts are distinct facts; a delicate balance of otherness and sameness characterizes the relation between a picture and what it depicts.

What does Wittgenstein's picture theory add to the present analysis? Not much, it may seem at first, since Wittgenstein intends his picture theory of meaning to apply to propositions in ordinary language as well as in artificial symbolic language. Moreover, he radically rejects the idea that some propositions have 'more meaning' than others, in virtue of a more accurate representation of their logical forms, as his angry reaction to Russell's introduction to the first edition of the *Tractatus* indicates (cf. Lopes dos Santos 1994). Russell had attributed to him the quest for an ideal, logically perfect language, and this is precisely what the *Tractatus* is **not** about: Wittgenstein is interested in the meaningfulness conditions of **all** meaningful propositions. For these reasons, his picture theory may at first seem not to be able to offer an illuminating account of the specificity of artificial languages.

But Wittgenstein does differentiate between what he calls 'everyday language' and an artificial sign-language; more importantly, just as Frege, he identifies inadequacies with respect to the former (even though his picture theory of meaning should apply to it too).

3.323 In everyday language it very frequently happens that the same word has different modes of signification – and so belongs to different symbols – or that two words that have different modes of signification are employed in propositions in what is superficially the same way. Thus the word 'is' figures as the copula, as a sign for identity, and as an expression for existence; 'exist' figures as an intransitive verb like 'go', and 'identical' as an adjective; we speak of something, but also of something's happening. (In the proposition, 'Green is green' – where the first word is the proper name of a person and the last an adjective – these words do not merely have different meanings: they are different symbols.)

3.324 In this way the most fundamental confusions are easily produced (the whole of philosophy is full of them).

In order to understand how the picture theory can apply to the 'imperfect' propositions of everyday language (a point which seems to have been misunderstood by Russell), we must consider the distinction between propositional sign (which corresponds to the concept of sign vehicle as used so far) and proposition properly speaking:

3.11 We use the perceptible sign of a proposition (spoken or written, etc.) as a projection of a possible situation. The method of projection is to think of the sense of the proposition.

3.12 I call the sign with which we express a thought a propositional sign. And a proposition is a propositional sign in its projective relation to the world.

Now, the projective method to be used in the case of propositional signs of everyday language may be more intricate than in the case of a language where the pictorial form of propositional signs corresponds to the logical form of the proposition being expressed. But there is always such a projective method if the propositional sign is meaningful at all. This being said, Wittgenstein is nevertheless a partisan of the introduction of special sign-languages in such a way that the projective method mapping its propositional signs to facts be as simple as possible, that is, where the pictorial form of the signs be as faithful a picture of the (possible) fact being represented as possible.

3.325 In order to avoid such errors we must make use of a sign-language that excludes them by not using the same sign for different symbols and by not using in a superficially similar way signs that have different modes of signification: that is to say, a sign-language that is governed by logical grammar – by logical syntax. (The conceptual notation of Frege and Russell is such a language, though, it is true, it fails to exclude all mistakes.)

That is, all meaningful propositional signs are (equally meaningful) pictures of facts, but it is as though some were more faithful pictures than others. Therefore, the main desideratum for an artificially created sign-language is that its propositional signs be depictions that are as similar as possible to what is being depicted (and therefore that the projective method required be as simple as possible). Naturally, in the case of formalizations of theories, what is represented are not facts and objects in the world properly speaking, but rather conceptual entities. Wittgenstein says that 'a proposition is a propositional sign in its projective relation to the world' (3.12), but when it comes to theoretical entities, there may not be this direct anchorage into the world (unless 'the world' is broadly understood so as to include the existence of theoretical entities). However, these conceptual entities are also organized in a certain configuration, and it is this configuration that a symbolization must seek to display.[458] Considering again Peirce's distinction between symbols, icons and indexes, it is

[458] True enough, if the propositional signs of an artificial language do not depict facts in the world, their status may be similar to that of numeric equations, which properly speaking do not express a thought

clear that what Wittgenstein defends in the *Tractatus* is an iconic foundation for the connection between propositions and reality. This should hold in particular of artificially created languages, since these are designed precisely with the purpose of remedying the imperfections of everyday language that often hinder the isomorphism between depictions (propositions) and the facts depicted.

4.3.2.2.2 Peirce and icons. Let me now turn to Peirce. That language in general, and logic in particular, can be iconic and not only symbolic is perhaps one of his boldest theses. Unlike the Wittgenstein of the *Tractatus* (and in line with the Wittgenstein of the *Investigations*), Peirce insists on the essentially conventional nature of ordinary words, not only with respect to words considered in themselves but also with respect to their disposition in propositions (as grammar can also be, to a great extend, conventional). Now, as already noted, ordinary language and the special notations of logic alike are systems of **symbols**. But besides creating a symbolic system of logic, he also created an **iconic** system of logic (cf. Shin 2002); more importantly, he considered his iconic logic to be in many respects superior to his symbolic logic (cf. Shin 2002, 2/3).

First, a remark on Peirce's conception of icon: even though the paradigmatic cases of icons are incomplex signs bearing a similarity to the objects they point at, Peirce sees as equally iconic signs that are similar to certain objects only insofar as their configuration is concerned, that is, along the lines of Wittgenstein's propositional signs. In other words, Peircean icons are not restricted to the Leibnizian kind of individual similarity between incomplex signs and objects. In effect, a given sign may be composed of symbols (such as the names of the subway stations, in the aforementioned example), but its configuration may represent **iconically** the configuration of the objects in question; such a representation would be a combination of symbols and of an iconic representation.

> For Peirce, icons included 'every diagram, even although there be no sensuous resemblance between it and its object, but only an analogy between the relations of the parts of each' (Peirce 1931–1958, 2.279). 'Many diagrams resemble their objects not at all in looks; it is only in respect to the relations of their parts that their likeness consists' (ibid., 2.282). (Chandler 1995, chap. 2)

The key concept here is that of diagram, which, as shown in (Shin 2002, 19–21), is not easily grasped within the Peircean framework; for Peirce, this term is clearly more broadly understood than in its ordinary meaning. Indeed, Peirce considers an algebraic equation, something most of us would probably not be prepared to call

---

(cf. *Tractatus* 6.2 and 6.22); they only show formal structural relations. Wittgenstein's radical thesis concerning the meaningless status of every proposition which is not empirical or contingent (such as tautologies and contradictions – cf. *Tractatus* 4.461) may attain the propositions of a theory's formalization. Be that as it may, that is, regardless of the ontological status of the entities being represented (whether only conceptual or as facts about the real world), the aim of an artificially created language would then be that of depicting these formal structures as graphically as possible; the concatenation of artificial symbols must be a picture, in the strict sense of the term, of what is being represented.

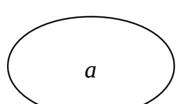
an icon or diagram, to be an icon.[459] A diagram, he says, ought to be as iconic as possible (cf. Peirce 1931–1958, 4.433), but it is clear to him too that icons alone are not sufficient for the expression of knowledge, in particular with respect to their lack of generality. For Peirce, a suitable formal system is a mixture of different kinds of signs: whenever possible, a diagram ought to be iconic, but many of the elements of a diagram are bound to be symbolic.

Indeed, what is perhaps most interesting and may even appear ironic concerning Peirce's theory of signs and its repercussions for logic is that, while he exerts himself to develop several taxonomies of signs, what transpires from many of his remarks is that he is a partisan of what are now called heterogeneous logical systems (cf. Shin 2002, 31–35), that is, logical systems that make use of different kinds of signs:

> I have taken pains to make my distinction of icons, indices and tokens clear, in order to enunciate this proposition: in a perfect system of logical notation signs of these several kinds must all be employed. (Peirce 1884, 181)

> If symbolic logic be defined as logic – for the present only deductive logic – treated by means of a special system of symbols, either devised for the purpose or extended to logical from other uses, it will be convenient not to confine the symbols used to algebraic symbols, but to include some graphical symbols as well. (Peirce 1931–1958, 4.372)

What is meant by a mixture of symbolic and iconic elements is made clearer with the help of an example (borrowed from Shin 2002, 26). Let us assume that the content to be expressed is that an individual, call it *a*, belongs to a class, call it *B* (notice that both '*a*' and '*B*' are *symbols*). Now, in a strictly symbolic representation of this content, a symbol must be introduced to express the relation of belonging, usually the Greek letter 'ε'; with these conventions, the symbolic representation of this content becomes '*a* ε *B*' (or combinations of these symbols in a different order, such as 'ε *aB*'). By contrast, in an heterogeneous representation of the same content, the relation of belonging can be expressed iconically: if the individual is represented by '*a*' again, and the class by a circle, the content can be represented as:



Peirce's emphasis on the pictorial character of signs and its potential uses in systems of logic is thus the main inspiration for the views defended here. To be sure, I am not only arguing that the disposition of symbols in an artificial language ought to be as iconic as possible; I am also suggesting, following Leibniz's idea, that the symbols themselves can be iconic.

Naturally, a tradeoff between iconicity and convenience must occur – for instance, Roman numerals may seem more iconic and thus more intuitive than Arabic numerals,

---

[459] 'In fact, every algebraical equation is an icon, in so far as it *exhibits*, by means of the algebraic signs (which are not themselves icons), the relation of the quantities concerned'.(Peirce 1931–1958, 2.282, quoted in Shin 2002, 25)

but experience shows that Arabic numerals are much more suitable for calculations and other purposes. But whenever possible, iconic elements must be used. As for the example above, while the representation of a class by a circle is essentially a symbol (since the class does not resemble the circle, properly speaking), the circle is still a more iconic representation of a class than, say, a line or a point. Similarly, the use of shading in Venn diagrams is a symbol, a convention to express emptiness, and not a representation of emptiness (Shin 2002, 32–33); nevertheless, the shading is arguably a more iconic representation of emptiness than, say, the drawings of cows or houses.

It may be argued that the view I am defending here, that artificial symbols ought to be like depictions, incurs the mistake of actually blurring the distinction, so carefully drawn by Peirce, between symbols and icons; indeed, what I am proposing is that artificial symbols be something in between icons and symbols. But aren't symbols and icons fundamentally different? Now, this is an interesting feature of Peirce's taxonomy: these distinctions are not intended to be clear-cut ones (cf. Shin 2002, 26). In fact, there is consensus among semioticians that, in every iconic representation, there is always a certain degree of cultural determination and conventionality;[460] in this sense, the distinction between symbols and icons is primarily a matter of degrees. Moreover, the historical development of writing basically consists of the evolution from indexical and iconic signs to symbolic ones (from mere ostension to hieroglyphs, and on to alphabets), which suggests that the passage from one kind of sign to the other is typically gradual. From this perspective, artificial symbols can be seen either as a step further from ordinary words, with increased conventionality and arbitrariness, or as a step backward, back to pictorial representations. Being pictorial representations (less conventional, more closely resembling their objects of representation), they may be more easily understood. And even though Peirce created two alternative systems of logic, a symbolic one and an iconic one, thus suggesting that there is a sharp distinction between these two kinds of signs, it is worth noticing that some of the most significant symbolic systems of logic, most notably Frege's ideography, are in fact quite iconic. Likewise, Peirce's own iconic systems contain many symbolic elements as well, as already argued, and as shown in (Shin 2002, chap. 3).

4.3.2.2.3 Iconic symbols. These considerations reveal that, within the symbolic category, some symbols may be closer to being icons than others. Now, since artificial notations do not have the legitimacy offered by years of usage, as ordinary languages do, it is advisable that their degree of 'naturalness' be increased. As already mentioned, the gradation symbol-icon-index corresponds to a decrease in conventionality and an increase in naturalness; therefore, if a symbol features iconic elements (even though it remains a symbol), it is more natural than a symbol that is in no way iconic. In sum, contrary to what prima facie may seem to be the case, a symbol that is to some extent iconic is not a incongruence, much the opposite; it is in fact what may make artificial notations more easily understood and incorporated into usage.

---

[460] 'Semioticians generally maintain that there are no pure icons – there is always an element of cultural convention involved'. (Chandler 1995, chap. 2)

If artificial symbols succeed in being, at least to some extent, iconic, then they will probably do a better job at conveying the corresponding content than ordinary words would, since these (as noticed by Wittgenstein) often provide 'distorted' pictures of facts (in which case complex methods of projection are required to link depiction and fact depicted). In truth, iconic artificial symbols may allow us to say more with less and to be more precise, as they may be more faithful pictures of the contents expressed.

Finally, the position I defend here may also seem to imply that, with respect to logic, since symbols that are iconic are superior to non-iconic symbols, full-fledged icons, like diagrams, are even more suitable. Indeed, even though I do not have an entirely developed view on the matter, I am inclined to believe that iconic logics such as Peirce's existential graphs may be better systems than their symbolic counterparts. Currently, the bias is still in favor of strictly symbolic logic,[461] but as argued by Shin (2002, chap. 1), this is mostly due to the (contingent) development of logic as (so far) predominantly symbolic, and to the still existing misapprehension of the properties of iconic logics (as opposed to the properties of symbolic logics), but not to the fact that logic ought to be, as a matter of principle, essentially symbolic. It might seem, though, in particular with respect to formalizations understood as abstraction of meaning, that the very intuitiveness of icons hinders formality and generality, which are, by contrast, enhanced by the essentially stipulational character of symbols. However, as argued by Shin (2002, 31–35), it might be a matter of improving 'iconic languages so that they maintain their strengths [intuitiveness] but rule out possible sources of fallacy'. (Shin 2002, 35)

4.3.2.3 Kinds of symbols: interpreted vs. uninterpreted languages

So far, I have argued that, to ensure their greater expressiveness, artificial signs ought to be like pictures; they may have some kind of resemblance to the object they represent, as suggested by Leibniz, but the pictorial feature must concern above all the representation of articulations, of the mutual configuration of objects, which may be obtained by isomorphism.[462] Now, some important aspects related to symbolizing are yet to be discussed. Firstly, in symbolizations and formalizations, there are usually two main kinds of signs: (i) those intended essentially as abbreviations and denoting one specific entity (concept, operation, thing etc.), and (ii) those that are placeholders for many entities (schematic letters and variables). This fundamental distinction must not be overlooked: if signs are to be pictures of the appropriate objects, it makes all the difference whether a given sign is intended to depict one specific entity, or rather many entities but none of them in particular. Secondly, this distinction applies exclusively to so-called meaningful logical languages, as opposed to uninterpreted

---

[461] Even though a great deal of interesting work has been done recently on the logic of graphs and diagrams, see (Shin and Lemon 2003) for an overview and references.

[462] The notion of 'structure of signature K' (where K is a signature, i.e., a set of individual constants, predicate symbols and function symbols) in model-theory is a highly regimented version of this idea of isomorphism between language and what it represents.

logical languages; concerning the latter, it is as if all their terms were placeholders. In this section, I discuss these two issues, namely the two kinds of symbols and the opposition between interpreted and uninterpreted languages.

The distinction between terms that are placeholders and terms with a definite meaning is familiar from mathematical notation, and has been put to a more general use by Frege:

> I adopt this basic idea of distinguishing two kinds of signs, which unfortunately is not strictly observed in the theory of magnitude, in order to apply it in the more comprehensive domain of pure thought in general. I therefore divide all signs that I use into those by which we may understand different objects and those that have a completely determined meaning. The former are letters and they will serve chiefly to express generality. (Frege 1879, 10/11)

Naturally, there isn't anything very esoteric about this distinction: in ordinary language the same distinction exists between words that designate many entities, such as common terms like 'man', or even more general terms like 'somebody' or 'something', and terms that are meant to designate one specific entity, such as proper names. Moreover, the boundaries between these two kinds are actually not entirely sharp: what to say of a word designating an operation, say, 'multiplication'? Does it denote one single entity, the operation, or does it denote every occurrence of the operation? This is an essentially metaphysical issue, related to principles of individuation, which we shall not go into here. But it must be acknowledged that the fundamental difference between terms that are not meant to denote any particular entity (albeit within a given class) and terms pointing at a determinate entity plays a crucial role in the design of an artificial language.

But which picture shall be used to denote a placeholder, that is, to depict many entities but none of them in particular? The very idea of picturing presupposes some kind of determination with respect to the object being pictured. Clearly, in such cases what is intended is precisely to convey generality, so the sign cannot be specific; in truth, these cases seem to speak against the idea that to symbolize is to depict. Another argument against this view is historical: one of the first uses of artificial notations in logic was precisely the use of schematic letters to replace terms by Aristotle in the *Prior* and*Posterior Analytics*;[463] well, clearly, Aristotle's schematic letters were not intended to depict, properly speaking, either the terms or the entities they replaced.

Indeed, at first sight it may not be obvious how single letters can display the existence of a gap, or of an indetermination, to be filled/substituted with entities of the appropriate kind, such as in the case of schematic letters or variables. But this is precisely what happens. In fact, while the notions of schematic letters and of variables are often conflated, the difference between them can be explained as follows: variables display an empty space, an insaturation in the expression where they stand (the same can be obtained with the use of other signs, such as '…', to indicate the gap, but it has become customary to use individual letters for this purpose); by contrast, schematic letters are intended to be placeholders of generality. This

[463] A few examples: (*Prior Analytics*) 68^a, 59^b; (*Posterior Analytics*) 86^b, 87^a, 80^a, among many others.

difference can be made intuitive by means of examples from ordinary language: the use of a schematic letter is roughly equivalent to the use of the expression 'somebody' in the sentence 'John loves somebody' – it is undetermined whom John loves, but the sentence in itself conveys a complete message. Alternatively, the use of a (unbound) variable corresponds to an incomplete phrase or sentence, such as 'John loves …' – the expression, as it stands, is simply not a complete sentence, and it indicates a gap to be filled appropriately.

In any case, in the strict sense of the term 'sign', schematic letters and variables are not signs, insofar as they are not associated to any specific object (or class of objects[464]), while at the same time being related to each one of the objects that can take their place. Paradoxically, these letters depict plurality, generality and void at the same time.

By contrast, signs intended to indicate a specific object, operation or concept, are of a different nature. In such cases, one may indeed attempt to find a graphic representation of the entity to be designated. But a lesson to be born in mind (learned from Wittgenstein and others) is that the depictive strength of a proposition, in ordinary language or otherwise, lies less in the individual resemblance of each term to an object than in the isomorphism between the spatial arrangement of the terms of the language and the relations between the objects depicted. In this sense, more than the individual symbols, what contributes most to the depictive strength of a language is its **syntax** or grammar, that is, the possibilities of concatenation of signs, and in particular whether they reproduce the arrangement of the things represented.

However, notice that the distinction between terms that are essentially placeholders and terms with a defined meaning only applies to meaningful logical languages. As is well known, since Hilbert proposed to consider formulas and terms as objects and not as meaningful expressions, and since Tarski introduced the idea of interpretation of a language on a model, it became customary to treat **all** terms of a logical language as meaningless placeholders. This passage from meaningful to uninterpreted logical languages is aptly described by Tarski himself.

> Already at an earlier stage in the development of the deductive method we were, in the construction of a mathematical discipline, supposed to disregard the meanings of all expressions specific to this discipline, and we were to behave as if the places of these expressions were taken by variables void of any independent meaning. But, at least, to the logical concepts we were permitted to ascribe their customary meanings. […] Now, however, the meanings of all expressions encountered in the given discipline are to be disregarded without exception, and we are supposed to behave in the task of constructing a deductive theory as if its sentences were configurations of signs void of any content. (Tarski 1959, 134)

On the one hand, signs of a (logical) language being taken in their pure materiality, that is, as devoid of meaning, make patent the point stressed by Wittgenstein, namely that, more than the individual terms, what is crucial in a proposition is their relative disposition. According to this approach, if *a, b, c* and *d* are terms of the same kind, and the same holds of *R* and *B*, then $a R b$ and $c B d$ are equivalent expressions insofar as

[464] Of course, the domain of the variable or schematic letter is often explicitly defined, but precisely because this information is not conveyed by the variable of schematic letter alone.

they can acquire the same meaning if *a* and *c, b* and *d*, and *R* and *B* are given the same interpretation, respectively. Moreover, the arbitrary stipulation of an interpretation on a model only emphasizes the arbitrariness and conventionality of symbols, features that are sometimes not entirely evident in the case of interpreted languages.[465]

On the other hand, if they are void of meaning, such signs are strictly speaking no longer signs, but merely potential sign vehicles (and remember that virtually anything is a potential sign vehicle); for this reason, these languages are no longer languages *stricto sensu*. They only become languages again when given an interpretation. At this point, the specific signs used in a logical language no longer matter, but only the syntax of the language insofar as it potentially describes a variety of structures.

Now, there seem to be two conflicting directions in which the concept of formalization is related to uninterpreted logical languages. On the one hand, if the semantic interpretation of a language (or, better put, of a system of sign vehicles) is its mapping into certain structures and objects, then one can say that formalization is the **dual** of interpretation, since formalizations consist essentially in the mapping of a structure of objects (the objects of the theory being formalized) into the appropriate formal language. In this sense, the language of a formalization is never uninterpreted, because a formalization is precisely the mapping of a structure or class of structures into a language (as opposed to the mapping of a language into a structure, which is the gist of interpretation), or of a language into another language – for example, the Gödelian encoding of formulas into numbers,[466] or, as is frequently done, the mapping of expressions of ordinary language into special notation; in these cases too, the language of the formalization is, of course, essentially interpreted, even if the purpose is to manipulate blindly the language of the formalization, disregarding its meaningful aspect.

On the other hand, there is a sense in which formalization means exactly the process by means of which the specific contents of the formulas of a theory are disregarded so that their mutual relations alone be considered (i.e., the 'formalization of mathematics' undertaken by Hilbert). In this case, formalization would be the **opposite** of interpretation, that is, the process of abstracting a language that does have an intuitive interpretation from the latter and of turning it into an uninterpreted system of sign vehicles (which may eventually be interpreted on different structures), thus considered as objects.

To be sure, these two procedures are not mutually exclusive: if an originally meaningful language *L* is encoded into a formal system or language L' (e.g., numbers or especially designed systems of sign vehicles), then the latter may be seen as meaningful insofar as it receives its meaningfulness from the original language *L*.

---

[465] 'Moreover, stipulation, part of the nature of symbols, makes formalization conceptually much easier than for other kinds of signs. As the meanings of symbols are conventionally stipulated, so are the rules of semantics and inference. The concept of formalization thus comes very naturally in the case of symbolic languages, which is why symbolic formal systems have been accepted without question'. (Shin 2002, 30)

[466] The encoding of formulae into numbers can also be seen as a case of conceptual translation, to be discussed below.

But once the encoding procedure is carried out, L' is often studied from a purely objectual perspective, that is, disregarding its meaning, and the conclusions drawn by the mechanical application of rules of inference may be translated back to L (compare Curry's remarks (1957, 18) on two different but in practice equivalent notions of formalization, Carnap's and his own). This two-step process is often applied when inference-drawing is more easily or more rigorously done in L'.

Notice though that an interpreted language does not amount to a language in which each expression has a defined meaning and/or a unique denotation: terms with no defined meaning, placeholders, will feature in virtually every logical language, given their convenience when it comes to generality and expressive power. Typically, in an interpreted language, the types of objects that constitute valid substitutional instances of placeholder are explicitly defined (as in constructive type-theory, cf. (Ranta 1994)), but their specific denotation can still vary.

At any rate, if the intention is to produce an interpreted language while performing a formalization/symbolization, one must seek the most suitable way of expressing 'gaps', possibly of different sorts – that is, the spaces to be occupied by the appropriate substitutional instances – and the best devices to represent each specific term with a defined meaning. But, even more importantly, the syntax of the language, that is, the possibilities of concatenation of its terms, should follow, as much as possible, the actual configurations of the objects denoted by these terms, be they purely conceptual objects or actual things in the world (assuming, of course, the realist tenet that there are such actual configurations, even if they are not immediately known).

It may be argued (as hinted in Shin 2002, 31–35) that the iconic approach may be fruitful in the case of languages that are intended to be meaningful, for example, in the case of formalizations of theories about specific objects, but not if the purpose is to study the properties of a system of uninterpreted sign vehicles, as is common in current logic. It might even be said that it is precisely because this latter approach has prevailed in the logic of the last seven or eight decades that symbolic logic has been so dominant, at the expense of iconic logic. Indeed, if icons have an intuitive way of conveying meaning, this may be inconvenient for the total abstraction of meaning and for the permutation of interpretations in models, which has become one of the cornerstones of recent research in logic. Moreover, if formality amounts to the exclusion of all extrinsic intuition, so that reasoning be conducted mechanically and only with elements explicitly stated, the fact that icons may have intuitive content may allow for implicit premises to 'sneak in'.

But this objection seems to overlook an important point: this might be the case if only individual signs were iconic; however, as I have stressed numerous times, even more than the iconic representation of **objects**, icons are particularly suitable to represent **facts** and **relations** between objects. Now, it is often said that uninterpreted languages concern the analysis of abstract structures; but if this is so, why couldn't these structures be represented iconically? To be exact, this does happen, but usually with heuristic purposes, for example, when one draws possible worlds and accessibility relations when reasoning about Kripke structures. Well, as argued by Shin (2002), it may be a matter of simply improving these heuristic devices so as to turn them into full-fledged logical languages, to be used for formal reasoning just as much as symbolic language, but with gain in intuitiveness of representation (of relations). In principle, an efficient iconic deductive system is not impossible. 'Therefore, iconic diagrammatization, that is, iconic formalization, should be possible'. (Shin 2002, 35)

In sum, what I propose is precisely to take the iconic dimension of artificial symbols into account; arguably, this dimension is already operational in actual logical languages, but not yet sufficiently discussed – indeed, it seems that their pictorial character is often responsible for the success and wide acceptance of some notations, at the expense of others, but this usually takes place on an implicit level. Ideally, it should be possible to combine the advantages of symbols (e.g., exclusion of unwanted intuitions) with those of icons (intuitiveness of meaning and potential simplification of the system of conventions) in designing powerful and effective logical languages. The pictorial principle may guide the choice of placeholder-terms as well as of terms with determined meaning (in the case of meaningful logical languages), but even more important is the recognition of the power of concatenation of symbols as a iconic device, for meaningful as well as for uninterpreted logical languages; in this sense, the pictorial principle is even more relevant for the determination of the syntax of an artificial language, and this holds just as well of interpreted and uninterpreted languages. Ultimately, uninterpreted formal languages describe structures,[467] articulations of objects, which may very well be represented iconically.

4.3.3 In what sense to symbolize is to formalize

So far, I have reached unexpected conclusions:

1. The process of symbolization as we usually understand it is not a symbolization, properly speaking. As there seems to be no principled distinction between ordinary words and artificial symbols, and, as according to Peirce's semiotic framework, they both belong to the same category of signs, namely symbols, if we were to take the meaning of the term literally, we could not call the process by means of which symbols are transformed into different symbols a **symbolization**.
2. The successful introduction of new artificial notations requires that attention be paid to the iconic aspect of these symbols. Hence, what is known as symbolization is in fact just as much '**iconization**'.

These foundational worries put aside, it is now time to investigate the ways in which what we usually call symbolization (basically the act of replacing ordinary words by special symbols – henceforth I will use this term in its usual acceptation, in spite of the doubts raised in the previous sections) is indeed a formalization, that is, in what sense to symbolize turns the object of symbolization into something (more) formal. Again, I will turn to the senses of formal defined in section 4.1.3 of this chapter.

---

[467] Compare the view known as structuralism in logic, cf. (Koslow 1992).

(1a) Symbolization and formality as accordance with rules. Insofar as it is essentially a matter of expressivity, the sense in which symbolization is related to the notion of the formal as regimentation is also to be seen as an effort towards the regimentation of a language's expressive power. As already mentioned, an important leitmotiv of the logical tradition stemming from Frege is that ordinary languages are too irregular, too imperfect to be used in scientific contexts; what is required is a regimented language where imperfections such as ambiguity and vagueness do not exist, and where the contents being expressed be represented in an orderly and controlled way. Now, formal languages, defined recursively from a finite alphabet, and with an explicit and regimented semantics (be it a model-theoretic formal semantics, or a semantics defined within the very language, as in Frege's *Begriffsschrift*, and, more recently, in constructive type-theory) seem the perfect candidates for the role of regimented languages for scientific purposes. The choice of symbols in itself is not the crucial part of a symbolization with respect to formality as regimentation; it is the explicit statement of the rules for the formation of legitimate sentences of the language (its syntax) and the regimented semantics given to these languages that enhance the formality of the object undergoing a symbolization.

(1b) Symbolization and the algorithmic notion of the formal. As already argued, algorithmicity does not require but may be enhanced by the symbolization of a language: in an artificial symbolic language,[468] the strings required to express certain contents tend to be shorter, which obviously facilitates computational operations with them. Indeed, the connection between algorithmicity and symbolic languages is particularly exemplified by Leibniz's project of a *lingua universalis*, as discussed in section 4.3.2.1. Furthermore, as argued by Frege, ordinary language often allows for unnoticed and unwanted premises to sneak into our reasoning, thus hindering purely algorithmic reasoning; to prevent this from happening, ordinary language has to be significantly twisted, and even then, at a certain level of complexity, it seems to become impossible to reason formally with ordinary language alone. But again, I can think of no argument **of principle** excluding the possibility of algorithmic reasoning being conducted solely within ordinary language, even though there is a clear tension between the desiderata of intuitiveness of meaning and computational manageability, as is attested by how 'unfriendly' programming languages can be, and how inefficient for algorithmic processes ordinary language tends to be.

The algorithmic notion of the formal also seems to speak against the iconic view on artificial notations; indeed, if reasoning is to be conducted with no interference of external, implicit intuitions, pictures may appear to be unsuitable for this task, as they may bring along with them several unwanted elements into the reasoning. For instance, if I am to prove a certain property of all triangles, and I do this on the basis of a drawing of a triangle which happens to be isosceles, I may be attributing to all triangles properties that in fact pertain exclusively to isosceles triangles

[468] Obviously, here I am not using the term 'symbol' in its Peircean sense, as, according to him, what we call 'non-symbolic' languages are also symbolic.

(cf. Shin 2002, 29). Here, a similar counter-argument as the one presented above can be offered: it is above all a matter of perfecting the pictorial language to be used and the algorithmic procedures to be employed, in such a way that certain fallacies of reasoning do not occur. But the main challenge for iconic or heterogeneous formal systems remains to guarantee that no accidental property of a sign (its size, its color) be taken to be essential for the purpose of the reasoning being conducted.

(2a) Symbolization and formality as structure and abstraction from content. While axiomatization enhances formality understood as structure on the level of entire theories, symbolization is particularly suitable to outline structures on the propositional/sentential level, that is, the logical form of propositions may be more accurately expressed by a special notation than by ordinary language sentences. Considering again Wittgenstein's idea of different projective methods connecting a propositional sign to the fact it represents, it is clear that a language composed of special notations and, more importantly, with a syntax other than that of ordinary language, may display the actual logical form of a proposition more accurately, in such a way that simpler projective methods can be used to connect the propositional signs to the corresponding facts (be they in reality or purely conceptual). To sum up, symbolization corresponds to formalization on the level of propositions or sentences insofar as it allows for a more faithful representation of their logical forms, which are their structures.

(2b) Symbolization and the 'absence of meaning' notion of the formal. Clearly, this notion of the formal is at the root of the concept of uninterpreted languages, or better put, uninterpreted systems of sign vehicles. It is also evident how and why symbolization enhances the absence of meaning: with the stipulation of a language with strictly arbitrary terms – artificial symbols – with no previous intuitive meaning, it becomes much easier to disregard any meaningful dimension of these so-called languages.

Now, this sense of formal also seems to be at the root of the wariness that still exists regarding the use of pictures and diagrams in logic. As discussed by Shin (2002, chap. 2), if the expressions of a formal language are to be taken purely as objects and not as expressions with meaning, the fact that pictures tend to convey an intuitive meaning (and this is precisely their strength) seems to make them unsuitable to be used in formal reasoning. To this objection, two replies are possible: either one eschews the whole approach to formality as absence of meaning (as in (Sundholm 2002b), (Sundholm 2003)), or one argues, as Shin does, that the objectual approach to pictures is in principle possible, and that it would simply be a matter of perfectioning the iconic languages in question (and notice that these two positions are not incompatible).

Similarly, as already discussed, this sense of formal seems to speak against the view that artificial notations ought to be as iconic as possible, as the purpose of formal languages would be precisely that of excluding meaning. To this it may be replied that, admittedly, iconicity may indeed be appropriate mainly for meaningful languages; but, as already argued, even when it comes to uninterpreted languages, iconicity may be used to give an intuitive expression to the salient **relations** and **articulations** between, and **structures** defined by, the objects in question, that is, the meaningless terms, in particular with respect to how they can be concatenated (i.e., the order of the symbols in well-formed formulas of the language, defined by its syntax).

Be that as it may, whether the symbols are iconic or arbitrary, the tight connection between this notion of the formal and symbolization is patent.

(2c) Symbolization and the variational notion of the formal. As already mentioned, the variational notion of the formal is historically very important; now, it is by no means a coincidence that the first uses of artificial notation in logic, in particular by Aristotle, were strictly related to this notion of the formal. In fact, one might even say that the variational notion was the original motivation for the first efforts towards symbolization. In effect, while it is possible to express the indetermination/insaturation on which this notion is based with ordinary words, using expressions such as 'something', 'some …' (where '…' is to be filled with the appropriate substitutional sort), this procedure quickly becomes exceedingly cumbersome, as the passage by Tarski in section 3.2.1 shows. Hence, from the variational vantage point, symbolization undisputably enhances formality.

(2d) Symbolization and formality as indifference to particular objects. The same holds of formality understood as indifference to particular objects, which is the objectual counterpart of the (essentially linguistic) variational/substitutional notion. In effect, symbolization is a powerful device for the correct definition of the concept of permutation invariance and for its applicability, in particular due to the use of letters as placeholders, indicating the gaps where permutation is to occur. One wonders whether the model-theoretic approach to logic introduced by Tarski, based on the idea of interpretation in models, and which has been so successful, would have been possible at all if it wasn't for the symbolization tools discussed here, as it seems very counterintuitive to give different interpretations to words with an already familiar meaning.

4.3.4 Conclusion

Even though I have outlined some of the conceptual difficulties concerning the concept of symbolization, I believe I have shown as well the various ways in which symbolization enhances formality, and thus why and how to symbolize is to formalize. In other words, that these two terms are treated virtually as synonymous by most logicians nowadays may be questionable, but it is certainly quite natural. It must be borne in mind, though, that formalizing amounts to more than solely symbolizing: in particular, the establishment of the deductive structure of an argument or theory – an axiomatization, in the loose sense of the term – is just as important.

Moreover, I have investigated what a successful symbolization consists of; my conclusion was that, the more pictorial a symbolization is, the more chances it has of being a successful symbolization, provided that no loss of generality occurs (as it can occur).

Now, with respect to the formalizations of fragments of medieval logic presented in the previous parts, symbolization seems to be one of their most salient aspects. As will be discussed in the next part, since I have often made use of 'ready-made' modern logical theories, in most cases I have simply borrowed their original notation (*pace* some occasional adaptations). So in this sense, in many aspects regarding the formalizations presented here, one cannot truly speak of a choice of the appropriate notations. An exception to this is the formalization of supposition, for which I have not used a previously existing uniform symbolic framework. In that part, I have attempted to express the relations between terms and their *supposita*, or between the different concepts of the theory, as much as possible, in a iconic way; but I have also tried to maintain a certain familiarity by means of the use of well-established conventions, so as not to tire the reader with tedious and unnecessary learning of too many new symbols.

In any case, the ultimate test for a symbolization is always whether the reader is able to grasp the meanings one tries to convey. Even though symbolizations are also typically intended to allow for efficient calculations, the symbolizations of fragments of medieval logic presented here have predominantly an expressive purpose, namely that of explicitly displaying the articulations, structures, concepts and operations implicitly involved in these logical theories.[469]

4.4 CONCEPTUAL TRANSLATIONS

Taking a brief glimpse at different formalizations in various topics, one quickly realizes that one of the most common procedures practiced to formalize a non-formal(ized) theory is to translate it, so to speak, into another, recognizably formal, theory or language. The indisputable champion among formal languages to be used in formalizations is still first-order predicate logic (FOL), as it is still frequently viewed as **the** logic; but, clearly, this predominance of FOL is not a matter of principle, and the application of alternative systems of logic in formalizations is in fact just as legitimate, or in some cases even more so, than the application of FOL.

There are several important issues surrounding the notion of translation of a theory into another. First of all, one cannot really speak of a mere translation of the **language** of a theory into the **language** of the formal theory used for the formalization. Indeed, when a translation from an ordinary language into another takes place (say English into French), a common conceptual framework is presumed to exist, and the translation of a word in another language is typically the corresponding word that is subsumed to the same concept, that is, that has the same meaning. Naturally, it often happens that there is no exact translation for a given word into another, or for a given syntactical structure, but in most cases this process is unproblematic: clearly the

[469] Here again, the main inspiration is Brandom's view of logic as 'making it explicit'.

words ‘homo’ in Latin, ‘man’ in English and ‘homme’ in French are all translations of each other.

But when it comes to the formalization of a theory, such a common conceptual framework is not readily available. True enough, the respective conceptual frameworks of each theory typically have similarities with one another – and this is in general the motivation for the choice of a given formal theory to formalize a non-formal(ized) theory; but these conceptual similarities are usually not entirely straightforward, that is, the concepts underlying each theory are not simply identical. Therefore, a **conceptual translation** between them must be carried out, that is, an analysis of the aspects in which similarity occurs (and those in which it does not), since their common ground cannot be simply assumed. Therefore, I call the use of a given formal theory to formalize another theory a **conceptual translation**, to stress the fact that more important than the mere translation of the language of one theory into that of the other is the adequate mapping of the basic concepts of the theory to be formalized into the basic concepts of the formal theory to be used. Clearly, this mapping is a necessary justification for the validity of the conclusions drawn about the formalized theory by means of the apparatus of the formal theory used, as I shall argue below. Of course, the use of the formal theory’s language in the formalization generally occurs as well, among other reasons precisely in order to outline conceptual analogies, but this linguistic translation is in fact a corollary of the more fundamental conceptual translation.

In what follows, I will first discuss some of the history of conceptual translations in logic as well as some foundational issues; I then examine the status of the results obtained with such translations; finally I turn to the particular conceptual translations undertaken in the present work. Indeed, while I have already motivated my choice of a given theory to formalize this or that fragment of medieval logic in the course of the formalizations properly speaking (in the previous part), a more systematic explanation of my choices seems to be a necessary addition to the foundational analyses of this part.

4.4.1 What is conceptual translation?

As the intention here is to give a foundational analysis of formalization, and as I claim that conceptual translations are often at the core of many formalizations, I here examine the grounds for undertaking such translations. As a preliminary appendage, I briefly examine the history of a very important conceptual translation, namely that of logic into mathematics, in order to clarify further what I mean by ‘conceptual translations’.

4.4.1.1 The history of conceptual translations

Logic as we now know it is a product of a major project of conceptual translation. As is well known, for many centuries logic was considered to be a subdivision of philosophy. In many languages, logic was also known as the art of reasoning (*Redeneerkunde*, in Dutch); in medieval Latin, *logica* and *dialectica* were synonymous. Naturally, insofar as it defined the principles permeating correct reasoning in all disciplines, and since knowledge and reasoning are always expressed in some language, logic as a discipline was considered to have many affinities with language studies; indeed, in the medieval academic curriculum it belonged to the *trivium* (logic, rhetoric and grammar), which, together with the *quadrivium* (music, astronomy, arithmetic and geometry) formed the program for the first (roughly seven) years of study.[470] In short, while it had an obvious foundational status, logic was also often treated on a par with language-related disciplines.

Naturally, nothing is more at odds with the present status of logic as a sub-branch of mathematics. To be sure, deductive reasoning has always been one of the defining traits of mathematics, and the rules of deductive reasoning are traditionally one of the main subjects of logic; moreover, even prior to the mathematical turn in the 19th century, there had been attempts of what one could call 'applications' of logic to the foundations of mathematics, as well as abundant use of mathematical examples to illustrate logical points (in both cases Aristotle is a noteworthy example, cf. (Mendell 2004)). So it would be incorrect to deny the existence of connections between the two disciplines before the 19th century. But never has the magnitude of these connections been as it is now.

The story of the transformation of logic as a discipline has been told in many places, and the passage from logic as philosophy to logic as mathematics becomes particularly conspicuous in overviews of its history such as (Kneale and Kneale 1984); what matters for the present investigation, though, is that this passage took place essentially by means of conceptual translations, which outlined the conceptual and structural similarities between portions of mathematics and logic.

Three authors are usually credited with the title of founder of logic as a mathematical discipline, to wit Leibniz, Boole or Frege (cf. Peckhaus 2004).[471] Scholars differ in their attribution of this title to either one of these three authors mainly in function of the logical trend of their preference (cf. Peckhaus 2004), and of the degree of mathematization of logic that they are willing to accept as the real start of the new discipline. Two of them, Boole and Frege, were mathematicians who took interest in logic, precisely because of the conceptual similarities identified by them; by contrast, Leibniz was one of the last pan-scholars who excelled in a variety of fields, including metaphysics, logic, mathematics and law. It has been argued, however, that logic and mathematics were really the core of Leibniz's doctrines (cf. Couturat 1901); but in the following I will only, and briefly, focus on Boole and Frege, and will not discuss Leibniz.

To be sure, what makes these authors the initiators of the mathematical approach to logic is not that the rigor and formal correctness of their systems were in all senses superior to those of preceding systems. In fact, as argued by Corcoran (2003), Aristotle's syllogistic is in many respects superior to Boole's logic (albeit of narrower

[470] Cf. (Kenny and Pinborg 1982), (Marenbon 1987).

[471] In this paper, Peckhaus focuses on Boole and Frege, and only mentions the trend that attributes the paternity of mathematical logic to Leibniz.

scope); it is, for example, sound and complete, while Boole's is not. But what is distinctive and innovative in the approach to logic proposed by these authors is the analogy between some key concepts of each discipline, logic and mathematics.

By the expression 'mathematical analysis of logic' Boole did not mean to suggest that he was analyzing logic mathematically or using mathematics to analyze logic. Rather his meaning was that he had found logic to be a new form of mathematics, not a form of philosophy as had been thought previously. More specifically, his point was that he had found logic to be a form of the branch of mathematics known as mathematical analysis, which includes algebra and calculus. (Corcoran 2003, 264)

Boole identified several important conceptual similarities between mathematics and logic. For example, he deemed that the main processes and operations in logic were, or in any case ought to be, analogous to some of those in mathematics, such as the substitution of equals for equals and the application of the same operation to both sides of an equation (cf. Corcoran 2003, 270). Moreover, for him the basic logical forms of propositions were not to be expressed in terms of quantifying expressions, negating terms and the copula, such as in Aristotelian syllogistics; rather, these forms were really equivalent to equations:

For Boole, the logical form of a proposition such as 'Every square is a polygon', treated by Aristotle as 'Polygon belongs-to-every square', is really an equation, two terms connected by equality, an equation in which nothing corresponds to 'belongs-to-every'. Here, equality is the strictest mathematic equality, 'is-one-and-the-same-as', also called numerical identity. (Corcoran 2003, 271)

That is, Boole intended to turn the expressions of logic into expressions as similar as possible to equations, so as to be able to apply to them the same operations usually applied to the latter.[472] Considering the extensive use of the mathematical sign of equality and of variables in current logic, among other elements, it is clear that many of Boole's ideas concerning this specific point have been fully incorporated into logic. Naturally, there are many other aspects concerning which Boole's application of mathematical concepts to logic has been influential; to name but one, we now call Boolean algebra the algebra of two-valued logic with only sentential connectives. In other words, Boole's conceptual translation has undoubtedly left a lasting mark on logic as a discipline.

As for Frege, his mathematization of logic also occurred in many aspects, but, regarding conceptual translations, one of them seems particularly important: his importation of the notion of **function** into logic. What Frege did was to cast fundamental logical concepts in terms of the concept of function (as is especially clear in his *Function and Concept* – Frege 1980): the crucial move was that of expanding the range of possible arguments for functions beyond the class of numbers, which were hitherto the only entities considered to be adequate arguments for functions. Here are some examples of his use of functions to account for logical concepts: according to

---

[472] Interestingly, Frege heavily criticized Boole's conceptual translation of logic into mathematics: according to Frege, there are significant disparities between the mathematical operations of multiplication and addition and the logical operations which Boole identified them with, to such an extent that Boole's conceptual translation becomes unwarranted. (cf. Frege 1972, 3/4)

Frege, the basic logical form of sentences corresponds to the dual function-argument; he accounts for the notion of true content with the idea that true contents are mapped onto the object 'True' by the function expressed by the horizontal stroke, while false sentences are mapped onto the object 'False'; the negation 'is the one [function] whose value is the False for just those arguments for which the value of —x is the True, and, conversely, is the True for the arguments for which the value of —x is the false' (Frege 1980, 34/5); and so forth.

Frege's definition of these logical concepts in terms of the (originally mathematical) concept of function is an excellent, perhaps even one of the most important, example(s) of conceptual translation of logic into mathematics. As already mentioned, Frege also imported some of the notational principles of mathematics into logic (cf. section 4.3.2.3 above). Interestingly, in the case of Frege, the conceptual translation between logic and mathematics in fact took place in both directions; once he had mathematized logic, Frege went on to use logic to build the foundations of arithmetic, in his *Grundgesetze*. In both cases, these conceptual translations are as of now still extremely influential: logic as a discipline of mathematics is, obviously, still thriving, and the program of giving mathematics logical foundations is, although not unanimously accepted, still carried out (see Shapiro 1991).

Another historically important case of a formalization consisting of a conceptual translation is Gödel's encoding of formulae of an already formalized language into arithmetic for the incompleteness theorem. He then used some of the properties of the natural numbers, such as their well-ordering, in his incompleteness proof. However, it has been noted[473] that numbers are just a convenient way of achieving the process of **naming** formulae and terms, which is really what is required for the proof; in fact, the latter can be constructed with non-arithmetical languages as well, and only minor adaptations are required. In this sense, with Gödel's theorem we may be essentially dealing with a case of symbolization rather than with a case of conceptual translation; but as far as Gödel's actual proof was construed, one is licensed to speak of a conceptual translation.

4.4.1.2 Foundation for conceptual translation: conceptual identity and conceptual similarity

Hence, the program of conceptual translations is undoubtedly fruitful. Nonetheless, some foundational issues concerning this notion must be addressed, and again it turns out not to be as simple a notion as might have been expected.

As already mentioned, translations of an ordinary language into another are made possible by the presumed sameness of meaning between words of different languages (but even this assumption is often criticized, as in Quine's thesis of indeterminacy of translation and his famous *gavagai* example – cf. (Quine 1960)). Now, when it comes to conceptual translation, what does this 'sameness of meaning' amount to? Evidently,

---

[473] I first heard this suggestion at a presentation by H. Gaifman, but during the discussion following the presentation it has been mentioned that R. Smullyan made a similar remark in one of his books (presumably, *Diagonalization and self-reference*).

clear criteria of 'sameness', or in any case of resemblance, between concepts are required, but these are far from being readily available; if fact, it is uncertain whether they exist at all. I examine here some suggestions as to what these criteria might look like, but I acknowledge in advance that, even if adequate definitions could in principle be reached (which may not even be the case), significantly more work would be required.

A minor clarification: here, I will mostly be concerned with conceptual similarity of compounds of concepts, namely the concepts and conceptual structures underlying whole theories, and less so with similarity between simple concepts.

1) **Extensional notion of conceptual similarity.** According to what we could call an extensional view on concepts, all there is to a concept are the individuals falling under it. According to this view, all co-extensional concepts are identical; the concepts 'cordate' and 'renate' are, on this view, identical.[474] Now, with respect to concepts within theories, what could correspond to their extension? Something like 'the set of the possible worlds in which the theory is true' is a good candidate for the job, but it is not very helpful for our current endeavor, as it seems just as problematic, if not more, to come up with a criterion of similarity between sets of possible worlds.

The concept of 'model', borrowed from model theory, seems a bit more helpful. For the sake of the argument, let us assume that there is a method to determine the model(s) underlying a given theory (obviously, a great deal of logic, mathematics and conceptual analysis would be required for this). Then, one may say that two theories are conceptually similar iff their extensions, that is, the respective underlying models, are similar. As a matter of fact, similarity between models is a much more tractable problem; there are technically adequate accounts of comparative criteria for models, such as isomorphism and bisimilation, which offer different grains of comparison.

For instance, consider two logics that are usually interpreted on Kripke-structures. If these logics have as their underlying models for example, S-4 or S-5 Kripke-structures, then they may be said to be conceptually similar, and thus conceptual translation between them is possible.

Nevertheless, there are some reasons why an extensional account of conceptual similarity is not satisfactory. Firstly, with the exception of highly formal theories, it is improbable that the underlying model(s) of every theory could be established; and since we are dealing precisely with theories not yet formalized (whose conceptual structure has to match that of an already formalized theory), this is a major obstacle. Moreover, the same objections that are usually brought up against a strictly extensional view of individual concepts apply here too: the extensional criterion is too coarse and does not discriminate between concepts or conceptual structures that are in fact very

[474] Of course, one may add a modal component to this definition, in which case these terms may not be identical, since even though it is the case in the actual world that all creatures with hearts also have kidneys, this might not have been the case.

dissimilar. Likewise, it may not recognize similarities that do exist but which do not amount to co-extensionality.

In any case, when all the requirements are available (the underlying models of two theories are determined, and there are clear criteria of identity and similarity between them), then extensional similarity may be a good indication that there is conceptual similarity, and thus that conceptual translation is possible.

2) **Similarity of conceptual structure.** With respect to simple concepts, more fine-grained, intensional criteria of identity and similarity between concepts have also been proposed. Stemming from Frege's distinction between sense and reference (Frege 1948), it has been suggested that concepts that share the same extension may nevertheless be differentiated from one another. In the case of 'cordate' and 'renate', clearly they have different senses, as the former is attributed to creatures insofar as they have hearts, while the latter is attributed to creatures with kidneys.

Now, the same idea can be applied to complex conceptual structures as well. Suppose that, in two different theories that are interpreted in the same model, two concepts, each belonging to one of the theories, have the same extension – for example, if they are relational concepts, the pairs of entities related by them are the same. According to the extensional notion, these would be similar, or even identical, concepts. Nonetheless, assume that, taking each of the theories as a whole, the respective places occupied by propositions featuring each of these concepts in the inferential structure of the theory is entirely different: while one is a fundamental concept in one theory, the other is a rather unimportant, derived concept in the other. In this case, clearly there is a sense in which these two concepts are very dissimilar.

In sum, another condition of sameness or similarity between concepts within theories may be that they ought to occupy similar positions in the general deductive structure of a theory; this idea is essentially borrowed from inferentialist accounts of meaning, and insofar as meaning is usually seen as an intensional property, this may be called an intensional criterion of conceptual equivalence and conceptual similarity. Again, this is just a suggestion, which would have to be considerably worked out before it becomes an effective criterion of conceptual similarity, but it provides an indication of the kind of criteria of conceptual similarity that I have used for the conceptual translations undertaken in previous chapters.

3) **Similarities of goals and procedures.** A less technical, more pragmatic, but perhaps equally promising approach is to take a global look at theories and to consider the goals to be attained by their applications and the procedures used in them to attain these goals, in order to pass judgement of sameness or similarity of concepts and thus of theories.

This seems to be the case, for example, with one of Boole's justifications for applying mathematics to logic: according to him, the calculating **procedures** used in analysis have valid logical procedures as their counterparts, as exemplified by the substitution of equals by equals, which is related to the notions of substitution *salva veritate* and to the criterion of substitution defining logical validity and formality.

In effect, while in logic the operation of substitution was originally conducted in a rather loose way, the application of the corresponding mathematical techniques by people like Bolzano, Boole and Tarski very much enhanced rigor in logic; now, in virtue of the conceptual similarity, this happened without any threatening conceptual misprojection.

Likewise, two theories may have similar goals, and this may be the foundation for their conceptual similarity. Going back once again to the example of mathematics and logic, both fields have a notorious concern with valid **deductive** reasoning, and, even more than specific techniques and operations, this may be the main point of similarity between them – presumably, they both investigate the same system(s) of valid deductive reasoning.

In any case, criteria of identity and similarity for any kind of entities are, generally speaking, among the thorniest issues in philosophy; naturally, when it comes to concepts, elusive entities themselves, this issue is bound to become even more complex. Therefore, at this point, I do not see how a uniform account of the notion of conceptual similarity could be given; seemingly, a case-by-case analysis is required in order to justify the process of conceptual translation, which I claim is an important component of various specific formalizations. That is, once more we encounter a foundational difficulty with respect to formalizations, suggesting that the application of this technique must be carried out with significantly more consideration than is usually done.

What is perhaps paradoxical concerning conceptual translation is that, while it is used to formalize theories, it is itself **not** a formal technique. Precisely because one of the theories involved is not in formalized form, the usual formal techniques cannot be applied. Conceptual translation is, as the name indicates, a conceptual procedure, just as much as axiomatization (the determination of the appropriate axioms and rules of inference of a theory) and symbolization (the choice of the appropriate language to express a theory); it presupposes a great deal of conceptual reflection before formal techniques can be approached. In other words, paradoxical though as it may seem, formalization seems to be essentially a conceptual procedure, not an algorithmic one, and therefore it could never be carried out by a machine.

4.4.2 The outcome of a conceptual translation

As of now, it seems impossible to define uniform methods for conceptual translation (or for formalization in general, for that matter); we only have a few guidelines orienting the case-by-case analysis. Naturally, with highly formalized languages, model-theoretical techniques can be applied to show equivalence or similarity between concepts and/or formulas; likewise, proof theoretical techniques can be applied in the case of fully axiomatized theories, to show, for example, that all the statements of a theory are theorems of another theory. But if a theory is not yet as worked out, and if the intention is precisely to produce a formalization of it by means of the application of a formal theory, tailor-made purely conceptual analysis seems to be required. Notwithstanding this difficulty, translations from one theory into another are indeed often carried out; we may suspect thus that significant gains are obtained with this technique for formalization, and we may do well to investigate what precisely is obtained by means of conceptual translations.

4.4.2.1 Formal semantics[475]

As I see it, the establishment of a formal semantics to a given language (be it formal or 'informal') is a case of conceptual translation. Currently, the typical approach to a 'logic' is first to introduce a (meaningless) formal language, that is, a finite alphabet and the rules to generate the sentences of the language from the alphabet, and then the 'semantics' of this language, that is, the objects and structures that are supposed to be the denotations of the originally meaningless 'terms' and 'sentences'. These structures are usually the objects of theories in their own right; set theory, for example, is a very popular tool for the establishment of the semantics of a language. In the case of first-order predicate logic, for instance, individual terms are assigned to individuals in the domain, and predicate terms are assigned to sets, in such a way that statements in the language become statements about the structure in question. Logical operators can then receive an intuitive interpretation in terms of sets (disjunction = intersection; conjunction = union etc.).

> One way of giving meaning to the various syntactical entities of a formal language is by modeling it on the way with which we are all familiar: the typical case would of course be the standard modeling of first order predicate logic. How does it proceed? […] To each individual term you assign an individual, that is, an element of the individual domain, and to each formula you assign a proposition. (Martin Löf 1987, 407)

Another kind of structure that is very popular for formal semantic purposes are possible-world Kripke structures. Applied to modal logic, such structures notoriously yield an intuitive interpretation of otherwise thorny concepts, such as 'necessary' and 'possible'; they are also applied to other logics of the same family such as epistemic logic, deontic logic, tense logic etc.

Why is the mapping of terms and sentences of a language into previously defined and rigorously studied structures said to be a **formal** semantics? As already noted, a frequent complaint against 'ordinary' languages used in scientific purposes is that they do not exclude phenomena such as ambiguity and vagueness, which arguably are out of place in scientific contexts; in other words, such languages are too chaotic and not sufficiently regimented. Now, given the notion of the formal as regimentation and accordance to well-defined rules, it becomes obvious that the procedure of mapping terms and sentences of a language unambiguously into certain structures, in order to confer (or clarify their) meaning is a formal approach to meaning, whence formal semantics.

Naturally, it is very convenient to use well studied structures to provide a formal semantics to a language, as the known properties of these structures can immediately be transferred into the language in question, and thus to the contents being expressed

---

[475] Notice that 'formal semantics' is understood here in its logical sense, that is, with respect to logical theories, and not in the sense of the study of the formal properties of so-called 'natural languages'.

by the language. But the choice of structures (and the theory accompanying them) to build a formal semantics for a language or theory cannot be aleatory, at the risk of disregarding much of its conceptual peculiarity. Therefore, a formal semantics is all the more adequate to a theory when there are significant conceptual and structural similarities between the theory receiving a semantics and the theory being used for this purpose – in other words, when it is a case of a legitimate conceptual translation. Hence, it is clear that the foundational issues discussed here regarding conceptual translations apply just as much to the construction of formal semantics for (logical) theories.

4.4.2.2 Transference of formality

In the early days of formal axiomatics and metamathematics, one of the preferred techniques for proving the consistency of a given theory was to show that it was reducible to another, presumed-to-be consistent, theory. This technique is based in one of the oldest axioms in logic, according to which the impossible (what is inconsistent) cannot follow from the possible (what is consistent): so if a theory T is consistent and implies T', then T' is consistent as well. At that time, the preferred presumed-to-be consistent theory for relative consistency proofs was arithmetic/analysis (cf. Zach 2003b on Hilbert's relative proof of the consistency of geometry), as its consistency was considered at first to be beyond any doubt. But Hilbert soon realized that the consistency of analysis had to be proved, and that a **direct** consistency proof was required, given the foundational status of analysis *vis-à-vis* other theories, thus inaugurating what became known as Hilbert's program (cf. Zach 2003b). Nowadays, especially after Gödel's results, there is consensus to the effect that, no matter how robust a theory appears to be, its consistency (or whichever one of the desirable properties of a theory, such as soundness, completeness etc.) cannot not be taken for granted. Nevertheless, if a given theory has been proved to be consistent, then a relative proof of consistency, reducing the theory to be proved consistent to the already proved to be consistent theory, is a very respectable and important technique.

Now, what is the connection between relative proofs of consistency and our current topic, conceptual translations? As I see it, formalizations that proceed by translating a (non-formalized) theory into another theory, recognized to be formal, are in a sense also **relative proofs of formality**. In the case of consistency proofs, what justifies the 'transference' of consistency is the fundamental logical property that the set of statements that follows from a consistent set of statements is itself consistent. By contrast, formality seems to require more than logical implication to be transferred; rather, what seems to be required is some kind of identity between theories, or at least isomorphism, so that a theory can be said to be formal insofar as it is isomorphic, in some relevant sense, to a theory known to be formal. Moreover, formality is a much less tractable notion than consistency, as is attested by the variety of notions of formal considered here (cf. section 4.1.3).

Hence, the role fulfilled by the appropriate notion of logical implication in the case of relative consistency proofs is, in the case of formalizations, fulfilled by the (admittedly elusive) notions of conceptual similarity and conceptual translation. But assuming that, in a given case, the conceptual translation is deemed to be adequate and that one of the theories involved is recognized to be formal, then it seems that the formalized theory rightly deserves the title of formal. In other words, when there is doubt as to whether a given theory may be said to be formal, a conceptual translation may be used as a relative proof of formality.

Naturally, such proofs may have limited scope as well, for example, a given theory may be said to be formal only with respect to certain of its aspects. Typically, non-formalized theories have formal as well as informal aspects, and a conceptual translation may serve the purpose of outlining its formal as well as (by opposition) its informal aspects (cf. the formalization of supposition theory presented in part 1).

In any case, this seems to me to be one of the main motivations behind various formalizations: to show that the theory being formalized is indeed 'formal', at least with respect to certain aspects and according to some notion of the formal. In effect, this is one of the explicit goals of the present investigation, namely to show that some fragments of medieval logic, whose general logicality is often questioned by those accustomed to the idea of logic as essentially a branch of mathematics, are indeed formal, and thus logical (even according to the current understanding of what logic is), in many important senses. Now, one of the most natural ways of showing that a theory is logical is, as I have attempted to do, to stress its resemblance with other theories recognized to be logical, that is, to carry out a conceptual translation of one into the other.

4.4.2.3 Dialogue

However, by stressing these similarities I have never intended to disregard the profound dissimilarities between modern and medieval logic. In fact, if the only purpose of studying medieval logic was to show to what extent it is a predecessor of modern logic, it is unlikely that valuable insights would be attained; naturally, according to the current standards of what is to count as logic, medieval logic is at best a primitive version of what logic was to become in the 19th and 20th century. From this perspective, the interest in studying medieval logic would not go beyond the observation of an interesting but dusty and inefficient museum piece. Now, one of the main assumptions of the present study is that medieval logic has much to contribute to modern logic, not so much in terms of techniques, but mainly on a conceptual level.

Still, for the dialogue between modern and medieval logic to be established at all, a certain common ground must exist. Considering the disparity of means of expression (regimented medieval Latin vs. symbolic languages supplemented by ordinary languages) and of general conceptual background, one can easily talk of two entirely different paradigms. Now, with respect to this particular case, Kuhn might be right in claiming that different paradigms cannot have a dialogue, if no work was done to establish minimal common grounds for the dialogue; as is certainly more than obvious by now, one of the main goals of the present work is to establish these minimal common grounds and thus to allow for a fruitful conversation between the medieval and the modern paradigms in logic.

Indeed, the conceptual translations presented here can also be seen as translations in the most common meaning of the term, as procedures to allow people speaking different languages to talk to each other. And this can hold of paradigms that differ not in historical period, but which are practiced simultaneously – perhaps one of the best examples is the lack of dialogue between the so-called analytic and continental traditions in philosophy.

Nevertheless, conceptual translations and formalizations are admittedly, at least to some extent, always modifications of the original theories and paradigms (just as much as any translation always winds up modifying the original). In the present case, it is clear that any formalization of medieval logic is inherently anachronistic; therefore, it is useless to try to avoid entirely a certain dose of anachronism in investigations such as this one. What must be sought after is for the anachronism to be well founded (on a significant conceptual similarity) and illuminating; indeed, my goal will have been attained if (i) scholars of medieval logic are able to recognize the original theories in the formalizations; (ii) logicians are able to understand better the medieval theories being formalized; and (iii) (the most ambitious of them all) the medieval ideas presented here serve as inspiration for new conceptual (and corresponding technical) developments in logic.

In any case, if a dialogue between two traditions as dissimilar as medieval and modern logic is to exist at all, a comprehensive work of translation must occur; in this particular case, since the translation needed is from medieval into modern logic, it naturally takes the form of formalizations, as the elements involved in formalizations – above all axiomatization and symbolization – are distinctive characteristics of the current paradigm.

4.4.3 Conceptual translation in the present work

Thus far, I have attempted to clarify what is meant by 'conceptual translation'; however, probably the best way to achieve this goal is by discussing examples of conceptual translation (I have already mentioned a few), and in particular the conceptual translations undertaken in the present work.

Moreover, with respect to axiomatization and symbolization, I have examined in what senses these practices enhance formality; here, by contrast, insofar as conceptual translation is a relative proof of formality, this practice will transfer to the object of formalization the specific kind of formality of the formal theory being used. Therefore, a discussion of specific notions of the formal is not appropriate when it comes to conceptual translations. We can speak here of relative, or derived, formality, dependent on the kind of formality of the theory into which the translation occurs.

I now turn to formalizations of medieval logic in general, and to the formalizations presented here in particular. What must be noted concerning previous formalizations is that most of them were essentially translations of fragments of medieval logic into first-order predicate logic (FOL). While this is understandable insofar as this logical language is still in many senses the canon of what is to count as logic, this choice is also questionable in many aspects. Two aspects seem particularly worth mentioning: (1) While FOL is essentially non-procedural and the emphasis is laid on syntax (in its proof-oriented approach), medieval logic is essentially procedural and semantic. Naturally, a semantic (i.e., model-theoretic) approach to FOL is also possible, and indeed quite common, but this is usually not the approach adopted in the formalizations of medieval logic using FOL. (2) There is a fundamental discrepancy between the medieval conception of the basic logical form of sentences and that of FOL. While medieval logicians typically worked with the triad subject-copula-predicate (sometimes the form subject-verb was also considered), FOL is famously based on the function-argument dichotomy, introduced by Frege precisely in opposition to the traditional subject-predicate dichotomy (cf. Frege 1879).

Clearly, FOL is not the best choice for conceptual translations of medieval logic into modern logic. Medieval logic is predominantly what we now call semantics (to be precise, formal semantics). Hence, according to the conception of logic predominant in the first half of the 20th century (centered on the notion of deductive systems), medieval logic stood less of a chance to be recognized as logic. But since the semantic, model-theoretic approach was introduced, its chances of reaching the logical Parthenon improved considerably. While proof-oriented formalizations of medieval logic (with FOL or other) seemed essentially misguided, it seems to me that the use of semantic techniques brings to light what is properly formal, and thus logical, in medieval logic. Hence, the choice of modern theories for the conceptual translations and formalizations undertaken here was guided by the medieval semantic focus. Techniques from model-theory and two-dimensional semantics proved to be particularly useful.

Moreover, insofar as medieval logic was fundamentally a *logica utens*, that is, it comprised instructions on how to handle a variety of issues, emphasis had to be laid on its **procedural** aspect; it was essentially a logic to be applied, even though higher levels of analysis (concerning the properties of these systems as such) also occurred (cf. Ebbesen 1991). The term ‘procedural’ is here understood as emphasis on instructions for the performance of actions, manipulations and calculations (in a loose sense of the term ‘calculation’, in the sense of regimented procedures having input and output). Furthermore, it invokes a dynamic aspect as well, that is, a certain order in which events and instructions are carried out, in particular insofar as an action may depend on the outcome of the previously executed actions (as is particularly clear in the case of *obligationes* and, to a lesser extent, of supposition).

I now discuss each of the three topics treated in the present work, namely supposition, *consequentia* and *obligationes*.

Supposition Theory. Unlike the other topics formalized here, I haven’t found a satisfactory counterpart in modern theories for what seemed to be going on with supposition theories. In this sense, the formalization of Ockham’s supposition theory presented here is the only case study where conceptual translation did not occur, but rather axiomatization and symbolization alone. What seems to come closest are theories of computational meaning generation, in the sense of what is now known as computational semantics[476] – hence the name of the chapter, 'Supposition theory as algorithmic hermeneutics'. What is clear is that supposition theory (at least in Ockham's version thereof) is procedural in that it comprises a set of rules on how to interpret (a certain class of) propositions based on syntactical and formal features, that is, on how to generate their possible reading(s): the input are propositions, and the output are their possible readings.[477]

Obligationes. The procedural nature of *obligationes* is even more patent, as it was indeed an activity performed by actual people – a real game, with rules, winners and losers. In this sense, the choice of the game-theoretical framework for the conceptual translation of *obligationes* was a rather obvious one, as it is (like games) essentially a rule-governed and goal-oriented activity. Moreover, the principles governing the game are essentially semantic: propositions known to be true must be accepted, propositions known to be false must be rejected etc. Even the instructions that seem at first to depend on syntactical logical concepts, namely based on the notions of 'follows from' and 'is repugnant to', can be formulated in semantic terms. It is for this reason that, alongside with the game-theoretical framework, model-theoretic notions were also used to formalize *obligationes*.

Consequentia. As is explained in the corresponding chapter, Buridan's notion of consequence is essentially semantic (as opposed to syntactic) and in many senses akin to the modern representational and model-theoretic notions of logical consequence. Therefore, the model-theoretic framework was particularly suitable for the formalization, that is, conceptual translation, of Buridan's theory of consequence. Model-theoretic tools were used in the section concerning the definition of consequence as well as in the proof of soundness of a specific fragment of Buridan's logic. Moreover, to account for Buridan's commitment to tokens as truth-bearers, the choice of two-dimensional semantics was also a natural one. In fact, the analogy between Buridan's treatment of tokens and two-dimensional semantics is one of the best illustrations of conceptual similarity as I understand it here.

In sum, what may have become clear is that medieval logic is probably best compared to what is now known as formal semantics, that is, the application of formal/logical tools for the analysis of semantic phenomena. To be sure, logical notions such as those of consequence and other logical connectives (disjunction and conjunction) were also discussed, but there was a predominance of semantic issues.

---

[476] '…computational semantics is that part of computational linguistics concerned with developing methods for computing the semantic content of natural language expressions …' (Blackburn and Kohlhase 2004, 117)

[477] In (Klima and Sandu 1991), an interesting parallel is drawn between the modes of personal supposition and game-theoretical semantics, with respect to numerical quantifiers. This parallel also confirms my claim that supposition theory must be viewed as essentially procedural (as is the case of game-theoretical semantics).

Accordingly, the choice of modern theories with which to formalize the medieval theories presented here followed this semantic penchant, as well as what I have termed their procedural nature.

4.5 CONCLUSION

It seems to me that the main conclusion to be drawn from the considerations in this part is that there are important foundational issues regarding formalizations, which must be addressed. By this I do not mean to say that formalization is an illegitimate practice, not to be carried out before we have found solid foundations for it; in fact, few theoretical constructions can afford the luxury of relying on rock-solid foundations. Clearly, if knowledge-producing practices could only be carried out in the presence of the appropriate foundations, we would probably be forced into catatonic skepticism. Nevertheless, it all seems to indicate that the very practices of formalizations can be greatly improved if these foundational issues are addressed by their practitioners.

Furthermore, very important is the recognition that, contrary to what is often assumed, formalization goes beyond symbolization; the organization of the statements of a theory in their correct deductive structure, which can be loosely referred to as 'axiomatization', is just as important. Moreover, if we are fortunate enough to encounter a thoroughly formalized theory that displays significant conceptual similarities with the theory T to be formalized, we may take a few shortcuts, that is, not having to address issues such as the determination of T's deductive structure and the choice of the appropriate symbolism, and perform what I have termed conceptual translation in order to formalize T. In sum, by putting forward the distinction between axiomatization, symbolization and conceptual translation, I have attempted to clarify different aspects involved in formalizations.

Moreover, what transpires from these considerations as well as from the formalizations in the previous chapters is that (suitable) formalizations, especially in the case of past philosophical theories, are above all a conceptual endeavor, more than a strictly logical/technical one. As I see it, the absence (and impossibility) of a uniform method for formalizing indicates that a case-by-case analysis is required, and that, more than technical skills, what is required from the formalizer is a profound conceptual understanding of the theory to be formalized: formalization is itself not a formal procedure.

CONCLUSION

1 RETROSPECT

If much must be said in a conclusion, this means that too little, with insufficient clarity, has been said throughout the main text. So here I do not intend to revisit all the results obtained in previous chapters, as they should speak for themselves. I will, though, offer a brief overview of the structure underlying this work, in order to outline the mutual connections between the different topics. But, most of all, I will attempt to draw a few conclusions on what can be said concerning the nature of logic from the analyses carried out so far.

This work presents formalizations of three of the most important logical topics in the medieval literature, namely supposition theory, the notion of consequence and the rules for the form of oral disputation known as *obligationes*. Obviously, with this choice I do not want to imply that these are the only topics in medieval logic, or even that they are the most important ones. They are simply those that attracted my attention for a variety of reasons, and which I deemed to be particularly permeable to formalizations. In other words, these three topics are best viewed as case studies attesting to the fruitfulness of the application of formalizations to medieval logic.

There are also interesting connections between each of these topics. Supposition was, as is widely acknowledged, a crucial concept for later medieval authors. It was used to account for a variety of issues (e.g., the truth-conditions of the traditional propositional forms, as defined in the first chapters of Ockham's *Summa II*); moreover, its range of applications was very vast, going from logic to theology[478] and to natural philosophy.[479] Significantly, the notions of supposition and consequence are clearly related in many interesting ways. In particular, supposition was employed to define some of the valid inferential patterns between propositions of a certain fragment of the language; the modes of personal supposition were used to define relations of ascent and descent between propositions and their singulars (propositions featuring singular terms, such as proper names and demonstrative pronouns), but also between

[478] For example: Ockham 1974b, p. 7, 11–13.

[479] For example: Ockham, *Expositio in libro* Physicorum Aristoteles (prologus 4) (*Opera Philosophica* IV).

non-singular propositions, as shown in section 2.4 of the present work. But in fact (even though I question the view according to which relations of ascent and descent were meant to provide definitions for the different modes of personal supposition), the order of conceptual priority between consequence and supposition is not entirely straightforward, sometimes approaching circularity: supposition is used to define some valid inferential patterns, but ascents and descents are used to define certain kinds of supposition.

As for supposition and *obligationes*, there are also interesting points of contact between the two topics. For a successful performance in an obligational disputation, it was mandatory to have a good mastery of the semantic apparatus, since many of the obligational puzzles were based in semantic intricacies. As is by now well known, supposition was one of the most fundamental concepts within the medieval semantic framework, so naturally, rules and principles of supposition often came in play during obligational disputations – Ockham even admitted '*distinguo*' as a valid response during a disputation,[480] and as we know, supposition theory was an important tool for the 'distinction' of propositions.

In the opposite direction, one regularly encounters references to the obligational framework in treatises on supposition.[481] In fact, this framework became a kind of general methodology, to be used for investigations in different logical topics (the relation between the obligational framework and *insolubilia* and *sophismata* is particularly evident, cf. Martin 2001). So, naturally, it was also used for investigations on the concept of supposition.

Concerning consequence and *obligationes*, there is an essential connection between the two topics, which has been noted in the chapter dedicated to *obligationes*: the moves in the obligational game are primarily defined by the existence or absence of inferential relations between propositions, so obviously the very rules of the game depend on the notions of inference and consequence that one adopts. In this sense, the concept of consequence is fundamental for the definition of the obligational framework. However, one interesting idea might be to reverse the order of conceptual priority and to define valid consequences as those corresponding to legitimate moves in the obligational game; admittedly, this proposal may have more conceptual appeal than historical justification, in fact I have not as yet found in the texts indication that medieval authors actually implemented this approach. But this is indeed the gist of much of the work currently being done in logic in connection with the game-theoretical framework (cf. van Benthem 2001, Hintikka and Sandu 1997), and seems prima facie entirely compatible with the general medieval approach.

As for the methodology adopted here for historical analysis, by now the reader is either already convinced of the fruitfulness of the formalization approach, or else nothing that I could say now would change this conviction, so I will be brief on this matter. Throughout the formalizations themselves and especially in chapter 4,

480 (Ockham, *Summa* III-3, chap. 39).
481 For example, (Buridan 2001, section 4.2.3).

dedicated to the philosophy of formalization, I have attempted to show that (i) the formalization method does justice to, and sheds new light on, the conceptual and historical elements of the medieval theories being formalized; (ii) it is successful with respect to making them more palatable to the modern reader and to allowing for philosophical reflection on the nature of these theories and of logic in general.

Furthermore, after having completed this work, I remain a firm believer in the fruitfulness of the historical vantage point for the philosophy of logic, not only with respect to the history of the philosophy of logic, but also plainly with respect to the philosophy of logic itself. Historical analysis makes us reflect on what is invariant and fundamental and what is different in the several distinct approaches to logic throughout its history. Accordingly, as a closure to this investigation, I now offer some reflections on the very nature of logic.

2 WHAT IS LOGIC?

With the present work, my intention is not solely that of offering a contribution to the history of logic, more specifically to the history of medieval logic; nor is its purpose only that of 'translating' medieval logic into a language that modern logicians and philosophers can understand, by means of formalizations. In fact, one of my most ambitious aims is to attempt an answer to a very important, simple and yet difficult question: **what is logic?** By raising, and attempting to give an answer to, this fundamentally philosophical question, it becomes evident that my aim is also that of offering a contribution to the philosophy of logic, as much as to the history of logic.

As just mentioned, I, along with other researchers, believe that the historical vantage point is particularly fruitful for investigations in the philosophy of logic. Indeed, logic as a discipline has undergone dramatic changes, especially in its recent history. Therefore, it is natural that issues concerning the univocal understanding of the term 'logic' should arise: can we speak of unity in the discipline through time, or is our use of the term 'logic' referring to current mathematical logic as well as to Aristotelian or medieval logic a case of mere equivocation? What, if anything at all, do these different traditions have in common? By considering the points of analogy and dissimilarity between two traditions as diverse as medieval logic and modern mathematical logic, one is naturally led to look for the very essence of logic – that is, if there is such a thing.

Far from wanting to offer definitive answers here, I merely wish to raise certain points, considerations that seem to me to transpire from the analyses carried out here. To be precise, I do not intend to establish efficient demarcating criteria for logicality, but only to reflect on some of the traits common to all, or many, of the different conceptions of logic encountered during the completion of this research.

a. Logic is not essentially symbolic.

As I have often stressed, one of the distinctive traits of modern mathematical logic is the use of specially designed notations for the expression of logical theories. This procedure has become so widespread that many seem to think that there is no logic other than symbolic logic, that is, that logic is in fact essentially symbolic. Here, we have seen that it is extremely difficult to characterize what a 'symbolic language' means in this context, as opposed to ordinary or so-called natural languages. This purported distinction is not clear-cut, and, more importantly, it is also not a distinction of principle, since so-called natural languages are also a product of conventions. So not only do we not know where to draw the line; we do not even know whether a line should be drawn at all.

Moreover, as suggested by logicians such as Tarski, it is not so much that logic must be expressed in tailor-made notation as a matter of principle; rather, what seems to be the case is that ordinary language is not the most adequate means to express formal concepts and structures, for a variety of reasons. Thus, the decision to use specially designed notation is, in a certain sense, a pragmatic decision, not a principled one; logic could, in theory, be expressed solely with ordinary words, but that would be detrimental to the level of complexity and rigor that one seeks to attain in this discipline.

This being said – that is, even though I claim that logic is not essentially symbolic – as is patent from the very nature of this project, I am convinced of the power of special notations for the purpose of rigor of expression and of enhancing the formal traits of theories. Symbolization is an extremely powerful tool in the formal sciences, as is attested by the great developments provoked in various formal disciplines, such as mathematics and logic, by the introduction of such tools. In other words, there is no doubt that symbolization enhances formality. However, once more I remind the reader that 'formal' is not tantamount to 'symbolic': there are arguments that are symbolic but not formal (e.g., when symbols are used for heuristic purposes), and others that are formal but not symbolic (e.g., when a formally correct proof is expressed solely with ordinary words).

More importantly, the analysis of medieval logic seems to speak both against and in favor of the view that logic is essentially symbolic: on the one hand, a remarkable degree of sophistication was often achieved by medieval logicians; on the other hand, they never attained the degree of technical complexity that is now characteristic of work in logic. Moreover, while the language they used can be viewed as an 'ordinary language', at that time Latin was no longer spoken in non-official contexts, for example, in everyday life; in fact, later medieval (academic) Latin is full of more or less ad hoc artificial conventions (cf. Klima 1991), to an extent that, even though it is not a full-blown 'symbolic' language, it is undoubtedly a highly regimented language (including extensive use of abbreviations and schematic letters).

What can we conclude concerning the nature of logic from the language used for logical investigations in the later medieval period? It seems to me that this element only confirms the view I defend here: logic is not symbolic as a matter of principle, but symbolization – or, in any case, regimentation – greatly enhances and facilitates its practice, for the reasons already explained. In other words, the fact that a highly regimented language was developed in that period for the expression of logical matters (and for other disciplines as well) indicates that plain ordinary language is not suitable for logic, but that different forms of regimentation, other than symbolization, may be adequate in logical contexts.

b. Logic: a subject matter or a method?

A frequent reaction to the question ‘What is logic?’ is to reply that logic is the study of notions such as entailment and consistency; in other words, what would characterize logic as a discipline are the topics it investigates, just as much as biology studies living beings and history the events of the past. However, it has also been argued that what is distinctive about logic is not so much its subject matter, but rather the methods and techniques it employs (see MacFarlane 2000, 2–3). More precisely: all disciplines are concerned with the inference of hitherto unknown information from what is already known, but these inferences are restricted to the specific subject matter of the discipline. By contrast, if logic is concerned with all valid inferences but with no subset of them in particular, there is no subject matter for logic properly speaking; rather, logic permeates all valid reasoning, as it is a body of methods determining the correct forms of reasoning.

Naturally, the claim that what defines logic as a discipline are its methods and techniques is not very informative if it is not supplemented by comprehensive definitions of which methods these are. In fact, the notion of logic as a discipline defined by its subject-matter may be easier to argue for than the method-based notion of logic, as it seems prima facie less problematic to formulate a list of distinctive logical **topics** rather than a list of distinctive logical **methods**. Indeed, the proposal of characterizing logic by emphasizing its methods, in particular the notion of algorithmic reasoning, stumbles upon the fact that the latter is just as widely used in mathematics; likewise, the method of permutation variation, which is often taken to be a defining characteristic of logic (although it is also used in mathematics) is also not sufficient in itself to demarcate logic (as argued in MacFarlane 2000, 6.6).

Hence, admittedly, the emphasis on methods and techniques as what characterizes logic is only a first step, and this approach also raises several complications; still, the method approach seems not only a more fruitful but also a more accurate approach to what is distinctive in logic. As the present study shows, in the Middle Ages, logical techniques were applied to a variety of subjects; the same can be observed nowadays. In other words, with respect to logic, there seems to be far more unity in the methods employed than in the subject matters covered.

To be sure, essentially logical subject matters also exist: investigations on the nature of implication, entailment, consistency etc. But it seems that most of us would view these investigations as pertaining to the realm of philosophy of logic, and not of logic *tout court*. This contrast is made clearer if one considers the opposition between philosophy of logic and philosophical logic: while the former consists in the investigation of the key concepts in logic, but not necessarily with distinctively logical methods, the latter consists in the examination of crucial philosophical concepts exclusively by means of the application of distinctively logical methods. Now, what I have done here is to apply logical tools, more specifically the technique of formalization, in order to examine some of the key concepts in medieval logic. In this sense, this kind of analysis pertains to the realms of both philosophy of logic and philosophical logic: it is the analysis of (historically important) logical concepts by means of logical tools.

Moreover, the analyses of fragments of medieval logic presented here seem to confirm the view that what is distinctive about logic, if anything, are its methods and techniques. Indeed, as already said, medieval logic was by and large a *logica utens*: the logical techniques, most of them pertaining to the general practice of semantic analysis, were applied to a variety of issues, and these techniques were used in domains as dissimilar as natural philosophy (physics) and theology. But while each specific application was governed by specific rules (rules of supposition for semantic analysis, rules of *obligationes* for obligational disputation), there was consensus as to the unity of logic as a discipline, albeit with multiple facets, and as to the foundational role it had, permeating correct reasoning in all disciplines. Logic as it was taught to medieval students was above all a body of methods and techniques, viewed as the foundation for all knowledge. In sum, logic was what students learned in order to be able to learn at all.

c. Medieval and modern logic: different enterprises, but with overlaps.

In order to deepen the comparison between medieval and modern logic, the distinction between subject matter and method is again crucial. If one considers only the subject matters dealt with by medieval and modern logicians, respectively, there seem to be significantly more dissimilarities than similarities, in a way that suggests hardly any common ground between the two traditions and, thus, an equivocal use of the term 'logic'. However, a glimpse at the methods used by each tradition gives us a different picture of the relations between them. To be sure, I am not contending that there are exact medieval counterparts for the techniques currently used in logic; naturally, investigations at that time did not mirror the degree of mathematical sophistication of modern logic, while medieval logicians had an impressive awareness of semantic and logical intricacies, which are now often neglected. But it seems to me that the gist of the procedures and methods currently used is also to be found in medieval logic.[482]

Indeed, in this work, the application of techniques that are currently recognized to be logical to formalize medieval logic was also intended to outline aspects of the latter that can be considered to be formal/logical from a modern viewpoint. Now, this is particularly clear concerning the methods used (broadly understood). Some of the methodological aspects common to medieval and modern logic, which emerged from this enterprise are: calculating procedures understood as the 'blind' application of rules (particularly conspicuous in the case of supposition theory); focus on the formal features of expressions, as opposed to the specific contents of each of them;

[482] However, the positivistic pitfall of looking for the ways in which their logic preceded ours should be avoided, as it usually leads to a lack of interest in the differences between the two traditions, and in particular disregard for the aspects in which they might have known better than us. As is clear from what I have attempted to do in this investigation, I am convinced that there are a few lessons we could learn from them, especially concerning accuracy with respect to semantic notions.

the use of the substitutivity technique to outline generality and the forms of expressions. Moreover, I have also identified a general procedural and dynamic approach to logic, as argued in section 4.4.3, which is akin to an increasingly influential trend in current logical developments. These elements lead me to conclude that, even though the specific methods of each tradition are fundamentally different – predominance of painstaking semantic analysis in medieval logic, as opposed to the application of techniques borrowed mainly from mathematics in modern logic – some significant common traits can be identified. Now, these are precisely the traits that I think characterize the essence of logic in general, independent of any specific tradition.

As for the subject matters treated; admittedly, many of the investigations that bore the title of logic in the Middle Ages would now at best be viewed as (formal) semantics or (formal) epistemology,[483] not as logic. The most conspicuous difference seems to be the fact that the heavy semantic/grammatical apparatus – for example, discussions on the properties of terms – was considered to belong to the realm of logic, whereas now this is certainly not the case (even though logical techniques are, of course, often employed for the study of language). Nonetheless, medieval philosophers also reflected on what we now view as some of the basic logical concepts, and these discussions could easily be seen as legitimately belonging to the realm of modern (philosophy of) logic. Some of these logical notions extensively discussed by medieval authors are:[484] the search for correct definitions for notions such as those of consequence (*consequentia*) and entailment,[485] truth,[486] the logical form of propositions,[487] logical connectives such as conjunction and disjunction,[488] the inferential relations between propositions (e.g., the modes of personal supposition),[489] modal notions,[490] among many others.

Naturally, medieval logicians also discussed topics that have as of now become at best interesting museum pieces, such as theories of fallacies and, to some extent, syllogisms. Nevertheless, even in these discussions, conceptual elements that are worth examining for a broader analysis of the nature of logic can be found, even though the theories in themselves have become somewhat obsolete.

But perhaps one of the most interesting points of proximity between medieval and modern logic is the position occupied by what were then and what are now, respectively, considered to be logical methods. While the first half of the 20th century witnessed a tight and almost exclusive connection between logic and mathematics, logical tools and techniques are now used for a variety of investigations, from language studies to decision theory, from computer science to legal contexts. At the

---

[483] Cf. discussions on the nature of scientific demonstration: (Ockham, *Summa* III-2); (Buridan 2001, part 8).

[484] I here refer only to Ockham and Buridan, but similar discussions can be found in several authors.

[485] (Ockham, *Summa* III-3), (Buridan TC).

[486] (Ockham *Summa* II).

[487] Discussions on *exponibilia*.

[488] (Buridan 2001, section 1.7.4 and 1.7.5); (Ockham, *Summa* II, chap. 32–33).

[489] Cf. section 2.4 of the present work.

[490] (Buridan 2001, 1.8); (Ockham, *Summa* II, chap. 9–10).

same time, the range of logical techniques has been expanded by the incorporation of elements originally belonging to different theories – a notable example is the use of game-theoretical notions in logic. Moreover, the focus on the procedural and dynamic aspects of logic has also increased significantly over the past decades, arguably under the influence of computer science.

Similarly, in the Middle Ages, (what was then considered to be) logic was the basis for the methodologies used in virtually all fields of investigation – in fact, much more so than now. Logic and semantics really provided unity to science then. As for the current situation, although fragmented and specialized knowledge does predominate and, if anything at all, it is the mathematical method that seems to be at the basis of science, a case can be made for the increasingly significant application of logical methods in a wide variety of fields. In some sense, it appears that we may be returning to the medieval situation of logic conferring unity to science.[491]

In sum, there is no doubt that medieval logic and modern logic are at best overlapping enterprises;[492] a lot of what the medievals considered to be logic is now out of the range of what we are prepared to call logic and, conversely, most of the investigations that we now see as logical did not even exist at that time. But there is certainly a significant overlap between the two traditions, in particular with respect to the gist of the methods that were applied then and are applied now, and with respect to the philosophical discussions accompanying them. Furthermore, we now seem to observe a general movement towards the expansion of the scope of logic, which is in many senses similar to the position of logic in the Middle Ages as permeating all correct reasoning.

Thus, this work ends with a confessed optimism about logic. In fact, this optimism underpins the whole project, as it is based on the double-sided conviction that the application of logical tools to the study of the history of logic is a fruitful enterprise, and that the history of logic is undoubtedly worth studying in order to advance current work in (philosophy of) logic.

---

[491] Naturally, this is not an entirely recent development. In the first half of the 20th century there have been significant attempts to 'logicize' the natural and even the social sciences, in particular with the Vienna Circle and Popper. But their understanding of logic was arguably too narrow, which may explain why these projects have been amply criticized and to a great extent abandoned. Now, with a broader understanding of what is to count as logical, which goes beyond first-order predicate logic, it is to be hoped that this project can be continued.

[492] 'Yet the influence of mathematics on logic has undeniably changed its character: mediaeval and modern logic are overlapping but distinct enterprises'. (King 2001, 140)

REFERENCES

Medieval texts in Latin

John Buridan 1998, *Summulae de Suppositionibus* (ed. R. van der Lecq). Nijmegen, Ingenium.

John Buridan 2004, *Summulae de practica sophismatum* (ed. F. Pironet). Turnhout, Brepols.

R. Green 1963, *The logical Treatise 'De Obligationibus': An Introduction with Critical Texts of William of Sherwood (?) and Walter Burley*. PhD dissertation, Université Catholique de Louvain.

H. Hubien 1976, *Iohannis Buridani tractatus de consequentiis*. Series *Philosophes Médievaux* Vol. XVI. Louvain, Université de Louvain.

John of St. Thomas 1929, *Ars Logica* (ed. B. Reiser). Rome, Taurini.

Marsilius of Inghen 1983, *Treatises on the Properties of Terms* (ed. and transl. E.P. Bos). Dordrecht, Reidel.

Paul of Venice 2002, *Logica Parva* (ed. A. Perreiah). Leiden, Brill.

Peter Abelard 1956, *Dialectica* (ed. L.M. de Rijk). Assen, van Gorcum.

Peter of Spain 1972, *Tractatus: Called Afterwards Summulae Logicales* (ed. L.M. de Rijk). Assem, van Gorcum.

Ralph Strode, *Obligationes* (ed. E.J. Ashworth). Unpublished edition, 43 pages.

Richard Billingham 2003, *Richard Billingham 'De Consequentiis' mit Toledo-Kommentar* (ed. S. Weber). Amsterdam, B.R. Grüner.

W. Seaton 1973, *An Edition and Translation of the 'Tractatus de Consequentiis' by Ralph Strode, Fourteenth Century Logician and Friend of Geoffrey Chaucer*. PhD dissertation, University of California, Berkeley.

P.V. Spade 1977, '*Roger Swyneshed's Obligationes*: *Edition and Comments*'. *Archives d'histoire doctrinale et littéraire du Moyen Age*. (Paris, 1978), pp. 243–285.

Walter Burley 1955, *De Puritate Artis Logicae – Tractatus Longior* (ed. Ph. Boehner). St. Bonaventure, The Franciscan Institute.

Walter Burley 1980, '*De Consequentiis*' (ed. N.J. Green-Pederson). *Fransciscan Studies* **40**, pp. 102–166.

William of Ockham 1974a, *Opera Philosophica* I. St. Bonaventure, The Franciscan Institute.

William of Ockham 1974b, *Opera Theologica* III. St. Bonaventure, The Franciscan Institute.

William of Sherwood 1995, *Einführung in die Logik* (ed. and transl. H. Brands and C. Kann). Hamburg, Felix Meiner Verlag.

Translations

Aristotle 1992, *On Sophistical Refutations*. In *Aristotle* (vol. 7 of 'Great Books'). Chicago, University of Chicago.

Aristotle 1995, *Prior Analytics*. In J. Barnes (ed.). *The Complete Works of Aristotle*, vol. 1. Princeton, Princeton University Press.

Aristotle 1995, *Posterior Analytics*. In J. Barnes (ed.). *The Complete Works of Aristotle*, vol. 1. Princeton, Princeton University Press.

John Buridan 2001, *Summulae de Dialectica* (transl. G. Klima). New Haven, Yale University Press.
John of St. Thomas 1955, *Outlines of Formal Logic* (transl. Francis C. Wade). Milwaukee, Marquette University Press.
G.W. Leibniz 1966, *Logical Papers* (edition and translation by G.H.R. Parkinson). Oxford, Clarendon Press.
Marsilius of Inghen 1983, *Treatises on the Properties of Terms* (ed. and transl. E.P. Bos). Dordrecht, Reidel.
Paul of Venice 1984, *Logica Parva* (transl. A. Perreiah). Munich, Philosophia Verlag.
Peter of Spain 1945, *The Summulae Logicales of Peter of Spain* (ed. and transl. J.P. Mullally). South Bend, University of Notre Dame Press.
Pseudo-Scotus 2001, 'Questions on Aristotle's *Prior Analytics*: Question X'. In Yrjönsuuri 2001 (ed.), 225–234.
Walter Burley 1988, *Obligations* (selection). In N. Kretzmann and E. Stump (eds.), *Logic and the Philosophy of Language*. The Cambridge Translations of Medieval Philosophical Texts. Cambridge, Cambridge University, pp. 369–412.
Walter Burley 2000, *On the Purity of the Art of Logic – The Shorter and the Longer Treatises* (transl. P.V. Spade). New Haven, Yale University Press.
William of Ockham 1998, *Summa Logicae Part I* (transl. M. Loux). South Bend, St. Augustine's Press.
William of Ockham 1998, *Summa Logicae Part II* (transl. A. Freddoso). South Bend, St. Augustine's Press.
William of Sherwood 1966, *Introduction to Logic* (transl. Norman Kretzmann). Minneapolis, University of Minnesota.

Secondary literature on medieval logic

E.J. Ashworth 1974, *Language and Logic in the Post-Medieval Period*. Dordrecht, Reidel.
E.J. Ashworth 1978, 'Multiple quantification and the use of special quantifiers in early sixteenth century logic'. *Notre Dame Journal of Formal Logic* **19**, 599–613.
E.J. Ashworth 1981, 'The problems of relevance and order in obligational disputations: some late fourteenth century views'. *Medioevo* **7**, 175–193.
E.J. Ashworth 1984, 'Inconsistency and paradox in medieval disputations: a development of some hints in Ockham'. *Franciscan Studies* **44**, 129–139.
E.J. Ashworth 1985, 'English *Obligationes* texts after Roger Swyneshed: the tracts beginning "Obligatio est quaedam ars" '. In P. Osmund Lewry O.P. (ed.), *The Rise of British Logic*. Papers in Medieval Studies **7**. Toronto, Pontifical Institute of Medieval Studies, pp. 309–333.
E.J. Ashworth 1991, ' "Nulla propositio est distinguenda": La notion d'*equivocatio* chez Albert de Saxe'. In J. Biard (ed.), *Itinéraires d'Albert de Saxe*. Paris, Vrin, pp. 149–160.
E.J. Ashworth 1993, 'Ralph Strode on inconsistency in obligational disputations'. In K. Jacobi (ed.), *Argumentationstheorie*. Leiden, Brill, pp. 363–386.
E.J. Ashworth 1996, 'Autour des *Obligationes* de Roger Swyneshed: la *Nova Responsio*'. *Les Etudes Philosophiques* **3**, 341–360.
E.J. Ashworth 2002, 'Review of M. Yrjönsuuri (ed.), *Medieval Formal Logic*'. *History and Philosophy of Logic* **23**, 296–297.
E.J. Ashworth 2003, ' "Tu curris" (vel "homo est asinus") "est tibi positum": l'histoire d'un sophisme dans les traités sur les *Obligationes*'. Forthcoming in the proceedings of the *Sophismata* Conference (Geneva, June 2003).
O. Bird 1961, 'Topic and consequence in Ockham's Logic'. *Notre Dame Journal of Formal Logic* **2**, 65–78.
I. Boh 1982, 'Consequences'. In N. Kretzmann, A. Kenny and J. Pinborg (eds.) 1982, pp. 300–314.
I. Boh 2000, 'Four phases of medieval epistemic logic'. *Theoria* **56**(2), 129–144.
I. Boh 2001, 'Consequences and rules of consequence in the post-Ockham period'. In M. Yrjönsuuri (ed.) 2001, pp. 147–181.
J. Boler 1963, *Charles Peirce and Scholastic Realism. A Study of Peirce's Relations to John Duns Scotus*. Seattle, University of Washington Press.

E.P. Bos 2004, 'Semantics and literal exposition in Henry of Ghent's *Regular Lecture on the Bible*'. In W. Janse (ed.), *Dutch Review of ChurchHistory*, vol. 84. Leiden, Brill, pp. 71–94.

H.A.G. Braakhuis 1998, 'Obligations in early thirteenth century Paris: the *Obligationes* of Nicholas of Paris (?) (Ms Paris, B.N. lat., 11.412)'. *Vivarium* **36**(2), 152–233.

S. Brown 1993, 'Medieval supposition theory in its theological context'. *Medieval Philosophy and Theology* **3**, 121–157.

N. Cocchiarella 2001, 'A logical reconstruction of medieval terminist logic in conceptual realism'. *Logical Analysis and History of Philosophy* **4**, 35–72.

A. de Libera 1982, 'The Oxford and Paris traditions in logic'. In N. Kretzmann, A. Kenny and J. Pinborg (eds.) 1982, pp. 147–187.

L.M. de Rijk 1962, *Logica Modernorum. A Contribution to the History of Early Terminist Logic. Vol. 1: On the Twelfth Century Theory of Fallacies.* Assen, Koninklijke Van Gorcum & Company.

L.M. de Rijk 1967, *Logica Modernorum. A Contribution to the History of Early Terminist Logic. Vol. 2. Part One: The Origin and Early Development of the Theory of Supposition.* Assen, Koninklijke Van Gorcum & Company.

L.M. de Rijk 1974, 'Some thirteenth century tracts on the game of obligations I'. *Vivarium* **12**, 94–123.

B. Dod 1982, 'Aristoteles latinus'. In N. Kretzmann, A. Kenny and J. Pinborg (eds.) 1982, pp. 45–79.

C. Dufour 1989, *Die Lehre der Proprietates Terminorum : Sinn und Referenz in mittelalterlicher Logik* . München, Philosophia Verlag.

C. Dutilh Novaes forthcoming, 'Theory of supposition vs. theory of fallacies in Ockham'. In J. Marenbon (ed.), *The Roots of Medieval Logic*. Forthcoming, Leiden, Brill.

C. Dutilh Novaes 2004a, 'A medieval reformulation of the de dicto/de re distinction'. In L. Behounek (ed.), *Logica Yearbook 2003*. Prague, Filosofia, pp. 111–124.

C. Dutilh Novaes 2004b, 'Ockham on supposition and equivocation in mental language'. Proceedings of the Society for Medieval Logic and Metaphysics vol. 3. Available at http://www.fordham.edu/gsas/phil/klima/SMLM/PSMLM3/PSMLM3.pdf.

C. Dutilh Novaes 2004c, 'The Buridanian account of inferential relations between doubly quantified propositions: a proof of soundness'. *History and Philosophy of Logic* **25**(3), 225–244.

C. Dutilh Novaes 2004d, Review of Klima 2004a. *Bulletin of Symbolic Logic* **10**(4), 592–593.

C. Dutilh Novaes 2005a, 'In search of the intuitive notion of logical consequence'. In L. Behounek and M. Bilkova (eds.), *Logica Yearbook 2004*. Prague, Filosofia, pp. 109–123.

C. Dutilh Novaes 2005b, 'Medieval *Obligationes* as logical games of consistency maintenance'. *Synthese* **145**(3), 371–395.

C. Dutilh Novaes 2005c, 'Buridan's *consequentia*: consequence and inference within a token-based semantics'. *History and Philosophy of Logic* **26**(4), 277–297.

C. Dutilh Novaes 2006a, 'Roger Swyneshed's *obligationes*: a logical game of inference recognition?'. *Synthese* **151**(1), 127–155.

C. Dutilh Novaes 2006b, 'Ralph Strode's *obligationes*: the return of consistency and the epistemic turn'. *Vivarium* **44**(2–3), 1–37.

C. Dutilh Novaes 2008, 'An intensional interpretation of Ockham's theory of supposition'. Forthcoming in the *Journal of the History of Philosophy*.

S. Ebbesen 1991, 'Is logic theoretical or practical knowledge?'. In J. Biard (ed.), *Itinéraires d'Albert de Saxe*. Paris, Vrin, pp. 267–283.

U. Eco 1989, 'Denotation'. In U. Eco and C. Marmo (eds.) 1989, *On the Medieval Theory of Signs*. Amsterdam, John Benjamins, 43–77.

N.J. Green-Pedersen 1984, *The Tradition of the Topics in the Middle Ages : The Commentaries on Aristotle's and Boethius' "topics"* . Munich, Philosophia Verlag.

E. Karger 1976, *A Study in William of Ockham's Modal Logic*. PhD dissertation, University of California, Berkeley.

E. Karger 1978, 'Conséquences et inconséquences de la supposition vide dans la logique d'Ockham'. *Vivarium* **16**, 46–55.

E. Karger 1984, 'Modes of personal supposition: the purpose and usefulness of the theory within Ockham's logic'. *Franciscan Studies* **44**, 87–106.

E. Karger 1993, 'A theory of immediate inferences contained in Buridan's logic'. In K. Jacobi (ed.), *Argumentationstheorie*. Leiden, Brill, pp. 407–429.

H. Keffer 2001, *De Obligationibus – Rekonstruktion einer spätmittelalterlichen Disputationstheorie*. Leiden, Brill.

A. Kenny and J. Pinborg 1982, 'Medieval philosophical literature'. In N. Kretzmann, A. Kenny and J. Pinborg (eds.) 1982, pp. 11–42.

P. King 1985, *Jean Buridan's Logic*. Dordrecht, Reidel.

P. King 1985a, 'Introduction'. In P. King 1985.

P. King 2001, 'Medieval proof theory'. In M. Yrjönsuuri (ed.) 2001, pp. 117–145.

P. King 2004, 'Peter Abelard'. *The Stanford Encyclopedia of Philosophy (Fall 2004 Edition)*. In E.N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2004/entries/abelard/>.

Klima 1991, 'Latin as a formal language – outlines of a Buridanian semantics'. *Cahiers de l'Institut du Moyen-Age Grec et Latin* **61**, 78–106.

G. Klima 1993a, ' "Debeo tibi equum": a reconstruction of the theoretical framework of Buridan's treatment of the *sophisma*'. In S. Read (ed.) 1993, pp. 333–347.

G. Klima 1993b, 'The changing role of "Entia Rationis" in mediaeval semantics and ontology: a comparative study with a reconstruction'. *Synthese* **96**, 25–59.

G. Klima 2004a, 'Consequences of a closed, token-based semantics: the case of John Buridan'. *History and Philosophy of Logic* **25**(2), 95–110.

G. Klima 2004b, 'John Buridan and the force-content distinction'. In A. Maierú and L. Valente (eds.), *Medieval Theories On Assertive and Non-Assertive Language*. Acts of the 14th European Symposium on Medieval Logic and Semantics. Rome, Olschiki, pp. 415–427.

S. Knuuttila and M. Yrjönsuuri 1988, 'Norms and actions in obligational disputations'. In O. Pluta (ed.), *Die Philosophie im 14. und 15. Jahrhundert*. Amsterdam, B.R. Grüner, pp. 191–202.

N. Kretzmann, A. Kenny and J. Pinborg (eds.)1982, *The Cambridge History of Later Medieval Philosophy*. Cambridge, Cambridge University Press.

N. Kretzmann and E. Stump 1985, 'The anonymous *De arte obligatoria* in Merton college MS. 306'. In E.P. Bos (ed.), *Medieval Semantics and Metaphysics: Studies Dedicated to L.M. de Rijk, PhD, Professor of Ancient and Medieval Philosophy at the University of Leiden on the Occasion of His 60th Birthday*. Nijmegen, Ingenium ("Artistarium," Supplementa, vol. 2), pp. 239–280.

H. Lagerlund and E. Olsson 2001, 'Disputation and change of belief – Burley's theory of *Obligationes* as a theory of belief revision'. In M. Yrjönsuuri (ed.) 2001, pp. 35–62.

R. van der Lecq and H.A.G. Braakhuis 1994, 'Introduction'. In *Johannes Buridanus Questiones Elencorum*. Nijmegen, Ingenium.

J. Marenbon 1987, *Later Medieval Philosophy (1150–1350) – An Introduction*. London, Routledge & Kegan Paul.

C. Marmo 1984, 'Guglielmo di Ockham e il significato delle proposizioni'. *Versus – Quaderni di Studi Semiotici* **38/39**, 115–148.

C.J. Martin 1986, 'William's machine'. *The Journal of Philosophy* **83**(10), 564–572.

C.J. Martin 1992, 'Obligations and liars'. In S. Read (ed.), *Sophisms in Medieval Logic and Grammar*. Kluwer, Dordrecht, pp. 357–381.

C.J. Martin 2001, 'Obligations and liars'. In M. Yrjönsuuri (ed.) 2001, pp. 63–94. (Revised version of Martin 1992)

G. Matthews 1973, '*Suppositio* and quantification in Ockham'. *Noûs* **7**, 13–24.

G. Matthews 1997, 'Two theories of supposition?'. *Topoi* **16**, 35–40.

M. Mullick 1971, 'Does Ockham accept material implication?'. *Notre Dame Journal of Formal Logic* **12**, 117–124.

C. Normore 1993, 'The necessity in deduction: cartesian inference and its medieval background'. *Synthese* **96**, 437–454.

C. Normore 1997, 'Material supposition and the mental language of Ockham's *Summa Logicae*'. *Topoi* **16**, 27–33.

C. Normore 1999, 'Some aspects of Ockham's logic'. In P.V. Spade (ed.) 1999, pp. 31–52.

G. Nuchelmans 1973, *Theories of the Proposition*. Amsterdam, North-Holland.

C. Panaccio 1992, *Les Mots, les Concepts et les Choses*. Paris, Vrin.

C. Panaccio 1999, 'Semantics and mental language'. In P.V. Spade (ed.) 1999, pp. 53–75.

C. Panaccio and E. Perini-Santos 2004, 'Guillaume d'Ockham et la *suppositio materialis*'. Vivarium **42**(2), 202–224.

T. Parsons 1997, 'Supposition as quantification versus supposition as global quantificational effect'. *Topoi* **16**, 41–63.

F. Pironet 2001, 'The relations between insolubles and obligations in medieval disputations'. In M. Yrjönsuuri (ed.) 2001, pp. 95–114.

G. Priest and S. Read 1977, 'The formalization of Ockham's theory of supposition'. *Mind* **86**, 109–113.

G. Priest and S. Read 1980, 'Merely confused supposition'. *Franciscan Studies* **40** (1980), 265–297.

G. Priest and S. Read 1981, 'Ockham's rejection of ampliation'. *Mind* **90**, 274–279.

A. Prior 1969, 'The possibly-true and the possible'. *Mind* **78**(312), 481–492.

S. Read (ed.) 1993, *Sophisms in Medieval Logic and Grammar*. Dordrecht, Kluwer.

S. Read 1999, 'How is material supposition possible?'. *Medieval Philosophy and Theology* **8**(1), 1–20.

S. Read 2001, 'Self-reference and validity revisited'. In M. Yrjönsuuri (ed.) 2001, pp. 183–196.

S. Read 2002, 'Medieval theories: properties of terms', *The Stanford Encyclopedia of Philosophy (Spring 2002 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2002/entries/medieval-terms/>.

I. Rosier-Catach 1994, *La parole comme acte: sure la grammaire et la sémantique au XIIIe siècle*. Paris, Vrin.

P.V. Spade 1978, 'John Buridan on the liar: a study and reconstruction'. *Notre Dame Journal of Formal Logic* **19**, 579–590.

P.V. Spade 1980, 'Synonymy and equivocation in Ockham's mental language'. *Journal of the History of Philosophy* **18**, 9–22.

P.V. Spade 1982a, 'Obligations: B. developments in the fourteenth century'. In N. Kretzmann, A. Kenny, J. Pinborg (eds.) 1982, pp. 335–341.

P.V. Spade 1982b, 'Three theories of *obligationes*: Burley, Klivington and Swyneshed on counterfactual reasoning'. *History and Philosophy of Logic* **3**, 1–32.

P.V. Spade 1988, 'The logic of the categorical: the medieval theory of descent and ascent'. In N. Kretzmann (ed.), *Meaning and Inference in Medieval Philosophy*. Dordrecht, Kluwer Academic Publishers, pp. 187–224.

P.V. Spade 1992, 'If obligations were counterfactuals'. *Philosophical Topics* **20**, 171–188.

P.V. Spade 1996, *Thoughts, Words and Things*. Available at www.pvspade.com.

P.V. Spade (ed.) 1999, *The Cambridge Companion to Ockham*. Cambridge, Cambridge University Press.

P.V. Spade 2000, 'Why don't medieval logicians ever tell us what they're doing? or, what is this, a conspiracy?' Available at www.pvspade.com/Logic/docs/Conspiracy.pdf.

P.V. Spade 2003, 'Obligationes'. In E. Zalta (ed.), *Stanford Encyclopedia of Philosophy*. Available at http://plato.stanford.edu/entries/obligationes/

P.V. Spade and E. Stump 1983, 'Walter Burley and the *obligationes* attributed to William of Sherwood'. *History and Philosophy of Logic* **4**, 9–26.

E. Stump 1981, 'Roger Swyneshed's theory of obligations', *Medioevo* **7**, 135–174.

E. Stump 1982a, 'Topics: their development and absorption into consequences'. In N. Kretzmann, A. Kenny and J. Pinborg (eds.) 1982, pp. 273–299.

E. Stump 1982b, 'Obligations: A. from the beginning to the early fourteenth century'. In N. Kretzmann, A. Kenny and J. Pinborg (eds.) 1982, pp. 315–334.

E. Stump 1985, 'The logic of disputation in Walter Burley's treatise on obligations'. *Synthese* **63**, 355–374.

J. Swiniarski 1970, 'A new presentation of Ockham's theory of supposition, with an evaluation of some contemporary criticisms'. *Franciscan Studies* **25**, 181–217.

M. Tweedale 1982, 'Abelard and the culmination of the old logic'. In N. Kretzmann, A. Kenny and J. Pinborg (eds.) 1982, pp. 143–158.

A. Willing 1991, 'Buridan's divided modal syllogistics'. *Notre Dame Journal of Formal Logic* **42**, 276–289.

S. Weber 2003, *Richard Billingham 'De Consequentiis' mit Toledo-Kommentar*. Amsterdam, B.R. Grüner.

J. Weisheipl 1956, *Early Fourteenth-Century Physics and the Merton 'School' With Special Reference to Dumbleton and Heytesbury*. PhD dissertation, Oxford University, Oxford.

M. Yrjönsuuri 1993, 'The role of casus in some fourteenth century treatises on *Sophismata* and obligations'. In K. Jacobi (ed.), *Argumentationstheorie*. Leiden, Brill, pp. 301–322.

M. Yrjönsuuri 1996, 'Obligations as thought experiments'. In I. Angelelli and M. Cerezo (eds.), *Studies on the History of Logic*. Walter de Gruyter, Berlin, pp. 79–96.

M. Yrjönsuuri 1998, 'The compossibility of impossibilities and *Ars Obligatoria*'. *History and Philosophy of Logic* **19**, 235–248.

M. Yrjönsuuri 2000, 'Disputations, obligations and logical coherence'. *Theoria* **56**(2), 205–223.

M. Yrjönsuuri (ed.) 2001, *Medieval Formal Logic*. Dordrecht, Kluwer.

M. Yrjönsuuri 2001a, 'Duties, rules and interpretations in obligational disputations'. In M. Yrjönsuuri (ed.) 2001, pp. 3–34.

General

J. Avigad and R. Zach 2002, 'The epsilon calculus'. *The Stanford Encyclopedia of Philosophy (Summer 2002 Edition)*, Edward N. Zalta (ed.), URL=<http://plato.stanford.edu/archives/sum2002/entries/epsilon-calculus/>.

S. Awodey and E. Reck 2002, 'Completeness and categoricity. part I: nineteenth-century axiomatics to twentieth-century metalogic'. *History and Philosophy of Logic* **23**, 1–30.

J. Barnes 1995, 'Metaphysics'. In J. Barnes (ed.), *Cambridge Companion to Aristotle*. Cambridge, Cambridge University Press, pp. 66–108.

J. van Benthem 2001, *Logic in Games*. Lecture Notes, University of Amsterdam, Amsterdam.

J. van Benthem and A. ter Meulen (eds.) 1985, *Generalized Quantifiers in Natural Language*. Foris, Dordrecht.

M. Beuchot and J. Deely 1995, 'Common sources for the semiotic of Charles Peirce and John Poinsot'. *The Review of Metaphysics* **48**, 549–566.

P. Blackburn and M. Kohnlhase 2004, 'Inference and computational semantics'. *Journal of Logic, Language and Information* **13**, 117–120.

B. Bolzano 1973, *Theory of Science*. Dordrecht, Reidel.

R. Brandom 1994, *Making It Explicit : Reasoning, Representing, and Discursive Commitment*. Cambridge (Mass.), Harvard University Press.

R. Brandom 2000, *Articulating Reasons: An Introduction to Inferentialism*. Cambridge (Mass.), Harvard University Press.

L.E.J. Brower 1923, 'On the significance of the principle of excluded middle in mathematics, especially in function theory'. In J. van Heijenoort (ed.) 1977, pp. 335–341.

F. Cajori 1928, *A History of Mathematical Notations* (vols. 1 and 2). London, The Open Court Company.

R. Carnap 1937, *The Logical Syntax of Language*. London, Routledge.

H. Caygill 1995, 'Logic'. In *A Kant Dictionary*. Oxford, Blackwell.

D.J. Chalmers 2003, 'The foundations of two-dimensional semantics'. Available at http://www.u.arizona.edu/~chalmers/papers/foundations.html.

D. Chandler 1995, *Semiotics for Beginners*. Available at http://www.aber.ac.uk/media/Documents/S4B/semiotic.html.

A. Church 1956, *Introduction to Mathematical Logic*. Princeton, Princeton University.

J. Corcoran 2003, 'Aristotle's *Prior Analytics* and Boole's *Laws of Thought*'. *History and Philosophy of Logic* **24**, 261–288.

J. Corcoran 2004, 'Schema'. *The Stanford Encyclopedia of Philosophy (Summer 2004 Edition)*, E.N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2004/entries/schema/>.

L. Couturat 1901, *La Logique de Leibniz d'apres des documents inedits*. Paris, Alcan.

H.B. Curry 1957, *A Theory of Formal Deducibility*. Notre Dame, University of Notre Dame.

H.B. Curry and R. Feys 1958, *Combinatory Logic*, vol. 1. Amsterdam, North-Holland.

R. Dedekind 1890, 'Letter to Keferstein'. In J. van Heijenoort (ed.) 1977, pp. 99–103.

M. Detlefsen (ed.) 1992, *Proof, Logic and Formalization*. London, Routledge.
R. Dipert 1995, 'Peirce's underestimated place in the history of logic: a response to Quine'. In *Peirce and Contemporary Thought*. New York, Fordham University Press, pp. 32–58.
K. Donnellan 1966, 'Reference and definite descriptions'. In *The Philosophical Review* **75**(3), 281–304.
J. Etchemendy 1990, *The Concept of Logical Consequence*. Cambridge, Harvard University Press.
G. Frege 1879, *Begriffsschrift*. In J. van Heijenoort (ed.) 1977, pp. 1–82.
G. Frege 1948, 'Sense and reference'. In M. Black (transl.), *The Philosophical Review* **57**, pp. 207–230.
G. Frege 1972, 'On the aim of the "conceptual notation" '. In T.W. Bynum (transl. and ed.), *Conceptual Notation and Related Articles*. Oxford, Clarendon Press, pp. 90–100.
G. Frege 1980, 'Function and concept'. In P. Geach and M. Black (eds.), *Translations from the Philosophical Writings of Gottlob Frege*. Oxford, Blackwell, pp. 21–41.
H. Gaifman 1992, 'Pointers to truth'. *Journal of Philosophy* **89**, 223–261.
P.T. Geach 1962, *Reference and Generality*. Oxford, Blackwell.
K. Gödel 1986, 'On formally undecidable propositions of *Principia mathematica* and related systems I'. In *Collected Works*, vol. I, Oxford, Oxford University Press, pp. 145–195.
K. Gödel 1995, *Collected Works vol. III: Unpublished Essays and Lectures*. Oxford, Oxford University Press.
M. Gómez-Torrente 1999, 'Logical truth and Tarskian logical truth'. *Synthese* **117**(3), 375–408.
M. Gómez-Torrente 2000, 'A note on formality and logical consequence'. *Journal of Philosophical Logic* **29**, 529–539.
I. Grattan-Guinness 2000, 'Mathematics and symbolic logics: some note on an uneasy relationship'. *History and Philosophy of Logic* **20**, 159–167.
S. Haack 1982, 'Dummett's justification of deduction'. *Mind* **91**, 216–239.
C.L. Hamblin 1970, *Fallacies*. Methuen, London.
S.O. Hansson 2000, 'Formalization in philosophy'. *Bulletin of Symbolic Logic* **6**(2), 162–175.
J. van Heijenoort 1967, 'Logic as calculus and logic as language'. *Synthese* **17**, 324–330.
J. van Heijenoort 1977, *From Frege to Godel – A Source Book in Mathematical Logic, 1879–1931*. Cambridge (Mass.) Harvard University Press, third printing.
D. Hilbert 1899, 'Grundlagen der geometrie'. In *Festschrift zur Feier der Enthüllung des Gauss-Weber-Denkmals in Göttingen*. Leipzig, Teubner, pp. 1–92, 1st ed.
D. Hilbert 1927, 'The foundation of mathematics'. In J. van Heijenoort 1977, pp. 464–479.
J. Hintikka and G. Sandu 1997, 'Game-theoretical semantics'. In J. van Benthem and A. ter Meulen (eds.), *Handbook of Logic and Language*. Elsevier/MIT, Amsterdam/Cambridge, Mass Press, pp. 361–410.
G.E. Hughes and M.J. Cresswell 1968, *An Introduction to Modal Logic*. London, Methueun.
D. Jacquette 1994, 'Formalization in philosophical logic'. *The Monist* **77**(3), 358–375.
D. Kaplan 1989, 'Demonstratives'. In J. Almog, J. Perry and H. Wettstein (eds.), *Themes from Kaplan*. Oxford, Oxford University Press, pp. 481–563.
L. Keiff, 'Dialogical Logic'. Forthcoming in E. Zalta (ed.), *Stanford Encyclopedia of Philosophy*.
S. Kleene 1951, *Introduction to Metamathematics*. Amsterdam, North Holland.
G. Klima and G. Sandu 1991, 'Numerical quantifiers in game-theoretical semantics'. *Theoria* **56**, 173–193.
W. Kneale and M. Kneale 1984, *The Development of Logic*. Oxford, Clarendon Press.
T. Korte 2001, *Frege and His Epigones*. PhD Dissertation, University of Turku.
A. Koslow 1992, *A Structuralist Theory of Logic*. Cambridge, Cambridge University Press.
S. Kripke 1975, 'Outline of a theory of truth'. *Journal of Philosophy* **72**, 690–716.
S. Kripke 1977, 'Speaker's reference and semantic reference'. In P. Ludlow (ed.), *Readings in the Philosophy of Language*. Cambridge (Mass.), MIT Press, pp. 383–414.
S. Kripke 1980, *Naming and Necessity*. Oxford, Blackwell.
S. Lesniewski 1992, *Collected Works*. Dordrecht, Kluwer.
L.H. Lopes dos Santos 1994, 'Apresentação'. In L. Wittgenstein, *Tractatus Logico-Philosophicus* (bilingual edition). São Paulo, Edusp, pp. 11–112.
P. Lorenzen 1961, 'Ein dialogisches Konstruktivitätskriterium', In *Infinitistic Methods*. Pergamon, Oxford, pp. 93–200.

J. Maat 1999, *Philosophical Languages in the Seventeenth Century: Dalgarno, Wilkins, Leibniz.* PhD dissertation, ILLC-UvA, Amsterdam.

J. MacFarlane 2000, *What does it mean to say that logic is formal?*. PhD Dissertation, University of Pittsburgh. Available at http://philosophy.berkeley.edu/macfarlane/Diss.pdf.

P. Martin-Löf 1987, 'Truth of a proposition, evidence of a judgemente, validity of a proof'. *Synthese* **73**, 407–420.

H. Mendell 2004, 'Aristotle and mathematics'. *The Stanford Encyclopedia of Philosophy (Summer 2004 Edition)*, Edward N. Zalta (ed.), URL = http://plato.stanford.edu/archives/sum2004/entries/aristotle.mathematics.

R. Montague 1970, 'English as a formal language'. In *Formal Philosophy: Selected Papers of Richard Montague*. Edited and with an introduction by Richmond H. Thomason (1974). New Haven, Yale University Press, pp. 247–270.

E. Nagel 1939, 'The formation of modern conceptions of formal logic in the development of geometry'. *Osiris* **7**, 142–224.

J. von Neumann 1925, 'An axiomatization of set-theory'. In J. van Heijenoort (ed.) 1977, pp. 393–413.

R. Parikh 2005, 'Logical omniscience and common knowledge; WHAT doe we know and what do WE know?'. Available at www.cs.gc.cuny.edu/∼kgb/docs/tark05-1.pdf.

T. Parsons 1996, 'Fregean theories of truth and meaning'. In M. Schirn (ed.), *Frege : Importance and Legacy*. Berlin, de Gruyter, pp. 371–409.

G. Peano 1889, 'The principles of arithmetic, presented by a new method'. In J. van Heijenoort (ed.) 1977, pp. 83–97.

V. Peckhaus 2004, 'Calculus raciocinator versus characteristica universalis? The two traditions in logic, revisited'. *History and Philosophy of Logic* **25**(1), 3–14.

C.S. Peirce 1931–1958, *Collected Writings* (8 vols.). Edited by Charles Hartshorne, Paul Weiss and Arthur W. Burks. Cambridge (Mass.), Harvard University Press.

C.S. Peirce 1884, 'On the algebra of logic'. In *American Journal of Mathematics* **7**, 180–202.

W. Quine 1960, *Word and Object*. Cambridge (Mass.), MIT Press.

W. Quine 1961, *From a Logical Point of View*. Cambridge (Mass.), Harvard University Press.

S. Rahman and L. Keiff 2005, 'On how to be a dialogician'. In D. Vanderveken (ed.), *Logic, Thought and Action*. Berlin, Springer, pp. 359–408.

A. Ranta 1994, *Type-theoretical Grammar*. Oxford, Clarendon Press.

S. Read 1994, 'Formal and material consequence'. *Journal of Philosophical Logic* **23**, 247–265.

S. Read 1995, *Thinking About Logic: An Introduction to the Philosophy of Logic*. Oxford, Oxford University Press.

M. Reimer 2003, 'Reference'. *The Stanford Encyclopedia of Philosophy (Spring 2003 Edition)*, E.N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2003/entries/reference/>.

G. Restall 2002, 'Carnap's tolerance, meaning and logical pluralism'. *Journal of Philosophy* **99**, 426–443.

R. van Rooij 2004, 'Signalling games select Horn strategies'. *Linguistics and Philosophy* **27**, 493–527.

B. Russell 1905, 'On denoting'. *Mind* **14**, 479–493.

M. Sainsbury 2005, *Reference Without Referents*. Oxford, Oxford University Press.

N. Salmon 1982, *Reference and Essence*. Oxford, Basil Blackwell.

J. Sebestik 1992, *Logique et Mathematiques chez Bernard Bolzano*. Paris,Vrin.

S. Shapiro 1991, *Foundations Without Foundationalism: A Case for Second-order Logic.* Oxford, Oxford University Press.

S. Shapiro 1998, 'Logical consequence: models and modality'. In M. Schirn (ed.), *Philosophy of Mathematics Today*. Oxford, Oxford University Press, pp. 131–156.

S.J. Shin 2002, *The Iconic Logic of Peirce's Graphs*. Cambridge (Mass.), MIT Press.

S.J. Shin and O. Lemon 2003, 'Diagrams'. In *The Stanford Encyclopedia of Philosophy (Winter 2003 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2003/entries/diagrams/>.

W. Sieg 1994, 'Mechanical procedures and mathematical experience'. In A. George (ed.), *Mathematics and Mind*. Oxford, Oxford University Press, pp. 71–140.

P. Simons 1987, 'Bolzano, Tarski, and the limits of logic'. Reprinted in P. Simons 2002, *Philosophy and Logic in Central Europe from Bolzano to Tarski*. Dordrecht, Kluwer, pp. 13–40.

R. Smith 1995, ‘Logic’. In J. Barnes (ed.), *The Cambridge Companion to Aristotle*. Cambridge, Cambridge University Press, pp. 27–65.

R. Stalnaker 1970, ‘Pragmatics’. *Synthese* **22**, 272–289.

R. Stalnaker 1999, *Context and Content*. Oxford, Oxford University Press.

M. Stokhof and J. Groenendijk 1997, ‘Questions’. In J. van Benthem and A. ter Meulen (eds.), *Handbook of Logic and Language*. Elsevier/MIT, Amsterdam/Cambridge, Mass. Press, pp. 1055–1124.

M. Stokhof 2005, ‘Hand or hammer?’. Talk presented at the ‘100 Years On Denoting’ meeting (June 24, 2005, Utrecht).

P.F. Strawson 1950, ‘On referring’. In *Mind* **59**, 320–344.

M. Stroinska and D. Hitchcock 2002, ‘Introduction’. In A. Tarski 2002, pp. 155–175.

B.G. Sundholm 1998a, ‘Inference, consequence, implication: a constructivist’s perspective’. *Philosophia Mathematica* **6**, 178–194.

B.G. Sundholm 1998b, ‘Inference versus consequence’. In Logica *Yearbook 1997*. Prague, Filosofia Publishers, pp. 26–35.

B.G. Sundholm 2001, ‘A plea for logical atavism’. In Logica *Yearbook 2000*. Prague, Filosofia Publishers, pp. 151–162.

B.G. Sundholm 2002a, ‘Varieties of consequence’. In D. Jacquette (ed.), *Blackwell Companion to Philosophical Logic*. Oxford, Blackwell, pp. 241–255.

B.G. Sundholm 2002b, ’What is an expression?’. In Logica *Yearbook 2001*. Prague, Filosofia Publishers, pp. 181–194.

B.G. Sundholm 2003, ‘Tarski and Lesniewski on languages with meaning versus languages without use: a 60th birthday provocation for Jan Wolenski’. In J. Hintikka, T. Czarnecki, K. Kijania-Placek, T. Placek and A. Rojszczak (eds.), *Philosophy and Logic. In Search of the Polish Tradition*. Dordrecht, Kluwer, pp. 109–128.

Tarski 1959, *Introduction to Logic and to the Methodology of Deductive Sciences*. New York, Oxford University Press, eighth printing.

A. Tarski 2002, ‘On the concept of following logically’. *History and Philosophy of Logic* **23**, 155–196.

R. Torretti 2003, ‘Nineteenth century geometry’. *The Stanford Encyclopedia of Philosophy (Fall 2004 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2004/entries/geometry-19th/>.

H. Wang 1955, ‘On formalization’. *Mind* **64**, 226–238.

L. Wittgenstein 1953, *Philosophical Investigations*. Transl. G.E.M. Anscombe. Oxford, Blackwell.

L. Wittgenstein 1963, *Tractatus Logico-Philosophicus*. Transl. D.F. Pears and B.F. MacGuiness. London, Routledge & Kegan Paul.

Zach 2003a, ‘The practice of finitism: epsilon calculus and consistency proofs in Hilbert’s program’. *Synthese* **137**(1–2), 211–259.

Zach 2003b, ‘Hilbert’s program’. *The Stanford Encyclopedia of Philosophy (Fall 2003 Edition)*, E.N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2003/entries/hilbert-program/>.

E. Zermelo 1908, ‘Investigations in the foundations of set theory I’. In J. van Heijenoort (ed.) 1977, pp. 199–215.

H. de Swart 1998, *Introduction to Natural Language Semantics*. Stanford, CSLI Publications.

INDEX

LOGIC, EPISTEMOLOGY, AND THE UNITY OF SCIENCE

Editors:
Shahid Rahman, *University of Lille III, France*
John Symons, *University of Texas at El Paso, U.S.A.*

1. S. Rahman, J. Symons, D. M. Gabbay and J. P. v. Bendegem (eds.): *Logic, Epislemology, and the Unity of Science*, Vol. 1. 2004 ISBN 1-4020-2807-5
2. D. Vanderveken (ed.): *Logic, Thought and Action*, Vol. 2. 2005 ISBN 1-4020-2616-1
3. J. v. Benthem, G. Heinzmann, M. Rebuschi and H. Visser (eds.): *The Age of Alternative Logics*. Assessing Philosophy of Logic and Mathematics Today, Vol. 3. 2006 ISBN 1-4020-5011-9
4. J. Faye, P. Needham, U. Scheffler and M. Urchs (eds.): *Nature's Principles*, Vol. 4. 2005 ISBN 1-4020-3257-9
5. B. v. Kerkhove and J. P. v. Bendegem (eds.): *Perspectives on Mathematical Practices*. Bringing Together Philosophy of Mathematics, Sociology of Mathematics, and Mathematics Education, Vol. 5. 2006 ISBN 1-4020-5033-X
6. A. Fagot-Largeault, S. Rahman and J. M. Torres (eds.): *The Influence of Genetics on Contemporary Thinking*, Vol. 6. 2007 ISBN 1-4020-5663-X
7. C. Dutilh Novaes: *Formalizing Medieval Logical Theories*. Suppositio, Consequentiae and Obligationes, Vol. 7. 2007 ISBN 978-1-4020-5852-3