Enrico Martino

Intuitionistic Proof versus Classical Truth

The Role of Brouwer's Creative Subject in Intuitionistic Mathematics

Springer

Enrico Martino
FISPPA Department
University of Padua
Padua
Italy

ISSN 2214-9775 ISSN 2214-9783 (electronic)
Logic, Epistemology, and the Unity of Science
ISBN 978-3-319-74356-1 ISBN 978-3-319-74357-8 (eBook)
https://doi.org/10.1007/978-3-319-74357-8

Library of Congress Control Number: 2017964580

Preface

I was stimulated by Göran Sundholm to collect some of my old papers on Intuitionism, which turn back to the eightieths and the ninetieths of the past century. In those years, there was a lively philosophical debate between classical and intuitionistic logicians and mathematicians. I was especially interested in the works of Dummett, Troelstra, van Dalen, Sundholm, Veldman and others.

A main peculiarity of my research is a deep analysis of the main tenets of the pioneers of Intuitionism: Brouwer and Heyting. In particular, it is analysed Heyting's explanation of the intuitionistic meaning of logical constants and Brouwer's idealisation of the creative subject as grounds of intuitionistic truth.

Besides, it is upheld the importance of the role of certain imaginary acts of choice performed by an ideal agent for explaining the notion of reference not only in intuitionistic but also in classical logic.

A crucial question, discussed in the work, is the following: to what extent succeeds the intuitionistic perspective to avoid the classical realistic notion of truth? My answer is that a form of realism is hidden in the idealisation of Brouwer's creative subject. In fact, in order to exploit the role of the creative subject, we need to think him as if he were a real being: the mere imagine of him in our mind would not be able to perform the actions required by his role.

Some papers of the present collection are written together with Daniele Giaretta and Gabriele Usberti.

I'm grateful to Göran Sundholm for his interest in my work on Intuitionism. I thank the Springer editor Shahid Rahman. Besides of thank Ali Mohammed, Stephen O'Reilly and Nisha Keeran. A particular thank to my colleague Vittorio Morato for preparing the manuscript in LATEX.

Padua, Italy — Enrico Martino
June 2017

Contents

1 Brouwer, Dummett and the Bar Theorem 1
1.1 Introduction 1
1.2 Terminology and Symbolism 1
1.3 Dummett's Argument 3
1.4 Critique of Dummett's Argument 4
1.5 Limits of the Eliminability of ζ-inferences 10
1.6 Final Considerations 12
References 14

2 Creative Subject and Bar Theorem 15
2.1 The Creative Subject 15
2.2 The Creative Subject and Existential Statements 19
2.3 Equivalence of (PIE) and (BI_M) 21
References 22

3 Natural Intuitionistic Semantics and Generalized Beth Semantics 23
3.1 Introduction 23
3.2 Generalized Beth-Models and Natural Models 23
3.3 Generalized Natural Models 24
References 26

4 Connection Between the Principle of Inductive Evidence and the Bar Theorem 27
4.1 Inductive Evidence 27
Reference 30

5 On the Brouwerian Concept of Negative Continuity 31
5.1 Introduction 31
5.2 The Negative Continuity Theorem 32
5.3 A Proof of NCP 33
5.4 Weak and Strong Negation 33

5.5 The Role of Time in Brouwer's Argument . . . 34
5.6 Brouwer's Argument and Solipsism . . . 35
5.7 NCP and Lawless Sequences . . . 37
5.8 Revising NCP with the Help of the Creative Subject . . . 40
5.9 Extensional Functions and Intensional Choice Sequences . . . 41
5.10 Troelstra's Abstraction Process and NCP . . . 42
5.11 Conclusions . . . 44
References . . . 45

6 Classical and Intuitionistic Semantical Groundedness . . . 47
6.1 Introduction . . . 47
6.2 Construction of Model M . . . 48
6.3 Axiomatisation of T . . . 49
6.4 The Aczel–Feferman Intensional Operator . . . 50
References . . . 51

7 Brouwer's Equivalence Between Virtual and Inextensible Order . . . 53
7.1 Introduction . . . 53
7.2 Reconstruction of Brouwer's Paper of 1927 . . . 54
7.3 Comment on Brouwer's Text . . . 55
7.4 How Brouwer Misinterpreted Himself . . . 57
7.5 A Minor Mistake in the Cambridge Lectures . . . 60
7.6 On Posy's Reconstruction . . . 61
References . . . 62

8 An Intuitionistic Notion of Hypothetical Truth for Which Strong Completeness Intuitionistically Holds . . . 63
8.1 Introduction . . . 63
8.2 Symbolism and Conventions . . . 64
8.3 The Failure of Strong Completeness for Natural Semantics . . . 65
8.4 Hypothetical Truth . . . 65
8.5 Remarks on Hypothetical Truth . . . 67
8.6 Generalized Beth Semantics . . . 69
8.7 Connection Between Hypothetical Semantics and GB-Semantics . . . 70
8.8 A Strong Completeness Proof for GB-Semantics . . . 71
References . . . 73

9 Propositions and Judgements in Martin-Löf . . . 75
9.1 Introduction . . . 75
9.2 Propositions and Judgements . . . 75
9.3 Truth and Evidence . . . 78
9.4 Metaphysical Realism . . . 80
References . . . 84

10 Negationless Intuitionism 85
10.1 Natural Semantics 85
10.2 Failure of Strong Completeness 86
10.3 Second-Order Negationless Semantics 89
10.4 Concluding Remarks 93
References 95

11 Temporal and Atemporal Truth in Intuitionistic Mathematics 97
11.1 Introduction 97
11.2 Tenselessness and Classical Truth 98
11.3 Potential Intuitionism as a Subsystem of Epistemic Mathematics 102
11.4 Temporal Truth 106
References 111

12 Arbitrary Reference in Mathematical Reasoning 113
12.1 Introduction 113
12.2 Some Objections to *TAR* 114
12.3 *TAR* as Embodied in the Logical Concept of an Object 117
12.4 The Ideal Agent 119
12.5 Arbitrary Reference and Impredicativity 122
12.6 Plural Reference Versus Sets 124
References 130

13 The Priority of Arithmetical Truth over Arithmetical Provability 133
13.1 Introduction 133
13.2 Orthodox Versus Non-orthodox Intuitionism 135
13.3 The Constructive Notion of a Process 138
13.4 Computational Realism 140
References 145

14 The Impredicativity of the Intuitionistic Meaning of Logical Constants 147
References 155

15 The Intuitionistic Meaning of Logical Constants and Fallible Models 157
15.1 Introduction 157
References 163

Origin of the Essays 165

Subject Index 167

Author Index 169

Chapter 1
Brouwer, Dummett and the Bar Theorem

with P. Giaretta

Abstract It is criticised Dummett's refutation of Brouwer's dogma. It is argued that his criticism rests on an erroneous interpretation of Brouwer's idea of "canonical proof".

1.1 Introduction

In Dummett (1977), *Elements of Intuitionism*, hereafter *E I*, Dummett gives a refutation of the "Brouwer's dogma", the famous assumption on which Brouwer based his proof of the bar theorem, by means of an original argument involving the intuitionistic notion of "proof containing inferences with infinitely many premisses". In the present article, we criticise Dummett's argument, which rests on what seems to us an incorrect interpretation of Brouwer's idea of "canonical proof", and we propose an alternative interpretation of it. Notwithstanding our critique, we thank Dummett for having stimulated our reflection on the foundation of Intuitionism.

1.2 Terminology and Symbolism

$\alpha, \beta, \gamma, \ldots$ are variables for choice sequences. $\overline{\alpha}(n)$ is the finite sequence of the first n terms of α:

$$\overline{\alpha}(n) = \langle \alpha(o), \ldots, \alpha(n-1) \rangle \quad (\overline{\alpha}(0) = \langle\rangle)$$

$\vec{u}, \vec{v}, \vec{w}, \ldots$ are variables for finite sequences of natural numbers. If $\vec{u} = \langle u_0, \ldots, u_n \rangle$, the length of $\vec{u}$ (the number of its terms), which we will denote by $l(\vec{u})$, is $n+1$. $l(\langle\rangle) = 0$

If $\vec{u} = \langle u_0, \ldots, u_n \rangle$ and $\vec{v} = \langle v_0, \ldots, v_m \rangle$, we let $\vec{u} * \vec{v}$ be $\langle u_0, \ldots, u_n, v_0, \ldots, v_m \rangle$. If n is a natural number, we let $\widehat{\vec{u}\, n}$ be $\vec{u} * \langle n \rangle$.

$\widehat{\vec{u}\, n}$, for any n, will be called a successor of $\vec{u}$ and $\vec{u}$ will be called the predecessor of every $\widehat{\vec{u}\, n}$.

We will use "$\vec{u}$ is extended by $\vec{v}$", in symbols $\vec{u} \geq \vec{v}$, to mean that $\exists \vec{w}(\vec{v} = \vec{u} * \vec{w})$.

We will use the term "spread" to refer to a function s defined on all finite sequences and with values in $\{0, 1\}$ such that:

$$s(\langle\rangle) = 0, \qquad s(\vec{u}) = 0 \to \exists k s(\widehat{\vec{u}\, k}) = 0, \qquad s(\widehat{\vec{u}\, k}) = 0 \to s(\vec{u}) = 0$$

If $s(\vec{u}) = 0$, we say that $\vec{u}$ is admissible for s, and s can be thought of as the tree whose nodes are the admissible finite sequences. In the following we will write indifferently $s(\vec{u}) = 0$ or $\vec{u} \in s$.

If $\forall \vec{u} s(\vec{u}) = 0$, s is the universal spread.

We will say that a sequence α belongs to a spread s, $\alpha \in s$, if $\forall n s(\overline{\alpha}(n)) = 0$.

A sequence α belongs to a finite sequence $\vec{u}$, $\alpha \in \vec{u}$, if $\exists n \overline{\alpha}(n) = \vec{u}$.

Let $\vec{u} \in s$ and R be a species of nodes of s. We say that R *bars* $\vec{u}$ if every choice sequence of s belonging to $\vec{u}$ meets R, formally if $\forall \alpha_{\alpha \in s}(\alpha \in \vec{u} \to \exists n \overline{\alpha}(n) \in R)$.

We say that R is *monotonic* if $\vec{u} \in R \to \forall k(\widehat{\vec{u}\, k} \in s \to \widehat{\vec{u}\, k} \in R)$.

We say that R is *decidable* if $\forall \vec{u}_{\vec{u} \in s}(\vec{u} \in R \vee \vec{u} \notin R)$.

Let R be a species of nodes of s. We call $\mathcal{F}$-closure of R (in s) the species $R^{\mathcal{F}}$ inductively defined as follows:

(i) $\vec{u} \in R \to \vec{u} \in R^{\mathcal{F}}$;
(ii) $\forall \vec{u}_{\vec{u} \in s}(\forall k \widehat{\vec{u}\, k_{\widehat{\vec{u}\, k \in s}}} \in R^{\mathcal{F}} \to \vec{u} \in R^{\mathcal{F}})$.

Following Kleene and Vesley (1965), hereafter referred to as *FIM*, we say that a node $\vec{u}$, belonging to s, is inductively barred by R if $\vec{u} \in R^{\mathcal{F}}$. It is evident, by induction on the construction of $R^{\mathcal{F}}$, that if R bars inductively $\vec{u}$, then R bars $\vec{u}$.

At this point bar theorem, more explicitly the bar-induction theorem (shortly BI), monotonic bar theorem, decidable bar theorem, with regard to a spread s and to a species of nodes R, can be expressed as follows:

- $BI(s, R)$: If R bars $\langle\rangle$, then R inductively bars $\langle\rangle$.
- $BI_M(s, R)$: If R is monotonic and R bars $\langle\rangle$, then R inductively bars $\langle\rangle$.
- $BI_D(s, R)$: If R is decidable and R bars $\langle\rangle$, then R inductively bars $\langle\rangle$.

In the following, for simplicity, we will refer mostly to the universal spread and simply say that R satisfies the bar theorem (and we will write $BI(R)$).

We will write simply BI instead of $\forall R BI(R)$ (analogously for BI_M and BI_D).

BI_M and BI_D are usually assumed as axioms in treatments of intuitionistic analysis. It is known that $BI_M \to BI_D$ and that, under the hypothesis of $\forall \alpha \exists x$-continuity, $BI_D \to BI_M$.

BI is instead certainly false, as the following *Kleene counterexample* (KC) shows.

Let $P(x)$ be an arbitrary decidable predicate on the natural numbers and R be the species of nodes defined by: $\langle k \rangle \in R \leftrightarrow P(k)$, $\langle\rangle \in R \leftrightarrow \neg \forall x P(x)$. Obviously R bars $\langle\rangle$, but until we have a proof of $\forall x P(x) \vee \neg \forall x P(x)$, we cannot assert that $\langle\rangle \in R^{\mathcal{F}}$.

Brouwer gave various formulations and proofs of the bar theorem or of particular cases of it in Brouwer (1924b, 1927, 1954), hereafter referred to as 1924b, 1927, 1954. The hypothesis of decidability or monotonicity of R never occurs explicitly in Brouwer's proofs. Yet the hypotheses of 1924b and of 1927 imply the decidability of R. On the other hand, in the formulation of 1954, which seems to be the most general, there is no hypothesis on the barring species and so, for KC, this formulation is certainly wrong.

All three proofs are based on the assumption, which we will call "Brouwer's dogma" (BD), that a proof of a node being barred can always be reduced to a "canonical proof" (c.p.), where only inferences of the following three forms occur:

- η-inferences: $\dfrac{\vec{u} \in R}{R \text{ bars } \vec{u}}$
- ζ-inferences: $\dfrac{R \text{ bars } \vec{u}}{R \text{ bars } \widehat{\vec{u}\, k}}$
- $\mathcal{F}$-inferences: $\dfrac{R \text{ bars } \widehat{\vec{u}\, 0}, \ldots, R \text{ bars } \widehat{\vec{u}\, k}, \ldots}{R \text{ bars } \vec{u}}$

(Indeed Brouwer does not mention explicitly the η-inferences.) Then the proof proceeds by acknowledging the 'eliminability of the ζ-inferences', i.e. the possibility of further reducing the canonical proof to one in which only η- and $\mathcal{F}$-inferences occur. Finally in Brouwer's argument, translated into our notation, the conclusion is obtained by observing that a so-reduced proof of 'R bars $\langle\rangle$' can be transformed into a proof of '$\langle\rangle \in R^{\mathcal{F}}$' by a simple substitution of '$\vec{u} \in R^{\mathcal{F}}$' for every occurrence of 'R bars $\vec{u}$'.

1.3 Dummett's Argument

Dummett describes a c.p., to which BD refers, as a proof "expanded into its 'fully analysed' form, that is, a form, in which every step has been broken down into a sequence of steps each of which is as short as possible" (Dummett 1977, p. 94). After he gives a more detailed description:

> Brouwer's notion of an analysis of a proof appears to be this: that whenever, in the course of the proof, we appeal to some operation as yielding a result of a certain kind, then, in the analysed form of the proof, that operation will actually be carried out. Thus the appearance of a universally quantified statement in a proof, for instance the statement
>
> $$\forall k\ R \text{ bars } \widehat{\vec{u}\, k}$$
>
> signifies our recognition that a certain operation will, when applied to any element of the domain (in this case, to any natural number k), yield a proof of the corresponding instance (here, of the statement 'R bars $\widehat{\vec{u}\, k}$'). In the fully analysed proof, therefore, the universal quantification does not appear: the operation is actually applied to each element of the domain, yielding a proof of the corresponding instance, and that which formerly was inferred from the universally quantified statement now appears as following from the individual instances taken together. (Dummett 1977, p. 96)

The proofs, in which, according to Brouwer, only η-, ζ- and $\mathcal{F}$-inferences occur, should be of this type. After having considered a c.p. in the form of a well-founded tree (i.e. of a finite paths tree), Dummett starts to expound Brouwer's proof by describing the process of elimination of the ζ-inferences. We quote here only the description of the case of a ζ-inference which is preceded by a η-inference:

$$\begin{array}{c} \vec{u} \in R \\ | \\ \eta \\ | \\ R \text{ bars } \vec{u} \\ | \\ \zeta \\ | \\ R \text{ bars } \widehat{\vec{u}\, k} \end{array}$$

Proof-tree 1

since this case, as we will see, turns out to be particularly critical.

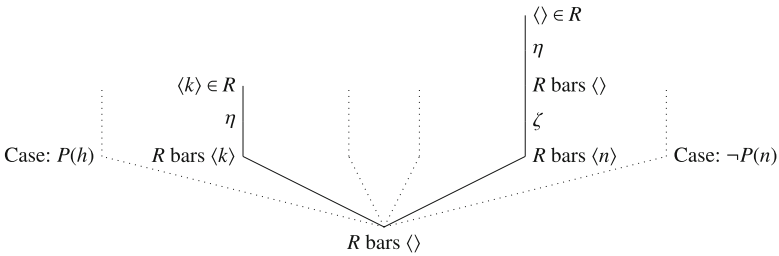
> […] we consider that path in the proof-tree which leads from the conclusion 'R bars $\widehat{\vec{u}\, k}$' of our ζ-inference to the conclusion 'R bars $\langle\rangle$' of the whole proof. Each of the statements occurring on this path is of the form 'R bars $\vec{v}$' for some finite sequence $\vec{v}$ which is either an initial segment or an extension of $\vec{u}$. Moreover, in the passage from one statement to the next, the length of the finite sequence mentioned is either increased or diminished by 1, and the length of the sequence mentioned at the end of the path is 0. Hence, somewhere along the path there must occur the statement 'R bars $\vec{u}$'. We accordingly replace the entire proof of that occurrence of 'R bars $\vec{u}$' by a derivation of it by means of an η-inference. (Dummett 1977, pp. 97–98)

Finally, after having expounded the obvious conclusion of the proof (which we mentioned at the end of the last paragraph), Dummett observes that since this proof makes no reference either to the decidability or to the monotonicity of R, it is certainly wrong because of KC. He concludes that what is wrong must be BD, the only unjustified assumption of the proof, and he considers KC a refutation of it.

> It comes as no surprise that consideration of this example shows that the flaw in Brouwer's proof lies in his unsupported assertion that any fully analysed proof that a species bars $\langle\rangle$ (or any other finite sequence $\vec{u}$) can contain only η-, ζ- and $\mathcal{F}$-inferences. (Dummett 1977, p. 98)

1.4 Critique of Dummett's Argument

According to Dummett's conclusion, there should not be only η-, ζ- and $\mathcal{F}$-inferences in a c.p of 'R bars $\langle\rangle$', where R is the barring species of KC. On the other hand the proof-tree:



Proof-tree 2

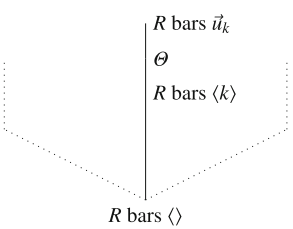
would seem to contain only inferences of the mentioned types. But Dummett rejects this obvious objection on the basis of a strange principle of constructivity.

He says that, since an intuitionistic proof must be understood as a mere mental act, "we cannot consider it as, so to speak, already having certain features, or displaying certain patterns, which we have failed to notice" (Dummett 1977, p. 100). He concludes that, with regard to the species R of KC, though it is obviously true that for a fixed k we can decide whether $P(k)$ or $\neg P(k)$ and so deduce 'R bars $\langle k\rangle$' according to the case, from '$\langle k\rangle \in R$' by means of an η-inference or from 'R bars $\langle\rangle$' by means of a ζ-inference, yet in the whole proof of 'R bars $\langle\rangle$', understood as a unique global mental act, the various $\langle k\rangle$ have not been individually analysed and therefore the various formulae 'R bars $\langle k\rangle$' are not deduced in it by Brouwer's inferences.

Dummett gives the name "Θ-inference" to the "analysed version" of an inference of the type:

$$\frac{R \text{ bars } \vec{u} \vee R \text{ bars } \widehat{\vec{u}\ k}}{R \text{ bars } \widehat{\vec{u}\ k}}$$

that is the inference which we obtain from it by replacing the premiss by a determinate one of the two disjoints, and maintains that all that can be said about the c.p. of 'R bars $\langle\rangle$', considered as global mental act, is that the formulae 'R bars $\langle k\rangle$' are deduced by means of Θ-inferences. Therefore the c.p. at issue would have, according to Dummett, not the aspect of the proof-tree 2 but the aspect:



where $\vec{u}_k \begin{cases} = \langle k\rangle & \text{if } P(k) \\ = \langle\rangle & \text{if } \neg P(k) \end{cases}$

Proof-tree 3

So the Θ-inferences would constitute a new type of inference which can occur in a c.p., and, according to Dummett, they would not be reducible to those of Brouwer.

We think, on the contrary, that the proof-tree 2 is intuitionistically no less correct than proof-tree 3. Certainly, in the mental act constituting the proof 2, the various $\langle k \rangle$ are not completely analysed; what is present in that mental act is only the mere possibility of analysing them and then of deducing 'R bars $\langle k \rangle$', according to the case, by a suitable Brouwer inference. But this is the best we can demand from an intuitionistic point of view! In accordance with the conception of the potential infinite, the fact that the proof of 'R bars $\langle\rangle$' consists of the infinitely many proofs of the various 'R bars $\langle k \rangle$' can only be understood in the sense that the knowledge of a generative process of the infinitely many proofs of the premisses of the last $\mathcal{F}$-inference is a constituent of the proof. This can be said with reference to the proof-tree 2 as well as to proof-tree 3.

Dummett himself explains very correctly the intuitionistic meaning of an inference with infinitely many premisses:

> Thus the only way of understanding the idea of an inference from denumerably many premisses $A(0), A(1), \ldots$ which is consistent with a constructivist outlook proves to coincide exactly with the intuitionistic interpretation of an inference from $\forall n A(n)$. (Dummett 1977, pp. 96–97)

But then he goes on:

> An intuitionistic proof involving inferences from universally quantified statements really is, therefore, what Brouwer maintains, a representation of a more analysed proof containing inferences from infinitely many premisses. (Dummett 1977, p. 97)

Now, it is not clear exactly what Dummett means. The expression "representation of a more analysed proof containing inferences from infinitely many premisses" and what he says about a "fully analysed proof" suggest the notion of a proof, where the infinitely many premisses $A(0), A(1), \ldots$ are all visibly present. But this notion does not have any sense from an intuitionistic point of view, according to which, as Dummett himself recognises, to draw a conclusion from the infinitely many premisses $A(0), A(1), \ldots$ can only mean to draw it from the premiss $\forall n A(n)$. There is no way in which a proof of $\forall n A(n)$ can have a "fully analysed version" such that the various proofs of the $A(k)$ occur explicitly in it. It may be, of course, that the knowledge of the generative process makes us a priori aware that the infinitely many proofs are very similar, in which case the "generic" proof of $A(k)$ would be visualisable, or that, on the contrary, the form of a proof of $A(k)$ essentially depends on k, in which case it would be difficult to conceive the global configuration of the infinitely many proofs. But this is a distinction which, though suggestive, rests on a concept of "visualisability" which, in our view, cannot satisfactorily characterise the soundness of an intuitionistic proof, both because it is too vague and, above all, because it seems to concern the intuitability of the actual infinite rather of potential infinite.

However, even if someone wanted to adopt such a standard of evaluation, it would not seem possible for him to consider the proof-tree 3 more acceptable than proof-tree 2. The same argument used by Dummett against proof-tree 2 can be used against proof-tree 3: in this proof, conceived as global mental act, it is not known what the various $\vec{u}_k$ are. Only after having analysed a single k, it is possible to establish

whether $\vec{u}_k = \langle k \rangle$ or $\vec{u}_k = \langle\rangle$ and so to execute the corresponding Θ-inference. But, without such an analysis, all that can be said is that inferences of the type:

$$\frac{R \text{ bars } \langle k \rangle \vee R \text{ bars } \langle\rangle}{R \text{ bars } \langle k \rangle}$$

and not their analysed forms, occur in that proof. It turns out that one can have a global intuition of the proof in question only if one gives up demanding a representation of its "fully analysed" form in the sense expounded by Dummett.

Dummett attributes to Brouwer the idea of "fully analysed proof" which he expounds. We do not know to which of Brouwer's articles he refers. However, it seems to us that at least in the articles 1924b, 1927, 1954, in which he discusses the bar theorem, Brouwer means by "analysis of a proof" something different. The intuition on which the dogma rests seems to be the following: the knowledge of the fact that a node $\vec{u}$ is barred must, after all, either be reduced to the knowledge that $\vec{u} \in R$ or be obtained through the knowledge that its predecessor or all its successors are barred. That is, it is always possible to transform any given proof P of 'R bars $\vec{u}$' into a new proof P' of one of the following types:

$\vec{u} \in R$

|

R bars $\vec{u}$

Proof-tree 4

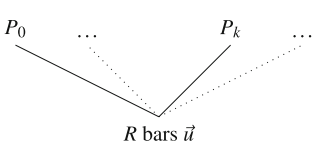
P''

|

R bars $\vec{u}$

Proof-tree 5

where P'' is a proof of 'R bars $\vec{v}$' and $\vec{v}$ is the precedessor of $\vec{u}$.



Proof-tree 6

where $P_0, \ldots, P_k, \ldots$ are proofs of 'R bars $\widehat{\vec{u}\,0}$', ..., 'R bars $\widehat{\vec{u}\,k}$', ...

P'' and $P_0, \ldots, P_k, \ldots$ are simpler proofs than P'. The analysis of P in order to obtain P' consists in explicating, by reflection on P, the implicit reference to the nodes surrounding $\vec{u}$ and not, as Dummett seems to believe, in actually carrying out those constructions which had been acknowledged feasible in P. In case of proof-tree 6 P' consists of a generative process of the proofs P_k, which are not necessarily canonical, and of the final $\mathcal{F}$-inference. In its turn, every P_k can be transformed into a P'_k which, if different from proof-tree 4, splits into simpler proofs, and so on.

These considerations suggest the following precise inductive definition of c.p.

Definition 1.1 (a) $\dfrac{\vec{u} \in R}{R \text{ bars } \vec{u}}$ is a c.p. of 'R bars $\vec{u}$';

(b) If P is a c.p. of 'R bars $\vec{u}$', then for every k, $\dfrac{P}{R \text{ bars } \widehat{\vec{u}\,k}}$ is a c.p. of 'R bars $\widehat{\vec{u}\,k}$';

(c) If, for every k, P_k is a c.p. of 'R bars $\widehat{\vec{u}\,k}$', then $\dfrac{P_0, \ldots, P_k, \ldots}{R \text{ bars } \vec{u}}$ is a c.p. of 'R bars $\vec{u}$'.

By generalising the Definition 1.1, it is possible to obtain the general concept of an intuitionistic proof-tree. To be precise, we state first of all the following definition of "inductively defined tree" (i.d.t.):

Definition 1.2 (a) $\{\langle\rangle\}$ is an i.d.t.

(b) If $T_0, \ldots, T_n$ is a finite sequence of i.d.t., then the species formed by $\langle\rangle$ and by $\langle k\rangle * \vec{u}$, where $\vec{u} \in T_k$ $(0 \leq k \leq n)$, is an i.d.t.;

(c) If $T_0, \ldots, T_k, \ldots$ is an infinite sequence of i.d.t., then the species formed by $\langle\rangle$ and by $\langle k\rangle * \vec{u}$, where $\vec{u} \in T_k$ $(0 \leq k)$, is an i.d.t.

Then the definition of "proof-tree" (p.t.) concerning a species $\mathcal{A}$ of axioms and a species $\mathcal{I}$ of inference rules is as follows:

Definition 1.3 A p.t., concerning $\mathcal{A}$ and $\mathcal{I}$, is a couple $\langle T, c\rangle$, where T is an i.d.t. (called the support of the p.t.) and c is an application which associates a formula with every node of T so that

(a) if $\vec{u}$ is a terminal node, then $c(\vec{u}) \in \mathcal{A}$;

(b) if $\vec{u}$ is not terminal, then $c(\vec{u})$ is the conclusion and the various $c(\widehat{\vec{u}\,k})$, for all k such that $\widehat{\vec{u}\,k} \in T$, are the premisses of an instance of one of the inference rules of $\mathcal{I}$.

We think that this is the precise and correct explication of the intuitionistic concept of p.t. A confirmation that this was, in substance, the concept of c.p. which Brouwer had in mind in his exposition of the bar theorem seems to be the fact that the Definition 1.2 is quite analogous to Brouwer's definition of well-ordered species and, in a footnote in 1927, Brouwer mentions explicitly the analogy between mathematical proofs and well-ordered species:

> Just as, in general, well-ordered species are produced by means of the two generating operations from primitive species [cf. (Brouwer 1926, p. 451)], so, in particular, mathematical proofs are produced by means of the two generating operations from null elements and elementary inferences that are immediately given in intuition (albeit subject to the restriction that there always occurs a last elementary inference). (Brouwer 1927, p. 460)

An alternative way of characterising a p.t. is to take as support a barred spread instead of an i.d.t. This is the way followed, in substance, by Kleene in FIM in the formalisation of Brouwer's proof of the bar theorem (FIM, pp. 65–67) and seems to be implicitly adopted by Dummett in his informal description of p.t. In particular Dummett does not care at all that the tree is inductively defined and demands only that it is well-founded:

> To be a proof, it must be well-founded: if the proof is conceived as arranged in tree form, every branch must be finite. (If it were possible to form an infinite sequence of propositions, beginning with the conclusion of the proof, each subsequent proposition being one of the premisses upon which the preceding one depended, then we should have no reason to accept as true any proposition in the sequence, including the conclusion of the 'proof'. This resembles Aquinas's denial of the possibility of an infinite regress in causes). (Dummett 1977, p. 95)

We want to point out a weak point of this approach. We define a *barred spread* (b.s.) as a couple $\langle s, R\rangle$, where s is a spread and R a species of nodes of s barring the vertex. We define a p.t. with support $\langle s, R\rangle$ in the following way:

Definition 1.4 A p.t. with support $\langle s, R\rangle$, concerning $\mathcal{A}$ and $\mathcal{I}$, is a term $\langle s, R, c\rangle$, where c is an application associating a formula with every node of s so that

(a) if $\vec{u} \in R$, then $c(\vec{u}) \in \mathcal{R}$;
(b) $c(\vec{u})$ is the conclusion and the various $c(\widehat{\vec{u}\ k})$, for all k such that $\widehat{\vec{u}\ k} \in s$, are the premisses of an instance of an inference rule $\mathcal{I}$.

We say that a p.t. (both according to the Definition 1.3 and the Definition 1.4) is sound if, whenever the axioms are true (under a certain interpretation) and the inference rules are truth-preserving, the conclusion (formula at the vertex) is true.

We say that an i.d.t. (a b.s.) is sound if every p.t. having it as support is sound.

Now, it follows immediately, by induction on its construction, that every i.d.t. is sound.

Is a b.s. always sound?

Let us suppose that $\langle s, R\rangle$ is sound. Then, if we take as $\mathcal{A}$ the species of the formulae '$\vec{u} \in R^{\mathcal{F}}$', where $\vec{u} \in R$, as $\mathcal{I}$ the species of the $\mathcal{F}$-inferences, as c the application associating the formula '$\vec{u} \in R^{\mathcal{F}}$' with every $\vec{u} \in s$, the p.t. $\langle s, R, c\rangle$ turns out to be sound, from which follows $\langle\rangle \in R^{\mathcal{F}}$. So $BI(s, R)$.

Let us suppose, vice versa, $BI(s, R)$. Let $\langle s, R, c\rangle$ be any p.t., with support $\langle s, R\rangle$, concerning a set $\mathcal{A}$ of true axioms (under a certain interpretation) and a set $\mathcal{I}$ of truth-preserving inference rules. Then, by induction on the construction of $R^{\mathcal{F}}$, it follows that, for every $\vec{u} \in R^{\mathcal{F}}$, $c(\vec{u})$ is true (under the interpretation in question). Since $\langle\rangle \in R^{\mathcal{F}}$ by hypothesis, we conclude that $c(\langle\rangle)$ is true. So $\langle s, R\rangle$ is sound.

Therefore

Corollary 1.1 *A b.s. $\langle s, R\rangle$ is sound if and only if $BI(s, R)$.*

Since, by KC, BI does not hold in general, it follows that a $\langle s, R\rangle$ is not generally sound.

Of course, soundness can be obtained by imposing suitable conditions on $\langle s, R\rangle$, but, by Corollary 1.1, the problem of fixing such conditions is exactly equivalent to the problem of determining the validity conditions of the bar theorem. Thus, in order to prove the last one, to use the notion of proof based on a b.s. leads to a *petitio principii*. In FIM it is explicitly demanded (translating into our terminology) that $\langle s, R\rangle$ satisfies $BI(s, R)$; equivalently, it is possible to consider a notion of p.t. defined as in Definition 1.4 but with reference to a support $\langle s, R\rangle$ inductively barred

rather than simply barred. But then, because of the inductive definability of $R^{\mathcal{F}}$ which the notion of inductive barredness refers to, the notion of proof-tree according to the Definition 1.4 is essentially reduced to that of the Definition 1.3.

Therefore, henceforth, a p.t. will always meant according to the Definition 1.3 and a c.p. according to the Definition 1.1. By assuming such references, Brouwer's dogma

$BD(R)$: if R bars $\langle\rangle$, there is a c.p. of 'R bars $\langle\rangle$'

acquires a precise meaning.

For convenience of exposition, we will state BD(R) in a slightly different way. Let us consider the species *IndR* inductively defined as follows:

Definition 1.5 (a) $\vec{u} \in R \rightarrow \vec{u} \in IndR$;
(b) $\vec{u} \in R \rightarrow \widehat{\vec{u}\, k} \in IndR$;
(c) $\forall k \widehat{\vec{u}\, k} \in IndR \rightarrow \vec{u} \in IndR$.

That is *IndR* is the closure of R by η-, ζ- and $\mathcal{F}$-inferences. In other words, *IndR* is the species of the nodes $\vec{u}$ for which there exists a c.p. of 'R bars $\vec{u}$'. Then $BD(R)$ can be expressed in the form:

$BD(R)$: if R bars $\langle\rangle$, then $\langle\rangle \in IndR$.

Now, since, for the species R of KC, it is quite self-evident that $\langle\rangle \in IndR$ (i.e. that the proof-tree 2 is a c.p. in our sense), BD(R) holds and so, contrary to what Dummett maintains, KC does not refute Brouwer's dogma.

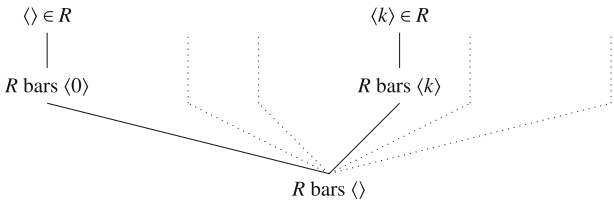
1.5 Limits of the Eliminability of ζ-inferences

Since, as we have seen, for the R of KC $BD(R)$ holds, but $BI(R)$ does not, the ζ-inferences certainly cannot be eliminated from the c.p. of 'R bars $\langle\rangle$'. In fact, the proof-tree 2 clearly shows that the process for eliminating the ζ-inferences described by Dummett (which we quoted in the Sect. 1.3) is not constructive: in order to apply it, we should replace the whole c.p. by the proof

$$\begin{array}{c} \langle\rangle \in R \\ | \\ R \text{ bars } \langle\rangle \end{array}$$

Proof-tree 7

if ζ-inferences occur in the original proof-tree 2, and we should obtain as c.p of 'R bars $\langle\rangle$'



Proof-tree 8

if ζ-inferences do not occur in proof-tree 2. Since proof-trees 7 and 8 differ in all nodes of length 1, it follows that, until we have proved $\forall x P(x) \vee \neg\forall x P(x)$ and so found out whether ζ-inferences occur in proof-tree 2, no node of length 1 of the c.p. can be built.

Therefore KC does not refute BD but refutes the unconditional eliminability of the ζ-inferences from a c.p.

The hypotheses of monotonicity or of decidability on R serve precisely to warrant such eliminability. In fact we have:

Lemma 1.1 *(a)* $\vec{u} \in IndR \wedge R\,is\,monotonic \rightarrow \vec{u} \in R^{\mathcal{F}}$.
(b) $\langle\rangle \in IndR \wedge R\,is\,decidable \rightarrow \langle\rangle \in R^{\mathcal{F}}$.

Proof (a) It suffices to prove that $R^{\mathcal{F}}$ is closed by ζ-inferences. In fact, by induction on $R^{\mathcal{F}}$, we can show that if $\vec{u} \in R^{\mathcal{F}}$, then $\vec{u} \in R$ or $\forall k\widehat{\vec{u}k} \in R^{\mathcal{F}}$. So, because R is monotonic, it follows in any case that $\forall k\widehat{\vec{u}\ k} \in R^{\mathcal{F}}$.

(b) We say that $\vec{u}$ is prebarred by R if there exists a $\vec{v} \geq \vec{u}$ such that $\vec{v} \in R$. Let us prove, by induction on *IndR*, that

$$\vec{u} \in IndR \rightarrow \vec{u} \in R^{\mathcal{F}} \vee \vec{u} \text{ is prebarred by } R.$$

Let us suppose that $\vec{u} \in IndR$. If $\vec{u} \in R$, the conclusion follows directly.

If $\vec{u} = \widehat{\vec{u}\ k}$ and $\vec{v} \in IndR$, then, by the inductive hypothesis, $\vec{v}$ is prebarred, in which case $\vec{u}$ is too, or $\vec{v} \in R^{\mathcal{F}} - R$, in which case $\forall k\widehat{\vec{v}\ k} \in R^{\mathcal{F}}$ and so, in particular, $\vec{u} \in R^{\mathcal{F}}$.

If, finally, $\forall k\widehat{\vec{u}\ k} \in IndR$, then we can decide, by virtue of the decidability of R, whether $\vec{u}$ is prebarred by R. In the positive case the conclusion follows directly. In the negative case we can show, by the inductive hypothesis, that, for every k, $\widehat{\vec{u}\ k} \in R^{\mathcal{F}}$ or $\widehat{\vec{u}\ k}$ is prebarred; in this last subcase, since $\vec{u}$ is not prebarred, $\widehat{\vec{u}\ k} \in R$ and so, in every one of the two subcases, $\widehat{\vec{u}\ k} \in R^{\mathcal{F}}$ whence $\vec{u} \in R^{\mathcal{F}}$.

It seems to us that this completely clarifies the relationship between KC and Brouwer's proof.

1.6 Final Considerations

Of course, the fact that KC does not refute BD does not say anything about the reliability of the latter. Yet it can easily be proved that BD is exactly equivalent to BI_M:

Theorem 1.1 $BD \leftrightarrow BI_M$

(where $BD = \forall R BD(R)$, $BI_M = \forall R BI_M(R)$)

Proof Let us assume BD and let R be a monotonic barring species. By BD $\langle\rangle \in IndR$ and so, by Lemma 1.1a, $\langle\rangle \in R^{\mathcal{F}}$. Vice versa, let us assume BI_M. If R is a species barring $\langle\rangle$, its monotonic closure R^{ζ} bars $\langle\rangle$ too, and so $\langle\rangle \in R^{\zeta\mathcal{F}} = IndR$.

Therefore, since BD and BI_M are equally reliable, Dummett's claim that it is convenient to assume BI_M directly as an axiom, in order to avoid the problematic character of BD, turns out to be quite groundless. Dummett makes a further observation in favour of his claim. Referring to the formalisation of the dogma in FIM, he says:

> In order to formalize it, it is necessary to assume, as an axiom, a formalization of the statement 'If R bars $\langle\rangle$, then there is a proof of "R bars $\langle\rangle$" which uses only η-, ζ- and $\mathcal{F}$-inferences'. Formalization of this statement is quite straightforward; a proof-tree can obviously be represented as a dressed spread. However, standing on its own, such an axiom appears quite ad hoc; especially when, in order to be able to exploit the well-founded character of the proof-tree, we should also have to assume axiomatically a principle of transfinite induction within such proof-trees (the general principle of transfinite induction can be proved equivalent to the principle of Bar Induction): it is therefore preferable to assume the principle of Bar Induction at the outset as an axiom or axiom schema. (Dummett 1977, pp. 102–103)

However, the assumption that the p.t.'s satisfy a principle of transfinite induction, which is essentially equivalent to the bar theorem, does not depend on the formalisation but on the notion of p.t. adopted by Kleene. This notion is explicitly intended to satisfy the above-mentioned principle, but, as we have already noted, there is no need to impose the same requirement on the notion of p.t. of Definition 1.3, since it is a natural consequence of its construction that the latter notion satisfies it.

We have two further remarks to make about the idea, suggested by Dummett, that BD would be an ad hoc axiom. First of all BD seems to express correctly Brouwers intuition, already mentioned above, that the knowledge that a node $\vec{u} \notin R$ is barred must be based on an examination of its surrounding nodes. On the contrary, this intuition is lost in the bar theorem, in which no explicit reference is made to the predecessor of $\vec{u}$. In the second place, even if no more evidence is attributed to BD than to BI_M, or one prefers to ignore the problem of evidence, BD seems to us interesting in itself. While, in fact, the bar definition involves the notion of infinite sequence, this notion does not occur at all in the definition of *IndR*, which is stated entirely in terms of finite sequences. Thus BD says, in effect, that the bar notion, even for an R such that $BI(R)$ does not hold, can be expressed in terms of finite sequences. In this sense, BD can be regarded as a generalisation of BI_M.

We want in conclusion to make some comments on Brouwer's original articles.

We shall consider first whether the mistaken idea that the ζ-inferences can, in general, be eliminated goes back to Brouwer. Brouwer summarises the elimination process in 1924a, and in 1954 without ever considering the case of a ζ-inference preceded by an η-inference, that is, the following situation:

$$\begin{array}{c} \vec{u} \in R \\ | \ \eta \\ R \text{ bars } \vec{u} \\ | \ \zeta \\ R \text{ bars } \widehat{\vec{u}\, k} \end{array}$$

Proof-tree 9

This case is not even considered in Heyting's description, concerning a finitary spread, in Heyting (1956, pp. 42–44). We can guess that in Brouwer (1924a, 1927) and in Heyting (1956), its treatment has been omitted because the implicit decidability of the barring species (besides, in Heyting, the finitary character of the spread) makes it obvious. But this is not so in Brouwer's (1954) in which no hypothesis on R occurs and thus, as we have seen, the elimination of the ζ-inferences is illicit. Brouwer's argument, paraphrased and translated into our terminology, is in substance the following:

> if in the given proof of 'R bars $\langle\rangle$' it has not been demonstrated that $\langle\rangle \in R$, then 'R bars $\langle\rangle$' has been deduced by an $\mathcal{F}$-inference; so, for every k, 'R bars $\langle k\rangle$' must have been proved before 'R bars $\langle\rangle$' and thus cannot have been deduced by a ζ-inference.

Here Brouwer seems indeed to overlook, probably because of the lack of precision of his statement, the possibility, pointed out by KC, that '$\langle\rangle \in R$' is not known but can nevertheless be used to construct, for some k, a proof of 'R bars $\langle k\rangle$'.

Our last observation concerns Brouwer's statement in 1954 that the barring species R is "not necessarily predeterminate". If, as we believe, this means that R can depend on choice sequence parameters, it seems to us that this possibility makes the validity of BD more problematical. Let us consider, in fact, the following example. We assume that the universe of choice sequences is built by a "creative subject" who, at every stage of knowledge n, conceives one and only one sequence α_n (the supposition that the creative subject conceives only one sequence per stage was put in question because of Troelstra's paradox, but this can be avoided by means of a suitable distinction of reference levels. Cf. Troelstra (1969)). Since such a universe, though numerable, is potentially as rich as the universe of all sequences, we conjecture that the bar theorem holds for it (within the limits within which it holds for the universe of all sequences).

Now, if R is the species of the nodes $\overline{\alpha}(n+1)$, for $n \geq 0$, then R obviously bars $\langle\rangle$ and is decidable (because a node $\vec{u}$ of length m belongs to R if and only if $m \neq 0$ and $\vec{u} = \overline{\alpha}_{m-1}(m)$, and yet $\langle\rangle \notin R^{\mathcal{F}}$, since $R^{\mathcal{F}} = R$).

In general, it seems to us that BD is fairly plausible, provided that R can be defined without reference to the concept of infinite sequence. In fact, as we have already noted, BD expresses the eliminability of this concept from the bar notion, i.e. the possibility of describing the species of the nodes barred by R without using the concept of infinite sequence. Now, this possibility seems to us rather unreliable (as well as not very interesting) if the concept in question is already essentially involved in the definition of R.

References

Brouwer, L. (1924a). Bemerkungen zum Beweis der gleichmässigen Stetigkeit voller Funktionen. *Verhandelingen der Koninklijke Nederlandsche Akademie van Wetenschappen te Amsterdam*, *27*, 644–646.

Brouwer, L. (1924b). Beweis, dass jede volle Funktion gleichmässig stetig ist. *Verhandelingen der Koninklijke Nederlandsche Akademie van Wetenschappen te Amsterdam*, *24*, 189–193.

Brouwer, L. (1926). Zur Begrüdndung der intuitionistischen Mathematik III. *Mathematische Annalen*, *96*, 451–488.

Brouwer, L. (1927). Über Definitionsbereiche von Funktionen. *Mathematische Annalen*, *97*, 60–75. English translation in From Frege to Gödel, Cambridge MA, 1967, pp. 446–463.

Brouwer, L. (1954). Points and spaces. *Canadian Journal of Mathematics*, *6*, 1–17.

Dummett, M. (1977). *Elements of intuitionism*. Oxford: Clarendon Press.

Heyting, A. (1956). *Intuitionism: an introduction*. Amsterdam: North-Holland Publishing Company.

Kleene, S. C., & Vesley, R. E. (1965). *The Foundations of intuitionistic mathematics*. Amsterdam: North-Holland Publishing Company.

Troelstra, A. (1969). *Principles of intuitionism*. Berlin: Springer.

Chapter 2
Creative Subject and Bar Theorem

Abstract In the present article, a reasonably precise description of Brouwer's notion of "creative subject" is proposed and an axiom is introduced which is conceptually equivalent to the bar theorem.

2.1 The Creative Subject

The idea of the creative subject occurs in various writings of Brouwer in a somewhat vague manner. Several logicians, among them Kreisel (1967), Troelstra (1969) and Dummett (1977), have recently tried to analyse this concept by proposing some axioms, which are however rather controversial, as Dummett's discussion shows.

The creative subject, which we will call Σ, carries out his mathematical activity in time, which we assume divided in ω distinct states of knowledge. The main feature of the theory of the creative subject consists in using the fact that knowledge increases in time by explicitly referring to the stage in which Σ gets to know a certain proposition. Such a reference is expressed by the propositional operator $\vdash_n$. If A is a formula, $\vdash_n A$ stands for: "at stage n, Σ has evidence for A".

The kind of axioms for $\vdash_n$ we put forward essentially depends on our idealisation of Σ.

In the literature, two essentially different conceptions of $\vdash_n A$ were proposed: the *strict* one, according to which Σ can at any stage know only a finite number of propositions, and the wider one, according to which at any stage Σ is allowed to know infinitely many propositions.

The wider conception is proposed, among others things, by Troelstra (1969) as a possible solution of a paradox of diagonalization in the strict theory (see Sect. 2.3), which he pointed out.

We begin with some critical observations on the wider conception. The following axioms are usually assumed both in the strict theory and in the wider one:

(1.1) $\forall n(\vdash_n A \vee \neg \vdash_n A)$
(1.2) $\forall n \forall m(\vdash_n A \rightarrow \vdash_{n+m} A)$
(1.3) $\exists n \vdash_n A \leftrightarrow A$

These axioms are not sufficient to yield a precise conception of the creative subject, but they seem to be minimal requirements; they are implicitly used by Brouwer himself in the well-known constructions of counterexamples for some classical theorems.

(1.1) is usually justified in the following manner: the predicate $\vdash_n A$ is decidable since at the stage n Σ knows whether he has evidence for A or not. We will shortly discuss such a justification.

(1.2) says that the knowledge of Σ is cumulative: at every stage, he knows again (or he maintains the knowledge of) what he knew at the preceding stages.

(1.3), from left to right, says that Σ reasons in an intuitionistically correct manner, that is, he knows only intuitionistically true propositions. From right to left, it seems to express the "solipsistic" conception of intuitionistic mathematics: the true propositions are only those which are known to the creative subject. If one does not want to commit oneself to this last statement, (1.3) can be weakened by replacing it by:

$(1.3')$ $\exists n \vdash_n A \rightarrow A$

and

$(1.3'')$ $A \rightarrow \neg\neg\exists n \vdash_n A$

$(1.3'')$ (called by Kreisel the Principle of Christian Charity) says that if A is true, it cannot be excluded that, in some stage, Σ comes to know it.

In the wider conception it is allowed that, if $\vdash_n A$ and if B is a "immediate consequence" of A, then $\vdash_n B$. Apart from the difficulty of generally characterising in a sufficiently precise manner the vague notion of immediate consequence, it seems that the following axioms for it are hardly objectionable (see Dummett 1977). Let $P(x)$ be a predicate on natural numbers:

(1.4) $\vdash_n \forall x P(x) \rightarrow \forall x \vdash_n P(x)$
(1.5) $\vdash_n P(\overline{m}) \rightarrow \vdash_n \exists x P(x)$
(1.6) $\vdash_n (A \vee B) \rightarrow \vdash_n A \vee \vdash_n B$

In (1.6), it is understood that A and B do not contain choice parameters nor any information which is not available at stage n (see Sect. 2.2).

Now, we see at once that from (1.1)–(1.6) we can deduce some intuitionistically incorrect propositions.

In fact, let $P(x)$ be a decidable predicate (without choice parameters), such that $\forall x(P(x) \vee \neg P(x))$. From (1.3) it follows that, for some n,

$$\vdash_n \forall x(P(x) \vee \neg P(x))$$

whence, by (1.4) and (1.6),

$$\forall x(\vdash_n P(x) \vee \vdash_n \neg P(x)).$$

It follows that, if $\exists x\, P(x)$, then there exists a k such that $\vdash_n P(\bar{k})$ and so, by 1.5, $\vdash_n \exists x\, P(x)$. Therefore, the following equivalence holds

$$\exists x\, P(x) \;\leftrightarrow\; \vdash_n \exists x\, P(x).$$

Thus, since, by (1.1), $\vdash_n \exists x\, P(x) \vee \neg \vdash_n \exists x\, P(x)$, we have $\exists x\, P(x) \vee \neg \exists x\, P(x)$. But this cannot be intuitionistically acceptable for an arbitrary $P(x)$.

In my opinion this incoherence arises from the fact that (1.1) is acceptable only if, in order to have $\vdash_n A$, it is required that Σ is aware of having evidence of A. Now, the evidence of $\forall x\, P(x)$ does not involve (unless we greatly stress the idealisation of Σ) the conscious evidence of all single instances $P(\bar{k})$. Thus, (1.4) holds only if evidence is meant as implicit, not necessarily conscious, evidence. Hence, the incompatibility of (1.1) and (1.4) results.

Since Σ is no real subject but an idealised subject, we could idealise him so that the evidence of $\forall x\, P(x)$ involves the conscious evidence of each single $P(\bar{k})$. But such an idealization would attribute to Σ superhuman powers which would allow him to know some classically but not intuitionistically true propositions. Since intuitionistic truths are to be humanly knowable, we must be very cautious in idealising Σ.

Therefore, since (1.4) seems to be a minimal requirement of the wider conception, we hold that it forces the axiom (1.1) to be abandoned.

Then, at the same stage, Σ can implicitly know infinitely many propositions. If, for instance, at stage n he knows the soundness of the axioms and of the inference rules of a certain system F, then at the same stage n he has implicit evidence of all theorems of F (but he is not necessarily able to decide whether a given formula is a theorem, as (1.1) requires). Indeed, one could even assume that every stage is closed with respect to "analytic" deductions and that the passage from one stage to another is characterised by an increase of external information (see Posy 1977). But we think that in that case $\vdash_n A$ should be more properly be read: "At the stage n, Σ can have evidence of A". The potential interpretation serves the purpose of avoiding that superhuman powers (as that one of deducing actually all what is deducible from the available information) are attributed to Σ. The closure with respect to the analytic deductions seems to be in agreement with Grzegorczyk's semantics, (cf. Grzegorczyk 1964).

Nevertheless, it seems to me that such a conception has not the effect of taking the idea of the creative subject seriously, but rather of replacing it by other notions, such as "deducibility (in an intuitive sense) from certain given information", (Van Dalen, 1978). On the contrary, we believe that the originality and the strength of the theory of the creative subject consist just in the possibility of exploiting the concept of conscious evidence, which Brouwer has already partially exploited by his implicit assumption of (1.1). We hold that new interesting results are obtainable not just by suppressing (1.1) but rather by adding to it new axioms intended to exploit the concept more deeply.

Therefore, we will abandon the wider conception in favour of the strict one, according to which the request of the conscious evidence is certainly plausible. This request can be made even more explicit by stating, for instance, that in order to

have $\vdash_n A$, Σ has to write a proof of A at the stage n. We here obviously pass over the inadequacy of natural language to express an intuitive proof; our statement serves only the purpose of clarifying metaphorically the meaning of the adjective "conscious". What is of importance is that, proving only finitely many propositions at each stage, Σ has the possibility of explicitly directing his attention at each one of them.

The axiom (1.2) results to be obvious. Without committing ourselves to the solipsistic conception, we will accept (1.3) with the following justification: since A is intuitionistically true, if it has been proved by someone, we suppose that, as soon as a proof of A is found (by someone), it is communicated to Σ who includes it among the propositions of which he has evidence (at the stage he is in at the moment of communication).

We now intend to propose a modification of the strict conception in order to obviate, at least in part, the drawback of the invalidity of (1.4) and of other axioms which seem to hold only in the large conception.

We replace (1.4) with the scheme

(1.4′) $\forall n(\vdash_n \forall x\, P(x) \rightarrow \vdash_n P(\overline{k}))$.

If (1.4′) is understood in the sense that all its instances should be simultaneously true, then it is certainly not acceptable in the strict conception, since it presents the same difficulties as (1.4). Therefore, we propose to understand the validity of an axiom scheme in the following sense: for each instance, it is possible "to programme Σ" at the stage 0 so that the instance in question is true. By programming Σ we mean "to instruct Σ so that if he happens to be in certain favourable conditions, he performs certain deductions which interest us". In this sense, (1.4′) is thoroughly acceptable: whenever we have fixed P and k, we can instruct Σ so that, if he deduces $\forall x\, P(x)$ at some stage, he takes care of deducing $P(\overline{k})$ at the same stage.

Similarly for (1.5), whenever P is fixed, Σ can be instructed to deduce $\exists x\, P(x)$ as soon as he has deduced $P(\overline{m})$ for some m. Likewise for (1.6) (with the necessary restrictions on the choice parameters).

In general, we admit the following principle:

(1.7) Whenever, for a fixed proposition A, we recognise that if we were in a certain state of knowledge s we could prove A, we can suppose that, whenever Σ is in the state s, he actually deduces A.

Some other axioms which turn out to be sound are the following:

(1.8) $\forall n(\vdash_n (A \rightarrow B) \rightarrow (\vdash_n A \rightarrow \vdash_n B)$
(1.9) $\forall n(\vdash_n (A \wedge B) \rightarrow \vdash_n A \wedge \vdash_n B)$

On the contrary, we observe that

(1.8′) $(A \rightarrow B) \rightarrow \forall n(\vdash_n A \rightarrow \vdash_n B)$

does not hold, since until we know a proof of $A \rightarrow B$ (or of $\neg(A \rightarrow B)$) we cannot instruct Σ to deduce B from A.

But from (1.8) and (1.3), we get:

$(1.8'')$ $(A \to B) \to \exists n(\vdash_n A \to \vdash_n B)$.

We observe that a formula $A(x)$ with a free variable x must be considered true if it is possible to programme Σ so that all instances $A(\overline{k})$ are simultaneously true. For example, the formula $\vdash_n x = x$ is not true: in fact, we deduce from it $\forall x \vdash_n x = x$ which would involve the evidence, at the stage n, of the infinitely many propositions $k = k$. On the other hand, the formula $\forall x \vdash_x x = x$ is true: Σ can be instructed so that, at every stage n, he knows $n = n$.

Finally, we observe that, according to our concept of programming, Σ can be given only positive instructions, i.e. instructions for performing (under favourable conditions) certain deductions, and not negative instructions, i.e. instructions for not performing certain deductions. At every stage, Σ must be free to perform at will some other deductions, in addition to the ones which are imposed on him. This guarantees that two programmings are always compatible. Therefore, if A and B are instances of two sound axiom schemata, there is a programme which makes them simultaneously true. It follows that since only a finite number of formulas occur in a formal deduction, every formula which is deduced by use of some sound axiom schemata is true with respect to a suitable programming. Thus, the usual axioms and inference rules are in accordance with our concept of soundness of a scheme. This would not be the case if we would admit negative instructions. Then, for instance, both formulas

$$\forall x(\vdash_x x = x) \qquad \text{and} \qquad \forall x(\neg \vdash_x x = x)$$

would be sound, since Σ could be instructed both to prove $n = n$ at stage n for every n and not to prove $n = n$ at stage n for all n. Hence, syntactic inconsistency would follow.

2.2 The Creative Subject and Existential Statements

The axiom

$(2.1')$ $\forall n(\vdash_n \exists m\, P(\overline{m}) \to \exists m \vdash_n P(\overline{m}))$

is usually accepted with the following justification: an intuitionistic proof of $\exists m\, P(n)$ is a proof of some instance $P\overline{k}$ of $P(x)$.

But we observe that it is possible to have evidence for $\exists m\, P(\overline{m})$ even if a proof of an instance of $P(x)$ is not available, but only a method for constructing such a proof is available. Therefore if, at the stage n, Σ has evidence for $\exists m\, P(\overline{m})$, it is the case that at stage n Σ has a procedure π for determining an instance of $P(x)$. The execution of π must be feasible in a finite number of stages but not necessarily at the same stage n, since the execution can require some information which can be

obtained only at future stages. Thus, what holds in general is not the axiom (2.1′) but the weaker axiom

(2.1) $\forall n(\vdash_n \exists m\, P(\overline{m}) \to \exists p \exists m \vdash_p P(\overline{m}))$.

For instance, let ρ be a free sequence such that the value $\rho(n)$ of ρ has been chosen (and known to Σ) at stage n. Then, if $m > n$

$$\vdash_n \exists x \rho(\overline{m}) = x \qquad \text{but not} \qquad \exists x \vdash_n \rho(\overline{m}) = x.$$

Observe that in this example, even though Σ cannot determine an instance of $P(x)$, at stage n, yet, at stage n, he knows at which stage (at the stage m) the execution of the procedure π will be completed.

But one sees at once that neither is this the general case, i.e. not always

$$\vdash_n \exists x\, P(x) \to \exists y(\vdash_n \exists x \vdash_y P(x)).$$

For example, if $P(x)$ is $\rho(\rho(\overline{m})) = x$ (and $n < m$), at the stage n Σ knows only at which stage (at the stage m) he can determine the stage (the maximum of m and $\rho(\overline{m})$) at which π will be completed.

In general, what seems reasonable to admit as implicitly involved in the evidence (of the fact) that the procedure π must terminate in a finite number of stages is that if $\vdash_n \exists x\, P(x)$, Σ should at least be able to point out, at the stage n, a stage at which he can obtain an important piece of information concerning π, i.e. a stage at which π is reduced to a procedure π', more elementary with respect to the complexity of the required information.

For a precise statement of this idea, let us consider, for every sentence $\exists x\, P(x)$, the species S of the stages at which "Σ has inductive evidence of $\exists x\, P(x)$". We define S inductively:

(i) if for some $m \vdash_n P(\overline{m})$, then $n \in S$;
(ii) if for some $m \vdash_n m \in S$, then $n \in S$.

Then the "principle of the inductive evidence" which we intend to propose can be stated in the following way:

(PIE) For every sentence of the type $\exists x\, P(x)$, $S_{\exists x P(x)} = \{n : \vdash_n \exists x\, P(x)\}$.

Therefore (PIE) says: "If Σ has, at the stage n, evidence of' $\exists x\, P(x)$, then he has inductive evidence of $\exists x\, P(x)$".

In particular, the above considerations can be extended to the axiom

$$\vdash_n (A \vee B) \to \exists m(\vdash_m A \vee \vdash_m B).$$

2.3 Equivalence of (PIE) and (BI_M)

Let α, β be variables for choice sequences, $\vec{u}, \vec{v}$ variables for finite sequences (nodes of the universal tree). The length of a sequence $\vec{u} = \langle u_0, \ldots, u_n \rangle$ is $\ell(\vec{u} = n+1$, where the length of the empty sequence is $\ell(\langle\rangle) = 0$. Initial segments of choice sequences are denoted by $\overline{\alpha}(n) = \langle \alpha(0), \ldots, a(n-1) \rangle$. We write $\alpha \in \vec{u}$ for $\overline{\alpha}(\ell(\vec{u})) = \vec{u}$. For a species of nodes R, we define inductively the $\mathcal{F}$-closure $R^{\mathcal{F}}$ of R by:

(a) $\forall u(\vec{u} \in R \to \vec{u} \in R^{\mathcal{F}})$
(b) $\forall u(\forall k \vec{u} * \langle k \rangle \in R^{\mathcal{F}} \to \vec{u} \in R^{\mathcal{F}})$

A species of nodes R will be called monotone if $\forall \vec{u} \forall \vec{v}(\vec{u} \in R \to \vec{u} * \vec{v} \in R$. We now state the monotone bar theorem in the form:

(BI_M) If R is a monotone species of nodes and $\forall \alpha \exists x \overline{\alpha}(x) \in R$, then $\langle\rangle \in R$.

Theorem 2.1 *(BI_M) $\leftrightarrow$ (PIE)*

Proof Let us assume (BI_M). For the sake of simplicity, we consider stage 0 and suppose that

$$\vdash_0 \exists x P(x)$$

We assume that it is possible to represent the information, on which the procedure π for determining an instance of $P(x)$ rests, by means of a sequence σ such that, for each n, $\overline{\sigma}(n)$ is a codification of the amount of information known to Σ at the stage n and, for every k, $\overline{\sigma}(n) * \langle k \rangle$ is a value which is a priori possible (i.e. at the stage n) for $\overline{\sigma}(n+1)$.

Let R be the species of the pieces of information $\vec{u}$ sufficient for the execution of π. Then

$$\forall \alpha \exists x \overline{\alpha}(x) \in R.$$

By induction on the construction of $R^{\mathcal{F}}$, let us prove that for each $\vec{u} \in R^{\mathcal{F}}$ one can programme Σ so that if $\vec{u} \in R^{\mathcal{F}}$, then $\ell(\vec{u}) \in S$:

(1) if $\vec{u} \in R$, then we instruct Σ so that if $\vec{u} \in \sigma$, at the stage $\ell(\vec{u})$ he performs π and gets the required instance of $P(x)$, hence $\ell(\vec{u}) \in S$ by (i);
(2) if $\forall k(\vec{u} * \langle k \rangle \in R^{\mathcal{F}})$ by the induction hypothesis one can instruct Σ so that if $\vec{u} * \langle k \rangle \in \sigma$, then $\ell(\vec{u}) + 1 \in S$. Then we can instruct Σ so that if $\vec{u} \in \sigma$ also $\ell(\vec{u}) + 1 \in S$ and Σ takes note of it at the stage $\ell(\vec{u})$, i.e. $\vdash_{\ell(\vec{u})} \ell(\vec{u}) + 1 \in S$, whence $\ell(\vec{u}) \in S$ by (ii).

Since R is obviously monotone, $\langle\rangle \in R$ by BI_M and so $0 \in S$.

Conversely let us assume (PIE). Let R be a monotone species such that $\forall \alpha \exists x \overline{\alpha}(x) \in R$. We programme Σ as follows: after having introduced a free sequence ρ with the property that $\rho(n)$ is chosen at the stage $n+l$, we get $\vdash_0 \exists x \overline{\rho}(x) \in R$ and we put S to be the species of the stages in which Σ has inductive evidence of $\exists x \overline{\rho}(x) \in R$. Let us prove, by induction on the construction of S, that $S \subseteq \{n : \overline{\rho}(n) \in R\}$.

(i) If for some $m \vdash_n \overline{\rho}(\overline{m}) \in R$, then if $n \geq m$, $\overline{\rho}(n) \in R$ because of the monotonicity of R; if $n < m$ then, since at the stage n Σ knows only the initial segment $\overline{\rho}(n)$ of ρ, it must hold that $\forall\alpha(\overline{\rho}(n) * \overline{\alpha}(m - n) \in R^{\mathcal{F}})$, whence, by induction on $m - n$, $\overline{\rho}(n) \in R^{\mathcal{F}}$.

(ii) If, for some m, $\vdash_n m \in S$, then, by the induction hypothesis, $\overline{\rho}(\overline{m}) \in R^{\mathcal{F}}$ and, as in (i), we recognise that $\overline{\rho}(n) \in R^{\mathcal{F}}$.

By (PIE) $0 \in S$ and so $\langle\rangle \in R^{\mathcal{F}}$.

References

Brouwer, L. (1975). Collected works. In A. Heyting (Ed.), *Philosophy and foundations of mathematics*. Amsterdam: Elsevier.

Dummett, M. (1977). *Elements of intuitionism*. Oxford: Clarendon Press.

Grzegorczyk, A. (1964). A philosophicalty plausibte formal interpretation of intuitionistic logic. *Indagationes Mathematicae*, *26*, 596–601.

Kreisel, G. (1967). Informal rigour and completeness proofs. In I. Lakatos (Ed.), *Problems in the philosophy of mathematics* (pp. 138–186). Amsterdam: North-Holland Publishing Company.

Posy, C. (1977). The theory of enrperical sequences. *Journal of Philosophical Logic*, *6*, 47–81.

Troelstra, A. (1969). *Principles of intuitionism*. Berlin: Springer.

Van Dalen, D. (1978). An interpretation of intuitionistic analysis. *Annals of Mathematical Logic*, *13*, 1–43.

Chapter 3
Natural Intuitionistic Semantics and Generalized Beth Semantics

Abstract In this chapter, the connection between the notion of truth in a generalized Beth model and the intuitive notion of truth according to the intuitionistic meaning of logical constants is analysed.

3.1 Introduction

Beth and Kripke semantics for intuitionistic predicate calculus (IPC) have been generalised by Veldman and de Swart in such a way to obtain completeness proofs within an intuitionistic metamathematics (de Swart 1976; Veldman 1976). The aim of this paper is to establish a connection between the notion of truth in a generalized Beth model (GB-model) and the intuitive notion of truth according to the intuitionistic meaning of logical symbols. It follows that "true in a GB-model" means "intuitively true under certain hypotheses", in a suitable understanding of such a locution.

3.2 Generalized Beth-Models and Natural Models

Include $\perp$ (absurdity) among the atomic closed formulas of IPC and define negation by putting $\neg A = A \rightarrow \perp$. The whole exposition is to be understood within an intuitionistic metamathematics.

A GB-model $M =< T, D, \Phi >$ consists of a spread T (tree with infinite paths), of a species $D \neq \varnothing$ and a binary relation Φ between T-nodes (indicated with $\vec{u}, \vec{v}, \ldots$) and atomic closed formulas of the IPC-language extended with individual constants for the members of D. We say that M *explodes* if there exists $\vec{u}$ such that $\Phi(\vec{u}, \perp)$. Extend Φ to the *forcing* relation $\vec{u} \Vdash A$ between nodes and closed formulas with the following inductive definition:

(i) for P atomic, $\vec{u} \Vdash P$ iff there is a bar $\mathcal{B}$ of $\vec{u}$ such that, for all $\vec{v} \in \mathcal{B}$, either M explodes or $\Phi(\vec{v}, P)$;
(ii) $\vec{u} \Vdash A \rightarrow B$ iff, for all $\vec{v} \geq \vec{u}$, if $\vec{v} \Vdash A$ then $\vec{v} \Vdash B$;
(iii) $\vec{u} \Vdash \forall x A(x)$ iff, for all $d \in D$, $\vec{u} \Vdash A(d)$ (similarly for $A \wedge B$);

(iv) $\vec{u} \Vdash \exists x A(x)$ iff there is a bar $\mathcal{B}$ of $\vec{u}$ such that, for all $\vec{v} \in \mathcal{B}, \vec{v} \Vdash A(d)$, for some $d \in D$ (similarly for $A \vee B$).

A is *true at* M ($M \Vdash A$) if $<>\Vdash A$. A is *GB-valid* if it is true in every GB-model. A *Beth model* (B-model)is a non-exploding GB-model.

A *natural model* (N-model) $\mathcal{M} =< D, V >$ consists of a species $D \neq \varnothing$ (domain of individuals) and a monadic relation V, defined on the closed atomic formulas of the extended language (as above), such that not $V(\bot)$ ($\bot$ is to be interpreted in a false proposition!). V is then extended to a relation $\mathcal{M} \Vdash A$ on all closed formulas, defined according to the intended meaning of logical constants.

The theory of free choice sequences yields a well-known relation between B-models and N-models (Troelstra 1977, p. 118). Precisely, if $M =< T, D, \Phi >$ is a B-model and α is a free choice sequence in T, one can associate to this an N-model $M_\alpha =< D, V >$ by defining V as follows: $V(P) \Leftrightarrow \exists n \Phi(\bar{\alpha} n, P)$ (where $\bar{\alpha} n = \langle \alpha_0, \alpha_1, \ldots, \alpha_{(n-1)} \rangle$).

Theorem 3.1 *Let M be a B-model and $\vec{u}$ a node. $\vec{u} \Vdash A$ iff, for all $\alpha \in \vec{u}$, $M_\alpha \Vdash A$.*

It easily follows that a formula is B-valid iff it is N-valid. Now, while GB-completeness is provable with the usual methods of intuitionistic analysis (Troelstra 1977, p. 177[3]), B-completeness (N-completeness) turns out to be intuitionistically equivalent to a controversial instance of Markov principle (Kreisel 1962). Excluding the latter from the available intuitionistic methods of proof, we can say that GB-validity fails to be intuitionistically equivalent to N-validity.

The problem arises to extend N-semantics to a GN-semantics corresponding to GB-semantics, according to a suitable counterpart of Theorem 3.1.

3.3 Generalized Natural Models

Call *sequence of hypotheses* a sequence $\langle Q_n \rangle_{n \in \mathbb{N}}$ of meaningful propositions (not mere syntactical formulas).

A *GN-model* $\mathcal{M} = \langle D, V, \langle Q_n \rangle_{n \in \mathbb{N}} \rangle$ consists (besides D and V, as above) of a sequence of hypotheses. The intention is that a formula A is true at $\mathcal{M}$ if it is true at the N-model $\langle D, V \rangle$ "under the hypotheses Q_n".This idea is made precise by means of the following inductive definition of the relation $\mathcal{M} \Vdash A$ (where $\mathcal{M}$ will be understood):

(i) if A is atomic and $V(A)$, then $\models A$;
(ii) if $\models A$ and $\models B$, then $\models A \wedge B$;
(iii) if $A \models$ or $\models B$, then $\models A \vee B$;
(iv) if $\models A$ implies $\models B$, then $\models A \rightarrow B$;
(v) if, for all $d \in D$, $\models A(d)$, then $\models \forall x A(x)$;

(vi) if, for some $d \in D$, $\models A(d)$, then $\models \exists x A(x)$;
(vii) if, for some $n \in \mathbb{N}$, Q_n implies $\models A$, then $\models A$;
(viii) $\models A$ only in virtue of (i) – (vii).

An N-model may be regarded as a GN-model at which all hypotheses are true.

For GN-semantics, soundness holds straightforwardly:

Theorem 3.2 *If $\vdash_{IPC} A$, then A is GN-valid.*

As to the connection with GB-semantics, let $M = \langle T, D, \Phi \rangle$ be a GB-model. Assuming Kripke scheme

$$\exists \chi (\exists x \chi x \neq 0 \leftrightarrow A)$$

(where χ is a choice sequence), one can find a sequence χ such that

(*) $\exists x \chi(x) \neq 0$ iff M explodes

Let us associate to any free choice sequence α in T the GN-model $M_\alpha = \langle D, V, < Q_n \rangle_{n \in \mathbb{N}} \rangle$ where Q_n is the proposition $\chi(n) = 0$ and V is defined by putting $V(P) \Leftrightarrow \exists n \Phi(\bar{\alpha} n, P)$ with $P \neq \bot$ ($V(\bot)$ is false at all GN-models).

One can prove the analogous of Theorem 3.1:

Theorem 3.3 *Let M be a GB-model and $\vec{u}$ a node. $\vec{u} \Vdash A$ iff, for all $\alpha \in \vec{u}$, $M_\alpha \models A$.*

It follows that every GN-valid formula is GB-valid. Therefore, GN-completeness follows from GB-completeness:

Theorem 3.4 *If A is GN-valid, then $\vdash_{IPC} A$.*

From GB-completeness and Theorem 3.2 it follows that, if A is GB-valid, then it is GN-valid. So

Theorem 3.5 *A is GB-valid iff it is GN-valid.*

Observe that Theorem 3.3 has been proved by using Kripke schema, a controversial principle involving the problematic theory of the creative subject. However, the proof of GB-completeness uses only GB-models with Φ decidable. For such models, a χ satisfying (*) can be defined without Kripke schema: namely, one can take an enumeration $\langle \vec{u}_n \rangle_{n \in \mathbb{N}}$ of nodes and put

$$\chi(n) = \begin{cases} 1 \text{ if } \Phi(\vec{u}_n, \bot) \\ 0 \text{ otherwise.} \end{cases}$$

Therefore, the Theorems 3.4 and 3.5 are provable with usual intuitionistic methods.

References

de Swart, H. (1976). Another intuitionistic completeness proof. *Journal of Symbolic Logic*, *41*, 644–662.

Kreisel, G. (1962). Weak completeness of intuitionistic predicate logic. *Journal of Symbolic Logic*, *27*, 139–158.

Troelstra, A. (1977). *Choice sequences*. Oxford: Oxford University Press.

Veldman, W. (1976). An intuitionistic completeness theorem for intuitionistic predicate logic. *Journal of Symbolic Logic*, *41*, 159–176.

Chapter 4
Connection Between the Principle of Inductive Evidence and the Bar Theorem

Abstract I introduced the "principle of inductive evidence" PIE in my paper "Creative subject and bar theorem" (Martino 1982). Because of a misunderstanding in my correspondence with the editors, the published version of the above paper is not the final revised draft, but a first outline of the article which needs some corrections and explications. I shall refer to the published version as CS. In CS, I asserted somewhat rashly the absolute equivalence of PIE and the monotonic bar theorem BI_M by means of all too sketchy proof in the course of which I introduced in passing a rather problematic assumption without explaining it properly. Therefore, I shall present here a more adequate treatment of the connection between PIE and BI_M. In fact, I shall assume acquaintance with Sects. 4.1 and Theorem 4.1 of CS and provide a revised version of Theorem 4.2.

4.1 Inductive Evidence

With respect to the activity of the creative subject Σ and an existential statement $\exists x\, P(x)$, we inductively define the species S of the stages of "inductive evidence" of $\exists x\, P(x)$:

(i) if, for some m, $\vdash_n P(m)$, then $n \in S$;
(ii) if, for some m, $\vdash_n m \in S$, then $n \in S$.

$$PIE : S = \{n \mid \vdash_n \exists x P(x)\}.$$

(In words: if Σ has evidence of $\exists x\, P(x)$, then he has *inductive* evidence of $\exists x\, P(x)$).

Let $\alpha, \beta, \ldots$ be variables for choice sequences, $\mathbf{u}, \mathbf{v}, \ldots$ variables for finite sequences (nodes of the universal tree). The length of a sequence

$$\mathbf{u} = \langle u_0, \ldots, u_n \rangle \text{ is } l(\mathbf{u}) = n + 1,$$

and the length of the empty sequence is $l(\langle\rangle) = 0$. Initial segments of choice sequences are denoted by $\overline{\alpha}(n) = \langle \alpha(0), \ldots, \alpha(n-1) \rangle$. We write $\alpha \in \mathbf{u}$ for $\overline{\alpha}(l(\mathbf{u})) = \mathbf{u}$.

For a species R of nodes, we inductively define the F-closure R^F of R by:

(i) $\forall \mathbf{u}(\mathbf{u} \in R \rightarrow \mathbf{u} \in R^F)$
(ii) $\forall \mathbf{u}(\forall k \mathbf{u} * \langle k \rangle \in R^F \rightarrow \mathbf{u} R^F)$.

R will be called monotonic if $\forall \mathbf{u} \forall \mathbf{v}(\mathbf{u} \in R \rightarrow \mathbf{u} * \mathbf{v} \in R)$. R is a bar of $\mathbf{u}$ if $\forall \alpha_{\alpha \in \mathbf{u}} \exists n \overline{\alpha}(n) \in R$

We state the monotonic bar theorem in the form

$$BI_M : \text{ if } R \text{ is a monotonic } bar\, of \langle\rangle, then \langle\rangle \in R^F.$$

Theorem 4.1 $PIE \rightarrow BI_M$

Proof Let R be a monotonic bar of $\langle\rangle$. We introduce a creative subject Σ and a lawless sequence ρ whose value $\rho(n)$ is chosen at the stage $n+1$. So Σ knows $\rho(n)$ at stage $n+1$ and cannot know it before.

By particularizing the available proof of $\forall \alpha \exists n \overline{\alpha}(n) \in R$, we get a proof of $\exists n \overline{\rho}(n) \in R$ which does not use any information on the values of ρ. Let us communicate such a proof to Σ and let him insert $\exists n \overline{\rho}(n) \in R$ among his theorems at stage 0. So $\vdash_0 \exists n \overline{\rho}(n) \in R$. Call S the species of stages of inductive evidence of $\exists n \overline{\rho}(n) \in R$.

By PIE, we have $0 \in S$. According to the inductive definition of S, a proof of $0 \in S$ is built up from the following elementary inferences:

$$\text{(a)} \frac{\exists m \vdash_n \overline{\rho}(m) \in R}{n \in S}$$

$$\text{(b)} \frac{\exists m \vdash_n m \in S}{n \in S}$$

Replacing every statement of the form $x \in S$ by $\overline{\rho}(x) \in R^F$, we obtain the inferences

$$\text{(a}') \frac{\exists m \vdash_n \overline{\rho}(m) \in R}{\overline{\rho}(n) \in R^F}$$

$$\text{(b}') \frac{\exists m \vdash_n \overline{\rho}(m) \in R^F}{\overline{\rho}(n) \in R^F}$$

We prove the correctness of (a′), (b′).

Since $R \subseteq R^F$, we may restrict ourselves to (b′). Suppose that, for some m, $\vdash_n \overline{\rho}(m) \in R^F$.

If $n \geq m$, observe that, since R is monotonic, R^F is monotonic too (by induction on the construction of R^F), whence $\overline{\rho}(n) \in R^F$.

If $n < m$, since at the stage n Σ knows only the initial segment $\overline{\rho}(n)$ of ρ, we get $\forall \alpha(\overline{\rho}(n) * \overline{\alpha}(m-n) \in R^F)$, whence, by induction on $m-n$, $\overline{\rho}(n) \in R^F$.

Thus, we can transform the available proof of $0 \in S$ into a proof of $\langle\rangle \in R^F$.

Remark. In the above proof, we have used the following principle: if $\vdash_n P(\overline{\rho}m)$ with $m > n$, then $\vdash_n \forall\alpha P(\overline{\rho}(n) * \alpha(m-n))$, where $\lambda\mathbf{u}P(\mathbf{u})$ is a property of nodes. This principle is evident provided the definition of P does not involve ρ (e.g. it does not hold for $P : u \in \rho$). In fact, we have tacitly assumed that, given the bar R, we can find a ρ not involved by the definition of R.

The key of the above proof consists in transforming, by means of the lawless sequence ρ, the bar statement into an existential statement. It does not seem quite evident that, conversely, one could always transform an existential statement into a bar statement. So, since PIE concerns all existential statements, it seems stronger than BI_M. In fact, I think that PIE cannot be logically reduced to other known principles. Perhaps PIE provides for existential statements a *strongly* constructive meaning which, though in accordance with the intuitionistic use of $\exists$, is not implicit in the usual intuitionistic *understanding* of $\exists$.

To bring out how BI_M helps to assert PIE, we introduce a further epistemic principle (perhaps more elementary than PIE) under which the implication $BI_M \rightarrow PIE$ holds.

Suppose that Σ, at the stage 0, has evidence of $\exists x P(x)$. Then he knows a procedure π to calculate an instance of $P(x)$. Though, at the stage 0, such an instance, because of a possible lack of information, may not yet be determined (see *CS* Sect. 2), nevertheless Σ knows a priori that certainly, at some future stage, the instance will be determined, no matter how the information proceeds. It seems to me quite reasonable to interpret such a state of knowledge as a bar assertion according to the following "principle of spreadlike information":

PSI: Any proof, at the stage 0, of $\exists x P(x)$ consists in

(a) arranging all a priori possible pieces of information (relevant for π) in a *lawlike* spread s;
(b) recognizing that s is barred by the species R of those pieces of information which are sufficient to execute π.

More explicitly, the above s represents the whole a priori possible information in the following sense:

(i) $\langle\rangle$ represents the piece available at the stage 0;
(ii) if $\mathbf{u} \in s$ represents the piece at the stage n, then the successors $\mathbf{u} * \langle k\rangle$ of $\mathbf{u}$ in s represent those pieces which, at the stage n, Σ knows as possible for the stage $n+1$.

Theorem 4.2 *$PSI \wedge BI_M \rightarrow PIE$.*

Proof According to our strict conception of creative subject (see CS, Sect. 2), we intend the above assertion in the following sense: assuming *PSI* and BI_M, we can instruct Σ so that PIE holds.

Suppose $\vdash_n \exists x P(x)$ and let s and R be as in PSI. Let $\sigma \in s$ be the "sequence of information". This means that, for all n, $\overline{\sigma}(n)$ is the real piece of information

available at the stage n. For the sake of simplicity, we assume s to be the universal spread. To every $\mathbf{u} \in R^F$, we assign an instruction $I_{\mathbf{u}}$. This will be executed by Σ at the stage $l(\mathbf{u})$ provided $\mathbf{u} \in \sigma$ and will have the effect that $l(\mathbf{u}) \in S$.

We define $I_{\mathbf{u}}$, by induction on the construction of R^F:

(i) $\mathbf{u} \in R$. $I_{\mathbf{u}}$: if $\mathbf{u} \in \sigma$, at the stage $l(\mathbf{u})$:

(a) carry out the procedure π by means of $\mathbf{u}$ and determine an instance $P(m)$ of $P(x)$;
(b) insert $P(m)$ among your theorems.

Thus, if $\mathbf{u} \in \sigma$, then $\vdash_{l(\mathbf{u})} P(m)$, whence $l(\mathbf{u}) \in S$.

(ii) $\mathbf{u} \in R^F$ and all $I_{\mathbf{u}*\langle k\rangle}$ are already defined.

$$I_{\mathbf{u}} : \quad \text{if } \mathbf{u} \in \sigma \text{ at the stage } l(\mathbf{u})$$

(a) observe that for some k (knowable at the stage $l(\mathbf{u}) + 1$) $\mathbf{u} * \langle k\rangle \in \sigma$ so that, in virtue of $I_{\mathbf{u}*\langle k\rangle}$, $l(\mathbf{u}) + 1 \in S$.
(b) Insert $l(\mathbf{u}) + 1 \in S$ among your theorems.

Thus $\vdash_{l(\mathbf{u})} l(\mathbf{u}) + 1 \in S$ whence $l(\mathbf{u}) \in S$.
Since, by BI_M, $\langle\rangle \in R^F$, the instruction $I_{\langle\rangle}$ makes the theorem true.

Reference

Martino, E. (1982). Creative subject and bar theorem. In D. V. Dalen & A. Troelstra (Eds.), *The L. E. J. Brouwer Centenary Symposium* (pp. 311–318). North Holland: Amsterdam. Reprinted here as chapter 2.

Chapter 5
On the Brouwerian Concept of Negative Continuity

Abstract In the present paper, we will discuss some of the known reconstructions of Brouwer's theorem of negative continuity and contend that the theory of the creative subject is the proper frame in which to understand Brouwer's argument. We will also point out a sense in which negative continuity is a cogent consequence of the general intuitionistic tenets while positive continuity is not.

5.1 Introduction

In a famous paper of 1927, Brouwer introduces the notion of negative continuity for real functions. At the end of Sect. 1, he proves the negative continuity theorem:

1.1 Every full function is negatively continuous (where a function is said to be full if its domain is the unit continuum [0, 1]).

In Sect. 3 of the same paper, Brouwer proves a much stronger result, the well-known uniform continuity theorem:

1.2 Every full function is uniformly continuous.

As a trivial consequence of this theorem, we get:

1.3 Every full function is (positively) continuous.

So at first sight Theorem 1.1 seems to be superseded by Theorem 1.3; in fact 1.1 does not occur in modern intuitionistic analysis.

However, we contend that Brouwer's direct proof of 1.1 has a remarkable foundational significance. Brouwer himself observes that 1.1 is an immediate consequence of the intuitionistic point of view. This is not the case for 1.2. Brouwer says he had knowledge of 1.1 since 1918 and though this result suggested to him the conjecture of 1.2, nevertheless he did not succeed in proving 1.2 until much later.

Yet, in my opinion, Brouwer's proof of 1.1 has not been fully understood, possibly because of its rather informal and elliptical exposition. Recently, several authors tried to reconstruct Brouwer's argument: see (Parsons 1967; Posy 1976; Troelstra 1982; Veldman 1982). Still, these reconstructions—though in some ways interesting—fail to shed light on the sense of the immediacy of 1.1 claimed by Brouwer. For instance,

this immediacy is completely obscured by Veldman. In fact his reconstruction rests on a continuity principle for choice sequences from which 1.3 itself follows, so that 1.1 and 1.3 are placed on the same level of evidence. Furthermore, Veldman explicitly opposes the contention that 1.1 is a more elementary result than 1.3.

Here we discuss some of such reconstructions of Brouwer's argument.

5.2 The Negative Continuity Theorem

To begin, we give a reformulation of the negative continuity theorem for choice sequences. From this Brouwer's theorem for real functions follows straightforwardly.

Let $\alpha, \beta, \ldots$ range over choice sequences, $m, n, p, \ldots$ over natural numbers. Let f be a function from choice sequences to natural numbers. We indicate by $D(f)$ the domain of f (a subset of the universe U of all choice sequences).

f is *extensional* if, for all $\alpha, \beta \in D(f)$, $\alpha = \beta \to f(\alpha) = f(\beta)$, where the "$=$" occurring between choice sequences means *extensional equality* ($\alpha = \beta \overset{\text{def}}{\leftrightarrow} \forall n \alpha(n) = \beta(n)$). From now on by "function" we always understand an extensional function from choice sequences to natural numbers.

f is *full* if it is defined on all choice sequences: $D(f) = U$. $\alpha \in D(f)$ is a *continuity point* for f if there is an n such that, for all $\beta \in D(f)$, $\overline{\alpha}n = \overline{\beta}n \to f(\alpha) = f(\beta)$ (where $\overline{\alpha}n = \langle \alpha(0), \ldots, \alpha(n-1) \rangle$). f is *continuous* if every $\alpha \in D(f)$ is a continuity point.

As observed by Veldman, Theorem 1.3 for real functions is derivable from the corresponding continuity principle for choice sequences

CP: Every full function is continuous.

Similarly, Theorem 1.1 is derivable from a "negative continuity principle" NCP for choice sequences which we are going to state.[1]

$\alpha \in D(f)$ is a *positive discontinuity point* (p.d.p.) for f if there is a sequence $\langle \beta_n \rangle_{n \in \mathbb{N}}$ of choice sequences $\in D(f)$ such that, for all n, $\overline{\beta}_n(n) = \overline{\alpha}(n)$ and $f(\beta_n) \neq f(\alpha)$. f is *negatively continuous* if it has no p.d.p.

NCP: Every full function is negatively continuous.

[1] To derive 1.1 from *NCP*, as well as 1.3 from *CP*, one has to assume that any rational approximation f of a real function is extensional on number generators. Though this fact is often accepted as unproblematic (Veldman tacitly assumes nothing less than the extensionality of every function on choice sequences!), it is far from being evident. All that follows from the definition of real function is only that $\alpha = \beta \to f(\alpha) \approx f(\beta)$ where $\approx$ is the relation "m touches n" between rational segments (coded by natural numbers) (see Veldman 1982, p. 14). However, replacing extensionality with the latter condition, the proof of *NCP*, *mutatis mutandis*, holds again. So the above assumption of extensionality is not needed.

5.3 A Proof of NCP

We translate Brouwer's proof of the negative continuity theorem into a proof of NCP.

Let f be a full function. Suppose α is a p.d.p. for f and let $\langle\beta_n\rangle_n$, be as above. Define a choice sequence γ as follows. We temporarily choose, for every natural number n that we have already considered, $\gamma(n) = \alpha(n)$, but reserve the right to determine at any time after the first, second, …mth value has been chosen, the choice of all further values (that is of the $(m+1)$th, $(m+2)$th and so on) in such a way that either $\gamma = \alpha$ or $\gamma = \beta_n$, for some n. Then the function f is not defined on γ, which is absurd.

5.4 Weak and Strong Negation

The most controversial point of the above proof is the conclusion

4.1 The function f is not defined on γ.

The standard reading of intuitionistic negation confers to $\neg A$ the "strong sense": a proof of A leads to a contradiction.

As Posy points out, sometimes Brouwer also used negation in the "weak sense": we have no evidence for A.

The problem arises whether the negation occurring in 4.1 is the strong or the weak one and, in the latter case, whether Brouwer's argument counts as an outright mathematical proof or as a simple plausibility argument.

In his first reconstruction of Brouwer's proof, Posy interprets negation in the strong sense and concludes that Brouwer violates the "protochristian charity principle" according to which, if A is a true proposition, then its truth must be known at some stage. According to Posy, γ would be so defined that, as a matter of fact, $\gamma = \alpha$ is a true proposition, but its knowledge is forever prevented by the virtual eternal right of making $\gamma = \beta_n$, at some future stage. So "f is defined on γ" would be absurd because at no stage could one have sufficient information to calculate $f(\gamma)$.

This interpretation seems to me untenable. For, if we cannot use the information $\gamma = \alpha$ to calculate $f(\gamma) = f(\alpha)$, then we cannot use it to derive a contradiction from "f is defined on y" either. Besides, there is no way to derive such a contradiction. If there were one, Brouwers argument would establish the following assertion:

4.2 If a function f has a p.d.p., then there is a point where f is not defined.

Now, 4.2 does not hold, as the following counterexample shows. Let f be defined as follows:

$$f(\xi) = \begin{cases} 0 & \text{if } \xi = 0 \\ 1 & \text{if } \xi \neq 0 \end{cases}$$

so that $D(f) = \{\xi | \xi = 0 \vee \xi \neq 0\}$ and 0 is a p.d.p. Assuming 4.2 there is a $\gamma \notin D(f)$, so $\neg(\gamma = 0 \vee \gamma \neq 0)$ which contradicts the general schema $\neg\neg(A \vee \neg A)$.

We conclude that the negation occurring in 4.1 cannot be understood in the strong sense. On the other hand, taking it in the weak sense, 4.2 says only that if f has a p.d.p., then we can construct a γ for which we have no evidence whether f is defined on it. So 4.2 is not a mathematical theorem but an assertion of the same kind as Brouwer's weak counterexamples.

Veldman reaches the same conclusion for NCP: taking 4.1 in the weak sense, we must regard Brouwer's argument as a mere plausibility argument for NCP. That is, Brouwer's argument would not lead to a contradiction the assumption that a full function has a p.d.p., but it would only show that for such a function it is plausible that a p.d.p. could be found.

A similar conclusion is reached by Troelstra (1982, p. 479) when he discusses Heyting's third interpretation of Brouwer's argument.

On the contrary, in the next section we will contend that, though—as we showed—4.1 has to be read in the weak sense, nevertheless Brouwer's argument is an authentic proof by contradiction.

5.5 The Role of Time in Brouwer's Argument

To rightly understand Brouwer's proof, we need to consider very carefully the role of time in his argument.

In the course of the proof, Brouwer explicitly refers to time more than once ("we temporarily choose …", "at any time …").

The role of time in intuitionistic proofs is pinpointed by the theory of the creative subject. Therefore, this theory seems to me the proper setting for Brouwer's proof.

In the theory of the creative subject, we divide time into countably many stages of knowledge and introduce the usual epistemic operator $\vdash_n$: if A is a proposition, $\vdash_n A$ means that at stage n we have evidence for A. This device serves the purpose of getting weak negation precise: statements of the form "we have no evidence for A" implicitly refer to some stage of our knowledge; this is made explicit by the more complete statement $\neg \vdash_n A$.

Now, to interpret within this framework the weak negation occurring in 4.1, we have to pinpoint a stage at which we cannot have evidence that f is defined by γ. Well, we can get an outright contradiction by arranging the construction of γ so that the stage at issue is one at which a proof of the fullness of f is available. This is shown by our following reconstruction of Brouwer's proof.

Suppose we have at some stage, say at stage 0, a proof of

5.1 "f is full and α is a p.d.p. for f".

At the same stage 0, we introduce a choice sequence γ whose nth value has to be chosen at stage $n + 1$ according to the following instructions (given at stage 0):

(1) Stage 0: choose no values and impose no restrictions.
(2) Stage $n+1$:

(a) If γ has yet been restricted at some foregoing stage, calculate $\gamma(n)$ according to the imposed restrictions.
(b) If γ is as yet unrestricted, choose among the following alternatives:
(i) put $\gamma(n) = \alpha(n)$ and impose no restriction.
(ii) Impose the restriction $\gamma = \alpha$ and calculate $\gamma(n)$ accordingly.
(iii) Impose the restriction $\gamma = \beta_n$ and calculate $\gamma(n)$ accordingly.

The instructions above are so devised that at stage 0 we cannot prove that γ will be at some later stage restricted, since they leave the possibility of indefinitely opting for alternative (i). And since we cannot—by the extensionality of f—determine $f(\gamma)$ as long as γ is unrestricted, it follows that, at stage 0, we cannot prove that $f(\gamma)$ is determinable in a finite number of stages, i.e. that f is defined on γ. On the other hand, a proof that f is defined on γ is derivable, at stage 0, by particularising the available proof that f is full.[2] An absurdity arises.[3]

Our reconstruction points out how Brouwer's argument essentially exploits the hypothesis of fullness on f, and hence, why the same argument cannot be used for proving the false statement 4.2.[4]

5.6 Brouwer's Argument and Solipsism

We discuss here a delicate question concerning the theory of the creative subject. We argued that this is the proper frame for Brouwer's argument as a tool for exploiting the time element. But the theory at issue is also regarded in the literature as a tool for exploiting the so-called "solipsistic conception" of mathematics. In fact, (Troelstra 1982, p. 474) points out that the principal new element in creative subject arguments is not so much the time element itself but the solipsistic exploitation, i.e. the use as a

[2]Perhaps it is worth noticing that a proof, at stage 0, that f is defined on γ does not necessarily provide, at the same stage 0, the value of $f(\gamma)$ because this may depend on some future information (see Martino 1982, p. 316). Such a proof gives only a method for calculating, *in a finite number of stages*, $f(\gamma)$. And this is enough for our purposes.

[3]Our reconstruction is similar to the second one of (Posy 1976, p. 113) but the latter is contaminated by misleading problems. In fact, Posy rightly recognises the possibility of getting an outright absurdity by opposing $\neg \vdash_0 \exists n f(\gamma) = n$ to $\vdash_0 \exists n f(\gamma) = n$. Notwithstanding, he does not realise that $\vdash_0 \exists n f(\gamma) = n$ straightforwardly follows from the a priori knowledge of the fullness of f. So he tries to justify it by conjecturing certain hypotheses of determinateness on f which Brouwer might have tacitly understood. By this way he goes off the rails and incurs Troelstra's criticism (see Troelstra 1982, note 10).

[4]The crucial importance in Brouwer's argument of the fullness hypothesis was not realised by Veldman. In fact, he "teases" Brouwer by remarking that, following the line of his reasoning, one could construct a real number at which the characteristic function $c_{\mathbb{Q}}$ of the set of rational numbers is not defined (Veldman 1982, p. 13). But Brouwer could well reply: "according to my reasoning, one could construct it if $c_{\mathbb{Q}}$ *were full*! This leads only to the right conclusion that $c_{\mathbb{Q}}$ fails to be full".

mathematical tool of the claim that all mathematical activity is the one carried out by the creative subject. Later (p. 479) Troelstra observes that a reconstruction of NCP based on the "solipsistic exploitation" would be somewhat anachronistic, because this device appears in Brouwer's writings only after 1948.

Let us examine the connection between solipsism and creative subject arguments. This connection arises from a certain interpretation of the crucial principle

6.1 $A \to \exists n \vdash_n A$.

Brouwer's counterexamples to classical mathematics based on the creative subject do implicitly use 6.1 or a weaker version of it (see Troelstra 1969, p. 96). Our proof of NCP uses the instance of 6.1 with A replaced by 5.1.

Identifying myself with the creative subject, I can read 6.1 as

6.2 If A is true, I will have evidence for it at some stage.

This may appear a solipsistic claim, since it seems to preclude that A might be true in virtue of having been proved by some other mathematician not in relation with me.

We want to argue, however, that, to justify 6.1 intuitionistically, we do not need any solipsistic insight. In fact, after observing that a proposition is intuitionistically true only when recognised as true by someone, the solipsistic claim arises by reading 6.2 in the more explicit form

6.3 If someone will have evidence for A, I myself will have evidence for it.

The point is that the very reason why 6.3 appears as a declaration of solipsism is that even an intuitionist is accustomed to understand implication 6.3 in the classical (realistic) sense in *everyday life*. This happens because 6.3 refers to empirical events (as someone having evidence for something) that are usually regarded as free from the intuitionistic criticism against realism. For this reason, an intuitionist may reasonably agree to the classical reading of 6.3. Nevertheless, as far as he is dealing with intuitionistic arguments, he has to read 6.1 within the intuitionistic framework. Now, according to Heyting's reading of implication, a proof of 6.1 consists in a method of transforming any proof of A into a proof of $\exists n \vdash_n A$. Well, such a method is certainly available to me trivially: if and when a proof of A will be given to me, I will have evidence for A and get a proof of $\exists n \vdash_n A$ by simply taking into account the stage at which I will be.[5] So I can assert 6.1 independently of any solipsistic assumption.

To read 6.1 intuitionistically as a solipsistic claim, we should read $\vdash_n A$ as ≪I myself have created a proof of A without any external information≫. In this way, by asserting 6.1, I would claim that if A will be proved, I will be the very artificer of such a proof. But this reading of $\vdash_n$ is of no use for any mathematical issue. All known applications of the creative subject need only the ordinary reading of $\vdash_n A$ as "I have evidence for A", where the means by which I have got my evidence are disregarded.

[5] Such a justification of 6.1, based on the mere time element, is also upheld by Dummett (1977, p. 349).

Thus, in my opinion, no "solipsistic exploitation" is involved in creative subject arguments. These serve the mere purpose of making the time element much more explicit than in usual intuitionistic arguments. In the case of NCP, this aspect consists—as we saw—in the connection of the time at which a proof of "f is full" is given with the stage of the construction of γ.

5.7 NCP and Lawless Sequences

Under the suggestion of Brouwer's argument, even beyond Brouwer's intentions, we want to investigate the possibility of reducing NCP to more elementary principles.

Suppose that our universe of choice sequences includes lawless sequences which will be indicated by $\rho, \sigma, \ldots$ Let u range over finite sequences and write $\rho \in u$ for "u is an initial segment of ρ". Call CP_l, the restriction of CP to the universe of lawless sequences. For these we assume, besides CP_l, the following principles:

7.1 $\forall u \exists \rho(\rho \in u)$,
7.2 $\forall m \neg \forall \rho \exists n(\rho(n) = m)$.

7.1 is the well-known *density axiom*: it expresses the genetic rule according to which, in generating a lawless sequence, we are allowed to fix an initial segment in advance (while the later values are to be chosen "at random").

7.2 is not taken as an axiom in the literature because it is derivable from the extension theorem or from the fan theorem (see Troelstra 1983, p. 211). But these theorems rest on very problematic insights, while 7.2 is an obvious consequence of the intended indeterminacy of lawless sequences. Thus, it is more convenient to assume 7.2 as primitive, for our purposes.

We want to connect NCP with the modern theory of lawless sequences by proving that

NCP is derivable from CP_l by means of 6.1 *and* 6.2, *without the help of the creative subject.*

Let f be full and α, $\langle \beta_n \rangle_n$, as above. We associate with every lawless sequence and ρ the choice sequence γ_ρ defined as follows:

$$7.3 \quad \gamma_\rho(n) = \begin{cases} \alpha(n) & \text{if } p_{\rho,n} \text{ is even} \\ \beta_{p_{\rho,n}}(n) & \text{if } \rho_{\rho,n} \text{ is odd} \end{cases}$$

where

$$p_{\rho,n} = \begin{cases} 0 & \text{if } \{m \leq n | \rho m = 0\} = \phi \\ \mu\{m \leq n | \rho(m) = 0\} & \text{otherwise.} \end{cases}$$

Suppose $f(\gamma_\rho) = f(\alpha)$. By applying CP_l to the function $\lambda\sigma.f(\gamma_\sigma)$, we get an m such that $\forall \sigma(\overline{\sigma}m = \overline{\rho}(m) \rightarrow f(\gamma_\sigma) = f(\gamma_\rho) = f(\alpha))$. It follows, by the extensionality

of f, that $\forall\sigma(\overline{\sigma}(m) = \overline{\rho}(m) \rightarrow \neg\exists n(p_{\sigma,n}$ is odd)). If $\forall n_{n<m}\rho(n) \neq 0$, we could construct by 7.1 a lawless sequence σ with $\overline{\sigma}(m) = \overline{\rho}(m)$ and $p_{\sigma,n}$ odd, which is absurd. So $\exists n_{n<m}\rho(n) = 0$.

Suppose $f(\gamma_\rho) \neq f(\alpha)$. By applying CP_l to the function $\lambda\sigma.f(\gamma_\sigma)$, we get an m such that $\forall\sigma(\overline{\sigma}(m) = \overline{\rho}(m) \rightarrow f(\gamma_\sigma) = f(\gamma_\rho) \neq f(\alpha))$. It follows, by the extensionality of f, $\forall\sigma(\overline{\sigma}(m) = \overline{\rho}(m) \rightarrow \neg\exists n(p_{\sigma,n}$ is even $> 0))$. (For, from $p_{\sigma,n}$ even and > 0 it would follow, by Definition 7.3, $\alpha = \gamma_\sigma$, and therefore $f(\gamma_\sigma) = f(\alpha)$.) If $\forall n_{n<m}\rho(n) \neq 0$, we could construct by 7.1 a lawless sequence σ with $\overline{\sigma}(m) = \overline{\rho}(m)$ and $p_{\sigma,n}$ even > 0, which is absurd. So $\exists n_{n<m}\rho(n) = 0$ again.

Hence, every lawless sequence has a zero, which is absurd by 7.2 (taking $m = 0$).

This proof provides a first sense in which *NCP* is more evident than *CP*. In fact if, as we believe, the evidence of 7.1 and 7.2 is unquestionable, we can say that *NCP* is reducible to CP_l. So, whoever judges CP_l more evident than *CP*, as is usually done, is compelled to judge *NCP* more evident than *CP*.

We can go further by analysing the intrinsic evidence of *CP* and CP_l. CP_l is supported by the following argument:

> at any stage of the construction of a lawless sequence ρ, the whole available information on ρ consists in the knowledge of an initial segment, so that $f(\rho)$ must be computable by using only an initial segment of ρ.

We contend that the above argument for CP_l is not so cogent as it seems at first sight. We have to touch on a rather subtle question concerning the nature of choice sequences. On the one hand, these are to be regarded as undetermined in the sense that their extensions may be never completely given (and actually never completely given in the case of lawless sequences).

On the other hand, if we want to deal with choice sequences as authentic entities and to refer to some specific one of them unambiguously, we have to think of them as well-determined objects perfectly distinguishable one from another. This second aspect is made explicit in the literature by introducing the decidable relation of *strict* or *intensional identity*.

Generally speaking, two objects are strictly identical if they are the same object and the *individuality* of an object is what makes it distinguishable from any other one.

Individuality is so basic a notion that it is hardly reducible to something more elementary. We will therefore assume it as primitive. The following remarks will be useful to properly understand the role of this notion in our context. First, we should recall that, according to the general intuitionistic ontology, mathematical objects exist only as mental constructions. Accordingly two objects can be the same only insofar as they are *thought of* as the same. In particular, two choice sequences α, β are strictly identical ($\alpha \equiv \beta$) if they are thought of as the same process of choices. In Troelstra's words:

> We shall regard two lawless sequences α, β, (strictly) identical ($\alpha \equiv \beta$) if they are given to us (we think of them) from the beginning as the *same* process. Thus, $\alpha \equiv \beta \vee \alpha \not\equiv \beta$. (Troelstra 1983, p. 208)

To save both determinacy and indeterminacy, we have to very carefully distinguish individuality from extension. The very same mental act of conceiving a specific choice sequence α confers on α its individuality. This is completely determined from the outset quite independently of the extension of α, whose values may be chosen step-by-step without any predetermination. Thus, the individuality of α counts as an essential datum beyond any information about its extension (i.e. initial segments and restrictions). Specifically, the data of a lawless sequence cannot merely consist in an initial segment.

Let us imagine, for instance, that the lawless sequences α and β are displayed to us step-by-step in such a way that at every stage n we know their initial segments $\overline{\alpha}n$ and $\overline{\beta}n$. If this were the whole available information, then we could not judge at any stage that $\alpha \equiv \beta$ and we could judge that $\alpha \not\equiv \beta$ only if and when we reach a stage n at which $\overline{\alpha}n \neq \overline{\beta}n$. So strict identity would fail to be decidable. We can therefore decide whether $\alpha \equiv \beta$ or not only in virtue of some further information on α and β: this is just the knowledge of their individualities.

Summing up, a lawless sequence is actually given only when all the following data are available: (1) its individuality, (2) the initial segment of all its already chosen values, (3) the initial segment specified in advance. (This third datum is needed to justifying the density axiom, but it is not involved in our discussion).[6]

We can conclude that the above argument for CP_l, as it stands, is incorrect, since it neglects datum (1). Whenever $\alpha \not\equiv \beta$, in virtue of this very information f could assign two different outputs to α and β. This possibility cannot be precluded, as we saw, by the nature of lawless sequences. Nor it can be precluded by the extensionality hypothesis on f, since, as it is well-known, *any* function on lawless sequences is extensional.

Thus, the argument for CP_l needs some stronger hypothesis on f to assure that datum (1), though certainly available, cannot be exploited by f. It would be interesting to look for some stronger notion of extensionality which would be adequate for this purpose. Here we want to show only that, in contrast, no stronger hypothesis on f than extensionality (in the usual sense) is required for establishing *NCP*, so that this turns out to be even more general than CP_l.

For, let us return to our deduction of *NCP* in the present section and observe that, instead of CP_l, we can use the axiom of *open data*

$$7.4 \quad A(\rho) \to \exists n \forall \sigma (\overline{\sigma}(n) = \overline{\rho}(n) \to A(\sigma))$$

taking for $A(\rho)$ the property $\neg\exists n(p_{\rho,n}$ is odd).

Of course, the general validity of Scheme 7.4 incurs the same difficulties as CP_l (which is derivable from 7.4). Nevertheless, our particular instance of 7.4 is not problematic at all: for our A, in establishing $A(\rho)$ no help can arise from the

[6] In Troelstra (1983), Troelstra is aware that the code of the initial segment specified in advance occurs among the data of a lawless sequence (p. 212). As to individuality, he seems to disregard it mostly. He is forced, however, to take it into account when formulating the general form of the axiom of open data (Troelstra 1969, p. 36). In this context, he seems to believe that one can exploit individuality only by means of the relation of equality. Perhaps he intends to restrict himself to properties and functions for which this is the case.

individuality of ρ, so that the only relevant datum is just an initial segment of ρ. So *NCP* is shown to be more general than CP_l.

Our argument against CP_l cannot be brought forward against *CP*. Indeed, dealing with the universe of all choice sequences, the extensionality hypothesis does certainly help in limiting possible references to the individualities of the inputs. There is no evidence, however, for concluding that such references are precluded at all by extensionality. In Sect. 5.9, we will try to show by a counterexample that this is not the case.

5.8 Revising NCP with the Help of the Creative Subject

The proof of *NCP* in the foregoing section does not use the creative subject explicitly. It may be interesting to revise that proof with the help of the operator $\vdash_n$. In this way, we get our final reconstruction.

Suppose we have at some stage, say at stage 0, a proof of 5.1. At the same stage 0, we introduce a lawless sequence ρ whose nth value will be chosen at stage $n+1$. The relevant feature of ρ is that, for all $m \geq n$, $\rho(m)$ is, at stage n, still completely undetermined. This feature immediately assures the truth of the following principle:

8.1 $\forall n \forall a(\overline{a}n = \overline{\rho}(n) \rightarrow \neg \vdash_n \exists m \rho(m) \neq a(m))$

where a ranges over lawlike sequences.

For, as, at stage n, only the first n values of ρ are known and all possibilities are open to future choices, it is impossible, at stage n, to prove that some future value of ρ will be different from the corresponding value of a.

Define the choice sequence γ as in 7.3 (omitting the index ρ, since we are dealing with only one ρ).

Let q be a stage, at which $f(\alpha)$ and $f(\gamma)$ are calculated. We have either $\vdash_q f(\alpha) = f(\gamma)$ or $\vdash_q f(\alpha) \neq f(\gamma)$.

In the first case, in virtue of extensionality, $\vdash_q \neg\exists n(p_n \text{ is odd})$. It follows from 8.1 that $\mu\{m < q | \rho(m) = 0\}$ exists and is even. For, if $\rho(m) \neq 0$, for all $m < q$, we could construct a lawlike sequence a with $\overline{a}q = \overline{\rho}q$ and having the first zero odd. So we would have, at stage q, a proof of $\exists m \rho(m) \neq a(m)$, against 8.1.

Similarly, in the second case, $\mu\{m < q \mid \rho(m) = 0\}$ exists and is odd.

Therefore, since $\vdash_0 (f(\alpha) = f(\gamma) \vee f(\alpha) \neq f(\gamma))$, we can conclude, at stage 0, that ρ has at least one zero, $\vdash_0 \exists m \rho(m) = 0$. By 8.1 again, taking for a a lawlike sequence without zeroes, an absurdity arises.

Observe that the operator $\vdash_n$ enables us to express the essential requisite of ρ, i.e. its being still completely undetermined when a proof of 5.1 is available.

In the reconstruction in Sect. 5.6, we made up for the lack of $\vdash_n$, by considering *all* lawless sequences. In such way, some lawless sequence whose stages of construction were *ad hoc* was automatically included. We can remark, however, that the evidence of the principles involved rested again on $\vdash_n$-considerations. For instance, to convince

ourselves of 7.2, we have to observe that *at any stage* of our activity we are able to introduce some new lawless sequence as yet completely undetermined so that *at that stage* we cannot know anything about its values. And this is just what we actually did when the single ρ at the appropriate stage was introduced. In this respect, the $\vdash_n$-argument is more straight and explicit.

It is also remarkable that the proof in Sect. 5.7 uses two distinct principles, 7.2 and 7.4, which—though both expressing undeterminacy of lawless sequences—are logically independent. On the contrary, the $\vdash_n$-argument exploits indeterminacy by means of the sole principle 8.1.

Summing up, we suggest that our final reconstruction is the most adequate to indicate the insights involved by *NCP*.

5.9 Extensional Functions and Intensional Choice Sequences

As we saw, the argument for *NCP* does not require any hypothesis on f except that of extensionality, while it is not evident that extensionality is sufficient to guarantee *CP*.

We are going to try to construct a model of choice sequences and an extensional function for which *NCP* holds while *CP* fails. This construction will not count as a disproof of CP since we do not claim that our models is the "intended" model of all choice sequences. Nevertheless, we think it may be useful for illustrating how one might find a strategy for exploiting the individuality of choice sequences, in spite of the extensionality hypothesis.

Let U be a universe of countably many choice sequences, enumerated as follows:

9.1 $\alpha_0, \alpha_1, \ldots, \alpha_n, \ldots$.

We assume that every finite segment occurs infinitely many times as initial segment of some choice sequences.

Let ϕ be a bijective map of U onto the set of ordinal numbers $< \omega^2$. We say that a sequence $\langle \beta_n \rangle_n$ of choice sequences is *bounded* if there is an ordinal $\tau < \omega^2$ such that $\phi(\beta_n) < \tau$ for all n. Let us restrict the notion of sequence of choice sequences to bounded sequences. So the enumeration 9.1 of the whole universe is not a sequence of our model.

We also impose the condition that any restriction on a choice sequence α is allowed to link α only to sequences of lower ordinal.

These conditions serve the purpose to assure that, for any sequence $\langle \beta_n \rangle_n$ of choice sequences, we can construct a "fresh" choice sequence not involved in the construction of the β_n's. This is all that is required for the validity of Brouwer's argument.

So *NCP* holds in our model. However, *CP* fails, as the following example shows.

Let us define the function f step-by-step with reference to the enumeration 9.1:

- Step 0: put $f(\alpha_0) = 0$.

- Step $n > 0$: suppose we have already defined $f(\alpha_0), \ldots, f(\alpha_{n-1})$;

$$\text{put } f(\alpha_n) = \begin{cases} f(\alpha_m) & \text{if, for some } m < n,\ \overline{\alpha_n}(n) = \overline{\alpha_m}(n) \\ \max\{f(\alpha_0), \ldots, f(\alpha_{n-1})\} + 1 & \text{otherwise.} \end{cases}$$

So f is extensional.

Now, it is easily proved that

9.2 $\forall n \exists \beta (\overline{\beta} n = \overline{\alpha_0} n \wedge f(\beta) \neq f(\alpha_0))$

For given n, let p be the least natural number $> n+1$ such that (1) $\overline{\alpha_p}(n) = \overline{\alpha_0}(n)$ and (2) $\overline{\alpha_p}(n+1) \neq \overline{\alpha_m}(n+1)$ for all $m \leq n+1$. Then 9.2 is satisfied by $\beta \equiv \alpha_p$.

Observe that 9.2 does not come into conflict with *NCP* because the function mapping n into β is not a sequence of our model.

5.10 Troelstra's Abstraction Process and NCP

In his discussion of *NCP*, Troelstra (1982, p. 478) declares himself favourable to a reconstruction based on the concept of "abstraction process". In my opinion, this is an interesting but highly problematic device, which Troelstra uses on several occasions. We devote the present section to discuss this reconstruction.

Following Troelstra (see, for instance, Troelstra 1983, p. 48), we describe the abstraction operator (*abstr*) as follows.

Given any choice sequence χ, *abstr* χ is the sequence obtained from χ by abstracting from (or forgetting) all restrictions on χ. More explicitly, *abstr* χ has the same values as χ, but at every stage n we consider the knowledge of the initial segment $\overline{abstr\ \chi}(n)$ $(= \overline{\chi} n)$ as the sole available information about *abstr* χ: we pretend (1) no restriction has yet been set and (2) any restriction is possible for the future. So, though $\chi = abstr\ \chi$ holds in a "metasense", we are not allowed to regard it as a mathematical assertion.

Now, returning to Brouwer's proof of *NCP* (in our notations), let us interpret γ as *abstr* α.

Since f is full, $f(\gamma)$ will be calculable at some stage n. Thus, $f(\gamma)$ is determined by the mere information $\overline{\gamma} n = \overline{\alpha} n$. And as this does not preclude any of the restrictions $\gamma = \alpha$ and $\gamma = \beta_m$ (for some $m > n$) for the future, an absurdity arises.

It would be very tempting to use the same argument for proving the false proposition 4.2 and to conclude that *abstr* is a contradictory operator. But that would be too ingenuous. I suppose Troelstra could easily defend his operator as follows.

Abstr has to be regarded not as a mathematical operator but as a tool for describing a "thought-experiment". Therefore, *abstr*α fails to be an authentic choice sequence. So from the failure of f being defined on *abstr*α, we cannot infer the existence of a choice sequence not belonging to $D(f)$ (as it would required for proving 4.2). On the other hand, as far as the calculation of f is concerned, *abstr*α is indistinguishable from a genuine choice sequence. Thus, if we know a priori that f is defined on *all*

choice sequences, then we must be able even to calculate f on *abstr* α (by using the permitted pieces of information of *abstr* α *as if* they were those of an authentic choice sequence).

This should explain why, by means of *abstr*, *NCP* can be deduced while 4.2 cannot.

However, the above charge against *abstr*, as well as the supposed reply, brings out an intrinsic ambiguity of *abstr*: according to the context, one can use or not the information $\alpha = abstr\,\alpha$ and it is not clear whether such dealing is lawful. To escape a contradiction, we were forced to maintain that *abstr* α is not an authentic choice sequence; but this leads to question what kind of entity it is.

The "forgetting" element suggested by Troelstra is not so innocent as it may seem at first sight. According to this suggestion, $\gamma(\equiv abstr\,\alpha)$ would actually be subject to the restriction $\gamma = \alpha$ but, when calculating $f(\gamma)$, we could forget such restriction. However, observe that the computation procedure may explicitly ask whether any possible restrictions have been imposed on the input. For this reason, forgetting the restrictions cannot simply amount to not using them. Rather, it amounts to using the *false* information that no restriction has been set; and false information may well yield a wrong output. So, if γ has actually the restriction $\gamma = \alpha$, we have no right to forget it.

On the other hand, if γ has no restrictions, it is very hard to grasp in which sense $\gamma = \alpha$ holds.

Troelstra (1983, p. 222) suggests that *abstr* describes a "thought-experiment", but he fails to enlighten the very sense of this expression. We could certainly imagine that for some choice sequence γ, though constructed without any restriction, it might be, *as a matter of fact*, $\gamma n = \alpha n$, for all n. But this "thought-experiment" would introduce a *factual element* which at no "metalevel" can be justified from an intuitionistic viewpoint. That seems to me a classical (realistic) way of thinking about an intuitionistic construction. It would also bring us close to Posy's first reconstruction which we discussed in Sect. 5.4.

We do not deny the possibility of exploiting the rough intuition (which seems to be behind *abstr*) that some unrestricted choice sequence "might" be equal to α. What we want to stress is the need of finding a clear—intuitionistically acceptable—interpretation of "might".

A possible way out would be that of understanding "might" in the sense that—as long as for a choice sequence γ no value has been chosen and no restriction has been imposed—we cannot have evidence that at some point γ will differ from α. But this would lead us back to principle 8.1.

For these reasons, we cannot be satisfied with the abstraction process, as a foundational support: as it stands, it still needs its own foundation.

5.11 Conclusions

Some final words. We have tried to focus what, in our opinion, is a point of foundational interest of *NCP* for the intuitionistic theory of real functions. In fact, a crucial philosophical problem in this area is the question whether Brouwer's main theorem that asserts the continuity of all full real functions is a cogent consequence of the general intuitionistic point of view on mathematics.

Well, our answer is that the general intuitionistic conception of mathematics leads certainly towards continuity but is not sufficient for establishing it. More precisely, the mere purpose of rendering real numbers and real functions intuitionistically intelligible leads cogently to negative continuity, not to positive continuity.

Indeed, as we saw, *NCP* holds for the most general notions of choice sequence and of extensional function in virtue of very elementary principles. On the contrary, the evidence of *CP* presupposes some understood limitations on the notion of function, such as that of prescinding from the individuality of inputs. Somebody might maintain that such a limitation is not substantial because no *mathematically interesting* function exploits the individuality of its inputs. This position is, in my opinion, not very convincing. After all, as already observed, the very importance of individuality arises just from the introduction of mathematical entities extensionally undetermined, such as choice sequences. It is just such indeterminacy that imposes the presence of individuality as something quite separate from extension. So upholding the mathematical irrelevance of individuality seems to me somewhat paradoxical. On the one hand, the universe of certain mathematical entities is so conceived that individuality is an essential feature of them; on the other hand, the right of prescinding on individuality, as mathematically irrelevant, is vindicated.

Even if there is no known example of a mathematically interesting function exploiting individuality, nevertheless such functions can rightly have their influence in establishing *global* mathematical properties of the universe of all functions.

In mathematics, this is a recurring situation. For instance, although most real numbers, taken in themselves, are of no mathematical interest, they nevertheless contribute to structuring the continuum as a whole, even from a classical point of view.

Finally a word on the assumption, we made in our final reconstruction (Sect. 5.8), of the controversial inclusion of lawless sequences in the universe of all choice sequences.

First, independently of Brouwer's opinion, I cannot see any good reason for expelling lawless sequences from the general concept of choice sequence.

Second—though inessential for reconstructing Brouwer's—argument the use of lawless sequences is helpful for analysing the underlying insights. Indeed, an interesting aspect of Brouwer's argument consists in the suggestion that continuity is not a mere consequence of freedom in choosing the values of choice sequences, but rather a consequence of the combined play of freedom and constriction. This aspect is evidenced by our reference to lawless sequences. The freedom component is singled out by principle 8.1. On the other hand, the failure of Brouwer's argument for the

universe of lawless sequences, where there is no room for bound choice sequences such as γ, shows the importance of the constriction component.

As a third point, it seems to me that the rejection of lawless sequences would not be in line with Brouwer's reasoning. For, the refusal of lawless sequences falls within a conception of choice sequence in which higher-order restrictions are not allowed (one may regard lawlessness as the second-order restriction preventing any restriction on values). Now observe that for his γ Brouwer "reserves the right" of choosing *only* between two particular kinds of restrictions; doing that seems just to impose a second-order restriction. I think therefore that, when he proved the negative continuity theorem, Brouwer thought of a notion of choice sequence from which lawlessness, though never mentioned, was not excluded.

References

Brouwer, L. (1927). Über Definitionsbereiche von Funktionen. *Mathematische Annalen*, *97*, 60–75. English translation in From Frege to Gödel, Cambridge MA, 1967, pp. 446–463.

Dummett, M. (1977). *Elements of intuitionism*. Oxford: Clarendon Press.

Martino, E. (1982). Creative subject and bar theorem. In D. V. Dalen & A. Troelstra (Eds.), *The L. E. J. Brouwer Centenary Symposium* (pp. 311–318). North Holland: Amsterdam. Reprinted here as chapter 2.

Parsons, C. (1967). Introduction to brouwer (1927). In J. Van Heijenoort (Ed.), *From Frege to Gödel* (pp. 446–457). Cambridge, MA: Harvard University Press.

Posy, C. (1976). Varieties of indeteminacy in the theory of general choice sequences. *Journal of Philosophical Logic*, *5*, 91–132.

Troelstra, A. (1969). *Principles of intuitionism*. Berlin: Springer.

Troelstra, A. (1982). On the origin and development of Brouwer's concept of choice sequence. In D.V. Dalen & A. Troelstra (Eds.), *The L. E. J. Brouwer Centenary Symposium* (pp. 465–486). North-Holland: Amsterdam.

Troelstra, A. (1983). Analysing choice sequences. *Journal of Philosophical Logic*, *12*, 197–259.

Veldman, W. (1982). On the continuity of functions in intuitionistic real analysis. Technical report, Nijmegen.

Chapter 6
Classical and Intuitionistic Semantical Groundedness

Abstract Kripke's notion of semantical groundedness for classical logic is developed in an intuitionistic framework. It is argued that semantical groundedness yields the most natural solution of the semantical paradoxes.

6.1 Introduction

The classical notion of truth has, as intuitionistic counterpart, a notion of (informal) provability. The latter fixes the intuitionistic meaning of logical constants, as the former fixes their classical meaning.

With reference to a first-order language, the classical meaning of logical constants is expressed by the Tarskian inductive clauses that fix the *truth conditions* of a complex sentence in terms of those of its components. Similarly, the intuitionistic meaning of logical constants is expressed by the Heyting inductive clauses that fix the (informal) *provability conditions* of a complex sentence in terms of its components.

We will consider any first-order classical or intuitionistic theory, capable of expressing first-order classical or intuitionistic arithmetic. With reference to an arithmetical codification of the syntax, we will indicate by $\ulcorner \phi \urcorner$ the code of a sentence ϕ.

The predicate of (informal) intuitionistic provability meets problems analogous to the classical truth predicate. In the sequel, we will indicate by T both the classical truth predicate and the intuitionistic provability predicate. This is suitable, because of the intuitionistic identification of truth with provability.

The Tarskian biconditional

$$\phi \leftrightarrow T\ulcorner \phi \urcorner$$

holds also intuitionistically, as well as Tarski's theorem, according to which no (enough powerful) theory can express its own truth predicate. In fact, otherwise, one could obtain the *liar* sentence χ of form $\neg T\ulcorner \chi \urcorner$, incompatible with the Tarskian biconditional. Observe that, in the intuitionistic case, the liar paradox cannot be avoided by the intuitionistic rejection of the *excluded middle*. For, the available knowledge that any hypothetical proof of χ, namely of $\neg T\ulcorner \chi \urcorner$, leads to contradiction

counts, in virtue of the intuitionistic meaning of negation, as a proof of $\neg T^{\ulcorner}\chi^{\urcorner}$, i.e. of χ. So, both χ and $\neg\chi$ turn out to be provable.

Kripke's notion of *semantical groundedness* (Kripke 1975) yields, in our opinion, the most enlightening approach to the semantical paradoxes.

Assume that T is a predicate of the language. Intuitively, a sentence ϕ is *grounded* if its truth conditions are inductively defined, starting from those of the basic sentences (i.e. atomic and without T), by using the Tarskian (or Heyting) clauses for the logical constants and a suitable clause for T. The latter, suggested by the Tarskian biconditional, defines the truth conditions of $T^{\ulcorner}\phi^{\urcorner}$, by identifying them with the (already defined) ones of ϕ.

Now, the liar sentence χ is ungrounded, since the truth conditions of χ would presuppose those of $T^{\ulcorner}\phi^{\urcorner}$.

So, semantical groundedness explains the reasons why T is inexpressible in the object language and puts forward a suitable set of sentences, the grounded ones, for which the truth predicate is expressible.

In Kripke's original work, as well as in the subsequent literature, semantical groundedness is treated only within a classical framework. The similar behaviour of the classical and intuitionistic predicates suggests, however, an intuitionistic solution of the liar paradox in terms of well-defined provability conditions.

We will sketch an approach, both classical and intuitionistic, to semantical groundedness.

Start from a model M_0 of a (classical or intuitionistic) theory Θ_0, relative to a first-order language L_0. Suppose, for the sake of simplicity, that L_0 has individual constants for all members of the domain $\mathbf{D}$ of M_0. Put $L = L_0 \cup \{T\}$, where T is a new monadic predicate constant, and extend the arithmetical codification to L. We want to extend M_0 to a model M of L, where T will be interpreted as the *grounded truth predicate*. Finally, we will extend Θ_0 to a theory Θ, conservative over Θ_0, by adding suitable axioms for T. M will be the *intended model* for T.

If Θ_0 and M_0 , as well as the metalanguage, are classical (intuitionistic), so are Θ and M.

In the classical case, T will be equivalent to Kripke's truth predicate, evaluated according to the Kleene weak schema.

6.2 Construction of Model M

Let us inductively define, relative to M_0, the grounded L-sentences and their *truth conditions* (to be understood, in the intuitionistic case, as Heyting *provability conditions*):

2.1

(i) If ϕ is an atomic L_0-sentence (including the falsum $\perp$), it is grounded. It is M-true if it is M_0-true.

(ii) If ϕ and ψ are grounded, so is $\phi \bullet \psi$, where $\bullet$ is any binary connective. The truth condition is defined according to the usual Tarski's or Heyting's clauses. (The case of negation is included by defining $\neg\phi =_{df.} \phi \rightarrow \bot$.)

(iii) If $\phi(x)$ is a formula with the free variable x and, for all $d \in \mathbf{D}$, $\phi(d)$ is grounded, so are $\forall x\phi(x)$ and $\exists x\phi(x)$. $\forall x\phi(x)$ is M-true if, for all $d \in \mathbf{D}$, $\phi(d)$ is M-true, $\exists x\phi(x)$ is M-true if, for some $d \in \mathbf{D}$, $\phi(d)$ is M-true.

(iv) If ϕ is a grounded sentence, so is $T\ulcorner\phi\urcorner$. This is M-true if ϕ is M-true.

Extend M_0 to the L-model M by the following truth conditions for T:

2.2 $T\ulcorner\phi\urcorner$ is M-true if ϕ is grounded and M-true (briefly *groundedly true*).

Observe that the truth conditions for $T\ulcorner\phi\urcorner$ are well-defined for all L-sentences ϕ, since 2.2 presupposes the hole inductive definition of groundedness. It turns out that, for all ungrounded ϕ, $T\ulcorner\phi\urcorner$ is false. Observe that this is the case even intuitionistically. For, groundedness is decidable; so, if ϕ is ungrounded, the knowledge, by 2.2, of the impossibility to prove $T\ulcorner\phi\urcorner$ counts, according to the intuitionistic meaning of negation, as a proof of $\neg T\ulcorner\phi\urcorner$.

One can easily verify the following propositions:

2.3 Any L_0-sentence is M-true iff it is M_0-true.
2.4 Any L-sentence ϕ is grounded iff $T\ulcorner\phi \rightarrow \phi\urcorner$ is M-true.
2.5 If ϕ is grounded, then $\phi \leftrightarrow T\ulcorner\phi\urcorner$ is M-true (restricted Tarski's biconditional).

Semantical groundedness may suggest the adoption of a *gap-semantics*, according to which ungrounded sentences are neither true nor false. But such a semantics does not offer a satisfactory solution of the *strengthened liar paradox*, which distinguishes falsity from failure of truth. In fact, if the liar sentence χ lacks any truth value, it is, in particular, not true; but, if so, $\neg T\ulcorner\chi\urcorner$, seems to say just that is not true, so that, after all, it should be true.

Our model M reconciles the intuition of the truth with that of the untruth of χ. By 2.2, since is ungrounded, $T\ulcorner\chi\urcorner$ is false, so $\neg T\ulcorner\chi\urcorner$, i.e. χ is true. But remember that T is the *grounded truth predicate*, so what $\neg T\ulcorner\chi\urcorner$ says is that fails to be groundedly true. Thus, no contradiction arises: the liar sentence is true but not groundedly true.

6.3 Axiomatisation of *T*

In accordance with 2.4, the *groundedness predicate* G is expressible in L by defining

$$G\ulcorner\phi\urcorner =_{df.} T\ulcorner\phi \rightarrow \phi\urcorner.$$

The model M suggests the following axiomatisation:

3.1 (a) $G^{\ulcorner}\phi^{\urcorner}$, for all atomic L_0-formula ϕ.
(b) $G^{\ulcorner}\phi \bullet \psi^{\urcorner} \leftrightarrow G^{\ulcorner}\phi^{\urcorner} \wedge G^{\ulcorner}\psi^{\urcorner}$, where $\bullet$ is any propositional connective.
(c) $G^{\ulcorner}\mathsf{Q}x\phi x^{\urcorner} \leftrightarrow \forall x G(\phi x)^*$, where Q is any quantifier and $(\phi x)^*$ is a term such that, for all $d \in \mathsf{D}$, $(\phi d)^* = {}^{\ulcorner}\phi d^{\urcorner}$.
(d) $G^{\ulcorner}T^{\ulcorner}\phi^{\urcorner\urcorner} \leftrightarrow G^{\ulcorner}\phi^{\urcorner}$.
(e) $T^{\ulcorner}\phi^{\urcorner} \rightarrow G^{\ulcorner}\phi^{\urcorner}$.
(f) $G^{\ulcorner}\phi^{\urcorner} \rightarrow (\phi \leftrightarrow T^{\ulcorner}\phi^{\urcorner})$.

Call Θ the theory obtained from Θ_0 by adding the axioms 3.1. M turns out to be a model of Θ.

3.2 *Theorem*. Θ is conservative over Θ_0.

Since, as we saw, every model of Θ_0 is extensible to a model of Θ, the theorem follows immediately, in the classical case, from the completeness theorem. Intuitionistically, completeness fails for Heyting's intended semantics. One can, however, recover intuitionistic completeness by adopting generalised Beth semantics or negationless semantics (see Martino 1998).

6.4 The Aczel–Feferman Intensional Operator

Aczel and Feferman (1980) proposed an intensional set theory with a comprehension principle of form:

4.1 $y \in \{x : \phi x\} \equiv \phi y$,

where $\equiv$ is a certain operator of *intensional equivalence*, introduced and axiomatized by the authors. They propose to read $\phi \equiv \psi$ as "ϕ and ψ are equivalent in virtue of given basic definitions" (see also Feferman 1984).

The operator $\equiv$ is interpretable through our grounded truth predicate T as follows:

$$\phi \equiv \psi =_{df.} (G^{\ulcorner}\phi^{\urcorner} \leftrightarrow G^{\ulcorner}\psi^{\urcorner}) \wedge (G^{\ulcorner}\phi^{\urcorner} \rightarrow (\phi \leftrightarrow \psi))$$

Let $S(x, y)$ be a term such that, if $d \in \mathsf{D}$, u is a variable and ϕ a formula, then $S(d, {}^{\ulcorner}\langle u, \phi(u)\rangle^{\urcorner}) = {}^{\ulcorner}\phi(d)^{\urcorner}$.

Put $\{x : \phi x\} =_{df.} {}^{\ulcorner}\langle x, \phi(x)\rangle^{\urcorner}$, $x \in y =_{df.} T(S(x, y))$.

With these definitions, 4.1 holds. Besides, in the classical case, all Aczel–Feferman axioms hold in Aczel's variant.

Our interpretation puts forward a precise meaning—both classical and intuitionistic—of the operator $\equiv$. It makes explicit the idea, suggested by Aczel, of *definitional equivalence*. For, our definition of $\phi \equiv \psi$ can be read: clauses 2.1 define truth conditions for ϕ iff they define truth conditions for ψ and, if this is the case, ϕ and ψ are logically equivalent.

In this way, we obtain a version of Aczel–Feferman theory that holds both classically and intuitionistically.

References

Aczel, P., & Feferman, S. (1980). Consistency of the unrestricted abstraction principle using an intensional equivalence operator. In J. P. Seldin & J. R. Hindley (Eds.), *To H. B. Curry: Essays in combinatory logic, lambda calculus and formalism* (pp. 67–98). Academic Press.

Feferman, S. (1984). Toward useful type-free theories. *Journal of Symbolic Logic*, *49*, 75–111.

Kripke, S. A. (1975). Outline of a theory of truth. *Journal of Philosophy*, *72*, 690–716.

Martino, E. (1998). Negationless intuitionism. *Journal of Philosophical Logic, 27*, 165–177. (reprinted here as chapter 10).

Chapter 7
Brouwer's Equivalence Between Virtual and Inextensible Order

Abstract Brouwer's theorem of 1927 on the equivalence between virtual and inextensible order is discussed. Several commentators considered the theorem at issue as problematic in various ways. Brouwer himself, at a certain time, believed to have found a very simple counterexample to his theorem. In some later publications, however, he stated the theorem in the original form again. It is argued that the source of all criticisms is Brouwer's overly elliptical formulation of the definition of inextensible order, as well as a certain ambiguity in his terminology. Once these drawbacks are removed, his proof goes through.

7.1 Introduction

The purpose of this paper is to clarify some controversial matters concerning Brouwer's characterisation of virtual orders as inextensible orders. The equivalence of the two notions of order was proved by Brouwer for the first time in the paper (1927). It is also proved in the Cambridge lectures (1981, 52–54).

Heyting (1975, 569–597), Posy (1980), van Dalen (1981, note 34) find Brouwer's definitions and proofs problematic in various ways. Brouwer himself observes in an unpublished note of 1933 that his 1927 proof that every virtual order is inextensible is invalid and supplies a very simple counterexample.

From a note in the margin of Brouwer's own copy of his 1927 paper, as well as from Brouwer's formulation of the matter in the Cambridge lectures, Heyting argues that Brouwer tried to save his theorem by changing his definition of virtual order and criticises the new definition as involving the metamathematical notion of derivability. The same conclusion is upheld by Posy and van Dalen. Posy tries to disentangle the matter by means of a suggestive reconstruction involving the creative subject. We claim that:

(1) The paper 1927 is correct;
(2) Brouwer never changed his notion of virtual order. The treatment of the topic is essentially the same in all Brouwer's papers. However, the definition of inextensible order in the Cambridge lectures needs a minor correction;
(3) Brouwer's unpublished counterexample is wrong;

(4) No metamathematical notion of derivability is involved in Brouwer's definitions;
(5) The creative subject is of no help for understanding Brouwer's argument.

7.2 Reconstruction of Brouwer's Paper of 1927

Brouwer's original paper is essentially correct. All objections of the commentators arise, in our opinion, from certain ambiguities in the symbolism and from the overly elliptical formulation of the definition of inextensible order.

Here, we shall try to give a reconstruction of Brouwer's paper free from these drawbacks. Our exposition differs from Brouwer's in the following aspects: we will introduce a symbolism more usual in today's literature and free from the ambiguities of the Brouwerian one; we will work out the notion of *addable pair* and will prove it to be equivalent to its double negation.

A partial order $(<, =)$ on a given species S consists of two binary predicates $<$, $=$ on S satisfying the following clauses:

2.1 $x = x$,
2.2 $x = y \to y = x$,
2.3 $x = y \wedge y = z \to x = z$,
2.4 $x = y \wedge z = u \wedge x < z \to y < u$,
2.5 $x < y \wedge y < z \to x < z$,
2.6 $x < y \to \neg y < x$.

A partial order is *virtual* if moreover

2.7 $\neg x < y \wedge \neg y < x \to x = y$,
2.8 $\neg x < y \wedge \neg x = y \to y < x$.

A partial order $(\prec, \approx)$ is an *extension* of a partial order $(<, =)$ (on the same species S) if $x < y \to x \prec y$ and $x = y \to x \approx y$. The pair (x, y) is (*positively*) *addable* to the predicate $<$ (to the predicate $=$) if there is some extension $(\prec, \approx)$ of $(<, =)$ such that $x \prec y (x \approx y)$. (x, y) is *negatively addable* to $<$ (to $=$) if its addability to $<$ (to $=$) is non-contradictory.

Lemma 7.1 *Let $(<, =)$ be a partial order on S and let $(a, b) \in S \times S$. The following statements are equivalent:*

(i) $\neg a < b \wedge \neg b < a$ $(\neg a < b \wedge \neg a = b)$;
(ii) (b, a) *is addable to* $=$ *(to* $<$*);*
(iii) (b, a) *is negatively addable to* $=$ *(to* $<$*).*

Proof

(i) $\to$ (ii). Suppose $\neg a < b \wedge \neg b < a$.

Define:

- $x \approx y := x = y \vee (x = a \wedge y = b) \vee (x = b \wedge y = a)$,

- $x \prec y := x < y \vee ((x < a \vee x < b) \wedge (y = a \vee y = b \vee a < y \vee b < y)) \vee ((x = a \vee x = b) \wedge (a < y \vee b < y))$.

A straightforward verification shows that $(\prec, \approx)$ is an extension of $(<, =)$ and $a \approx b$. Suppose $\neg a < b \wedge \neg a = b$. Define:

- $x \approx y := x = y$,
- $x \prec y := x < y \vee ((x < b \vee x = b) \wedge (a < y \vee y = a))$.

A straightforward verification shows that $(\prec, \approx)$ is an extension of $(<, =)$ and that $b \prec a$.

(ii) $\rightarrow$ (iii). Trivial.

(iii) $\rightarrow$ (i). Suppose, by reduction, $a < b$. Then, for any extension $(\prec, \approx)$ of $(<, =)$, $a \prec b$ and hence, by 2.4 and 2.6, $\neg a \approx b$. This contradicts the negative addability of (b, a) to $=$. Thus, $\neg a < b$. Similarly, we get $\neg b < a$ and therefore $\neg a < b \wedge \neg b < a$.

A partial order $(<, =)$ is *inextensible* if, for all $x, y \in S, x < y (x = y)$, provided (x, y) is addable to $<$ (to $=$). The lemma immediately implies the following theorem:

Theorem 7.2 *Any partial order is virtual if, and only if, it is inextensible.*

7.3 Comment on Brouwer's Text

For Brouwer "relations" are not the predicates $<$ and $=$ (subspecies of $S \times S$), as it is usual in today's literature, but the formulas of type $x < y$ and $x = y$, with $x, y \in S$. In this terminology, to give a partial order amounts to interpreting the above formulas in such a way as to satisfy clauses 2.1–2.6.

The species of *existing relations* ("*bestehenden Relationen*"), which Posy finds somewhat mysterious, is simply the species of formulas true in the given interpretation. This species is perfectly defined, provided a specific partial order on S is given.

With regard to the notion of inextensible order, Brouwer's definition is the following: a partial order is called *inextensible* if every relation $p < q$ or $p = q$, which can be added to the existing ones in a non-contradictory manner ("*welche sich den bestehenden Relationen…widerspruchsfrei hinzufügen lässt*"), is an existing relation.

Brouwer does not explicitly define the meaning of the expression "addable in a non-contradictory manner". However, from the context there can be no doubt that it is to be understood according to our definition of negative addability.

First of all, we observe that the alternative interpretation of "$p < q (p = q)$ is addable in a non-contradictory manner" as $\neg\neg p < q (\neg\neg p = q)$, which might appear plausible at first sight, is certainly to be rejected. The terms "add" and "inextensible" seem to involve the idea of enlargement, whereas the interpretation at issue is insensitive to this idea. For, since the assertion $p < q$ certainly means "the

relation $p < q$ is an existing one", the assertion $\neg\neg p < q$ must be understood as "the *existence* of $p < q$ is non-contradictory". Thus, the interpretation in question identifies "addable" with "existing".

Besides, that interpretation trivially falsifies Brouwer's theorem that every inextensible order is virtual. For, let S be a species with at least two distinct elements, let $=$ be the real identity and define $<$ by putting, for all $p, q \in S$, $p < q$ if, and only if, $0 = 1$, so that $\neg p < q$ for all $p, q \in S$. $(<, =)$ is obviously an inextensible order which fails to be virtual.

On the contrary, our interpretation, as we saw in Sect. 7.2, underlies the idea of enlargement and verifies the equivalence between virtual and inextensible orders. Moreover, and more importantly for upholding its fidelity to Brouwer's thought, it is in perfect harmony with Brouwer's argument.

For, in order to prove that any inextensible order satisfies clause 2.8, Brouwer shows that from $\neg s < r \wedge \neg s = r$ the possibility follows of adding the relation $r < s$ to the existing ones in a non-contradictory manner. To show that, Brouwer constructs a species γ including the species a of existing relations and the relation $r < s$; then he shows that γ satisfies the clauses of partial orders. This means just that $r < s$ holds in some partial order extending the given one, i.e. that the pair (r, s) is addable to $<$, according to our definition of (positive) addability. Similarly for Brouwer's proof that any inextensible order satisfies clause 2.7.

Still there arises the question of why Brouwer formulated his definition of inextensible order in terms of *negative* addability. This may seem strange, since in the proof that any inextensible order is virtual he actually uses the notion of *positive* addability, as we have already seen. With regard to the converse, Brouwer's reasoning holds *a fortiori* for the *positive* notion.

Observe that the assertion that every virtual order is inextensible is (*a priori*) stronger if inextensibility is formulated in terms of negative addability. On the other hand, positive addability strengthens the converse assertion.

Now, in dealing with the virtual order of the continuum, Brouwer primarily conceived of inextensibility as a property of virtual orders, rather than of virtuality as a property of inextensible orders. So, in order to exhaustively express inextensibility of virtual orders, he wanted to determine a minimal addability condition assuring the existence of a given relation. This is, in our opinion, the reason of his option for negative addability.

Finally, observe that the equivalence of the two notions of inextensibility follows immediately from Brouwer's proof. For call *strong* (*weak*) inextensibility the notion expressed in terms of negative (positive) addability. Then, his arguments show that every *weakly* inextensible order is virtual and that every virtual order is strongly inextensible. Thus, every *weakly* inextensible order is *strongly* inextensible, whence the claimed equivalence.

7.4 How Brouwer Misinterpreted Himself

In a note on Brouwer (1927), Heyting tells us that, in a handwritten note of 1933, Brouwer observed:

> In [1927] it was proved that every inextensible order is a virtual order. But the proof of the converse theorem found there is invalid. For the incompatibility of the simultaneous non-contradictory existence of an arbitrary two of the relations $r = s, r < s, s < r$ does not entail the incompatibility of the simultaneous non-contradictory addability of an arbitrary two of these relations. Indeed, the species consisting of the four elements a, b, c, d with the order relations $a < b, a < c, a < d, b < d, c < d$ is virtually but not inextensibly ordered.

However, Brouwer's counterexample is patently wrong. Namely the order concerned fails to be virtual, since $\neg b < c$, $\neg c < b$, $\neg b = c$ simultaneously hold. Strangely, the commentators do uphold Brouwer's counterexample and regard it as an authentic difficulty for the theorem in question. Van Dalen, in a note to the Cambridge lectures, defends the counterexample as follows (in Brouwer 1981, note 34):

> Classically, one could say $b = c$ and $b < c$ do not hold in the structure, so that $b \neq c$ and $\neg b < c$ hold, hence by (8) [our 2.8] $b > c$ should hold. However, intuitionistically $b = c$ cannot be claimed, as this would require enough information to reduce $b = c$ ad absurdum. So intuitionistically, (8) and (9) [our 2.8 and 2.7] are vacuously true, and hence the ordering is virtual. But, as Brouwer observes, the order can consistently be extended, e.g. by adding $b < c$.

Van Dalen's argument is, however, fallacious.

First of all, it seems evident to me that the order described by Brouwer is to be understood as *completely* given by the relations $a < b$, $a < c$, $b < d$, $c < d$, $a < d$. In other words, we believe that the mentioned relations are to be considered, *by definition*, as the only existing ones in the given order. If Brouwer had wanted to leave the possibility of the existence of $b < c$ or $c < b$, e.g. by making it depending upon some unsolved problem, he would certainly have said so explicitly, as he always did in such cases. So, a proof of $x <$y consists in the trivial verification that $x < y$ occurs among the given relations. Hence, $\neg b < c$ and $\neg c < b$.

Secondly, let us agree for a moment to van Dalen's interpretation according to which no information would be available about any possible relation between b and c. The temporary lack of information to prove the antecedent of an implication cannot assure, of course, the truth of the implication. To vacuously assert an implication, one has to prove the *absurdity* of the antecedent. So, to vacuously assert

$$\neg b < c \wedge \neg b = c \rightarrow c < b,$$

we would have to prove $\neg(\neg b < c \wedge \neg b = c)$, which in turn would require some non-available information. So in no way are the virtuality conditions vacuously satisfied.

Now, if, as we maintained, Brouwer defined his counterexample with the intention to preclude the existence of any relation between b and c, how could he believe that order to be virtual? We think that the reason for Brouwer's mistake is of a different nature from van Dalen's. It lies in Brouwer's imprecise use of symbolism and

terminology. In our symbolism of Sect. 7.2, where $<$ and $=$ are used to indicate the given order, an assertion $r < s$ ($r = s$) has the precise meaning that the pair (r, s) falls under the predicate $<$ ($=$). In Brouwer's symbolism, where $r < s$ ($r = s$) is an uninterpreted formula, the same assertion should be expressed in the form "$r < s (r = s)$ holds in the given order" or, in Brouwer's terminology, "$r < s (r = s)$ is an existing relation". It is quite natural, however, to write simply $r < s (r = s)$ as an abbreviation of the above assertion, relying upon the context for the distinction between mentioning a relation and asserting its existence. Brouwer used this convention without stating it explicitly. Unfortunately, he was working within a framework where intuition is not very stable in governing that tacit convention. In fact, he stated clause 2.7 and 2.8 in the form:

4.1 From the simultaneous incompatibility of relations $r < s$ and $s < r$ follows $r = s$.
4.2 From the simultaneous incompatibility of relations $r < s$ and $r = s$ follows $s < r$.

In a context where there is an interplay between existence and addability of relations, to say that a relation $r < s$ is incompatible may be ambiguous, without the specification of whether what is incompatible is the existence or the addability of $r < s$.

According to the above convention, in 4.1 "incompatibility" refers to existence, since $r < s$ is short for "$r < s$ is an existing relation". Brouwer certainly understood clauses 4.1 and 4.2 in this sense, when writing his paper. This is also confirmed by the fact that Brouwer also formulated the virtuality clauses in earlier papers where addability was not in question. But it happened that, when thinking his theorem over, he confused existence with addability and referred to the latter the "incompatibility" in 4.1 and 4.2. In other words, Brouwer, when writing the unpublished note, interpreted 4.1 as

4.3 From the simultaneous incompatibility of adding the relations $r < s$ and $s < r$ (in a non-contradictory manner) follows $r = s$.

Similarly for the interpretation of 4.2.

Keeping in mind this misinterpretation, we are able to understand the content of Brouwer's note completely. Given any virtual order, try to prove it is inextensible. Let $r = s$ be addable in a non-contradictory manner. Then, we can suppose without contradiction that $r = s$ holds in some extension of the given order. From the incompatibility of the simultaneous existence of an arbitrary two of relations $r = s$, $r < s$, $s < r$ it follows that neither $r < s$ nor $s < r$ can belong to such extension, so that none of them can hold in the given order either. This is all that is needed for concluding $r = s$ by means of 4.1, in its correct interpretation. But, understanding 4.1 in the wrong form 4.3, we are not able to reach the desired conclusion. The reason is expressed just in Brouwer's remark that the incompatibility of the simultaneous non- contradictory existence of an arbitrary two of the relations $r = s$, $r < s$, $s < r$ does not entail the incompatibility of the simultaneous non-contradictory addability

of an arbitrary two of these relations, i.e. the antecedent of implication 4.3. This is what is shown by Brouwer's counterexample, where each of relations $b = c$, $b < c$, $c < b$ is addable in a non-contradictory manner.

If our reconstruction is correct, Brouwer's error reduces itself to a trivial confusion. So we should expect that Brouwer became aware of his error very soon. In fact, not only did he never publish any correction to the incriminated theorem, he even asserted it again in some later papers. This fact does not agree, however, with Heyting's conclusion in the above-mentioned note to Brouwer (1927). Indeed Heyting argues that, in order to avoid the difficulty shown by his counterexample, Brouwer tried to change the definition of virtual order. Heyting's argument would seem to be supported by another unpublished document. In the margin of his copy of 1927, Brouwer reformulated the virtual clauses as follows:

4.4 "From the impossibility of deriving either one of the relations $r < s$ and $r > s$ from the definition of the ordered [species], it follows $r = s$".

4.5 "From the impossibility of deriving either one of the relations $r < s$ and $r = s$ from the definition of the ordered [species], it follows $r > s$".

He added the following remark:

4.6 "In this way is ruled out the circumstance in which, for some pair (a, b), it is impossible to derive any of the three relations $a = b$, $a < b$ and $a > b$ from the definition of the ordered [species]".

According to Heyting, clauses 4.4 and 4.5 would constitute a new definition of virtual order, in virtue of which Brouwer's theorems hold. He criticizes, however, the new definition, because it would seem to involve the metamathematical notion of derivability. Heyting also observes that in the Cambridge lectures Brouwer explicitly says that the virtuality conditions are to be interpreted according to 4.4 and 4.5.

Heyting's comment, however, is wrong. It is clear from our analysis that 4.4 and 4.5 do not define any new notion of virtual order, but merely explain the old one. When Brouwer realized his error, he tried to reformulate the virtuality clauses in such a way to avoid that ambiguity which caused his error. Indeed, 4.4 and 4.5 amount just to 4.1 and 4.2 with moreover the warning of referring incompatibility to existence. For, the existence of a relation $r < s$, i.e. its holding in the given order, can be established intuitionistically only by deducing $r < s$ from the definition of the given order, so that the impossibility of doing that amounts just to the incompatibility of the existence of $r < s$. Our account is in accordance with Brouwer's remark 4.6: if 4.1, 4.2 are understood in the correct way, for any pair (a, b) the impossibility of deducing $a = b \vee a < b \vee b < a$ is contradictory.

Brouwer's intention to stress the distinction between existence and addability is even more clear in the Cambridge lectures. There Brouwer points out that to give an order on a certain species means to create *criteria* for the relations $=$, $<$ by virtue of which the proper axioms hold. The validity of these is to be understood, Brouwer continues, in the sense that "one or more possibilities or impossibilities of deducing one of these relations from the criteria imply another possibility or impossibility of deducing one of these relations from the criteria".

The possibility or impossibility of deducing a relation from the criteria expresses, as already observed, the intuitionistic meaning of its existence or non-existence. In contrast with existence, Brouwer also makes addability explicit. Indeed, after defining a partial order to be inextensible if no further relations can be added to the existing ones in a non-contradictory manner, he adds between brackets: "which in this case means in such a way as to make it impossible to deduce contradictions from the old and the new relations under application of the properties (1)–(9) [the axioms of partial order]".

Observe that this explanation is an explicit formulation of our definition of negative addability. For, the existence of an extension $(\prec, \approx)$ of the given order $(<, =)$ such that $r \prec s$ is non-contradictory just when it is impossible to get a contradiction from the assumption that, for some predicates $\prec$, $\approx$ extending $<$ and $=$ with $r \prec s$ the axioms of partial orders hold.

We conclude that both in the pencil note 4.4–4.6 and in the Cambridge lectures the notion of virtual order is the same as in the original paper 1927, but a few remarks were added in the new formulation, for the sole purpose of avoiding possible misunderstandings. Accordingly, these remarks, which were put in brackets in the Cambridge lectures, were altogether removed from the later version of 1950.

At this point, it should be clear that, in spite of Heyting's criticism, no meta-mathematical or proof-theoretical notion of derivability is involved by Brouwer's explanations. The "deducibility" to which Brouwer refers is to be understood in the sense of the intuitionistic concept of informal proof. This concept underlies any intuitionistic mathematical notion and comes out explicitly whenever we try to fully explicate that.

7.5 A Minor Mistake in the Cambridge Lectures

We observe in passing that Brouwer's definition of inextensible order in the Cambridge lectures is incorrect. In fact, an inextensible order is defined there as a partial order to whose existing relations no further relations can be added in a non-contradictory manner. Now, with respect to the existing relations, a further relation is a non-existing one. And the impossibility of adding a non-existing relation means that any addable relation cannot be a non-existing one. This does not entail the existence of any addable relation, as would be required for proving the virtuality of any inextensible order (take, e.g. a relation whose existence depends upon an unsolved problem).

7.6 On Posy's Reconstruction

Posy tries to reconstruct Brouwer's (1927) by means of the notion of creative subject. In fact, he proposes to consider "existing relations" of type $p < q$ (or $p = q$) as those relations for which

6.1 $\exists n \vdash_n p < q$ (or $\exists n \vdash_n p = q$)

and "relations addable to the existing ones in a non-contradictory manner" as the ones for which

6.2 $\neg\exists n \vdash_n \neg p < q$ (or $\neg\exists n \vdash_n \neg p = q$)

In this way, inextensibility is expressed by the clauses

6.3 $\neg\exists n \vdash_n \neg p < q \rightarrow p < q, \neg\exists n \vdash_n \neg p = q \rightarrow p = q.$

By the principle of "christian charity" (accepted by Posy), 6.2 is equivalent to $\neg\neg p < q$ (or $\neg\neg p = q$), whence those in 6.3 are equivalent to

6.4 $\neg\neg p < q \rightarrow p < q, \neg\neg p = q \rightarrow p = q$

6.4 are the so-called stability clauses of the order. Thus, according to Posy, a partial order is inextensible if, and only if, it is stable.

Now, it is clear that a stable order is not necessarily virtual. (Take Brouwer's counterexample discussed above where the 6.4 vacuously holds for the pair (b, c)). Thus, his equivalence between virtuality and inextensibility does not hold in Posy's interpretation. Posy tries to prove it, but his proof is wrong. In fact, he infers virtuality from inextensibility by means of the following argument: The creative subject can prove a negative statement $\neg p < q$ or $\neg p = q$ only by means of the axiom

$$p < q \rightarrow \neg p = q \wedge \neg q < p$$

since this is the only axiom for partial orders which establishes a negation. However, the possibility of deducing a relation or its negation depends not on the general axioms of partial orders but on the definition of the given specific order (the *criteria* of the Cambridge lectures). Posy himself explicitly assumes that such a definition is known to the creative subject from the beginning (stage zero). Nor could it be otherwise. If only the general axioms of order were available to the creative subject, then, according to Posy's definition of existing relation, no relation of type $p < q$ would be existing, since the only relations deducible from the axioms are those of the form $p = p$.

It is true that in the Cambridge lectures Brouwer expresses addability in terms of deducibility from the axioms (and not from the *criteria*). But, as we saw, the reason of this is that Brouwer's addability means existence in some *extension* of the given order, while deducibility from *criteria* means existence in the given order itself. Brouwer's addability, however, is incompatible with Posy's one which, as we

observed, makes the addability of $p < q$ equivalent to $\neg\neg p < q$. Thus, Posy's argument is incorrect.

Posy also tries to distinguish two theories of order in Brouwer's work, one tensed (that of 1927 and of (1950) and the other intuitionistic (that of the Cambridge lectures). But all his discussion rests on a misunderstanding. He interprets Brouwer's "deducibility from *criteria*" as deducibility at the stage 0 of the activity of the creative subject, in contrast with deducibility along the whole course of time. So the version of the Cambridge lectures would be untensed, because of its reference to one and the same stage (stage 0). But, as we have shown, the relevant distinction is that between relations holding in the given order (deducible from the *criteria*) and relations holding in some extension of it (deducible from the axioms and the existing relations). And this distinction is neutral with respect to the time element, since it does not involve the very nature of such *criteria*. These may fully predetermine the existing relations, or they may very well make them depend upon some as yet unknown information. Brouwer himself gives in 1950 some examples of *criteria* referring to future information.

We conclude that the notions of virtual and inextensible order are the same everywhere in Brouwer's work. They are certainly tensed in Posy's sense that the negation occurring in all axioms is to be intended as strong negation and not as non-evidence at the present. An untensed theory would be of no use and, as Posy himself observes, would conflict with Brouwer's weak counterexamples to the linearity of the order of the *continuum*.

The time element is certainly present insofar as it is implicit in the intended meaning of intuitionistic logical constants, in particular of negation (in the usual strong sense). But it does not need any explication by means of the creative subject, neither in the formulation of the two notions of order nor in the proof of their equivalence.

References

Brouwer, L. (1927). Virtuelle ordnung und unerweiterbare ordnung. *Journal für die reine und angewandte Mathematik*, *157*, 255–257, also in Brouwer (1975, pp. 406–408).

Brouwer, L. (1950). Remarques sur la notion d'ordre. *Comptes Rendus de l'Académie des Sciences de Paris*, *230*, 263–265, also in Brouwer (1975, pp. 499–500).

Brouwer, L. (1975). Collected works. In A. Heyting (Ed.), *Philosophy and foundations of mathematics*. Amsterdam: Elsevier.

Brouwer, L. (1981). In van Dalen, D. (Ed.), *Cambridge lectures in intuitionism*. Cambridge: Cambridge University Press.

Heyting, A. (1975). *Notes in Brouwer*. Amsterdam: Elsevier.

Posy, C. J. (1980). On Brouwer's definition of unextendable order. *History and Philosophy of Logic*, *1*, 139–149.

van Dalen, D. (1981). *Notes in Brouwer*. Cambridge: Cambridge University Press.

Chapter 8
An Intuitionistic Notion of Hypothetical Truth for Which Strong Completeness Intuitionistically Holds

Abstract An intuitionistic notion of truth under a set of hypotheses is introduced in this chapter. By means of that, intuitionistic semantics is extended to a new semantics for which validity turns out to be equivalent to generalized validity. Strong completeness is proved intuitionistically.

8.1 Introduction

We call models *natural intuitionistic models* for *IPC* (intuitionistic predicate calculus) those formally defined as classical models, but interpreting logical constants intuitionistically. In the literature, such models are also called *intuitive models* (Troelstra 1977) or *internal models* (Dummett 1977).

Natural intuitionistic semantics is to be developed within an intuitionistic metamathematics. In this framework, including lawless sequences, it is provable that any formula of *IPC* is valid for natural semantics iff it is valid for Beth semantics. It follows that the problem of natural completeness is equivalent to that of Beth completeness within intuitionistic metamathematics.

It is well known (see Dummett 1977 or Troelstra 1977) that natural (as well as Beth) completeness is equivalent to Markov's principle. And since the latter goes beyond usual intuitionistic constructivism, natural completeness fails to hold intuitionistically. Furthermore, as we shall show, natural strong completeness is incompatible with the theory of lawless sequences.

However, Veldman and de Swart introduced certain generalizations of Kripke and Beth semantics respectively, for which intuitionistic completeness holds (De Swart 1976a, b; Veldman 1976). The peculiarity of such generalized models consists in allowing the absurdity $\perp$ to be true at some node.

The question arises whether these completeness results are of some significance for natural semantics. The problem is to give a plausible account for the possibility of being $\perp$ true is a natural model. Of course, an essential requirement for such an account is agreement with the general conception of intuitionistic truth. And since Intuitionism identifies truth with (informal) provability, the problem amounts to explaining what a proof of the absurdity is.

Now, there are two cases in which the working mathematician may prove the absurdity: when his reasoning is erroneous or when he argues under some false assumptions. The first case is of no theoretical interest, because the highly idealized notion of proof, on which intuitionistic truth is based, is free of errors. The second case, on the contrary, has a very important role also in the activity of the idealized mathematician, whose hypothetical reasonings may well rest on some unproved assumptions.

These considerations suggest the attempt to pursue the desired interpretation by relativizing truth (in a natural model) to a certain *set of hypotheses* and defining validity as truth in every model under every set of hypotheses.

At first sight, this way may seem to be inadequate to strengthening validity. For, if a sentence is true *tour court*, it seems to be a fortiori true under any set of hypotheses.

However, the force of this objection depends upon the way of understanding the notion of *truth under a set H of hypotheses*. Indeed, as we shall see, such a notion, unless H is finite, is not fully determined by the usual intuitionistic acceptation of the term *hypothesis*.

In the present paper, we will propose an intuitionistic definition of *hypothetical truth* which escapes the above objection and succeeds to reach the desired goal. By means of that we will strengthen the relation of semantical consequence for natural semantics, so as to make it equivalent to the corresponding relation for generalized Beth semantics.

8.2 Symbolism and Conventions

L is an intuitionistic first-order language without function symbols.

A, B, C are L-sentences (closed formulae), Γ, Δ, H sets of L-sentences. H will be used when its members play the role of *hypotheses* (in the sense of this paper).

$\bot$ (absurdity) is included among logical constants and negation is defined by $\neg A := A \to \bot$.

All proof-theoretical terms and symbols as *derivation*, $\vdash \cdots$ refer to *IPC* (intuitionistic predicate calculus) and all semantical terms and symbols, as model, truth, validity, $\models \cdots$, when lacking further specifications, refer to natural semantics.

We indicate by M (with possible subscripts) a natural model and by $|M|$ its domain. $L(M)$ is the language extended by individual constants for all members of $|M|$.

All our definitions, theorems, proofs are to be read intuitionistically.

Sets are understood according to the Brouwerian concept of *species*. In particular, a set of sentences is given provided we know what counts as a proof that a sentence belongs to it. So sets of sentences may fail to be decidable or even enumerable.

We use $\subset$ for weak inclusion: $\Gamma \subset \Delta$ iff$_{df}$, for all A, $A \in \Gamma \Rightarrow A \in \Delta$.

A set is *finite* when there is a (constructive) map from an initial segment of ω (the set of natural numbers) onto it. So, if Γ is finite, one has to be able to write down all sentences belonging to it.

For finite Γ, we indicate by $\wedge\Gamma$ the conjunction of all its members, if any, and put $\wedge\phi := \neg\bot$.

$u, v, w, \ldots$ are used for finite sequences of natural numbers. If $u = \langle u_0, u_1, \ldots, u_{n-1}\rangle$, $lth\ u := n$ and, if $k \in \omega$, $u\ k := \langle u_0, \ldots, u_{n-1}, k\rangle$

$\alpha, \beta, \gamma, \ldots$ are used for lawless sequences with abbreviations $\overline{\alpha}n := \langle \alpha_0, \alpha_1, \ldots, \alpha_{(n-1)}, \alpha \in u := \overline{\alpha}(lth\ u) = u$.

We use $\equiv$ for *strict* (intensional) equality.

8.3 The Failure of Strong Completeness for Natural Semantics

We show that the compactness of relation $\Gamma \models A$ of semantical consequence is inconsistent with the theory of lawless sequences.

Theorem 8.1 *Not for all* Γ, A, *if* $\Gamma \models A$ *then, for some finite* $\Gamma' \subset \Gamma$, $\Gamma' \models A$.

Proof For every lawless sequence α, define $\Gamma_\alpha = \{A : A \equiv \bot \wedge \exists x \alpha x = 0\}$. Since $\forall\alpha\neg\neg\exists x\alpha x = 0$, for all α, we have $\neg\neg(\bot \in \Gamma_\alpha)$. So, for all α and all M, $M \not\models \Gamma_\alpha$ and hence, vacuously, $\Gamma_\alpha \models \bot$. But, since $\neg\forall\alpha\exists x\alpha x = 0$ and hence $\neg\forall\alpha\bot \in \Gamma_\alpha$, not for every α there is a finite $\Gamma'_\alpha \subset \Gamma_\alpha$ such that $\Gamma' \models \bot$.[1]

Corollary 8.1 (Failure of strong completeness) *Not for all* Γ, A, *if* $\Gamma \models A$, *then* $\Gamma \vdash A$.

8.4 Hypothetical Truth

Let M be an L-model and H a set of $L(M)$-sentences, which will be called *hypotheses*. We define the main notion of the paper, i.e. the truth of an $L(M)$-sentence A in M under the set H of hypotheses, briefly the H- *truth* of A in M, for which we use the notation $M \models^H A$.

Definition 8.1 $M \models^H A$ is defined by induction on A:

(a) for atomic A, $M \models^H A$ iff, for some finite $H' \subset H$, if $M \models \wedge H'$ then $M \models A$;

[1] Prof. Troelstra suggested to me by letter the following interesting variant of the proof, which shows how compactness for natural semantics would imply the validity of the excluded middle principle. Given any sentence A and any model M, define $\Gamma_A = \{B : B \equiv \bot$ and $M \models A \vee \neg A\}$. Since $M \models \neg\neg(A \vee \neg A)$ and hence not $M \not\models A \vee \neg A$, we have $\neg\neg(\bot \in \Gamma_A)$. Thus, for all model M', $M' \not\models \Gamma_A$, whence vacuously, $\Gamma_A \models \bot$. Assuming compactness, we get $\bot \in \Gamma_A$ and therefore $M \models A \vee \neg A$.

(b) $M \models^H A \wedge B$ iff $M \models^H A$ and $M \models^H B$;
(c) $M \models^H A \vee B$ iff, for some finite $H' \subset H$, if $M \models \wedge H'$, then $M \models^H A$ or $M \models^H B$;
(d) $M \models^H A \to B$ iff $M \models^H A \Rightarrow M \models^H B$;
(e) $M \models^H \forall x A(x)$ iff, for all $d \in |M|$, $M \models^H A(d)$;
(f) $M \models \exists x A(x)$ iff, for some finite $H' \subset H$, if $M \models \wedge H'$, then for some $d \in |M|$, $M \models^H A(d)$.

The intuitive idea underlying the above definition is that, when H-proving an interpreted sentence A, at every stage we are allowed to assume finitely many sentences of H. In clauses (b), (d), (e) explicit reference to H is omitted, since it would be redundant.

Definition 8.2 Let H, Γ be sets of L-sentences and A an L-sentence.

(a) A is *H-consequence* of Γ, $\Gamma \models^H A$, iff$_{df}$, for all M, $M \models^H A$ provided $M \models^H \Gamma$ (i.e. $M \models^H B$ for all $B \in \Gamma$).
(b) A is *H-valid*, $\models^H A$, iff$_{df}$ $\phi \models^H A$.
(c) A is *hypothetical consequence* of Γ, $\Gamma \models^H A$, iff$_{df}$, for all H, $\Gamma \models^H A$.
(d) A is *hypothetically valid*, $\models^H A$, iff$_{df}$ $\phi \models^H A$.

When all hypotheses are decidable in M, the definition of H-truth can be simplified.

Definition 8.3 H is *decidably built* in M iff, for all $A \in H$, $M \models A \vee \neg A$.

The following proposition is easily verified, by induction on A:

Proposition 8.1 *Let H be decidably built in M.*

(a) For atomic A, $M \models^H A$ iff either $M \models A$ or, for some $B \in H$, $M \models \neg B$
(b) $M \models^H A \wedge B$ iff $M \models^H A$ and $M \models^H B$;
(c) $M \models^H A \vee B$ iff $M \models^H A$ or $M \models^H B$;
(d) $M \models^H A \to B$ iff $M \models^H A \Rightarrow M \models^H B$;
(e) $M \models^H \forall x A(x)$ iff, for all $d \in |M|$, $M \models^H A(d)$;
(f) $M \models^H \exists x A(x)$ iff, for some $d \in |M|$, $M \models^H A(d)$.

Corollary 8.2 *If H is decidably built in M, there is a model M' such that, for all $\perp$-free A, $M \models^H A$ iff $M' \models A$.*

Proof Define M' by $|M'| = |M|$ and, for all atomic $A \neq \perp$, $M' \models A$ iff$_{df}$ $M \models^H A$.

Corollary 8.3 *Let H be decidably built in all models and let A and all sentences of Γ be $\perp$-free. Then $\Gamma \models^H A$ iff $\Gamma \models A$.*

Formal derivation in IPC respects hypothetical consequence:

Theorem 8.2 (Soundness) *If $\Gamma \vdash A$ then $\Gamma \models^H A$.*

Proof Let M be a model and H a set of L-sentences. If Γ is a set of L-formulae, A an L-formula and a an assignment of elements of $|M|$ to the free variables of Γ and A, we indicate by Γ^a and A^a the set of $L(M)$-sentences and the $L(M)$-sentence obtained from Γ and A by means of a. It is easily verified that, if $\Gamma \vdash A$, then, for every assignment a, $\Gamma^a \models^H A^a$. The proof runs straightforwardly by induction on the derivation of A from Γ.

8.5 Remarks on Hypothetical Truth

Some comments on our notion of hypothetical truth are in order. It is worthwhile to compare $\models^H A$ with the popular relations $H \models A$ and $H \vdash A$, since all three express some sort of deducibility of A from H.

Notation 8.1 $[H \models A]_M$ for $M \models H \Rightarrow M \models A$.

Lemma 8.1 *If* $M \models H$, *then* $M \models^H A$ *iff* $M \models A$.

Proof Since, for all finite $H' \subset H$, $M \models H'$, the definition of $M \models^H A$ reduces to that of $M \models A$.

Proposition 8.2 *If* H *is finite, then* $M \models^H A$ *iff* $[H \models A]_M$.

Proof Take $H' = H$ in the inductive definition of $M \models A$.

Proposition 8.3 *If* $M \models^H A$, *then* $[H \models A]_M$.

Proof Apply Lemma 8.1.

Using lawless sequences, we can show that the converse fails in a very strong form:

Proposition 8.4 *There are* A, H *and* M *such that* $H \models A$ *but not* $M \models^H A$.

Proof Suppose L has infinitely many individual constants, say numerals $\overline{0}, \overline{1}, \overline{2}, \ldots$ Let α be a lawless sequence and P a unary predicate symbol. Define $H = \{P(\overline{n}) : n \in \omega\} \cup \{\neg\forall x P(x)\}$ and $A \equiv \neg\forall x P(x)$, so that trivially $H \models A$.

Let M be a model, with domain ω, in which P is interpreted by: $M \models P(\overline{n})$ iff$_{df.}$ $\exists x > n\, \alpha x = 0$. We have trivially $M \models^H \forall x P(x)$. From $\neg\forall x \exists y > x \alpha y = 0$ it follows $M \models \neg\forall x P(x)$ and, as $\forall x \neg\neg \exists y > x\, \alpha y = 0$, $M \models \neg\neg P(\overline{n})$, for all n.

Hence, for every finite $H' \subset H$, $M \models \neg\neg \wedge H'$. Thus $M \not\models^H \bot$ and therefore $M \not\models^H \neg\forall x P(x)$.

We conclude that

Proposition 8.5 $\models^H A$ *is strictly stronger than* $H \models A$.

For H, A as in the proof of Proposition 8.4, it holds also $H \vdash A$. It follows that relation $H \vdash A$ is not stronger than $\models^H A$. On the other hand, $\models^H A$ is not stronger than $H \vdash A$, otherwise, for $H = \phi$, it would hold completeness for natural semantics. Thus

Proposition 8.6 *Relations* $\models^H A$ *and* $H \vdash A$ *are incomparable.*

In a certain sense, however, formal derivability may be regarded as the formal counterpart of hypothetical truth. For, since $H \vdash A$ expresses the derivability of A from some *finite* subset of H, from Proposition 8.2 it follows that

Proposition 8.7 *If* $H \vdash A$, *then, for some finite* $H' \subset H$, $\models^{H'} A$.

The above propositions explain the main difference between the two concepts of deducibility represented by

$$M \models^H A \tag{8.1}$$

and

$$[H \models A]_M \tag{8.2}$$

respectively.

For finite H, both relations express deducibility according to the usual intuitionistic sense of implication: you can assert both $M \models^H A$ and $[H \models A]_M$ when you have a method of transforming any proof of $\wedge H$ into a proof of A. This concept of deducibility from a finite set of (interpreted) sentences is extrapolated for general sets of sentences into different directions by the two relations.

(8.2) interprets deducibility of A from H as an implication whose antecedent is the universal quantification on all sentences of H and whose subsequent is A. Accordingly, you can assert $[H \models A]_M$ when you have a method of transforming any proof of $(\forall B \in H)B$ into a proof of A.

In contrast, (8.1) does not involve the idea of simultaneous provability of all sentences of H, but only of finitely many of them at a time. The H-truth of A is conceived as the provability of A, when all provability conditions of A and of its subsentences are relativized to suitable assumptions of the form $\wedge H'$, with finite $H' \subset H$. In other words, the assumption of the whole H is not conceived as the assumption of its sentences all at once, but rather as a (possibly infinite) process, in the course of which some sentences from H may be assumed step by step.

This difference is quite remarkable with respect to the two notions of *contradictory set of (interpreted) sentences* expressed by

$(1')$ $[H \models \bot]_M$

and

$(2')$ $M \models^H \bot$

According to $(1')$, H is regarded as contradictory provided $\neg\forall B \in H(M \models B)$, i.e. provided it is impossible to prove the sentences of H all at once, even if we know that no sentence of H is refutable, as for the H and M_α in the proof of Proposition 8.4.

In contrast, we can assert $(2')$ only after having found in H finitely many incompatible sentences.

On this respect, (8.1) seems to be more interesting than (8.2), especially when dealing with *undetermined* sets of hypotheses.

For instance, if H is a lawless sequence of sentences, we can assert $(1')$ even without knowing any sentence of H. So $H \models A$ trivially holds for every A. On the contrary, the set $\{A : M \models^H A\}$ is far from being trivial.

The example in the proof of Proposition 8.4 points out also another peculiar aspect of hypothetical truth: it may happen that $\not\models^H A$, even though $A \in H$. This feature may seem an oddity, since one would expect A to be trivially deducible from itself. But, on reflection, one should agree that this feature is in agreement with the intuitionistic intended meaning of implication and, in particular, of negation.

For, the truth of $\neg B$ expresses the absurdity of finding an *absolute* proof of B. But by no means it entails the absurdity of finding a proof of B with the help of some suitable set of hypotheses. Therefore you have no right to assert the H-truth of a sentence on the basis of its mere truth. The more a set of hypotheses helps to prove a sentence, the more it hampers to prove its negation. Now, the meaning of $A \in H$ is that, when proving H-truths, we are allowed to suppose to have an *absolute* proof of A, but this supposition in general does not enable us to find an H-proof of A. What the example in Proposition 8.4 shows is just that, as a matter of fact, at least when dealing with lawless parameters, such a situation can be realised.

These considerations also explain how our notion of *hypothetical truth* can escape the initial objection, according to which true sentences would be a fortiori true under any set of hypotheses.

Later on, we will be concerned with relation $\Gamma \models^H A$. In Sect. 8.7, we will prove it to be equivalent to the relation of semantical consequence for generalised Beth semantics. Here we want to illustrate the action of the "understood hypotheses", by showing as our counterexample to compactness of $\Gamma \models A$, in the proof of Theorem 8.1, does not apply to $\Gamma \models^H A$.

Take Γ_α as in the proof of Theorem 8.1. Suppose $\Gamma_\alpha \models^H \bot$. This means that, for all M, H, if $M \models^H \Gamma_\alpha$ then $M \models \bot$. Taking $H = \Gamma_\alpha$, we have trivially $M \models^{\Gamma_\alpha} \Gamma_\alpha$ and hence $M \models^{\Gamma_\alpha} \bot$. It follows that $\bot \in \Gamma_\alpha$, so that $\{\bot\}$ is the finite subset of Γ_α required by compactness.

8.6 Generalized Beth Semantics

We recall here the main facts about generalised Beth models, for short *GB-models*.

A GB-model $\mathcal{M} = \langle D, T, I\rangle$ consists of an inhabited set D, a spread T and a binary relation I between nodes of T and atomic $L(D)$-sentences ($\bot$ included).

We use $u, v, w, \ldots$ for nodes of T and write $u \leq v$ when u is an initial segment of v.

Definition 8.4 The *forcing* relation between T-nodes and $L(D)$-sentences is inductively defined as follows:

(a) for atomic A, $u \Vdash A$ iff$_{df}$ there is a bar $\mathcal{B}$ of u such that, for all $v \in \mathcal{B}$, either $I(v, A)$ or $I(v, \bot)$;
(b) $u \Vdash \forall x A(x)$ iff$_{df}$, for all $d \in D$, $u \Vdash A(d)$ (likewise for $A \wedge B$);
(c) $u \Vdash \exists x A(x)$ iff$_{df}$ there is a bar $\mathcal{B}$ of u such that, for all $v \in \mathcal{B}$, $v \Vdash A(d)$, for some $d \in D$ depending on v (likewise for $A \vee B$);
(d) $u \Vdash A \rightarrow B$ iff$_{df}$, for every $v \geq u$, if $v \Vdash A$ then $v \Vdash B$.

Definition 8.5 A node u of a GB-model *explodes* iff$_{df}$ $u \Vdash \bot$.

It is easily seen that if u explodes then $u \Vdash A$, for all A. A Beth model may be viewed as a GB-model without exploding nodes.

Definition 8.6

(a) $\mathcal{M} \Vdash A$ iff$_{df}$ $\langle\rangle \Vdash A$ (where $\langle\rangle$ is the top of $\mathcal{M}$).
(b) $\Gamma \Vdash A$ iff$_{df}$, for every $\mathcal{M}$, if $\mathcal{M} \Vdash \Gamma$ then $\mathcal{M} \Vdash A$.
(c) $\Vdash A$ iff$_{df}$ $\phi \Vdash A$.

Theorem 8.3 (Soundness) *If* $\Gamma \vdash A$, *then* $\Gamma \Vdash A$.

Theorem 8.4 (Strong completeness) *If* $\Gamma \Vdash A$, *then* $\Gamma \vdash A$

Soundness is verified straightforwardly. For enumerable Γ, strong completeness has been proved by Veldman (1976), De Swart (1976a) and, in a simpler way, by Friedman Troelstra and Van Dalen (1988). I do not know whether the general form Theorem 8.4 (without restrictions of Γ) is available in the literature. Anyway we shall prove it in Sect. 8.8.

8.7 Connection Between Hypothetical Semantics and GB-Semantics

Hypothetical semantics is connected with GB-semantics in a similar way as natural semantics is connected with Beth semantics.

Let $\mathcal{M} = \langle D, T, I\rangle$ be a GB-model. We denote by $\alpha, \beta, \gamma, \ldots$ lawless sequences in T. With every α, we associate the natural model M_α, such that $|M_\alpha| = D$ and, for every atomic sentence $A \not\equiv \bot$, $M \models A$ iff$_{df}$, for some n, $I(\overline{\alpha}n, A)$. Besides, we associate with α the set of hypotheses $H_\alpha = \{A : A \equiv \bot$ and, for some n, $I(\overline{\alpha}n, \bot)\}$.

Theorem 8.5 *Let u be a node of $\mathcal{M}$. $u \Vdash A$ iff, for all $\alpha \in u$, $M_\alpha \models^H A$.*

Proof By induction on A. We develop two cases (the others being similar):

(a) A atomic. $u \Vdash A \Leftrightarrow$ for all $\alpha \in u$, there is n such that $I(\overline{\alpha}n, A)$ or $I(\overline{\alpha}n, \bot) \Leftrightarrow M_\alpha \models^H A$ for all α.

(b) $A \equiv B \to C$. Suppose $u \Vdash B \to C$. If $\alpha \in u$ and $M_\alpha \models^H B$, then there is $m \geq lth\ u$ such that $M_\beta \models^H B$, for all $\beta \in \overline{\alpha}m$ (principle of *open data*). By the induction hypothesis, $\overline{\alpha}m \Vdash B$ and therefore $\overline{\alpha}m \Vdash C$ whence $M_\alpha \models^H C$. Thus, $M_\alpha \models^H B \to C$.

Conversely, suppose $M_\alpha \models^H B \to C$ for all $\alpha \in u$. If $v \geq u$ and $v \Vdash B$, then, by the induction hypothesis, for all $\beta \in v$, $M_\beta \models^H B$ and therefore $M_\beta \models^H C$, whence $v \Vdash C$. Thus, $u \Vdash B \to C$.

Corollary 8.4 *If* $\Gamma \models^H A$ *then* $G \Vdash A$.

The converse follows immediately from Theorems 8.4 and 8.3. Thus

Theorem 8.6 $\Gamma \models^H A$ *iff* $\Gamma \Vdash A$.

From Theorems 8.6 and 8.4, we get strong completeness for hypothetical semantics:

Theorem 8.7 $\Gamma \models^H A$ *iff* $\Gamma \vdash A$

Observe that H_α-s used in Theorem 8.5 are decidably built, since their only possible member is $\bot$. So Corollary 8.4, Theorems 8.6 and 8.7 follow also restricting hypothetical semantics to decidably built set of hypotheses. Thus

Proposition 8.8 $G \models^H A$ *iff, for all* M *and all* M*-decidably built* H*, if* $M \models \Gamma$*, then* $M \models A$.

From Proposition 8.8 and Corollary 8.3, it follows strong completeness for natural semantics restricted to $\bot$-free sentences:

Theorem 8.8 *Let* A *and all sentences of* Γ *be* $\bot$*-free. Then* $\Gamma \models A$ *iff* $\Gamma \vdash A$.

8.8 A Strong Completeness Proof for GB-Semantics

We prove here strong completeness for *GB*-semantics, by refining Friedmans's completeness proof presented by Troelstra and Van Dalen (1988, Chap. 13, Sect. 2)

Let $c_1, c_2, \ldots, c_n, \ldots$ be a sequence of infinitely many individual constants not occurring in L. Define the hierarchy of languages $L_0 = L$, $L_1 = L_0 \cup \{c_1\}, \ldots, L_{n+1} = L_n \cup \{c_n\}, \ldots; L_\omega = \cup_{n \in \omega} L_n$.

Let $A_0, A_1, \ldots, A_n, \ldots$ be an enumeration of all disjunctive or existential L-sentences such that, for all n, A_n is an L_n-sentence with infinitely many repetitions.

We denote by T the full binary spread and by $u, v, w, \ldots$ its nodes (finite $0-1$ sequences).

Let Γ be a given set of L-sentences. We will assign to every u a set of Γ_u of L_n-sentences, where $n = lth\ u$. Γ_u is defined by induction on $lth\ u$, as follows:

(i) $\Gamma_{\langle\rangle} = \Gamma$

(ii) $A \in \Gamma_{uk}$ iff$_{df}$ either

(a) $A \in \Gamma_u$, or

(b) $A_n \equiv \exists x B(x)$, $\Gamma_u \vdash A_n$ and $A \equiv B(c_{n+1})$; or

(c) $A_n \equiv B \vee C$, $\Gamma_u \vdash A_n$ and $A = \begin{array}{ll} B & \text{if } k = 0 \\ C & \text{if } k = 1 \end{array}$

Lemma 8.2 *Let A be an L_n-sentence, with $n = lth\ u$. If, for all $k = 0, 1$, $\Gamma_{u\,k} \vdash A$, then $\Gamma_u \vdash A$.*

Proof Suppose $A_n \equiv \exists x b(x)$. By hypothesis, there are derivations D_0, D_1 of A from $\Gamma_{u\,0}$ and $\Gamma_{u\,1}$, respectively. If $B(c_{n+1})$ does not occur among the assumptions of D_0, then D_0 is a derivation of A from Γ_u. Otherwise $B(c_{n+1}) \in \Gamma_{u\,0}$, so that either $B(c_{n+1}) \in \Gamma_u$ and therefore $\Gamma_{u\,0} = \Gamma_u$ or $\Gamma_u \vdash \exists x B(x)$. In the latter case, since c_{n+1} does not occur neither in Γ_u nor in A, A is derivable from $\Gamma_u \cup \{\exists x Bx\}$ and hence from Γ_u.

Similarly for $A_n \equiv B \vee C$.

Corollary 8.5 *Let A be an L_n-sentence with $n = lth\ u$. $\Gamma_u \vdash A$ iff there is a bar $\mathcal{B}$ of u such that, for all $v \in \mathcal{B}$, $A \in \Gamma_v$.*

Proof If $\Gamma_u \vdash A$ take $m \geq n$ such that $A_m = A \vee A$. For all $v \geq u$, with $lth\ v = m$, $A \in \Gamma_{v\,k}$.

For the converse, apply the lemma and the fan theorem.

Define the GB-model $\mathcal{M} = \langle D, T, I\rangle$, where D is the set of all individual constants of L and, for every atomic L_ω-sentence A, $I(u, A)$ iff$_{df}$ $A \in \Gamma_u$.

Theorem 8.9 *Let A be an L_n-sentence, with $n = lth\ u$. $u \Vdash A$ iff $\Gamma_u \vdash A$.*

Proof By induction on A:

(a) A atomic. Apply Corollary 8.5.

(b) $A \equiv \exists x B(x)$. Suppose $u \Vdash A$ and let $\mathcal{B}$ be a bar of u such that, for all $v \in \mathcal{B}$, there is $d \in D$ such that $v \Vdash B(d)$. Then, for all $w \geq v$ with $B(d) \in L_{lth\ w}$, $w \Vdash B(d)$ and, by the induction hypothesis, $\Gamma_w \vdash B(d)$ and hence $\Gamma_w \vdash \exists x B(x)$. Thus, $\Gamma_u \vdash \exists x B(x)$ by Corollary 8.5.
Viceversa, suppose $\Gamma_u \vdash A$. Take $m \geq n$ such that $A_m \equiv \exists x B(x)$. If $v \geq u$ with $Ith\ v = m$, then $B(C_{m+1}) \in \Gamma_{v\,k}$ for all k. By the induction hypothesis, $v\,k \Vdash B(c_{m+1})$, whence $u \Vdash \exists x B(x)$.

(c) $A \equiv B \vee C$. Similar to (b).

(d) $A \equiv \forall x B(x)$. Suppose $u \Vdash \forall x B(x)$. Let C be any L-tautology and take $m \geq n$ such that $A_m \equiv C \vee C$. If $v \geq u$ with $ltbv = m$, we have $v\,k \models \forall x B(x)$, from which $v\,k \models B(c_{m+1})$. By the induction hypothesis, $\Gamma_{v\,k} \vdash B(c_{m+1})$ and, since $\Gamma_{v\,k} = \Gamma_v \cup \{C\}$, $\Gamma_v \vdash \forall x B(c_{m+1})$. Therefore, as c_{m+1} does not occur in Γ_v, $\Gamma_v \vdash \forall x B(x)$. By Corollary 8.5, $\Gamma_u \vdash \forall x B(x)$.
Viceversa, suppose $\Gamma_u \vdash \forall x B(x)$. Given $d \in D$, for all $v \geq u$ with $d \in L_{lth\ v}$, we have $\Gamma_v \vdash B(d)$ and hence $v \Vdash B(d)$. Therefore $u \Vdash B(d)$.

(e) $A \equiv B \wedge C$. Similar to (d).

(f) $A \equiv B \to C$. Suppose $u \Vdash B \to C$. Take $m \geq n$ such that $A_m \equiv B \vee E$, where E is an L-tautology. For all $v > u$, with $lth\ v = m$, $B \in \Gamma_{v\,0}$ so that $v\,0 \Vdash B$ and hence $v\,0 \Vdash C$. By the induction hypothesis $\Gamma_{v\,0} \vdash C$ and, since $\Gamma_{v\,0} = \Gamma_v \cup \{B\}$, $\Gamma_v \vdash B \to C$. Thus $\Gamma_u \vdash B \to C$. The converse is trivial.

Corollary 8.6 *For every L-sentence A,* $\Gamma \vdash A$ *iff* $\mathcal{M} \Vdash A$.

References

De Swart, H. (1976). Another intuitionistic completeness proof. *Journal of Symbolic Logic*, *41*, 644–662.

De Swart H (1976) Intuitionistic logic in intuitionistic metamathematics. Ph.D. thesis, Katholieke Universiteit, Nijmegen.

Dummett, M. (1977). *Elements of intuitionism*. Oxford: Clarendon Press.

Troelstra, A. (1977). *Choice sequences*. Oxford: Oxford University Press.

Troelstra, A., & Van Dalen, D. (1988). *Constructivism in mathematics* (Vol. II). Amsterdam: North-Holland.

Veldman, W. (1976). An intuitionistic completeness theorem for intuitionistic predicate logic. *Journal of Symbolic Logic*, *41*, 159–176.

Chapter 9
Propositions and Judgements in Martin-Löf

with G. Usberti

Abstract It is considered Martin-Löf's distinction between propositions and judgements. It is argued that propositions can be regarded as the only fundamental entities of logic, since all mathematical activity may be analysed in terms of the creation and demonstration of propositions.

9.1 Introduction

In a number of papers Martin-Löf argues that at the basis of an appropriate foundation of logic a strong distinction should be made between two primitive notions: the notion of proposition and the notion of judgement. These notions, which he uses systematically in his intuitionistic type theory, are amply illustrated in Martin-Löf (1985), where they are taken as fundamental to an intuitionistic explanation of the logical constants, and in Martin-Löf (1987), where the distinction between proposition and judgement is given within as more general conceptual framework. We will be considering only these two papers since it is here that Martin-Löf attempts to justify the distinction and to characterise the two notions philosophically, independently of their role in any specific formal system.

9.2 Propositions and Judgements

At the beginning of Martin-Löf (1985), Martin-Löf expresses the need to distinguish propositions from judgements in the following terms. If A and B are propositions, an inference rule, for example the rule of conjunction introduction, is usually stated in the form

$$\frac{A \quad B}{A \wedge B} \tag{9.1}$$

However, this formulation is incorrect, says the author, since the rule does not take us from the *propositions* A, B to the proposition $A \wedge B$, but it takes us from the

affirmations of A and of B to the affirmation of $A \wedge B$. Using Frege's assertion sign, he continues, the rule is correctly stated as

$$\frac{\vdash A \qquad \vdash B}{\vdash A \wedge B} \tag{9.2}$$

This means that the rules of inference operate not on propositions but on assertions, and this, according to Martin-Löf, necessitates a distinction being made between entities which the logical operations operate on, i.e. propositions, and those which inference rules apply to, i.e. *assertions* or *judgements*. However, he does not adhere to Frege's notation but prefers to write the conjunction introduction rule in the form

$$\frac{A \text{ true } B \text{ true}}{A \wedge B \text{ true}} \tag{9.3}$$

where "A true" (short for "A is true") is one of the fundamental forms of judgement.

It would seem at this point that the basic reason for the incorrectness of (1) and the correctness of (2) or (3) rests, for Martin-Löf as for Frege, in that while judgements have assertoric force, propositions do not, if we conceive them as mere thoughts which, while they may be true or false, in themselves do not assert their own truth or falsity. But this is not so. To see why let us take a closer look at Frege's position.

To judge, for Frege, is to recognise a thought as true; the judgement is this act of recognition. Frege distinguishes several different types of act having a thought as their common (intentional) object (Fig. 9.1).

Note that Dummett insists that expressing is not a linguistic act, if by this we mean the result of accompanying the utterance of a thought with a certain force (assertoric, interrogative, imperative, etc.): to express a thought does not have any force (see Dummett 1981, p. 494)

The act of recognising the truth of the thought that A is expressed in a natural language by the act of asserting the sentence expressing it:

> We express acknowledgement of truth in the form of an assertoric sentence (Frege 1984, p. 356).

Mental Act	Linguistic Act	Ideographic Expression
to grasp	(to express)	$-A$
to judge	to assert	$\vdash A$
(to take as a hypothesis)	to assume	
(to ask)	(to ask)	
to connect	to connect	$\overset{A}{\top}, \overset{B}{\llcorner A}$
⋮	⋮	⋮

Fig. 9.1 Acts not explicitly named by Frege are indicated in brackets

In the ideographic language, this is indicated by the sign $\vdash$, composed of the content stroke $-$ and the judgement stroke |. The content stroke applies to names (remember that for Frege sentences too are names) and functions thus:

- if A is the name of a judgeable content (in practice, a sentence), then $-A$ is true iff A denotes the truth;
- if A is a different name, then $-A$ is false.

The judgement stroke is a *sui generis* expression which applies to sentences and indicates their assertoric use. It is important to appreciate the significance of this Fregean choice, which at first sight might appear strange. What does it mean that the judgement stroke is *sui generis*? Essentially that it is *not* a predicate. In this way, Frege goes beyond the position he had previously taken in the *Begriffsschrift*, where | was considered as "the common predicate of all judgements"; and the reason is that if the judgement stroke is conceived as a predicate, then it can be used to construct formulas which shall necessarily be able to be used *without* assertoric force (for instance, as disjuncts in a disjunction or antecedents in an implication).

It is therefore the need to preserve the assertoric force of the verbal expression of judgements which induces Frege not to allow that judgements be expressed by formulas of the language which the propositions being judged belong to.

In conclusion, for Frege there is only one type of *entities*: thoughts or, in Russell's terminology, propositions. This, however, does not prevent him from retaining (9.2) as the correct formulation of the inference rules, rather than (9.1), since the propositions which appear in the rules are asserted rather than simply expressed as when two propositions are connected by means of a connective. On the other hand, not even the style of (9.3) would be seen as correct by Frege, if "A true" is considered as a formula of the object language. (The assertion stroke seems equivalent to a metalinguistic or pragmatic predicate, although this is not said by Frege).

Let us now turn to Martin-Löf. Bearing in mind that in the Scholastic tradition judgements were acts of judging and propositions the verbal expressions of judgements, he observes that beginning with Kant and successively with Frege and Russell there have been considerable changes in terminology; the old propositions are now called judgements. In this way, it happened that "the term judgement became ambiguous between the act of judging and that which is judged, or the judgement made […]" (Martin-Löf 1985, p. 207). And this is the use of the term "judgement" which Martin-Löf essentially adheres to. For him judgements, in so far as they are objects of knowledge, are not necessarily evident, but are only candidates for becoming evident under favourable circumstances. Martin-Löf explicitly poses the problem of whether a judgement is such even before it is recognised or proved. He answers that there is a sense in which it is and offers the following example. Let G be Goldbach's conjecture, according to which every even number greater than 2 is the sum of two prime numbers. In this case, says Martin-Löf, "G true" is not a judgement in the sense that we do not know if G is true, but it certainly is a judgement in the sense that we understand perfectly what it means that G is true. Thus, Martin-Löf explicitly admits the existence of judgements which are not evident and introduces the term "evident judgement" for a judgement that has been proved.

It seems to us important to point out that the judgements of Martin-Löf, unlike those of Frege, do not carry assertoric force by themselves but are only capable to acquire it when they have been proved. This undermines Frege's motivation for distinguishing judgements from propositions as the appropriate objects of inference in virtue of their assertoric force. Without this, judgements are reduced to Fregean thoughts which in turn can be identified, presumably, with Martin-Löf's propositions. In the above example, if "G true" expresses a judgement for the mere reason that we understand what it means that G is true, equally G itself must be considered a judgement, since we can understand what it means that every even number greater than 2 is the sum of two prime numbers. G does not seem to be an entity intrinsically different from the judgement that G is true.

It is not clear how far Martin-Löf is aware of this considerable difference between his conception of judgement and Frege's. With regard to this, the following passage is significant:

> The closest possible correspondence between the analysis that I am giving and Frege's notation for a judgement
>
> $$\vdash A$$
>
> is obtained by thinking of the vertical, judgement stroke as carrying the epistemic force
>
> I know...
>
> and the horizontal, content stroke as expressing the affirmation
>
> ...is true
>
> Then it is the vertical stroke which is superfluous, whereas the horizontal stroke is needed to show that the judgement has the form of an affirmation. But this can hardly be read out of Frege's own account of the assertion sign; you have to read it into his text (Martin-Löf 1985, pp. 226–227).

Here Martin-Löf suggests that his conception of judgement may be read between the lines in the work of Frege. But from what we have said above it follows that for Frege there is no sense in speaking of judgements which are not evident, and for him the horizontal stroke alone does not express any judgement.

9.3 Truth and Evidence

Having seen that Martin-Löf's judgements and propositions cannot be distinguished in terms of assertoric force, let us try to find some other distinguishing characteristic.

Martin-Löf speaks of the *truth of a proposition* and the *evidence of a judgement*. He seems to want to characterise propositions as entities which *may be true* and judgements as entities which *may be evident*. The distinction between these two types of entities thus takes us back to the distinction between *truth* and *evidence*. Now what confers evidence on a judgement is a *proof*, while that which confers truth on a proposition is a *verification*. So let us see how Martin-Löf defines these last two notions.

In Martin-Löf (1985), the assertability conditions of categoric judgements are outlined in the following way (we should remember that Martin-Löf makes a distinction between two forms of categoric judgement: ‘*A* true’, which we have already seen, and ‘*A* prop’, which asserts that *A* is a proposition).

“*A* prop” may be asserted when we know *what* to do to verify *A*, what counts as a verification of *A*. (Using Dummett’s terminology, “*A* prop” may be asserted whenever we know the meaning of *A*).

“*A* true” may be asserted when we know *how* to verify *A*, when we know a *method* for the verification of *A*.

From this we can conclude that a proof of “*A* true” is a method of verification for *A*. (This is explicitly stated in Martin-Löf 1985, p. 41.) The definition of the notion of proof of a judgement thus presupposes the definition of the notion of verification of a proposition. Prior to examining this, we should look at what constitutes a proof of a hypothetical judgement, that is a judgement with hypotheses that have not been discharged. In Martin-Löf (1985), only two forms of hypothetical judgement are taken into consideration, which we can simplify (reducing the hypotheses to one) in the following way:

$$A \text{ true}/B \text{ prop}$$

and

$$A \text{ true}/B \text{ true}$$

A proof of a judgement of the first form is a *hypothetical proof* of “*B* prop” from the hypothesis “*A* true”.

> The notion of hypothetical proof, [...] which is a primitive notion, is explained by saying that it is a proof which, when supplemented by proofs of the hypotheses, or antecedents, becomes a proof of the thesis or consequent (Martin-Löf 1985, p. 252).

Such a hypothetical proof is symbolised:

$$\vdash \begin{array}{l} A \text{ true} \\ B \text{ prop} \end{array}$$

In the same way, a proof of the second form of hypothetical judgement will a hypothetical proof of the type:

$$\vdash \begin{array}{l} A \text{ true} \\ B \text{ true} \end{array}$$

Let us now see how the notion of verification is defined:

A VERIFICATION OF	IS
$A \wedge B$	a proof of “*A* true” and a proof of “*B* true”
$A \vee B$	a proof of “*A* true or a proof of “*B* true”
$A \subset B$	a proof of “*A* true / *B* true”
$\bot$	Nothing

It will be seen that the notions of proof of a judgement of the form "A true" and of verification of the proposition A are defined with a simultaneous induction. (The basis is missing, namely the definition of a verification for an atomic proposition; the problem is open, but we are not concerned with this here.)

It should be noted that a proof of "A true" is nothing other than what would be, in a traditional system interpreted intuitionistically, a *proof* of A; and that a verification of the proposition A is what would be, in a traditional system, a *canonical proof* of A.

But at this point it seems that, as both proofs and canonical proofs prove the same entities, that is propositions, so the same entities, whether they are called propositions or judgements, must be capable to be evident as much as verified.

Thus, we have not so far found any acceptable reasons for differentiating, from an intuitionistic point of view, truth and evidence.

9.4 Metaphysical Realism

Martin-Löf appears however to firmly believe in the possibility of making a distinction between truth and evidence, for reasons which, although he connects them closely to the proposition/judgement distinction, are in fact practically independent of this. Let us look at the following passage:

> There is absolutely no question of a judgement being evident in itself, independently of us and our cognitive activity, That would be just as absurd as to speak of a judgement as being known, not by somebody, you or me, but in itself. To be evident is to be evident *to* somebody, as inevitably as to be known is to be known *by* somebody. That is what Brouwer meant by saying, in *Consciousness, Philosophy, and Mathematics*, that there are no non-experienced truths, a basic intuitionistic tenet. This has been puzzling, because it has been understood as referring to the truth of a proposition, *and clearly there are true propositions whose truth has not been experienced*, that is, propositions which can be shown to be true in the future, although they have not been proved to be true now. But what Brouwer means here is not that. He does not speak about propositions and truth: *he speaks about judgements and evidence, although he uses the term instead of the term evidence*. And what he says is then perfectly right: there is no evident judgement whose evidence has not been experienced, and experience is what you do when you understand, comprehend, grasp, or see it. There is no evidence outside our actual or possible experience of it. The notion of evidence is by its very nature subject related, relative to the knowing subject, that is, in Kantian terminology. (Martin-Löf 1985, pp. 223–224, ours italics)

Here it seems that the existence of true propositions, whose truth has not yet been experienced, is not intended in the sense that for some propositions, although a method of verification is known, the verification has in fact not been executed, as might be the case for the infinite disjunctive instances $A(n) \vee B(n)$ of a verified universal proposition of the type $\forall x(A(x) \vee B(x))$; for in Martin-Löf's terminology such instances, even though they are not verified, are recognised as true thanks to the procedure of verification provided by the verification of the universal proposition. It seems rather that he refers to propositions A for which the judgement "A true",

even though it is not at this moment evident, *may* become evident in the future. But such a concept of possibility is unacceptable to an intuitionist. The same reasons which lead him to deny propositions true in themselves lead him to deny possibilities whose existence cannot be recognised through the actual possession of a method. Such possibilities would lead to an intuitionistic modal justification of the principle of the excluded middle; for example, if G is Goldbach's conjecture, $G \vee \neg G$ could be interpreted as "G is verifiable or G is not verifiable". Anyhow, Martin-Löf's interpretation of the passage from Brouwer quoted above is certainly erroneous. The fact that Brouwer uses the term "truth" in place of "evidence" is not due to an improper terminological use, as Martin-Löf seems to suggest, but to the intentional identification of truth and evidence. Brouwer in fact criticises the classical conception of truth in the following words:

> …classical mathematics […] believes in the existence of unknown truths, and in particular applies the *principle of the excluded third* expressing that every mathematical assertion (i.e. every assignment of a mathematical property to a mathematical entity) either is a truth or cannot be a truth (Brouwer 1975, p. 488)

If truth and evidence were for Brouwer distinct notions and if he were referring to evidence using incorrectly the term "truth", his criticism would be expressed better by substituting the word "truth" in this passage with the word "evidence". On the contrary, any such substitution completely falsifies the sense of Brouwer's remarks, making his criticism trivially unfounded. No classical mathematician believes in the existence of unknown *evidences* nor interprets the principle of the excluded third by saying that every mathematical assertion either is an evidence or cannot be an evidence! What the classical conception admits is the existence of unknown *truths*, clearly distinguishing truth from evidence. And precisely, this belief in a truth distinct from evidence is what Brouwer criticises.

Besides it is not only in the interpretation of Brouwer's passage quoted above that Martin-Löf appears to rather yielding to a realist conception of truth. Against Aristotelian truth he argues that a judgement of the type "A true" is not correctly asserted if A is true but only provided that the truth of A is *known*. Regarding this he notes that, even if Goldbach's conjecture is true, it is an error to assert it until it has been proved. In this way, he appears to implicitly accept the possibility that G is true although it has not been proved: the lack of a proof prevents its assertability, not its *being* true. His argument, rather than leading to the denial of classical logic in favour of intuitionistic logic, is in fact in perfect agreement with classical logic. No classical mathematician would claim that he could assert a truth which had not been proved! The principle of the excluded middle is rightly assertible from a classical point of view, since within a Platonistic perspective, it is evident. Nor, from the classical point of view, does the evidence of $A \vee \neg A$ require the evidence of A or the evidence of $\neg A$, if it is understood, as Martin-Löf says, that what are combined with connectives are propositions and not judgements, so that to judge $A \vee \neg A$ does not mean at all to judge A or to judge $\neg A$; it only means to recognise the necessity of the truth (not necessarily known) of A or of $\neg A$. Perhaps Martin-Löf, although he insists that the notion of evidence conceptually precedes that of truth and that there

is no truth independently of the knowledge of it, is afraid of the identification of truth and evidence. The reason of this is presumably that he sees in such an identification the danger of an incurably subjective conception of truth, which would destroy the objectivity of mathematics, and render pointless any discussion about the correctness of a proof, which would be intolerable even for an intuitionist. Concerning this point, it is significant that, towards the end of Martin-Löf (1987), he attempts to delineate a type of realism, that he calls *metaphysical realism*, which would be compatible with Intuitionism:

> …the knowledge theoretical idealism which is so characteristic of the explanation of the notion of truth of a proposition, that is, the intuitionistic explanation of the notion of truth of a proposition, is entirely compatible with realism, if by realism you mean the philosophical position which takes the notion of truth or reality for granted, realism, of course, signifying reality here. And what is the opposite of that position? That is a position for which the notion of truth or reality in this sense does not exist, which means that the most that I can say about a judgement, for instance, is that it is evident to me: it may not be evident to you, and to you something may be evident which is perhaps even in conflict with what is evident to me, and there is no way of resolving that conflict because there is no notion of correctness to appeal to. So the ordinary discussion as to who is right that we mathematicians embark upon in that situation simply cannot arise. […] Maybe this kind of realism could be called metaphysical realism to distinguish it on the one hand from knowledge theoretical realism, that is, the view that the world exists independently of us and our cognitive activity, which is the opposite of the knowledge theoretical idealism characteristic of the intuitionistic analysis of the notion of truth of a proposition, and on the other hand from the realism with respect to the existence of universals which figured in the medieval debate about the nature of universals. If you agree to use the word realism also for this third position, namely, the position which simply takes the notion of truth or reality for granted, then we mathematicians, whether intuitionists or not, all seem to be realists. […] (Martin-Löf 1987, pp. 419–420)

It may be that the identification of truth and evidence is incompatible with the metaphysical realism which Martin-Löf speaks of, as this seems to attribute objectivity to truth but not to evidence. However, if this is the case, we find it difficult to understand where the objectivity of truth comes from, once truth is conceived as a product of evidence.

Anyway, in Brouwer's Intuitionism the objectivity which Martin-Löf takes to heart pertains to the very notion of evidence. In fact, the evidence which Brouwer places at the foundation of mathematics is not the evidence of the empirical mathematician but of the *creative subject*, a highly idealised mathematician. The discussions of empirical mathematicians on the correctness of deductions are therefore perfectly meaningful since they can be interpreted in terms of the agreement or disagreement of their deductions with the knowledge of the creative subject. The realist aspect of Intuitionism lies precisely in the presence of the creative subject, who, although he is obviously an imaginary character, is treated on the theoretic level as if he were real.

Given the difficulty in distinguishing between truth and evidence, Martin-Löf's distinction between propositions and judgements seems to us intuitionistically unintelligible, whereas it is intelligible from a classical point of view, even if it is not in agreement with Frege. We would also like to point out that if, in a classical framework, a clear distinction between truth and evidence is necessary, it is certainly not

necessary to distinguish things which may be true from things which may be evident. The characteristic that propositions have of being true or false does not in any way prevent them from becoming evident once they have been proved. The introduction of special entities as candidates solely for being evident seems to us pointless and artificial. As we have seen in illustrating Frege's approach, what is important is to distinguish different types of acts on propositions: these may be either asserted or simply considered, or even be the objects of many other mental acts such as to believe, to conjecture, to suppose. This is in perfect agreement with the use of propositions in the *Principia*, where such entities as judgements do not appear, but the assertion stroke is simply attached to a proposition when it is asserted by the authors. Russell's position on this subject is clearly expressed in his letter to Frege of 24 May 1903:

> In all cases, both imagination and judgement have an object: what I call a "proposition" can be the object of judgement, and it can be the object of imagination. There are therefore two ways in which we can think of an object, in case this obiect is a complex: we can imagine it, or we can judge it; yet the object is the same in both cases (e.g., when we say "the cold wind" and when we say "the wind is cold"). To me, the judgement stroke therefore means a different way of being directed towards an object. Complexes are true or false; in judging, we aim at a true complex; but we may, of course, miss our aim (Frege 1980, p. 159).

Sharing Russell's viewpoint, we are in a position to reinforce our criticism of Martin-Löf's argument that the appropriate objects of inference are judgements. We have already seen that this argument presupposes that judgements carry assertoric force, contrary to Martin-Löf's conception. Now we can add that, even if judgements are conceived as carrying assertoric force (i.e. even if only those which the author calls "evident judgement" are taken into consideration), Martin-Löf's criticism of the use of propositions as objects of inference seems to be hardly convincing. The thesis according to which the objects of inference are propositions in no way claims that the conclusion of an inference is created by the inference itself. An inference simply *associates* the conclusive propositions to the premised propositions, associating (provided it is correct) a true conclusion to true premises. In this sense, rules of inference can perfectly well operate on entities without assertoric force. The activity of proving is certainly creative, but it is not for this reason that it has to create its own objects of proof. Its creativity consists in giving evidence to which have previously been created, without transforming them by this into entities different from what they previously were.

As for the mental acts which create propositions, it seems to us that they are not properly part of the activity of judging (or of recognising as true), as Martin-Löf, suggests by placing "*A* prop" among the fundamental forms of judgement. The conception of a proposition, understood intuitionistically, does not consist in the recognition of something; it consists rather in the formation of a concept: to conceive the proposition *A* means to form the concept of proof of *A*. A mental act of this kind does not recognise anything but is presupposed by acts of recognition, since the evidence of a proposition consists in the knowledge that the concept of proof related to it is not empty.

It seems to us that it can be concluded that propositions can be regarded as the only fundamental entities of logic, since all mathematical activity may be analysed in terms of the creation and demonstration of propositions.

References

Brouwer, L. (1975). Collected works. In A. Heyting (Ed.), *Philosophy and foundations of mathematics*. Amsterdam: Elsevier.

Dummett, M. (1981). *The interpretation of Frege's philosophy*. London: Duckworth.

Frege, G. (1980). *Philosophical and mathematical correspondence*. Oxford: Blackwell.

Frege, G. (1984). *Collected papers in mathematics*. Blackwell: Logic and Philosophy.

Martin-Löf, P. (1985). On the meaning and justification of logical laws. In E. Bernardi & R. Pagli (Eds.), *Atti degli Incontri di Logica Matematica* (Vol. II, pp. 291–340). Siena: Università di Siena.

Martin-Löf, P. (1987). Truth of a proposition, evidence of a judgement, validity of a proof. *Synthese*, *73*, 407–420.

Chapter 10
Negationless Intuitionism

Abstract The present paper deals with natural intuitionistic semantics for intuitionistic logic within an intuitionistic metamathematics. We show how strong completeness of full first-order logic fails. We then consider a *negationless semantics* à la *Henkin* for second-order intuitionistic logic. By using the theory of *lawless sequences* we prove that, for such semantics, strong completeness is restorable. We argue that lawless negationless semantics is a suitable framework for a constructive *structuralist* interpretation of any second-order formalisable theory (classical or intuitionistic, contradictory or not).

10.1 Natural Semantics

By a *natural* interpretation of intuitionistic logic we mean what in the literature is also called an *intuitive* (Troelstra and Van Dalen 1988) or an *internal* (Dummett 1977) interpretation.

A natural interpretation $\mathfrak{I}$ of first-order predicate logic (IPC) assigns to every individual constant a member of a certain inhabitated domain D of *individuals* and to every predicate letter a *property* or a *relation* on D of appropriate degree. Such an interpretation provides *proof-conditions* for every atomic sentence (of the language extended with constants for all members of D): assigning a property to the predicate letter P amounts to giving proof-conditions for sentences of form $P(d)$ (with $d \in D$). Proof-conditions for compound sentences, as well as for *absurdity* $\bot$ (which is taken as a primitive logical constant), are determined according to Heyting's explanation of the intuitionistic meaning of logical constants.

(1) *nothing* is a proof of $\bot$;
(2) a proof of $\forall x A(x)$ is a method of proving all $A(d)$'s ($d \in D$) [likewise for $A \wedge B$];
(3) a proof of $\exists x A(x)$ is a method of proving some $A(d)$ ($d \in D$) [likewise for $A \vee B$];
(4) a proof of $A \rightarrow B$ is a method of transforming every proof of A into a proof of B;

Negation is defined by $\neg A =_{df.} A \rightarrow \bot$.

A sentence A is true (in $\mathfrak{I}$) if there is a proof of A (where the existence of a proof, as well as all metamathematical arguments and definitions, is to be understood *intuitionistically*).

A sentence A is *valid* (with respect to natural semantics) ($\vDash A$) if it is true in all natural interpretations. A sentence A is a *logical consequence* of a set Γ of sentences ($\Gamma \vDash A$) if every (natural) model of Γ is a model of A.

IPC is *complete* if, for all sentences A, $\vDash A \Rightarrow \vdash A$. IPC is *strongly complete* if, for all Γ and all A, $\Gamma \vDash A \Rightarrow \Gamma \vdash A$.

As it is well known, the completeness of IPC for natural semantics (within an intuitionistic metamathematics) is equivalent to a certain form of Markov's principle (see Dummett 1977). Since the latter is far from being intuitionistically evident, this result provides an argument against completeness. On the other hand, since Markov's principle seems to be compatible with the general principles of Intuitionism, the above argument for incompleteness is not conclusive.

What is cogent, however, is the failure of strong completeness, as we will show in the next section.

10.2 Failure of Strong Completeness

Strong completeness for natural semantics can be rejected by using nothing beyond the intuitionistic meaning of logical constants.

Theorem 10.1 *Strong completeness is contradictory*

Proof Assume, by way of *reductio*, strong completeness. Let M be an arbitrary natural structure and A an arbitrary sentence. Consider the set of sentences $\Gamma = \{\bot : M \vDash A \vee \neg A\}$. Since $\neg\neg(A \vee \neg A)$ is logically valid, $M \vDash \neg\neg(A \vee \neg A)$ and hence $\neg\neg(M \vDash A \vee \neg A)$, so that $\neg\neg(\bot \in \Gamma)$. Since $\bot$ cannot be true in any structure, it follows that Γ has no models so that, vacuously, $\Gamma \vDash \bot$. By strong completeness $\Gamma \vdash \bot$. As IPC is consistent and Γ cannot have any other members but $\bot$, it follows that $\bot \in \Gamma$ and hence that $M \vDash A \vee \neg A$. Since M and A are arbitrary, we get the validity and, by the supposed completeness, the derivability of the excluded middle principle, which is absurd.[1]

Observe that in the above proof an essential role is played by $\bot$ (and hence by negation). In fact, as we shall see, the negationless fragment of first-order logic without identity turns out to be strongly complete. However, if identity occurs in the language and is taken as a logical primitive constant to be interpreted as *intensional* identity, then even the negationless fragment is incomplete.

Theorem 10.2 *Strong completeness of the negationless fragment of IPC with intensional identity is contradictory*

[1] This proof was suggested to me by letter by prof. Troelstra.

Proof Assume, by way of contradiction, strong completeness for the negationless fragment. Define $\hat{\bot} =_{df.} \forall x \forall y(x = y)$, $\sim A =_{df.} A \to \hat{\bot}$, $E_2 =_{df.} \exists x \exists y \sim (x = y)$.

Given an arbitrary structure M and an arbitrary sentence A, put $\Gamma =_{df.} \{\hat{\bot} : M \vDash A \vee \sim A\} \cup \{E_2\}$. It is easily seen that $\neg\neg(M \vDash A \vee \sim A)$, so that $\neg\neg(\hat{\bot} \in \Gamma)$. Let N be a (possible) model of Γ. Since $N \vDash E_2$, there are $a, b \in N$ such that $\sim (a = b)$. If $a \neq b$, then $\neg(N \vDash \hat{\bot})$, against $\neg\neg(\hat{\bot} \in \Gamma)$. As intensional identity is decidable, we get $a = b$ and hence $N \vDash \hat{\bot}$. Thus $\Gamma \vDash \hat{\bot}$ and, by strong completeness, $\Gamma \vdash \hat{\bot}$. Since $\neg(E_2 \vdash \hat{\bot})$, it follows that $\hat{\bot} \in \Gamma$, from which $M \vDash A \vee \sim A$. We conclude that $\vDash A \vee \sim A$ and therefore $\vdash A \vee \sim A$, which is absurd.

The above proofs reject strong completeness merely in virtue of the intended meaning of logical constants. They do not provide, however, any counterexample to strong completeness, i.e. any Γ and A such that $\Gamma \vDash A$ but not $\Gamma \vdash A$. We think that no such counterexample could be found without exploiting deeper features of Intuitionism. A counterexample to strong completeness has been found by Charles McCarty (1991) under the assumption of Church's thesis. But we are not inclined to accept the intuitionistic version of Church's thesis: it imposes excessively strong limitations of a mechanistic nature on the intuitive notion of proof, which seem to be inappropriate to Brouwerian Intuitionism. We prefer to use lawless sequences and exploit Brouwer's conception of a universe consisting of indeterminate entities. We give here a counterexample to strong completeness by using a set of sentences depending on a lawless parameter.

Let HA be Heyting's arithmetic extended with a new 1-place predicate symbol P and a new individual constant c. With reference to a lawless sequence α, define the sets of sentences

- $\Gamma = \{P(n) :$ the least $m > n$ such that $\alpha(m) = 0$ exists and is odd$\}$
  $\cup \{\neg P(n) :$ the least $m > n$ such that $\alpha(m) = 0$ exists and is even$\}$,
- $\Delta = \{c > \underline{n} : n \in \omega\}$,
- $\Sigma = HA \cup \Gamma \cup \Delta$.

Σ has no models:

By way of contradiction, let M be a model of Σ.

From the induction schema (extended to P), we get $M \vDash \neg\neg\forall x < c(P(x) \vee \neg P(x))$. On the other hand, if $M \vDash \forall x < c(P(x) \vee \neg P(x))$, then $M \vDash P(n) \vee \neg P(n)$ for all $n \in \omega$ and it follows that $\forall x \exists y > x \alpha(y) = 0$, against lawlessness. Hence $M \vDash \neg\forall x < c(P(x) \vee \neg P(x))$, which is absurd. Thus $\Sigma \vDash \bot$.

On the other hand, every finite subset of Σ is interpretable in the standard model of ω, therefore not $\Sigma \vdash \bot$.

Incompleteness for natural semantics shows that the formal inference rules are inadequate for capturing the intended meaning of logical constants, as expressed by Heyting's explanation. The nature of this gap can be better understood in the light of the completeness results with respect to so-called *fallible* models. As it is well known, Veldman, de Swart and others of the Nijmegen school got, in the seventies, intuitionistic completeness proofs for IPC with respect to modified versions of

Kripke and Beth models (De Swart 1976; Troelstra and Van Dalen 1988). The main feature of such models consists in allowing $\bot$ to be true at some node. For this reason they are called sometimes *fallible* models. The simplest completeness proof for single formulas with respect to fallible models was found by Friedman and rearranged by Troelstra (Troelstra and Van Dalen 1988). In that proof a universal fallible Beth model is constructed, in which exactly the derivable sentences of IPC are true. In Martino (1988), Friedman's proof has been generalised to a strong completeness proof for fallible Beth models.

Though Beth models are of no immediate significance for natural semantics, the theory of lawless sequences provides an important connection between Beth semantics and natural semantics: with every lawless path of a Beth model, it is possible to associate a natural model in such a way that truth in the Beth model is equivalent to truth in all its paths (see Dummett 1977). So Beth validity turns out to be equivalent to natural validity.

The problem arises to seek an intuitive counterpart of fallible Beth semantics, so as to recover strong completeness for some intuitionistically acceptable extension of natural semantics. The difficulty is that allowing $\bot$ to be true in a natural interpretation amounts to allowing its possible intuitive provability, in disagreement with Heyting's clause (1), which characterises $\bot$ as the *unprovable par excellence*. In Martino (1988) the possibility of interpreting the truth of $\bot$ as provability under a set of "hidden-hypotheses", which may turn out to be contradictory, is investigated.

Here we want to pursue the possibility of interpreting second-order intuitionistic logic (with negation and identity) within a semantics without negation and without intensional identity.

The fact that intuitionistic logic is sound with respect to fallible semantics shows that formal deduction is inadequate to capturing the intended meaning of $\bot$. In fact, $\bot$ is characterised, in the intuitionistic system of natural deduction, by the single rule '*ex falso quodlibet*'

$$\frac{\bot}{A}$$

This rule states that a proof of $\bot$ amounts to a proof of every proposition of the language. So to prove $\bot$ is absurd iff it is impossible to prove all propositions. Now, if negation is taken as primitive, so that a proof of $\neg A$ is nothing but the recognition of the unprovability of A, then $A \wedge \neg A$ is certainly unprovable and therefore not all propositions are provable. But if negation is defined, as above, by means of $\bot$, then the unprovability of $A \wedge \neg A$ rests in turn on the unprovability of $\bot$. Formal deduction seems therefore inadequate for capturing the intended falsehood (i.e. the intuitive unprovability) of $\bot$: it is not able to distinguish the authentic $\bot$ from any other proposition (provable or not) entailing all propositions of the language.

A similar situation occurs with identity. The formal rule of *substitutivity of identicals* is inadequate to capturing the intended meaning of intensional identity, understood as *absolute sameness*, since such rule holds also for the relation of *indiscernibility* with respect to the properties expressible in the language.

On the other hand, formal deduction is adequate to formalise mathematics. This suggests that the intended meaning of $\bot$ is inessential to the development of mathematics and that any formalised theory is interpretable in a negationless semantics. This is realisable in a perspicuous way within second-order predicate logic interpreted *à la* Henkin.

10.3 Second-Order Negationless Semantics

Consider second-order intuitionistic logic with *full comprehension schema* and *intensional identity*.

We use $x, y, \ldots$ as individual variables, $F, G, \ldots$ as monadic second-order variables.

Observe that, by comprehension,

$$\bot \leftrightarrow \forall F \forall x F(x)$$
$$x = y \leftrightarrow \forall F(F(x) \leftrightarrow F(y)).$$

This suggests the possibility of adopting a semantics *à la* Henkin, in which $\bot$ and $=$ are not interpreted according to their intended meanings but are *defined* as

$$\bot =_{df.} \forall F \forall x F(x)$$
$$\neg A =_{df.} A \rightarrow \bot$$
$$x = y =_{df.} \forall F(F(x) \leftrightarrow F(y)).$$

Precisely, a *negationless interpretation* (of the full second-order language) is meant to interpret

(i) first-order variables (and constants) into an inhabitated domain D of individuals;
(ii) second-order variables (and constants) into *any* domain of (intensional) *properties* and *relations* on D such that, provided $\bot$ and $=$ are defined as above, second-order intuitionistic logic is *sound*.

The soundness condition is satisfied, it should be noted, as soon as comprehension and $x = y \vee \neg x = y$ hold.

In particular, the domain of properties may consist of the sole universal property (under which every individual falls); in this case the model is said to be *exploding*. In such a model every sentence, in particular $\bot$, is true (intuitively provable). So, with respect to negationless semantics, every set of sentences (even contradictory) is satisfiable.

We want to show that, if *lawless models* (i.e. models depending on lawless parameters) are included in our negationless semantics, then strong completeness is restorable.[2]

Let Γ be any set of sentences expressible in a language L of second-order intuitionistic logic. With every lawless sequence α of the full binary spread s, we will associate a negationless model M_α of Γ.

Introduce infinitely many new constants c_i^j $(i = 1, 2, \ldots; j = 0, 1, \ldots)$ for first- and second-order entities: c_i^0 is an individual constant while, for $j > 0$, c_i^j is a constant for relations of degree j. Define the hierarchy of languages

$$L_0 = L, L_{n+1} = L_n \cup \{c_{n+1}^j : j \in \omega\}, L_\omega = \bigcup_{n\in\omega} L_n.$$

Let $A_0, A_1, \ldots, A_n, \ldots$ be an enumeration of all *disjunctive* and *existential* L_ω-sentences, each of them occurring infinitely many times.

We indicate by $u, v, w, \ldots$ nodes of s (finite $0-1$-sequences). If $u = \langle u_1, \ldots, u_n\rangle$, n is the *length* of u $(lth(u))$. If u is as above and m is a natural number, $u\hat{}m$ stands for $\langle u_1, \ldots, u_n, m\rangle$.

$\alpha, \beta, \gamma, \ldots$ indicate 0–1-lawless sequences. $\overline{\alpha}(n)$ denotes the initial segment of α of length n. If u is an initial segment of α, we write $\alpha \in u$.

We shall use the following axioms for lawless sequences:

- *Density:* $\forall u \exists \alpha \alpha \in u$.
- *Open data:* $P(\alpha) \rightarrow \exists n \forall \beta_{\beta\in\overline{\alpha}(n)} P(\beta)$

(for any predicate P on lawless sequences).

A bar $\mathfrak{B}$ of u is a set of nodes such that $\forall \alpha_{\alpha\in u} \exists n \overline{\alpha}(n) \in \mathfrak{B}$.

Fan Theorem: if $\mathfrak{B}$ is a bar of u, then $\exists m \forall \alpha_{\alpha\in u}\ \exists n_{n<m} \overline{\alpha}(n) \in \mathfrak{B}$.

Later on, ξ stands for a variable of any order and degree. If A is a formula, $A[c_i/\xi]$ is the result of replacing ξ in A by the constant c_i^j (introduced above) of the appropriate degree j.

To every node u of s we assign a set Γ_u of L_n-sentences, where $n = lth(u)$, as follows:

(i) $\Gamma_{\langle\rangle} = \Gamma$;
(ii) for every L_{n+1}-sentence A, $A \in \Gamma_{u\hat{}k}$ if$_{df.}$ either

(a) $A \in \Gamma_u$ or
(b) $A_n := \exists \xi B$, $\Gamma_u \vdash A_n$ and $A := B[c_{n+1}/\xi]$ or

---

[2]Observe that, if negationless semantics is adopted for *classical* second-order logic (within a classical metamathematics), the presence of exploding models shows that even classical logic is inadequate for capturing the classical meaning of $\bot$, as false *par excellence*. The reason why this inadequacy does not affect classical strong completeness, for usual Henkin semantics, is that usual validity amounts to negationless validity *restricted* to non-exploding models and, classically, such a restriction is immaterial, since *every* model is exploding or non-exploding and every sentence holds in an exploding model.

(c) $A_n = B \vee C$, $\Gamma_u \vdash A_n$ and $A := \begin{cases} B & \text{if } k = 0. \\ C & \text{if } k = 1 \end{cases}$

Lemma 10.3 *Let A be an L_n-sentence, with $n = lth(u)$. If, for all $k = 0, 1$, $\Gamma_{u^\frown k} \vdash A$, then $\Gamma_u \vdash A$.*

Proof Suppose $A_n := \exists\xi B$. Let p_0, p_1 be formal proofs of A from $\Gamma_{u^\frown 0}$ and $\Gamma_{u^\frown 1}$ respectively. If $B[c_{n+1}/\xi]$ does not occur among the assumptions of p_0, then p_0 is a derivation of A from Γ_u. Otherwise $B[c_{n+1}/\xi] \in \Gamma_{u^\frown 0}$, so that either $B[c_{n+1}/\xi] \in \Gamma_u$ and therefore $\Gamma_{u^\frown 0} = \Gamma_u$ or $\Gamma_u \vdash \exists\xi B$. In the latter case, since c_{n+1} occurs neither in Γ_u nor in A, A is derivable from $\Gamma_u \cup \{\exists\xi B\}$ and hence from Γ_u.

Similarly for $A_n := B \vee C$.

Theorem 10.4 *Let A be an L_n-sentence, with $n = lth(u)$. $\Gamma_u \vdash A$ iff there is a bar $\mathfrak{B}$ of u such that, for all $v \in \mathfrak{B}$, $A \in \Gamma_v$.*

Proof If $\Gamma_u \vdash A$, take $m \geq n$ such that $A_m := A \vee A$. For all $v \geq u$, with $lth(v) = m$, $A \in \Gamma_{v^\frown k}$. The converse immediately follows from the lemma and the fan theorem.

For every lawless sequence α, define the natural model M_α of L as follows:

(i) the domain D of individuals is the set of all individual constants of L_ω;
(ii) for every second-order constant C of degree m of L_ω, let C_α be the n-place relation on D which holds among $d_1, \ldots, d_m$ if$_{df.}$ $C(d_1, \ldots, d_m) \in \Gamma_{\overline{\alpha}(n)}$, for some n. Take the C_α's as the properties and relations of the model.

L_ω is interpreted in M_α by mapping every individual constant into itself and every second-order constant C into C_α.

Theorem 10.5 *Let A be an L_n-sentence and u be a node of length n. $\Gamma_u \vdash A$ iff, for all $\alpha \in u$, $M_\alpha \vDash A$.*

Proof By induction on A:

(i) A atomic. The thesis follows immediately from Theorem 10.4.
(ii) $A := \exists\xi B$.

Suppose that, for all $\alpha \in u$, $M_\alpha \vDash A$. Given α, there is a constant c such that $M_\alpha \vDash B[c/\xi]$. By *open data* there is $m \geq n$ such that, for all $\beta \in \overline{\alpha}(m)$, $M_\beta \vDash B[c/\xi]$ and, by the induction hypothesis, $\Gamma_{\overline{\alpha}(m)} \vdash B[c/\xi]$ and hence $\Gamma_{\overline{\alpha}(m)} \vdash \exists\xi B$. It follows that there is a bar $\mathfrak{B}$ of u such that, for all $v \in B$, $A \in \Gamma_v$ and, by Theorem 10.4, $\Gamma_u \vdash A$.

Conversely, suppose that $\Gamma_u \vdash A$. Take $m \geq n$ such that $A_m := A$. If $v \geq u$ and $lth(v) = m$, then $B[c_{m+1}/\xi] \in \Gamma_{v^\frown k}$, for all $k = 0, 1$. By the induction hypothesis $B[c_{m+1}/\xi]$, and hence A, is true in M_α, for all $\alpha \in u$.

Similarly for $A := B \vee C$.

(iii) $A := \forall\xi B$.

Suppose that, for all $\alpha \in u$, $M_\alpha \vDash A$. Let T be any logically valid L-sentence and take $m \geq n$ such that $A_m := T \vee T$. If $v \geq u$ and $lth(v) = m$, for all $\beta \in v\hat{}k$ $(k = 0, 1)$, $M_\beta \vDash B[c_{m+1}/\xi]$ so that, by the induction hypothesis, $\Gamma_{v\hat{}k} \vdash B[c_{m+1}/\xi]$. Since $\Gamma_{v\hat{}k} = \Gamma_v \cup \{T\}$, $\Gamma_v \vdash B[c_{m+1}/\xi]$; and since $c_{m+1} \notin \Gamma_v$, $\Gamma_v \vdash A$.

Similarly for $A := B \wedge C$.

(iv) $A := B \to C$.

Suppose that, for all $\alpha \in u$, $M_\alpha \vDash A$. Take $m \geq n$ such that $A_m := B \vee T$, where T is a logically valid L-sentence. If $v \geq u$ and $lth(v) = m$, then $B \in \Gamma_{v\hat{}0}$, so that, for all $\beta \in v\hat{}0$, $M_\beta \vDash B$ and hence $M_\beta \vDash C$. By the induction hypothesis, $\Gamma_{v\hat{}0} \vdash C$, and since $\Gamma_{v\hat{}0} = \Gamma_v \cup \{B\}$, $\Gamma_v \vdash B \to C$. By Theorem 10.4 $\Gamma_u \vdash A$.

Conversely, suppose that $\Gamma_u \vdash A$. Let $\alpha \in u$ and assume that $M_\alpha \vDash B$. In virtue of *open data*, there is $m \geq n$ such that $M_\beta \vDash B$, for all $\beta \in \overline{\alpha}(m)$. By the induction hypothesis, $\Gamma_{\overline{\alpha}(m)} \vdash B$, from which $\Gamma_{\overline{\alpha}(m)} \vdash C$ and therefore $M_\alpha \vDash C$.

Call $\mathfrak{M}(\Gamma)$ the class of *lawless models* M_α's of Γ defined above.

Theorem 10.6 (Strong Completeness) *For every set Γ of sentences and every sentence A, $\Gamma \vdash A$ iff, for all $M \in \mathfrak{M}(\Gamma)$, $M_\alpha \vDash A$.*

A set Γ of sentences will be called *classical* if$_{df.}$ it contains the *excluded middle sentence*

(EM) $\quad \forall F \forall x (F(x) \vee \neg F(x))$.

Intuitionistic derivability from a classical set Γ amounts to classical derivability. Thus Theorem 10.6 includes a theorem of strong completeness for classical logic with respect to intuitionistic semantics. Concerning this, it is worth noticing that the commitment to lawless sequences is inessential.

For, let $\mathfrak{C}$ be any class of *choice* sequences (in particular of *lawlike* sequences) satisfying the *density* principle and the *fan* theorem; let $\mathfrak{M}_{\mathfrak{C}}(\Gamma)$ be the class of models defined as $\mathfrak{M}(\Gamma)$ but relative to $\mathfrak{C}$. By slightly modifying the above completeness proof, we get the following theorem:

Theorem 10.7 (Intuitionistic strong completeness of classical logic) *For every classical set Γ of sentences and every sentence A, $\Gamma \vdash A$ iff, for all $M \in \mathfrak{M}_{\mathfrak{C}}(\Gamma)$, $M \vDash A$.*

Proof One can rearrange the proof of Theorem 10.5 by using, instead of *open data*, the fact that, for every $\alpha \in \mathfrak{C}$ and every A, there is n such that $A \in \Gamma_{\overline{\alpha}(n)}$ or $\neg A \in \Gamma_{\overline{\alpha}(n)}$. One proves, by induction on A, that:

$$\text{if } \Gamma_u \vdash A [\Gamma_u \vdash \neg A] \text{ then, for all } \alpha \in u,\ M_\alpha \vDash A [M_\alpha \vDash \neg A].$$

As an example, we develop the case $A := B \to C$.

Suppose that $\Gamma_u \vdash A$, $\alpha \in u$. Assume $M_\alpha \vDash B$ and let $m \geq lth(u)$ be such that $B \in \Gamma_{\overline{\alpha}(m)}$ or $\neg B \in \Gamma_{\overline{\alpha}(m)}$. In the first case $\Gamma_{\overline{\alpha}(m)} \vdash C$ and, by the induction

hypothesis, $M_\alpha \models C$. In the second case $M_\alpha \models \neg B$, so that M_α is exploding. Thus, in any case, $M_\alpha \models A$.

Suppose that $\Gamma_u \vdash \neg A$ and $\alpha \in u$. Since Γ is classical, $\Gamma_u \vdash B$ and $\Gamma_u \vdash \neg C$ and, by the induction hypothesis, $M_\alpha \models B$ and $M_\alpha \models \neg C$, i.e. $M_\alpha \models \neg A$.

10.4 Concluding Remarks

It is often remarked that the models constructed in completeness proofs are *ad hoc* and hence of little philosophical significance.

Of course none of the M_α's is in itself of any mathematical interest, in the sense that none of them is the intended model of a mathematical theory. However, lawless semantics seems to be of remarkable interest from the *structuralist* point of view. According to this point of view, a mathematical theory does not describe *any* privileged structure. Though a certain intended structure may have an essential *premathematical* role in suggesting and motivating the axioms, once the axiomatic system has been built up, the mathematical discourse concerns any arbitrary structure satisfying the axioms. So any class of models, for which soundness and strong completeness hold, constitutes an appropriate semantics. The intuitionistic interest of our lawless models is that they are intuitionistically intelligible and that one reasons about them according to the intuitionistic meaning of the primitive logical constants. Soundness and completeness assure that, by reasoning intuitionistically in any arbitrary M_α, one gets exactly all logical consequences of Γ. In this sense, lawless semantics provides an adequate characterisation of *abstract intuitionistic negationless reasoning*, of which second-order intuitionistic logic turns out to be a *formal* adequate counterpart.

The possibility of defining $\neg A$ as "A implies every proposition", without loss of any theorem provable by means of standard negation, was already noted by Russell (1906).[3] He failed however to realise that this means that the authentic notion of falsehood is not needed for abstract mathematical reasoning. He rejected the definition by arguing that it would prevent to prove that anything is false. "If any man is so credulous as to believe that anything is true", Russell observes, "then the method in question [i.e. the adoption of the above definition of negation] is powerless to refute him" (p. 201). The quantification over all propositions was understood by Russell in an *absolute* sense as referring to *all possible propositions*. Therefore, the impossibility of proving that any proposition is false was rightly for him a symptom of inadequacy, since it is plain that there are false propositions. Not so, however, in the perspective of Henkin semantics, where, as we saw, there is room for interpretations in which all propositions are actually true.

Independently of the purpose of recovering strong completeness, negationless semantics is in itself of special significance for constructivism. Indeed the classical notion of truth can be understood only together with its opposite notion of falsehood, so that the two notions are conceptually inseparable. This is not the case, however,

[3] I am grateful to the referee for suggesting to me this reference to Russell.

for their intuitionistic counterparts. In order to understand the intuitionistic notion of truth of a proposition A, one has to grasp what a proof of A is, while, in order to understand the intuitionistic notion of falsehood, one has to understand what it means to say that a hypothetical proof of A leads to absurdity. So the intuitionistic concept of falsehood of A rests on the primitive notion of *absurdity* (or of *impossibility* of constructing a proof of A), which is by no means involved in the mere concept of proof of A (provided negation does not occur in A). Thus not even an implicit understanding of intuitionistic falsehood is presupposed by intuitionistic negationless semantics. For this reason negationless Intuitionism expresses a more restrictive conception of constructivism than traditional Brouwerian Intuitionism. Negationless Intuitionism seems to be in agreement with Griss' well-known criticism to intuitionistic negation (see Franchella 1992; Griss 1946; Heyting 1956). According to Griss, a proof by *reductio ad absurdum* is not in agreement with the general intuitionistic perspective. If the assumption that a proof of A is given leads to contradiction, then that assumption itself cannot be regarded as a clear mental act and is therefore inappropriate to intuitionistic reasoning; therefore absurd assumptions should be banished from Intuitionism. Brouwer's reply to Griss rests on the consideration that negation is essential to the development of mathematics and that negative assertions cannot be always rephrased in positive terms. However, though a structuralist view of mathematics is certainly far from Griss' (and Brouwer's) conception, negationless semantics shows that structuralism, together with the indeterminateness of lawless sequences, makes Griss' ideal of a *positive* mathematics realisable in a very general setting. Observe that, on this respect, the role of lawlessness is essential. In fact, the mere lack of absurdity (and negation), as a primitive logical constant, does not prevent, in itself, hypothetical arguments with arbitrary assumptions, in particular with absurd assumptions. For instance, in the standard model of negationless arithmetic, $0 = 1$ is certainly an absurd assumption. But any lawless model M_α has the crucial feature of being "*potentially*" exploding, in the sense that, at every stage of construction of α, the possibility is left that some future choices of values for α make the model exploding. It follows that, as long as one reasons on M_α according to the intuitionistic meaning of the primitive logical constants involved, it can never happen that an assumption turns out to be absurd. To prove $\neg A$ amounts to recognising that every possible choice to the effect of rendering A true will make the model exploding. So, since any sentence is always *potentially* capable of becoming true, any assumption describes a state of knowledge about the model which is effectively possible. Thus, within the framework of lawless semantics, hypothetical reasoning is free of any danger of involving absurd assumptions. In this sense Griss' requirement is satisfied.

Besides, even the (non structuralist) intuitionistic metamathematics, in which lawless semantics is defined and strong completeness is proved, is negationless: in our treatment no notion of absurdity or of impossibility (to carry out some construction) has been exploited. This suggests that negationless Intuitionism is a suitable framework for the development of metamathematics.

Another remarkable aspect of negationless semantics is the intuitionistic interpretability of classical logic. Negationless semantics provides an intuitionistic justification of classical reasoning, even within a lawlike conception of Intuitionism, as

Theorem 10.7 shows. Though the excluded middle principle is not intuitionistically acceptable as a *logical* principle, nevertheless classical reasoning is intuitionistically intelligible: the classical consequences of Γ are the intuitionistic ones relative to a suitable restricted stock of intuitionistic models.

Lawless semantics is also in agreement with predicativism. For, the theory of lawless sequences, however problematic in some respects, is trivially predicative, since the construction of any lawless sequence does not involve anything but free choices of values (not so for the notion of lawlike sequence (and, more generally, of *choice sequence*), which involves the highly impredicative notion of *arbitrary law*). Furthermore, in any M_α the properties and relations, whose existence is asserted by the comprehension principle, are not constructed by reference to the totality of properties and relations; they are defined in terms of the *syntactical* notion of formal derivability, which, from a logical point of view, can be thought of as *pre-existent* to any interpretation of the language, so that no violation of Russell's *vicious circle principle* is involved. In this sense we get a predicativistic justification of the second-order comprehension principle.[4]

As it is well known, Brouwer devised his theory of choice sequences for rejecting classical mathematics and constructing alternative mathematical theories, as the intuitionistic continuum, in striking conflict with their classical counterparts. Besides, he was a sworn enemy of formal logic, which he charged of rendering mathematics meaningless. In spite of that, the formal analysis of mathematical language is just what opens the way towards exploiting Brouwer's insights for guaranteeing constructive and predicative meaning to *any* mathematical argument.

References

De Swart, H. (1976). Another intuitionistic completeness proof. *Journal of Symbolic Logic*, *41*, 644–662.

Dummett, M. (1977). *Elements of intuitionism*. Oxford: Clarendon Press.

Franchella, M. (1992). *The Griss-Brouwer debate on negation*. Technical report. University of Milan, CUEM Preprint, Milan.

Griss, G. F. C. (1946). Negationless intuitionist mathematics. *Indagationes Mathematicae*, *8*, 675–681.

Heyting, A. (1956). *Intuitionism: An introduction*. Amsterdam: North-Holland.

Leblanc, H. (1975). That Principia Mathematica, first edition, has a predicative interpretation after all. Indagationes Mathematicae (Vol. 8, pp. 675–681).

Martino, E. (1988). An intuitionistic notion of hypothetical truth for which strong completeness intuitionistically holds. Teoria (Vol. 8, pp. 131–144), reprinted here as chapter 8.

McCarty, D. C. (1991). Incompleteness in intuitionistic metamathematics. *Notre Dame Journal of Formal Logic*, *32*, 323–358.

Russell, B. (1906). The theory of implication. *American Journal of Mathematics*, *28*, 159–202.

Troelstra, A., & Van Dalen, D. (1988). *Constructivism in mathematics* (Vol. II). Amsterdam: North-Holland.

[4] As observed by the referee, a similar argument has been brought forward by Leblanc (1975), in order to show that even Russell's theory of simple types is capable of a predicative interpretation.

Chapter 11
Temporal and Atemporal Truth in Intuitionistic Mathematics

with G. Usberti

Abstract In Sect. 11.2, we argue that the adoption of a tenseless notion of truth entails a realistic view of propositions and provability. This view, in turn, opens the way to the intelligibility of the *classical* meaning of the logical constants and consequently is incompatible with the antirealism of orthodox Intuitionism. In Sect. 11.3, we show how what we call the "potential" intuitionistic meaning of the logical constants can be defined, on the one hand, by means of the notion of atemporal provability and, on the other hand, by means of the operator K of epistemic logic. Intuitionistic logic, as reconstructed within this perspective, turns out to be a part of epistemic logic, so that it loses its traditional foundational role, antithetic to that of classical logic. In Sect. 11.4, we uphold the view that certain consequences of the adoption of a *temporal* notion of truth, despite their apparent oddity, are quite acceptable from an antirealist point of view.

11.1 Introduction

Nowadays, the most widespread view about constructive truth is that truth should be conceived as a tenseless notion. Prawitz (1987, 153–154) writes:

> [A] mathematical sentence is true if there exists a proof of it, in a tenseless or abstract sense of exists […]. Or we may express the same idea by saying that a sentence A is true if 'we can prove A' […]. That we can prove A is not to be understood as meaning that it is within our practical reach to prove A, but only that it is possible in principle to prove A [...]. Similarly, that there exists a proof of A does not mean that a proof of A will be constructed but only that the possibility is there for constructing a proof of A. […] I see no objection to conceiving the possibility that there is a specific method for curing cancer, which we may discover one day, but which may also remain undiscovered.

Martin-Löf (1991) distinguishes between actual and potential truth of a proposition. These notions would be explained intuitionistically by the notions of *actual* and *potential existence* of a proof. A proof of a proposition A exists actually if, as a matter of fact, A has been proved; it exists potentially if A *can* be proved. Here possibility is not understood in the traditional intuitionistic sense as knowledge of a method to prove A, but as "knowledge-independent and tenseless" possibility. Accordingly, a proposition that has been proved becomes actually true, but it was potentially true

even before having been proved, and it would be true even if, in fact, it had never been proved. In this way, according to Martin-Löf, the intuitionist can overcome the well-known objection that saying that a proposition *becomes* true just when it is proved is counterintuitive and in conflict with the standard use of the truth predicate: potential truth is not open to that objection.

Dummett's position on this point seems to be rather oscillating. On the one hand, he argued for the need of a notion of truth—of *some* notion of truth—within the constructivist conceptual framework and for some "necessary concession to realism", and he was probably the first who suggested conceiving intuitionistic truth as tenseless. On the other hand, especially in recent years, he has manifested some perplexities about the compatibility of a tenseless notion of truth with the antirealism of the intuitionists.

11.2 Tenselessness and Classical Truth

As explained above, Martin-Löf (1991) maintains that potential truth is knowledge-independent, in the sense that a proposition may be potentially true even if nobody knows (nor will know) that it is. Nevertheless, he considers this notion intuitionistically meaningful for two reasons. First, it is *conceptually* dependent upon the notion of knowledge; for "to say that A is potentially true is to say that A can be actually true", and to say that A is actually true is to say that A is *known* to be true. So potential truth depends on knowledge in the sense that it is definable in terms of knowledge (in conformity to the Aristotelian idea that *actus est prior potentia*). Second, potential truth obviously satisfies the constructivist requirement that every true proposition can be proved.

However, the definition of potential truth makes essential reference not only to the notion of knowledge (or proof) but also to the notion of possibility. Therefore, the mere conceptual priority of the notion of knowledge over that of potential truth does not guarantee the definability of the latter in intuitionistic terms (or even its intuitionistic acceptability). Analogously, the thesis that every true proposition can be proved will be an intuitionistic thesis only if the modality involved is intuitionistic.

Now, the concept of possibility, understood in the manner described above, is undoubtedly extraneous to Brouwer and Heyting's Intuitionism. According to this, the possibility of a construction is always to be conceived as an *epistemic state*, i.e. as a state in which the knowing subject acknowledges that he is able to perform a certain construction. Prawitz and Martin-Löf's possibility, on the contrary, is *not* characterised as an epistemic state, but as merely factual accessibility to an epistemic state.

A passage of Dummett (1977, p. 19) seems to be responsible for some confusion on this point. Dummett writes:

> It would be possible for a constructivist to agree with a Platonist that a mathematical statement, if true, is timelessly true: when a statement is proved, then it is shown thereby to have been true all along. To say this is, in effect, to equate "A is true" with "We can prove A"

rather than with "*A* has been proved", and "*A* is false" with "We cannot prove *A*". Such an interpretation of 'true' and 'false' remains faithful to the basic principles of Intuitionism only if "We can prove *A*" ("*A* is provable") is not interpreted to mean either, at one extreme, that, independently of our knowledge, there exists something which, if we became aware of it, we should recognise as a proof of *A*, nor, at the other, that as a matter of fact we either have proved *A* or shall at some time prove it. In the former case, we should be appealing to a platonistically conceived objective realm of proofs; in the latter, we should be entitled to deny that *A* was provable on non-mathematical grounds (e.g. if the obliteration of the human race were imminent). "We can prove *A*" must be understood as being rendered true only by our actually proving *A*, but as being rendered false only by our finding a purely mathematical obstacle to proving it.

But if "*A* is true" is equated with "We can prove *A*" and "We can prove *A*" is rendered true only by our actually proving *A*, then *A* was not true before having been proved, and the notion of truth is not tenseless. Therefore, if we want truth to be tenseless, it is necessary that also the notion of possibility, in terms of which it is defined, is conceived of as atemporal. So, as long as we identify truth with provability and agree with traditional Intuitionism that the provability of *A* obtains only in virtue of our actually proving *A*, we cannot agree with a Platonist that a mathematical statement, if true, is timelessly true.

Prawitz and Martin-Löf want to reconcile Intuitionism with the belief that mathematical truths are eternal. Being aware of the above difficulty, they profess the conviction that the provability of a proposition can be understood as tenseless even within an intuitionistic perspective.

Prawitz (1987), referring to Dummett's passage above, writes that he has no objection to understanding "the existence of a proof with reference to an objective realm of proofs", since "in such an objective realm of proofs there can be no question of the existence of a proof that is not in principle recognisable by us." (p. 154).

Observe that the eternal actual existence of proofs is by no means in conflict with their potentiality: the latter does not concern existence but the capacity of being grasped by the (idealised) human mind: briefly, proofs are *actually* existent but only *potentially* known.

On his part, Martin-Löf (1991) rejects any identification of provability with knowledge of a proof:

> If something has been, is being or will be done, then it can be done, but not in the converse direction. In the case of proving a proposition, this means that, if a proposition has been, is being or will be proved, then certainly it can be proved, that is, it is potentially true, but there is absolutely no reason to believe that we can go in the opposite direction. The principle just spelled out is again a principle which had a succinct scholastic formulation: it is the principle, *Ab esse ad posse valet consequentia* (*illatio*). (p. 193)

In other terms, from the availability of a proof of *A* we can of course infer the provability of *A*, but not conversely, as it would be the case if the provability of *A* subsisted only in virtue of our knowledge of a proof of *A*.

Martin-Löf does not commit himself to Prawitz's position, which could be called "proof platonism". However, once a tenseless notion of provability has been espoused, the commitment to an objective realm of *propositions* is unavoidable. For, if the possibility to prove a proposition *A* is conceived as atemporal, then *A* itself

becomes an atemporal entity. If a proposition existed only *qua* created by the mind of the creative subject (as the orthodox intuitionist maintains), it would have an exclusively temporal existence, so that, if it is provable, its provability would be temporal as well, since it could not subsist before the very creation of the proposition. Thus, potential truth presupposes at least the existence of an objective realm of propositions, and consequently, whether an objective realm of *proofs* is accepted or not, it leads inevitably to a sort of realism.

Independently of the questions at issue, an ontology of proofs may be in some respects problematic (especially for reasons of impredicativity). We are inclined to believe that a commitment to such an ontology is not necessarily presupposed by the intuitionistic conception of truth as (temporal or atemporal) provability: the mathematical activity of proving does not seem to require any reification of proofs. Therefore, in what follows we will not assume that the adoption of an atemporal notion of provability entails an ontology of proof-objects.

Let us call "potential Intuitionism" the view described above, supporting an atemporal notion of possibility, and "orthodox Intuitionism" the original, Brouwerian conception of mathematics. The following question arises: is the potential intuitionist still in a position to reject classical logic in favour of intuitionistic logic?

Atemporal possibility is certainly incompatible with Brouwer's theory of free choice sequences. The tenselessness of provability presupposes that the objects of discourse are well-determined. For example, if α is a lawless sequence the possibility (or impossibility) of proving $\alpha(\overline{n}) = \overline{m}$ is determined when the nth choice is made. Once this proposition has been proved, we may certainly say that it had the possibility of being proved even before the choice of its nth value, in the sense that nothing could prevent that the nth choice was just m. But this is not the kind of possibility we are concerned with here: it is a temporal possibility which would have been lost if a number different from m had been chosen as $\alpha(\overline{n})$, so that in no sense could we assert that $\alpha(\overline{n}) = \overline{m}$ was true before having been proved. Therefore, the potential intuitionist must reject choice sequences and confine himself to a lawlike ontology.

In any case, choice sequences do not play a central role in the intuitionistic criticism of classical logic. Brouwer's rejection of the excluded middle does not rest on any alleged indeterminateness of mathematical objects, but merely on his particular conception of the link between truth and knowledge (although he exploited choice sequences to give strong counterexamples to classical theorems). So our problem can now be restated as follows: is the potential intuitionist still in a position to defend Brouwer's negative attitude about classical logic? The answer, we hold, is necessarily negative. The reason can be found, *in nuce*, in a passage of Dummett (1987):

> There is a well-known difficulty about thinking of mathematical proofs […] as existing independently of our hitting on them, which insisting that they are proofs we are capable of grasping or of giving fails to resolve. Namely, it is hard to see how the equation of the falsity of a statement (the truth of its negation) with the non-existence of a proof or verification can be resisted: but, then, it is equally hard to see how, on this conception of the existence of proofs, we can resist supposing that a proof of a given statement either exists or fails to exist. We shall then have driven ourselves into a realist position, with a justification of bivalence. If we refuse to identify falsity with the non-existence of a proof, we shall be little better off, because we shall find it hard to resist concluding that there are statements which are

> determinately neither true nor false, there being no proof of them or of their negations: we shall then have a quasi-realist denial of bivalence. (p. 285)

Let us elaborate Dummett's remark. From the fact that possibility is conceived as tenseless it follows that the following principle of *Potential Excluded Middle*:

(PEM) *A* is potentially true or *A* is not potentially true

becomes *intelligible*, and valid, in its *classical* reading. For, on this reading, it simply means that all propositions, as they are conceived by the potential intuitionist, are atemporally *determinate*, and this is clearly true: if it were indeterminate whether *A* is provable or not, the provability of *A* would be for ever prevented since, according to the conception at issue, a proposition cannot *become* provable. Therefore, such a hypothetical state of indeterminateness of *A* could be nothing but a state of well-determined unprovability of *A*. Whether *A* is provable or not is a fact concerning the immutable world of propositions, where there is no room for any indeterminateness. Of course, the potential intuitionist might refrain from interpreting classically the logical constants occurring in PEM; he could interpret them intuitionistically and deny PEM consequently. But this move would not affect the fact that from his point of view the classical interpretation of PEM is intelligible. While the orthodox intuitionist can reject the classical interpretation of PEM because of his refusal of certain categories involved (such as an objective realm of propositions), the potential intuitionist must acknowledge that those categories are the ones which he himself is committed to. If he insisted that the classical construal of PEM is meaningless, he would simply renounce the possibility of expressing his own conviction that provability is atemporally determinate, but he would not have gained an argument against the realist.[1]

The potential intuitionist is therefore forced to admit that the realist interpretation of the logical constants is meaningful and appropriate for expressing certain facts about potential truth.

Of course the potential truth of a sentence, interpreted according to the intuitionistic meaning of the logical contents, does not coincide at all, in general, with its classical truth. However, the potential intuitionist is able to *reconstruct* the notion of classical truth within his own conceptual framework, at least in the case of a language in which the classical and the intuitionistic meaning of atomic sentences is the same, in particular when atomic sentences are decidable. In that case, Tarski's inductive definition of classical truth is perfectly intelligible to the potential intuitionist, since

---

[1] In some respects, this seems to be the move implicit in Martin-Löf's type theory, in particular in his distinction between propositions and judgements. According to it, "*A* is potentially true" is a judgement, and a general feature of judgements is that the logical operations cannot be applied to them; as a consequence, a judgement expressing the content of PEM does not exist. However, this answer is acceptable only to someone who subscribes to (1) the distinction between judgements and propositions, and (2) the reasonability of the interdiction to apply negation to a judgement. In Martino and Usberti (1991), we stated some reasons not to accept either (1) or (2). Observe that if PEM were meaningless for the reason at issue, so would be Martin-Löf's claim that "If *A* is actually true, then it is potentially true", where the inapplicability of the logical constants to judgements is violated too.

the Tarskian clauses are based on the notion of potential truth for *atomic* sentences, and the classical meaning of the metalinguistic logical constants (involved in the Tarskian clauses) is appropriate to the realistic objectivity of potential truth. Let us illustrate this point in a particular case.

Let $P(x)$ be a decidable predicate on natural numbers. Given a number n, $P(\overline{n})$ is classically (as well as intuitionistically) true iff it is provable according to the decision procedure. So the classical truth of $\forall x P(x)$ is defined in terms of the potential truth of all the $P(\overline{n})$'s and the classical metalinguistic quantifier "every". Since, for every n, it is objectively determined whether $P(\overline{n})$ is provable or not, it is also objectively determined whether every $P(\overline{n})$ is provable or some is not. Therefore, the classical meaning of the metalinguistic quantifier is accessible to the potential intuitionist and he can understand perfectly well the Tarskian account of "$\forall x P(x)$ is true".

In contrast, such an account is not available to the orthodox intuitionist. For him, every $P(\overline{n})$ is provable or not in the mere sense that, in virtue of the decision procedure, given any n, he knows how to verify or falsify $P(\overline{n})$. But from his point of view verifiability or falsifiability are not intrinsic properties of the $P(\overline{n})$'s subsisting independently of his knowledge. Consequently, the provability of all the $P(\overline{n})$'s is not an objective fact that subsists or not in the realm of propositions; it is a fact that comes to being only if and when the knowing subject becomes aware, through an abstract reasoning, that the decision procedure, applied to an arbitrary n, will verify $P(\overline{n})$. This point of view is essential for the intuitionistic rejection of the principle of excluded middle; this asserts, according to that view, that we know a method to decide every proposition, and it can therefore be criticised as a principle of omniscience. But the potential intuitionist cannot refuse the classical interpretation of the excluded middle, according to which $\forall x P(x) \vee \neg \forall x P(x)$ expresses the obvious *factual* truth that either all the $P(\overline{n})$'s are tenselessly provable or some of them are not, a truth that has nothing to do with the human ability to decide which alternative holds.

As a consequence, the potential intuitionist, unlike the orthodox, is forced to admit the intelligibility of classical logic. However, he can still defend intuitionistic logic by arguing that it, unlike classical logic, satisfies certain desirable requisites, in particular the basic constructivist requirement that every truth is provable in principle, whereas classical truth, although definable in the framework of potential Intuitionism, is irremediably separated from provability. In the next section, we examine this line of argument.

11.3 Potential Intuitionism as a Subsystem of Epistemic Mathematics

Consider a logically complex proposition, for instance $\forall x A(x)$, and suppose we want to explain its meaning in terms of some key notion such as truth or proof. The standard format of the explanation is the following. If we choose truth as the key notion, we assume that we know what the truth conditions of A are, and we define the truth

conditions of $\forall x A(x)$; in order to do this, we have to use the metalinguistic logical constant "all" (or some mathematical notion in terms of which it is definable), whose meaning we therefore assume as known. Analogously, if we choose proof as the key notion, we assume that we know what a proof of A is, and we define what a proof of $\forall x A(x)$ is; in order to do this, we have to use the metalinguistic logical constant "all" (or some mathematical notion in terms of which it is definable), whose meaning we therefore assume as known. In any case, we must assume that we know, at the metalinguistic level, the meaning of essentially the same constants we are explaining; we seem to have fallen into an infinite regress. Does this compromise the efficacy of our explanation? This is a very general and interesting problem we shall not discuss here; we mentioned it only because an important difference between the orthodox and the potential intuitionist concerns the answer they can give to it.

According to the *naïve*—realist—view, proofs are arguments to get evidence that certain states of affairs obtain or that certain objective facts subsist. From this point of view, it is quite natural to characterise proofs in terms of the facts they prove; for instance, a proof of $\forall x A(x)$ can be defined as a proof of the fact that all the facts expressed by $A(d_1/x), A(d_2/x), \ldots$ subsist. Now, the potential intuitionist *can* understand the naive notion of proof, since he can give sense to the notion of objective fact: the objective fact that A is the fact that the proposition "A" is atemporally provable. Consequently, he can inductively define conditions of a compound proposition in terms of objective facts concerning the truth of the subpropositions. Precisely, given provability conditions for atomic sentences ($\neq \bot$), the *potential meaning* of the logical constants can be expressed as follows:

(1) $\bot$ is not true;
(2) $A \wedge B$ is true if both A and B are true;
(3) $A \vee B$ is true if either A is true or B is true;
(4) $A \supset B$ is true if it is provable that B is true provided A is true;
(5) $\forall x A(x)$ is true if it is provable that, for every individual d, $A(d/x)$ is true;
(6) $\exists x A$ is true if, for some individual d, $A(d/x)$ is true.[2]

Observe that the logical constants occurring in the metalanguage are classically interpreted in the world of propositions. Consequently, the potential intuitionist is able to give an explanation of the potential meaning of the logical constants that is really *reductive*, and he does not fall into the infinite regress mentioned above. Moreover, the above clauses are equivalent, for him, to Heyting's ones, in the sense that, if a proposition A is true according to the potential interpretation, there is a proof of A according to Heyting's explanation. Compare for instance clause (4) above with Heyting's version:

(4′) A proof of $A \supset B$ is a method to transform any proof of A into a proof of B.

Suppose that $A \supset B$ is true according to (4). Given a proof of A, by using a proof that B is true provided A is true, we come to know that B is provable, and this very

[2] A similar interpretation of the intuitionistic logical constants has been proposed by Carlo Dalla Pozza (1991).

knowledge must be the outcome of a proof of B (since *to know* that a proposition is provable amounts to having proved it). So a proof of $A \supset B$ provides a method to transform any proof of A into a proof of B. Similar considerations hold for the other clauses.

The orthodox intuitionist is in a completely different position since, owing to his antirealistic conception of mathematics, he *cannot* understand the *naïve* notion of proof. The *raison d'être* of Heyting's revision of classical logic is not simply the constructivist aim to stress the central role of provability in mathematical truth, but the need to reform the very notion of intuitive proof. The lack of objective reference forces him to abandon the usual realistic understanding of proofs and raises the question of how intuitionistic proofs are to be understood. In the absence of objective facts to prove, what do intuitionistic proofs prove? His explanation of the meaning of the logical constants is just an attempt to answer this question. By inductively defining what a proof of a compound sentence is in terms of what is a proof of its components, Heyting tries to overcome the difficulty of explaining *what* an intuitive proof of a proposition *is*, without saying *which facts* it has to prove.

Sometimes this aspect of Heyting's explanation has not been adequately appreciated in the literature. Consider for example Heyting's clause for universal quantification:

(5′) A proof of $\forall x A(x)$ is a method to get, for every individual d, a proof of $A(d/x)$.

Some commentators have proposed to add a "second clause" requiring a proof that the method in question yields, for every d, a proof of $A(d/x)$. The need of such a proof is brought about by a misleading classical view of the matter. According to this, it may happen that a method yields, as a matter of fact, a proof of $A(d/x)$ for all d, even if it is not known that it does. It follows that a proof of $\forall x A(x)$, besides giving the required method, must *show* that it works as desired. For instance, let G be Goldbach's conjecture that every even number > 2 is the sum of two primes. If G is true, there is a trivial method satisfying (5'): calculate, for any even number n, all the sums of pairs of primes $\leq n$, until a pair is found whose sum is just n. Of course this method is not itself a proof of G: in order to prove G, we must show that, by applying the method to any even number, we will always reach the desired pair. But, according to orthodox Intuitionism, there are no knowledge-independent facts, so that there is no room for distinguishing between the mere fact that the method works and a proof of this fact. Therefore, for the orthodox intuitionist the second clause is utterly useless: its introduction is nothing but a misleading attempt to characterise a proof of $\forall x A(x)$ by saying what it has to prove.

Of course, since the metalanguage, in which Heyting's clauses are formulated, is to be understood intuitionistically, his explanation is *not* reductive, and it does fall into the infinite regress mentioned above.

What an intuitionistic proof of a universal quantification really is, lies hidden in the intuitionistic primitive notion of (general) method.

In sum, the potential intuitionist not only has at hand a notion of proof more akin to the ordinary use, but also can define the potential meaning of the logical constants,

thereby overcoming an essential difficulty connected with Heyting's explanation. However, the price to pay for these advantages (maybe with no regret) is the impossibility of *criticising* classical mathematics.

It may be instructive to compare potential Intuitionism with *epistemic mathematics*.[3] Given a first-order language L, let L_e be the epistemic language obtained by adding an epistemic operator K with the formation clause: if A is a formula, then KA is a formula. We shall read KA as "It is (atemporally) knowable (or provable) that A". The intuitionistic translation A^* of a formula A of L is inductively defined as follows:

(7)
- $A^* = KA$ for A atomic $\neq \bot$
- $\bot^* = \bot$
- $(A \wedge B)^* = (A^* \wedge B^*)$
- $(A \supset B)^* = K(A^* \supset B^*)$
- $(\forall x A(x))^* = K(\forall x A^*(x))$
- $(\exists x A(x))^* = \exists x A^*(x)$

It is known that a formula is derivable in intuitionistic logic iff its intuitionistic translation is derivable in epistemic logic (see Shapiro 1985). But we are concerned here with meaning-theoretical considerations. Given a *classical* interpretation of L, extend it to L_e by interpreting KA as explained above. By speaking of the *epistemic meaning* of a sentence of L_e, we shall refer to this interpretation.

Now, suppose that the interpretation of the atomic sentences is intuitionistically acceptable (i.e. that the meaning of such sentences is given by proof-conditions). Then, as the argument developed in Sect. 11.2 shows, the whole interpretation of L_e (where the logical constants have their classical meanings) is accessible to the potential intuitionist. Moreover, if A is a formula of L, the epistemic meaning of A^* is identical to the potential intuitionistic meaning of A (as shown by a straightforward induction on the complexity of A). Thus, within the very framework of potential Intuitionism, the potential meanings of the logical constants are definable, with the help of the epistemic operator K, in terms of the classical meanings.

We can conclude that the potential interpretation of the logical constants is devoid of any logical significance. It cannot be proposed as a basis for a logic alternative to the classical one. Classical logic is not only intelligible to the potential intuitionist but also perfectly adequate to treating his notion of knowability. Unlike the orthodox, the potential intuitionist has no need to revise classical logic: as far as logic is concerned, he reasons just as the classical mathematician does. Moreover, he has a very good reason to use classical logic, since it enables him to give a reductive explanation of the intuitionistic meaning of the logical constants. In general, what characterises his position is his interest in certain propositions involving the notion of knowability, realistically conceived as knowability of objective facts. This is the *naïve* notion of intuitive provability, shared also by the classical mathematician, which has little to do with the very sophisticated notion of provability of orthodox Intuitionism.

[3]For a nice introduction to this topic see (Shapiro 1985).

11.4 Temporal Truth

In the preceding section, we reached the conclusion that potential Intuitionism is a realistic position whose constructivist aim does not lead to the rejection of classical logic but to an epistemic mathematics based on classical logic and naive provability.

It is certainly an interesting position as a way of reconciling, to some extent, classical and intuitionistic mathematics. However, as our considerations should have shown, the philosophical position of the potential intuitionist is very far from Brouwer and Heyting's view of mathematics. The essential difference between orthodox Intuitionism and classical mathematics turns out to lie in the opposition between temporal and atemporal provability rather than in the opposition between transcendent truth and provability. Epistemic logic, with the operator K interpreted as "atemporally provable", provides a very natural, but essentially classical, interpretation of intuitionistic logic. For these reasons, if one is interested in Intuitionism as an antirealistic philosophy of mathematics and in intuitionistic logic as the basis of antirealistic reasoning, he must reject atemporal provability and resist the temptation of introducing into Intuitionism any surrogate of classical truth. He must rather accept certain consequences of his antirealistic view, even when they are not in agreement with the *naïve*, realistic, attitude of the working mathematician. Temporal truth, the notion according to which a proposition becomes true only when it is proved, is certainly counterintuitive insofar as it conflicts with the usual *naïve* notion of proof as recognition of preexistent facts. But that is as it should be, since the antirealist must, of course, refuse any intuition of a realistic nature. We want to show now that, within the framework of orthodox Intuitionism, certain recurrent objections against temporal truth are groundless.

First of all, there is an important question about temporal truth which needs some clarification. Sometimes it has been maintained that a proposition, in order to be true, does not need to have been effectively proved; it needs only to be *provable*, in the *epistemic* sense that a method is known to prove it. This notion of provability is certainly congenial to an antirealistic view. However, it is quite natural to admit that a method to prove A is itself a proof of A: we cannot know that we are able to know A, unless we know A. For this reason, as far as temporal truth is concerned, there is no room for distinguishing between potential and actual truth.

The first objection we want to discuss is the paradox of inference.[4] As Dummett writes, "The justifiability of deductive inference [...] requires *some* gap between truth and its recognition". The validity of an inference rule is usually expressed by saying that it preserves truth, in the sense that if the premises are true, so is the conclusion. Thus, the objection goes, the identification of truth with its recognition entails that if the premises of a valid inference are known, so is the conclusion. Consequently, valid deduction cannot provide any increase of knowledge and is therefore useless.

However, the validity of an inference rule is to be interpreted, in the orthodox perspective, according to Heyting's meaning of implication: an inference rule is valid

[4]For a formulation of this paradox and the problems it raises see also the introduction to the present issue.

in so far as we know how to get from the knowledge of the premises to the knowledge of the conclusion. This does not mean that the knowledge of the premises amounts to the knowledge of the conclusion. In order to know the conclusion, we must perform a further mental act consisting of putting together the knowledge of the premises and that of the validity of the rule. Between the knowledge of the premises and that of the conclusion there is always a (however small) epistemic distance. Observe that this is not in conflict with our stipulation that the knowledge of a method to prove a proposition counts itself as a proof of that proposition: the transition from the knowledge of the premises to the knowledge of the conclusion is realised by the mere recognition that the available method of transforming any proof of the premises into a proof of the conclusion, together with a proof of the premises, yields a method of proving the conclusion; no effective application of the method is required. For example, once we have recognised the validity of the axioms and rules of Heyting's arithmetic, we can certainly claim that all theorems of Heyting's arithmetic are true. But by this claim, we do not mean the *empirical* (and obviously false) fact that we have proved every arithmetical theorem. That claim is more appropriately expressed by the (metalinguistic) assertion

(8) For every arithmetical sentence A, if A is a theorem, then A is true,

which is to be interpreted according to the intuitionistic meaning of implication and universal quantification: we know that, given A and a formal proof of it, we are able to come to know A. But such a specific A does not become true as soon as we get evidence of (8). A becomes true if and when, having become acquainted with it and with a proof of it, we explicitly realise that our knowledge of (8) confers evidence to A.

Another objection is the following. Consider the propositions

(9) If somebody proved or will prove that there are infinitely many twin primes, then he knows or will know a great deal about prime numbers;
(10) If there are infinitely many twin primes, then someone knows or will know a great deal about prime numbers.

It is objected that (10) does not follow from (9), as it should be if their antecedents had the same meaning. So, the argument goes, we need a notion of truth such that "It is true that A" does not mean the same as "It has been or will be proved that A".

But the reason why the argument seems compelling is that the above propositions are read according to the classical meaning of implication. As far as (9) is concerned, such a reading may be intelligible to the orthodox intuitionist. For, since both the antecedent and the consequent of (9) are *empirical* assertions, they may be understood as descriptions of certain (possible) events of the world where the knowing subject lives, a world which can be thought of *realistically* even in the orthodox perspective. The case of (10) is completely different; its antecedent describes, in the classical reading, a (possible) fact of the objective world of numbers, which is refused by the orthodox intuitionist. He can interpret (10) only according to Heyting's meaning of implication: a method is known to transform any proof that there are infinitely many

twin primes into the knowledge of a great deal about prime numbers. Therefore, in the antirealistic perspective, the objection vanishes.

Finally, we want to discuss the paradox of knowability. It arises from the *verificationist thesis* that every truth is knowable. The thesis can be formalised as

(VT) $\quad (A \supset \Diamond K_0 A)$

where K_0 is an epistemic operator to be read as "it is (or will be) known that".

By using very weak principles concerning $\Diamond$ and K_0, one can intuitionistically derive from (VT) that there are no unknown truths:

(PK) $\quad \neg(A \wedge \neg K_0 A)$.

Informally, the argument runs as follows. Assume $A \wedge \neg K_0 A$; by applying (VT) we get $\Diamond K_0 (A \wedge \neg K_0 A)$ and therefore $\Diamond (K_0 A \wedge \neg K_0 A)$, which is absurd.

The conclusion seems to be paradoxical since, intuitively, the mere possibility that the truth of A becomes known does not seem to imply that as a matter of fact A will be known. In the orthodox perspective what is problematic is how to interpret the modal operator $\Diamond$, since the most natural interpretation as atemporal possibility is, as we know, precluded. In any case, whatever the interpretation of $\Diamond$ may be, the solution of the paradox is trivial: for the orthodox intuitionist (PK) is perfectly acceptable. In fact, he can uphold even the (intuitionistically) stronger claim

(PK*) $\quad A \supset K_0 A$

For, as we have already argued with regard to (10), (PK*) can be interpreted only according to the intuitionistic meaning of implication, so that it expresses the trivial observation that, as soon as a proof of A is given, A becomes known.

Williamson (1988) tries to reject the cogency of (PK*), even from an antirealistic point of view. He proposes to think of intuitionistic proofs as types of mental constructions and observes that, while we have a trivial method of transforming a proof-token of A into a proof-token of $K_0 A$, there is no method, in general, of transforming a proof-type of A into a proof-type of $K_0 A$.

The trouble is that a proof of (PK*) cannot be understood as a method which operates on types: from a constructive point of view it can be applied only to tokens. Williamson seems to be aware of this difficulty and indicates a way to overcome it:

> The notion of proof-type sounds suspiciously Platonist. Fortunately for intuitionists, all they really need is the ontologically neutral concept of sameness in type, where two proof-tokens of the same type are required to have identical conclusions and structure, but need not occur at the same time. For a function from proof-types to proof-types need be nothing more than a function f of a special kind from proof-tokens to proof-tokens: one that is *unitype* in the sense that if p and q are proof-tokens of the same type then so are $f(p)$ and $f(q)$. To speak of proof-types is then just to speak of proof-tokens in ways not sensitive to differences between proof-tokens of the same type; there need be no suggestion that the existence of a type is anything over and above the existence of tokens of that type. […] [T]he result is that a proof of $P \supset KP$ [$A \supset K_0 A$ in our symbolism] is a unitype function that evidently takes any proof-token of P to a proof-token, for some time t, of the proposition that P is proved at t. Need there be such a function?
>
> For any proof-token p of P, let $t(p)$ be the time at which p, as a particular mental construction, was carried out; we can allow that reflection on p can yield a corresponding proof-token $f(p)$

> that p was proved at $t(p)$ […] Now if we know of such a proof-token p, we can consider a function f_p which maps any proof-token q of P to a proof-token of KP of the same type as $f(p)$; f_p is obviously unitype and thus constitutes a proof of $P \supset KP$. If P has not yet been decided, however, the best we can do is to consider the function f itself, for the hypothetical proof-token of P is the only one we have to play with in attempting to construct a proof-token of KP. But f is not a unitype function: for there can be proof-tokens p and q of P of the same type but carried out at different times, in which case $f(p)$ and $f(q)$ are proof-tokens of distinct types, for their conclusions differ: $f(p)$ is a proof that P is proved at $t(p)$, $f(q)$ is a proof that P is proved at $t(q)$, and $t(p) \neq (t(q)$. Thus if we have no decision procedure for P, we have no way of constructing a proof of $P \supset KP$. (Williamson 1988, 430–431)

We do not believe that Williamson's proposal is suitable for the orthodox conception. The distinction between proof-types and proof-tokens may be appropriate if proofs are conceived as outright objects, so that a proof of $A \supset B$ is an outright mathematical function f from proofs of A to proofs of B (as in the case of Martin-Löf's type theory, where propositions are treated as sets of proofs). In such a framework, it is certainly legitimate to question when two proofs are equal and to require that a function from proofs to proofs preserves equality. But, according to orthodox Intuitionism, intuitive proofs are nothing but acts of knowing whose structures are of no mathematical interest (in contrast with formal proofs, whose structures are the subject matter of proof theory). Heyting's explanation of logical constants does not even require any reification of proofs: a proof of $A \supset B$ consists in the conscious knowledge of how to prove B as soon as A will be proved, and such a knowledge does not involve any function from proofs to proofs. This interpretation is also in agreement with Kolmogorov's interpretation of logical constants in terms of problems, according to which $A \supset B$ is the problem of reducing problem B to problem A. To accept Williamson's proposal would amount to denying that the knowledge of how to prove B as soon as A is known counts as evidence for $A \supset B$; and this is highly counterintuitive.

But let us concede, for the sake of argument, that a proof of an implication is an outright function from proofs to proofs. Williamson's argument still seems unconvincing. For he admits that, if a proof p of A is known, then there is a unitype function f_p from proofs of A to proofs of K_0A, i.e. the constant function which maps every proof of A into the proof that A has been proved at time $t(p)$. He claims, however, that the situation is quite different if no proof of A is known, since in this case what is available to us is only a *hypothetical* proof-token of A, with which no specific time is associated. But the required function f is not expected to operate on the hypothetical proof-token: such an object does not exist! Its arguments cannot be anything but given proof-tokens; as long as no proof of A is known, f has nothing to map. So we can still define f as the constant function which, once a proof p of A is known, maps every proof q of A into the proof that K_0A is known at time $t(p)$.

We conclude that (KP*) is a cogent consequence of the orthodox perspective, for which the paradox of knowability is not at all paradoxical. The apparent paradoxicality arises from the fact that the intuitionistic tenet that every truth is provable is often misunderstood in terms of the usual, knowledge-independent, notion of possibility. But the orthodox intuitionist refuses any metaphysical notion of possibility; for him the possibility to do something is an epistemic state in which it is known how to

do that. As soon as possibility is understood in this way, (VT) and (PK*) become indistinguishable.

The paradox is much more significant for potential Intuitionism, where there is no identification between truth and knowledge of truth. In this framework, the truth of A consists in the tenseless existence of a proof of A, and (VT) expresses an essential peculiarity of proofs: they are capable of being grasped by the (idealised) human mind. Of course this feature does not entail by itself the actual grasp of a proof by the mind. So (PK*) is certainly to be rejected. But what about (PK)?

According to the potential intuitionistic meaning of the logical constants, (PK) says that there is a proof that there cannot be both a proof p of A and a proof q that p will never be grasped. This is certainly correct. For, if both p and q existed, their conjunction would not be capable of being grasped by the mind. The intuitionistic reading of (PK) is by no means in conflict with the existence of unknown truths. What the paradox shows is that the possibility of such existence cannot be expressed intuitionistically. However, the classical meaning of the logical constants is accessible to the potential intuitionist, as we saw, and he can express his interpretation of mathematical propositions within classical logic by means of the operator K. Consequently, even if A is intended to be an intuitionistic proposition, he can understand classically (VT), (PK) and the whole derivation of the latter from the former. And since (PK) is classically equivalent to (PK*), he can recognise a clear sense in which (VT) entails the wrong conclusion (PK*) and is therefore to be rejected, at least in its whole generality. Indeed the paradox arises by applying (VT) to the proposition $A \wedge \neg K_0 A$, which involves the factual operator K_0 and is therefore to be regarded as *empirical*.[5] Now, the truth of $K_0 A$ cannot consist in the tenseless existence of a certain proof; otherwise, for every proposition it would be completely determined whether it will be known or not. But such a deterministic conception of factual knowledge would not be compatible with the view that for no proof the possibility of being grasped by the human mind is ever precluded: this view entails that for no provable proposition it is determined that it will remain for ever unknown. Thus, a proposition of the form $K_0 A$ cannot be atemporally true: it becomes true just in the moment (if any) in which a proof of A is grasped.

In conclusion, the potential intuitionist must distinguish carefully between *mathematical* and *factual* propositions and restrict to the former (not involving the operator K_0) his conception of atemporal truth, as well as his verificationist thesis (VT). But, as far as factual propositions are concerned, he cannot maintain the thesis that every truth is knowable.

[5] Observe that, even if our epistemic operator K may be analysed as $K = \Diamond K_0$, it is not to be regarded as empirical, because of the presence of the modal ingredient: the mere knowability of a proposition holds quite independently of any empirical event of actual knowledge.

References

Dalla Pozza, C. (1991). Un'interpretazione pragmatica della logica proposizionale intuizionistica. In G. Usberti (Ed.), *Problemi fondazionali nella teoria del significato* (pp. 49–75). Firenze: Leo Olschki Editore.

Dummett, M. (1977). *Elements of intuitionism*. Oxford: Clarendon Press.

Dummett, M. (1987). Reply to dag prawitz. In B. Taylor (Ed.), *Michael Dummett: Contributions to philosophy* (pp. 281–286). The Hague: Nijhoff.

Martin-Löf, P. (1991). A path from logic to metaphysics. In G. Sambin & G. Corsi (Eds.), *Atti del Congresso Nuovi Problemi della Logica e della Filosofia della Scienza* (pp. 141–149), Vol. II. Bologna: CLUEB.

Martino, E., & Usberti, G. (1991). Propositions and judgements in Martin-Löf. In G. Usberti (Ed.), *Problemi fondazionali nella teoria del significato* (pp. 125–136). Firenze: Leo Olschki Editore. Reprinted here as chapter 9.

Prawitz, D. (1987). Dummett on a theory of meaning and its impact on logic. In B. M. Taylor (Ed.), *Michael Dummett: Contributions to philosophy* (pp. 117–165). Dordrecht: Springer.

Shapiro, S. (1985). Epistemic and intuitionist arithmetic. In S. Shapiro (Ed.), *Intensional mathematics* (pp. 11–46). Amsterdam: North Holland Publishing Company.

Williamson, T. (1988). Knowability and constructivism. *Philosophical Quarterly*, *38*, 422–432.

Chapter 12
Arbitrary Reference in Mathematical Reasoning

Abstract It is claimed that the ideal possibility of picking up any object of the universe of discourse is essential not only in intuitionistic but also in classical logic and mathematics.

12.1 Introduction

Mathematicians use very often in their reasoning expressions of the kind "let *a* be an arbitrary object of the universe of discourse", for instance "let *a* be an arbitrary real number". Observe that there is no link between the letter "*a*" and the number which it is supposed to be indicating. Taking such expressions at face value, mathematical reasoning would seem to presuppose, at least ideally, the possibility of indicating any object of the universe of discourse, even when, as in the case of real numbers, not every object has a name in the language. Within the intuitionistic conception of mathematics, such a presupposition is made quite explicit by the doctrine according to which an object exists only as a mental construction of an *ideal mathematician*, so that any object is capable of being exhibited, and therefore indicated, by acquaintance with it. The ideal mathematician can refer to any object, in virtue of his direct access to his own mental constructions. In contrast, within the *classical (realist)* conception of mathematics, it would seem, from the literature, that the mathematical treatment of a domain of objects by no means requires the ideal possibility of individual reference to every object of the domain. It seems to be a widespread opinion that, once the objective existence of e.g. the real numbers has been accepted, the general theory of real numbers is developable, by using the device of quantification, without any need that each number be capable of individual reference.

On the contrary, we claim that the *ideal possibility* of referring individually to any object of the universe of discourse is essential even in the realist perspective. We will call this claim *TAR* (*thesis of arbitrary reference*).

12.2 Some Objections to *TAR*

There are a number of arguments in the literature, which seem to be in disagreement with our thesis. The most obvious is perhaps due to a misunderstanding of the notion of arbitrariness. One can argue that considering an arbitrary number is nothing but a way of speaking, which by no means involves the possibility of actually singling out such a number, since, for the very same arbitrariness, it is irrelevant which number one is speaking of. Indeed, when reasoning about an arbitrary number, there is no need to know it. Furthermore, the argument goes, the ignorance of which number one is referring to has the desired effect to grant generality to the reasoning: what is provable for a completely unknown number holds necessarily for all numbers. However, the lack of information about a cannot avoid the *assumption* that the letter "a" designates a precise number: lacking that assumption, it would make no sense talking about a, not even to say that it is unknown.

Some remarks on the notion of *possibility* involved by *TAR* are in order. Such a possibility is not to be understood as the ability of the speaker to exhibit or describe the arbitrary object he is referring to. Of course we do not need such ability in order to speak of an arbitrary number or of an arbitrary (possibly extraterrestrial) living being. What is needed is merely to *imagine* that the object in question has been in some way fixed. The verb "to fix" often recurs in the course of informal mathematical talk. For instance, the well-known $\epsilon - \delta$-definition of limit sounds: "however a positive number ϵ has been fixed, you can find a positive number δ such that …". In such a context, you do not have to worry about *how* ϵ has been fixed, but you must *imagine* that *in some way* it has been fixed and that it may be *any* positive number. We want to hold that locutions of this kind are not to be regarded as a mere way of speaking, but that they play an essential role in mathematical reasoning. The use in natural language of the indefinite article "a" may erroneously suggest that, in order to talk about an arbitrary object, there is no need to think of it as well-determined. You can realise that this suggestion is deceptive by reflecting on the use of pronouns, which do refer to a well-determined object, even when this has been introduced by means of the indefinite article. Consider e.g. the talk: "Take an arbitrary real number. …Suppose *it* is irrational …" Which number does the pronoun "it" refer to? Of course the correct answer is not "to any real number", but "to *the* number under consideration". In this answer, the *definite* article is used just for referring to *the* number introduced by the *indefinite* article. The puzzle "How can the definite article be appropriate, since, as a matter of fact, no number has been fixed?" has, we maintain, the answer: "To consider an arbitrary number means to *imagine* that a number has been fixed. *Imagination* is all is required for this kind of reference". Arbitrary reference rests on our ability of imagining that an object of the universe of discourse has been fixed.

The importance of imagination in the platonist conception of mathematics has been emphasized by Bernays in his famous paper "On platonism in mathematics":

> The value of platonistically inspired mathematical conceptions is that they furnish models of abstract imagination. These stand out by their simplicity and logical strength. They form representations which extrapolate from certain regions of experience and intuition. (Bernays 1964)

If, as we believe, *TAR* is correct, it is of remarkable interest for the philosophy of mathematics. It poses the problem of supplying a more definite content to the act of imagining involved by arbitrary reference, as a constituent of mathematical realism. Before addressing this problem, we want to discuss some further possible objections to our thesis.

An argument against *TAR* may contend that, though arbitrary reference occurs in informal reasoning, it is not essential, since it may be avoided by the use of quantifiers, which do not refer individually to any object of the quantification domain. This argument rests on the confusion between the locutions "any" and "each": talking about any object may seem to amount to talking about *each* object. This is not the case, however. Russell was clearly aware of the difference:

> The general enunciation tells us something about (say) all triangles, while the particular enunciation takes one triangle and asserts the same thing of this one triangle. But the triangle taken is any triangle, not some one special triangle; and thus, although, throughout the proof, only one triangle is dealt with, yet the proof retains its generality. If we say: "Let *ABC* be a triangle, then the sides *AB* and *AC* are together greater than the side *BC*", we are saying something about *one* triangle, not about *all* triangles; but the one triangle concerned is absolutely ambiguous, and our statement consequently is also absolutely ambiguous. We do not affirm any one definite proposition, but an undetermined one of all the propositions resulting from supposing *ABC* to be this or that triangle. This notion of ambiguous assertion is very important, and it is vital not to confound an ambiguous assertion with the definite assertion that the same thing holds in *all* cases.
>
> The distinction between (1) asserting any value of a propositional function and (2) asserting that the function is always true is present throughout mathematics, as it is in Euclid's distinction of general and particular enunciations. In any chain of mathematical reasoning, the objects whose properties are being investigated are the arguments to any value of some propositional function […] For this reason, when *any* value of a propositional function is asserted, the argument […] is called a *real* variable, whereas, when a function is said to be *always* true, or to be *not always* true, the argument is called an *apparent* variable […]
>
> If ϕx is a propositional function, we will denote by "$(x) \cdot \phi x$" the proposition "ϕx is always true". […] Then the distinction between the assertion of all values and the assertion of any is the distinction between (1) asserting $(x) \cdot \phi x$ and (2) asserting ϕx where x is undetermined. The latter differs from the former in that it cannot be treated as one determinate proposition.
>
> The distinction between asserting ϕx and asserting $(x) \cdot \phi x$ was, I believe, first emphasised by Frege (1893, 31). His reason for introducing the distinction explicitly was the same which had caused it to be present in the practice of mathematicians, namely, that deduction can only be effected with *real* variables, not with apparent variables. In the case of Euclid's proofs, this is evident: we need (say) some one triangle ABC to reason about, though it does not matter what triangle it is. The triangle ABC is a *real* variable; and although it is *any* triangle, it remains the *same* triangle throughout the argument. But in the general enunciation the triangle is an apparent variable. If we adhere to the apparent variable, we cannot perform any deduction, and this is why in all proofs real variables have to be used. (Russell 1908)

In today's formal logic Russell's distinction between *real* and *apparent* variables is faithfully reproduced, with a sheer change in terminology, by the well-known distinction between *free* and *quantified* variables. Singular reference plays an essential role in quantification theory. This fact is made quite perspicuous by the meaning of the quantification rules in natural deduction. According to the elimination rule for the existential quantifier, in order to derive a conclusion A from an existential

assumption $\exists x P(x)$, one has to assume $P(a)$ (where a is a fresh free variable) and derive A from $P(a)$ (with the due restrictions). This rule is justifiable only if it is granted that, under the existential assumption, one can consider an arbitrary object a such that $P(a)$. A similar observation holds for the introduction rule of the universal quantifier. The soundness of classical natural deduction rests therefore on the hidden assumption that every object of the domain is capable of being the referent of some act of reference. Thus formal logic does justice to the informal locution "let a be an object such that $P(a)$". What the textbooks of logic fail to tell us is what the act of referring to an arbitrary object consists in. About this issue, the formal theory is silent and, as far as we know, the philosophical remarks in the literature seem to be somewhat confusing. Even Russell's passage, quoted above, though enlightening the need of reasoning about *single* arbitrary objects, seems to give a rather misleading explanation of the nature of arbitrary reference. To say that *any* triangle is not some one special triangle, but that it is absolutely ambiguous, might erroneously suggest that the triangle concerned is a strange object enjoying the strange property of being absolutely ambiguous. But, of course, an ontology of ambiguous objects would be far from desirable. Alternatively, Russell's explanation might suggest that what is ambiguous is the act of referring to any arbitrary triangle, in the sense that it is undetermined which triangle it refers to. Indeed Frege, after rejecting the first alternative, seems to hold the second:

> [Mr. E. Czuber] …defines a variable as an indefinite number. But are there indefinite numbers? Must numbers be divided into definite and indefinite? Are there indefinite men? Must not every object be definite? 'But is not the number n indefinite?' I am not acquainted with the number n. 'n' is not the proper name of any number, definite or indefinite. Nevertheless, we do sometimes say'the number n'. How is this possible? Such an expression must be considered in a context. Let us take an example. 'If the number n is even, then $\cos n\pi = 1$'. Here only the whole has a sense, not the antecedent by itself nor the consequent by itself. The question whether the number n is even cannot be answered; no more can the question whether $\cos n\pi = 1$. For an answer to be given, 'n' would have to be the proper name of a number, and in that case this would necessarily be a definite one. We write the letter n in order to achieve generality. This presupposes that, if we replace it by the name of a number, both antecedent and consequent receive a sense.
>
> Of course we may speak of indefiniteness here; but here the word 'indefinite' is not an adjective of 'number', but 'indefinitely' is an adverb, e.g., of the verb 'to indicate'. We cannot say that 'n' designates an indefinite number, but we *can* say that it indicates numbers indefinitely. (Frege 1904)

But, besides being quite obscure what an act of *indicating indefinitely* is, such indefiniteness would be incompatible with the essential fact, clearly stressed by Russell, that, in the whole course of the proof, the referent is always the same. This means that the letter "a" indicates the same object in all its occurrences in the proof. If the referents of such occurrences were not well-determined objects, it would be meaningless to say that they are the same. Perhaps Frege believed this objection to be superseded by his doctrine of functions as *unsaturated entities*. In "Function and Concept", he says:

> […] people who use the word 'function' ordinarily have in mind expressions in which a number is just indicated indefinitely by the letter x, e.g. '$2x^3 + x$' …[But] x must not be considered as belonging to the function; this letter only serves to indicate the kind of supplementation that is needed; it enables one to recognize the places where the sign for the argument must go in. (Frege 1891)

It seems that, according to Frege, "to indicate *indefinitely*" really means "not to indicate anything at all". The variables occurring in a functional expression do not denote anything; they are mere placeholders marking the gaps to be filled by individual names. According to this view, a reasoning about an arbitrary object x may be regarded as a schema of reasoning, i.e. as a function which maps every object into the reasoning obtained from the *schema* by replacing "x" with a name of such object. It is plain, however, that such a schema can have all the desired instances only if every object is capable of being singled out and named. Thus regarding "x" as a schematic letter is of no help to avoid *TAR*.

We think therefore that the ambiguity shown by Russell is to be understood in a purely *epistemic* sense. Referring to an arbitrary object a amounts to supposing that "a" designates an *unknown*, though *well-determined*, object. Being well-determined justifies the behaviour of "a", in the course of the reasoning, as a name designating the same object in all its occurrences. On the other hand, being unknown guarantees that all that is established for it holds as well for any other object of the domain.

12.3 *TAR* as Embodied in the Logical Concept of an Object

The foregoing considerations show that arbitrary reference is essential for the proof theory of classical logic. Now, one can wonder if *TAR* is already implicit in the *semantics* of classical logic. Of course, the answer is certainly affirmative if one agrees that the meaning of logical constants is determined by the inference rules, since, as we have seen, arbitrary reference is involved in the quantification rules. We think, however, that such a thesis is not appropriate to mathematical realism. We assume therefore that the understanding of the semantic notions is prior to that of inference rules, which are justified a posteriori insofar as they are *recognised* as *truth preserving*. Now, most working mathematicians do agree that the inference rules are truth preserving. They are inclined to recognize as *intuitively* correct the metamathematical formal proof of the soundness theorem. In particular they feel the cogency of the argument that, after recognising $P(a)$ for a certain object a, without any assumption about a, one can rightfully conclude $(x)P(x)$. But, since the proof of $P(a)$ clearly exploits the *nameability* of a, the generalisation is justified only under the assumption that every object is *nameable*. The fact that no mathematician feels the need of making this assumption explicit seems to suggest that the possibility in principle of referring to any object individually is implicit in the very same general concept of object.

The involvement of *TAR* is not made explicit by Tarski's definition of truth either. In fact, this rests on the definition of satisfaction of a formula, relative to an *assignment* of *arbitrary* members of the universe of discourse to the free variables, an assignment being understood as a *set-theoretical function*. So, it may seem that the task of assigning arbitrary objects to the free variables is accomplished by a set-theoretical function. But that is illusory, since the problem of referring to an arbitrary object of the given domain shifts to that of referring to an arbitrary function from variables to objects. Thus, in order to avoid a regress *ad infinitum*, one must take, at some stage, arbitrary reference as primitive.

The commitment to *TAR* is quite evident in Hintikka's *game theoretical semantics* for first-order logic (Hintikka 1997). Here the meaning of logical constants is explained in terms of *acts of choice*. With every sentence of a first-order language, Hintikka associates a game between two ideal players, the *verifier* and the *falsifier*, who are trying, respectively, to verify and to falsify the sentence. He then defines the truth and the falsity of a sentence as the existence of a *winning strategy*, respectively, for the verifier and for the falsifier. The game rules are defined in terms of arbitrary choices of individuals by the players and introduction of names for the chosen individuals. The definition proceeds by induction on the complexity of the sentence. In particular, the clause for the existential quantifier is the following:

A play relative to $\exists x S(x)$ *starts with a choice of an individual b by the verifier. Then the plays continues as for* $S(b)$.

(The clause relative to the universal quantifier is similar, with the choice made by the falsifier).

Hintikka observes that the name "*b*" does not necessarily belong to the given language but that, since the length of a sentence is finite, any play requires only finitely many new names. Hintikka takes for granted the ideal possibility of choosing any individual and giving it a name. He stresses the constructive flavour of his semantics, arising, in his view, from the fact that the meaning of logical constants is grounded on the notion of *action*. He proves the soundness of classical logic for his semantics and concludes that, in spite of the intuitionistic tenets, classical logic is constructively justified. Indeed Hintikka's game rules are perfectly intelligible from the intuitionistic viewpoint. However, his proof of soundness rests on a tacit *realistic* attitude concerning the existence of a winning strategy: once the game rules have been established, he regards as a well-determined objective fact the existence or non-existence of a winning strategy for the verifier or for the falsifier. So Hintikka's proof that for every game there is a winning strategy for one of the two players fails intuitionistically. This explains why Hintikka's semantics turns out to be equivalent to the usual Tarskian semantics. Hintikka's "constructivism", based on the *action of choosing*, has nothing to do with intuitionistic *antirealism*.

We think, however, that Hintikka's philosophical perspective is coherent and that it supplies a faithful analysis of the usual mathematical reasoning. It makes explicit a "constructive" aspect hidden even behind the classical conception of mathematics. In particular it seems to regard the possibility in principle of choosing any individual as implicit in the logical concept of object.

The idea of *an act of choice* seems to provide an appropriate framework for understanding the notion of arbitrary reference. What is the content of the assumption that an object has been arbitrarily assigned to letter "*a*"? Since, as a matter of fact, neither the mathematician who makes the assumption nor any other real human being has assigned any object to "*a*", the assumption must concern an *imaginary assignment*. The precise content of such an assumption might seem to be irrelevant, since no mathematical reasoning about *a* needs taking into account the way the supposed act of assigning has been done. Any talk about *a* does exploit, however, the counterfactual *possibility* of such an assignment. Therefore, a careful analysis of what is implicitly assumed in mathematical reasoning must face the problem of explaining how to understand the possibility at issue. How can we imagine that *a* has been fixed? At first glance, one could imagine that every object is capable of being characterised by some property (possibly non-expressible in the formal language), so that it would be fixed by means of a definite description. But then the describing property should meet the condition of being satisfied by a unique object; and this condition involves a quantification over all objects. Now, since, as we saw, the game-theoretical explanation of the meaning of quantifiers rests on the assumption that any object can be chosen, one can recognise that an object can be fixed by means of a characterising property only under the assumption that it can be chosen. This suggests that arbitrary reference is more primitive than reference by description. We will pursue the idea that arbitrary reference is a sort of *direct reference* based on an imaginary *act of choice*.

12.4 The Ideal Agent

Let us imagine that we have *direct access* to an *ideal agent*, who in turn has *direct access* to every object: he can *choose* any object at will (here, we are identifying ourselves with the *working* mathematician carrying on the mathematical reasoning). We can explain the locution "Let *a* be an arbitrary object" as follows: we ask the agent to choose an object at his will (without communicating us anything about the chosen object) and call it "*a*". It is clear that the adjective "arbitrary" does not concern the nature of the chosen object, but the freedom of the choice act. Accordingly, we will assume the following *choice act principle*:

(*CAP*) *Every object of the universe of discourse is capable of being chosen by the ideal agent.*

Of course, *CAP* faces the problem of providing an account of what the act of choosing a mathematical object consists in. What does it mean to choose an *infinite* entity such as a real number or a set? It is hard to give a general answer, since any answer depends essentially on how the entities in question are conceived. *CAP* is to be seen as a constraint, which must be taken into account by any conception of mathematical objects. The structuralist development of mathematics has shown that mathematical theories do not determine the specific nature of the entities they are talking about.

Therefore, given any mathematical theory, the possibility is left of searching for models built up from objects whose accessibility to the ideal agent is perspicuous. For first-order arithmetic, a suitable model is Hilbert's model of numerals thought of as finite strings of strokes. These are, in Charles Parsons' words, *quasi-concrete objects*, i.e. types of spatio-temporal objects, the accessibility to which requires only a minimal idealisation of the agent, needed for dealing with the mathematical infinite. All that the agent is expected to be able to do is to write down any finite (however long) string of strokes. It is worth noticing that, at least from a logical point of view, this idealisation is also sufficient for interpreting any mathematical theory. For, as is well known, it follows from the Löwenheim–Skolem theorem that every consistent theory is interpretable in first-order arithmetic. Although arithmetical models are, in general, far from being the "intended" model of the given theory (say set theory), it is surprising that, as soon as one has accepted the idealisation of natural numbers, he can interpret within his framework any mathematical talk. Among other things, arithmetical models assure that, as soon as a theory is consistent, *CAP* is certainly satisfiable. Later, we will show how to use *CAP* for justifying the constraint of predicativity for Russell's intensional logic and for defending and developing Boolos' interpretation of second-order logic based on plural quantification. For the moment, we want to further explain some general aspects of *CAP*.

Our ideal agent, unlike the Brouwerian idealized mathematician, has no other job than that of performing arbitrary acts of choice. He is not expected to have the capacity of restricting his choices to objects satisfying some required condition. The inferential step from $\exists x F(x)$ to "let a be an arbitrary object such that $F(a)$" is justified by referring to a *completely arbitrary* choice of the agent, calling "a" the chosen object and *assuming* $F(a)$. Though this assumption may certainly be false, any logical consequence B we can draw from it must be true under the hypothesis $\exists x F(x)$ (provided that a does not occur in B). For, the existential hypothesis and *CAP* assure that the agent *could* have chosen (though unconsciously) an object satisfying F; and since we do not know anything about a, even in that counterfactual case our reasoning would be correct. And as the truth or falsity of B is quite independent of the effective choice of the agent, B is true anyway (under the existential assumption).

It may be instructive to compare *CAP* with the celebrated set-theoretical axiom of choice. *CAP* does not assert the existence of any mathematical object; it explains the meaning of free variables and justifies their inferential role. In contrast, the axiom of choice states that, given any set α of non-empty pairwise disjoint sets, *there is* a set β (call it the *choice set*) sharing a unique element with each member of α. From a logical point of view, no act of choice is involved in the understanding of the content of this axiom (besides that, with which we are concerned, implicit in the general concept of an object). It is a purely existential statement expressible in the language of set theory: as all sets, a choice set exists, in a realistic perspective, quite independently of any human action. On the other hand, the act of choice seems to constitute the intuitive ground for the existence of a choice set. It is usually agreed that the axiom serves the purpose of granting the existence of a choice set even when its elements are not singled out by any propositional function. The existence of such a set seems to be intuitively justified by thinking of its members as *arbitrarily chosen*.

This aspect was just the main source of the well-known dogged opposition to the axiom: it was charged of introducing into mathematics indefinable sets (a set being *definable* if it is the extension of a propositional function without parameters). It may be puzzling that, if α is finite, the axiom of choice is not needed for the existence of α, even when it is indefinable. In particular, given any single non-empty set α, no axiom of choice is needed in order to guarantee the existence of a subset β of α with a unique member. The reason is to be found in a hidden application of *CAP*, implicit in quantification theory. For, the proof runs as follows: *let a be a member of* α; by the pairing axiom there is a set β whose unique member is a. It is clear that this kind of reasoning (formalisable in axiomatic set theory) is correct only under the hidden assumption that any object of the universe of sets is capable of being chosen.

The role of choice in mathematics, contrary to a widespread belief, is far from being restricted to the use of the axiom of choice; it is pervasive of the whole mathematics and logic. The axiom of choice seems to exploit the idea of choice in a more problematic way, since it involves the possibility of a *simultaneous choice* of infinitely many objects. Later we will argue, however, that this possibility is already implicit in the usual notion of a set as constituted by its members (in contrast to the logical notion of a class as extension of a property).

One may object, against the need of *CAP*, that the mathematical language can be understood by *direct extrapolation* from the ordinary talk about concrete objects (for which reference is not problematic), so that, in particular, arbitrary reference to mathematical objects would be immediately intelligible without any need of further explanations. According to this objection, the familiar understanding of a talk about any man, any horse or any pencil would make a talk about any real number immediately meaningful. This opinion seems to be shared by Shapiro. He notes the elusiveness of reference in mathematics, but does not seem to find it very problematic:

> Probably the most baffling, and intriguing, semantic notion is that of *reference*. The underlying philosophical issue is sometimes called the "'fido'-fido problem". How does a term come to denote a particular object? What is the nature of the relationship between a singular term ("Fido") and the object that it denotes (Fido), if it denotes anything? Notice that model theory, by itself, has virtually nothing to say on this issue. In textbook developments of model theory, reference is taken as an unexplicated primitive. It is simply *stipulated* that an "interpretation" includes a function from the individual constants to the domain of discourse. This is a mere shell of the reference relation.
>
> […] As far as the model-theoretic scheme goes, it does not matter how this reference is to be accomplished or whether it can be accomplished in accordance with some theory or other. There is nothing problematic in the abstract consideration of models whose domains are beyond all causal contact. As far as model theory goes, reference can be *any* function between the singular terms of the language and the ontology. […]
>
> It is fair to say that when it comes to mathematics and theories of other *abstracta*, realism in ontology often falters over reference (about as much as it falters over epistemology). If we assume that ordinary languages are understood and if we accept the premise that model theory captures the structure of ordinary interpreted languages, then we can do better. There is, of course, no consensus on how reference to ordinary physical objects is accomplished. The theories are legion. I do presume, however, that reference to proverbial medium-sized physical objects is accomplished. […] Understanding how to use ordinary language involves an understanding, at some level, of reference (however it works). (Shapiro 1997)

Here Shapiro seems to hold that the comprehension of the notion of reference, acquired from the use of natural language, is sufficient for understanding the reference to the objects of any abstract mathematical structure. Our reply is that the understanding of the general notion of reference, "*however it works*", rests necessarily on the presupposition that *in some way* (though it does not matter which) it *must* work. Therefore the question: is there any way of referring to an arbitrary real number? Shapiro is an upholder of the so-called *realism ante rem*, according to which mathematical objects are conceived of as *positions* in *abstract structures*, whose existence is prior to their possible specific instances. But, aside from the difficulty of a non-metaphorical understanding of what such positions are, if they are to be treated as objects of quantification, one cannot avoid the question: what does it mean to single out an arbitrary position?

12.5 Arbitrary Reference and Impredicativity

In order to better enlighten the reasons that may have obscured the importance of *TAR* and *CAP*, it is worthwhile to discuss certain observations by Ramsey and Gödel concerning Russell's ramified type theory. Ramsey criticises the doctrine of the *Principia Mathematica* according to which every class is defined by a propositional function. He observes that, since it is impossible to list all members of an infinite class, there is no evidence that, in general, such a class is definable by a propositional function. He continues:

> To this it will be answered that a class can only be given by enumeration of its members, in which case it must be finite, or by giving a propositional function which defines it. So that we cannot be in any way concerned with infinite classes or aggregates, if such there be, which are not defined by propositional functions. But this argument contains a common mistake, for it supposes that, because we cannot consider a thing individually, we can have no concern with it at all. Thus, though an infinite indefinable class cannot be mentioned by itself, it is nevertheless involved in any statement beginning 'All classes' or 'There is a class such that', and if indefinable classes are excluded the meaning of all such statements will be fundamentally altered. (Ramsey 1925)

Clearly Ramsey does not take into account the problem of arbitrary reference, of which, as we saw, Russell was aware. We want to suggest that one of Russell's reason for adopting the logicist notion of class as extension of a propositional function arises from the question: how can one choose an infinite class? Russell's answer was: through the choice of a propositional function. Russell's option seems to be justified by the consideration that propositional functions, because of their *intensional* nature, are, at least in principle, directly accessible to the human mind, whereas sets, understood as entities built up by their members, are not. An alternative option is that of fixing a set through a *simultaneous choice* of its elements. This will be considered later.

Ramsey's argument has been resumed by Gödel in his paper "Russell's mathematical logic". Gödel criticises Russell's *vicious circle principle*, according to which

no totality can contain members definable *only* in terms of the totality itself. Gödel observes that classical mathematics does not respect such a principle; and since classical mathematics can be reconstructed on the basis of *Principia*, this work itself cannot respect that principle either "if 'definable' means 'definable within the system' and no methods of defining outside the system (or outside other systems of classical mathematics) are known except such as involve still more comprehensive totalities than those occurring in the systems". He adds:

> I would consider this rather as a proof that the vicious circle principle is false than that classical mathematics is false, and this is indeed plausible also on its own account. For, first of all one may, on good grounds, deny that reference to a totality necessarily implies reference to all single elements of it or, in other words, that "all" means the same as an infinite logical conjunction. (Gödel 1984)

Then Gödel observes that, even if "all" were intended as an infinite conjunction, the vicious circle principle would be tenable only within a constructive perspective:

> In this case [i.e. if the entities in question are constructed by us] there must clearly exist a definition (namely the description of the construction) which does not refer to a totality to which the object defined belongs, because the construction of a thing can certainly not be based on a totality of things to which the thing to be constructed itself belongs. If, however, it is a question of objects that exist independently of our constructions, there is nothing in the least absurd in the existence of totalities containing members, which can be described (i.e., uniquely characterised) only by reference to this totality. (*ibid*)

In a note, he points out that "an object a is said to be described by a propositional function $\phi(x)$ if $\phi(x)$ is true for $x = a$ and no other object".

Certainly by "definable" Russell does not mean "definable in the system". For, it is well known that those predicative propositional functions whose existence is granted by the axiom of *reducibility* fail, in general, to be definable in the object language of type theory: the range of the variables includes functions indefinable in the object language. But, as we saw, the values of quantified variables must be capable of singular reference. It would therefore be circular to accept as range of quantification a universe with some members definable *only* by quantifying over the universe itself. Gödel's remarks do not take into account the sense of "definition" as ideal singular reference. By contrast, *TAR* supplies a plausible justification of the vicious circle principle, provided that "definable" is interpreted as "capable of singular reference". For, as we saw, quantification over a universe U presupposes the possibility of arbitrary reference to any member of U. So an act of arbitrary reference cannot involve in turn, on pain of vicious circularity, quantification over U. Precisely, we can restate the vicious circle principle as follows:

(VCP) *No universe of discourse can contain a member such that the agent can refer to it only by means of quantification over the universe itself.*

VCP leads to the rejection of the *impredicative comprehension principle* of second-order (and higher-order) logic

(CP) $\exists F \forall x (F(x) \leftrightarrow A(x))$,

provided second-order entities are understood, *à la* Russell, as *intensional* entities.

For, let F be the property expressed by the propositional function $A(x)$ (where x is a free individual variable). Because of the intensionality of F, there is no access to it but through its linguistic expression. In other words, a choice of F can be understood only as the thought of the formula $A(x)$ with its intended interpretation. A second-order quantification occurring in $A(x)$ would be therefore a violation of VCP.

For instance, take for $A(x)$ the propositional function

(*) $\forall G G(x)$.

The property F, whose existence is assured by CP, is, according to the intensional interpretation, the property of enjoying every property. There is no way of grasping this property without using a quantification over all properties. But that presupposes, as we saw, the a priori possibility of choosing *any* value of G. It follows that the property at issue cannot be a value of second-order variables. So, CAP supplies an explanation of why one cannot take the universe of all individual properties as the range of second-order variables.

Besides, VCP is compatible with Russell's reducibility axiom. For, one can imagine that the agent has direct access to a certain universe of *primitive* properties (possibly non expressible in the formal language). If second-order variables are restricted to such properties, CP can be accepted, without any circularity, as a *richness* assumption: it says intuitively that the universe of primitive properties is so rich that every *extension* of a second-order propositional function is the extension of some primitive property.

Thus, our analysis supplies a new reason for adopting a ramified hierarchy, when dealing with intensional properties and relations.

12.6 Plural Reference Versus Sets

Plural quantification is a reinterpretation of second-order monadic logic, proposed by Boolos (1984, 1985). In Boolos' perspective second-order monadic logic is ontologically innocent: contrary to the most accredited view, it does not entail any commitment to classes or to properties but only to individuals. According to Boolos, second-order quantification differs from first-order quantification only in that it refers to individuals *plurally*, while the latter refers to individuals *singularly*.

Boolos' view, though very attractive, is highly controversial. It has met the criticism of several philosophers of mathematics (see Resnik 1988 and Parsons 1990). Quine's old claim that second-order logic is "set theory in disguise" does not seem to have lost its advocates.

We want to show how the theory of arbitrary reference can throw new light on the theory of plural quantification.

Let us examine Resnik's criticism of Boolos' proposal.

Boolos argues that Quine's slogan "to be is to be the value of a variable" does not entail that the value of a second-order variable must be a set (or a property) of individuals. The slogan is compatible, Boolos claims, with the *plural interpretation*, according to which the value of such a variable is a *manifold* of individuals. To the purpose, he restates the Tarskian truth definition for second-order logic by modifying the notion of *assignment*.

Precisely, given a domain D of individuals, he defines as an assignment any binary relation R between variables and individuals which correlates a unique individual with every first-order variable, while it is subject to no constraint for second-order variables. So R may correlate a second-order variable with no, one or (possibly infinitely) many individuals. The satisfiability relation is inductively defined as usual, with the following clauses for atomic formulas and second-order existential quantification:

(1) R satisfies the atomic formula Fx iff the correlate of x is one of the correlates of F;
(2) R satisfies $\exists FA$ iff there is a relation R', differing from R at most for the correlates of F, such that R' satisfies A. (The universal quantifier is defined in terms of the existential one).

Truth is then defined as usual in terms of satisfaction. So, the *set* of the correlates of F is not involved in the definition of truth.

This makes the notion of plural quantification precise and shows how it yields an alternative semantics for second-order logic. This semantics turns out to be equivalent to the usual one, according to which the values of second-order variables are all sets of individuals. And since the notion of *value of a variable* can be made precise only by the definition of assignment, the proposed reformulation shows that Quine's slogan does not commit second-order logic to any entities but individuals.

It is clear, however, that Boolos' device is, in itself, inadequate for the conclusion that plural quantification does not implicitly involve the notion of class. The problem is simply turned into the following: does the new definition of assignment presuppose the notion of set of individuals? The answer is certainly affirmative, of course, if relations are understood set-theoretically. But a relation can in turn be understood in terms of *plural reference* to certain ordered pairs (taking for granted the notion of ordered pair). So the definition of assignment becomes: certain ordered pairs R are an assignment if their first components are variables, their second components are individuals and every first-order variable occurs in exactly one of the R's. However, the use of plural reference in the metalanguage begs the crucial question, whether plural reference involves surreptitiously the notion of set. Boolos is aware of this difficulty and does not attempt to convince the opponents that plural reference is free of any commitment to sets. He only remarks that who is inclined to see plural reference as a genuine alternative to classes will certainly appreciate the possibility of recovering within his view the Tarskian definition of truth.

Indeed, several authors have raised some doubts about the alleged ontological innocence of plural reference. Resnik observes that the use of plural reference in natural language is ambiguous and that, at least in certain contexts, there is no

evidence that it is free of any commitment to classes. The locution "there are some objects such that …" sometimes simply means "there is at least an object such that …", so that it is expressible in first-order language. Sometimes, however, it has a meaning which one can hardly make explicit without invoking the notion of class. For instance, the famous Geach–Kaplan's example "some critics admire only one another" is paraphrased by Boolos as "there are some critics such that each of them admires a critic only if the latter is one of them different from the former". This proposition, not formalizable in first-order language, seems, according to Resnik, hardly interpretable without resorting to classes. How could we understand "one of them" without referring to a certain class and agreeing that the referent of "one" belongs to it? In general, while, according to Boolos, the use of plural reference in natural language would testify to the ontological innocence of second-order logic, according to Resnik the use of second-order logic for formalising those plural references non-expressible in first-order logic would bring to light certain ontological commitments hidden behind natural language. A similar criticism has been made by Parsons, although he attributes to Boolos the merit of throwing new light on the old notion of manifold:

> Boolos has not, in my view, made a convincing case for the claim that his interpretation of second-order logic is ontologically non-committal. The great interest of his reading, in my view, is that he breathes new life into the older conception of pluralities or multiplicities. As a source of second-order logical forms, the plural and plural quantification are rightly distinguished from what was so much emphasized by Frege, predication and, more generally, expressions with argument places. In particular, if it is the idea of generalization of predicate places that we appeal to in making sense of second-order logic, then the most natural interpretations will be relative substitutional or by semantic ascent, and these will not license impredicative comprehension, and it is hard to see how that will be justified. But if one views examples such as Boolos's as involving 'pluralities', they are more like sets as understood in set theory in that no definition by a predicate is indicated, so that one need not expect them to be definable at all. Thus no obstacle to the acceptance of impredicative comprehension is removed.
>
> An advocate of Boolos' interpretation in an eliminative structuralist setting could grant my claims about ontological commitment, but then take a position analogous to the Fregean: second-order variables indeed have pluralities as their values, but these are not objects. It does not seem to me to have the same intuitive force as Frege's position, since there is no analogue to the regress argument that can be made if one views the reference of a predicate as an object. There will still be, just as with Frege's concepts, the irresistible temptation to talk of pluralities as if they were objects, as we have already noted above. The only gain this interpretation offers over the Fregean is a more convincing motivation of impredicativity. (Parsons 1990)

Certainly the use of plural reference in natural language does not guarantee, in itself, its ontological innocence. Plural reference to individuals often seems nothing but a sloppy reference to a class of individuals. The attempts to paraphrase the language of classes by using locutions of natural language avoiding explicit reference to singular values of second-order variables cannot dispel the doubt that classes are only concealed. We believe, however, that the theory of arbitrary reference can support the claim that the role of classes as referents of second-order variables is inessential.

Parsons observes, in the quoted passage, that, though Boolos' interpretation does not reach the goal of ridding second-order logic of any commitment to classes, nevertheless it gives new evidence that classes can be thought of as pluralities in the set-theoretical sense, in contrast to classes in the logicist sense as extensions of predicates. This interpretation would have, according to Parsons, the merit of justifying the *impredicative* comprehension principle. For, a set as a plurality of individuals exists quite independently of any description of its members and, therefore, describing it by quantifying over all sets by no means yields any circularity.

We want to argue, however, that, in virtue of the doctrine of arbitrary reference, the very same notion of a set as constituted by its members rests on the notion of plural reference, so that the latter turns out to be more fundamental than the former.

Assume that second-order variables range over sets of individuals. According to *CAP*, every such set must be capable of being chosen by the agent. The problem arises how to conceive the act of choosing such a set (taking for granted the accessibility to any individual). Now, all we know about sets is that they are entities determined by their members. Although we regard a set as a single object, we lack any insight about its individuality. Once the logicist notion of a class as extension of a concept has been rejected, one has no longer any intuition of what should keep together the members of a set. This fact has been clearly pointed out by Black in his famous paper "The elusiveness of sets":

> […] Cantor's formula, stripped to essential, runs quite simply: "A set is an assembly into a whole of (well-defined) objects". Here, the phrase "assembly into a whole" certainly suggests that something is *to be done* to the elements, in order for the "whole" or "the unified thing", which is the set to result. But *what* is to be done, if not merely thinking about, the set? . . . What kind of unification is in point? …The truth is that once the elements of a set have been identified, *nothing* needs or can be done to produce the corresponding set. (Black 1971)

But then it seems that there is no other way of access to a set than through its members. So a choice of a set must consist in the choice of its members. Now, the choice of infinitely many individuals may be thought of either as an infinite process of choosing a single individual at a time or as a simultaneous choice of all the individuals in question. The first alternative would allow the choice of only countably many individuals, whose totality would be undetermined (an infinite process of choices being forever *in fieri*). In this perspective, a set could be thought of as a well-determined entity only by identifying it *intensionally* with the process itself of choosing its members. But the introduction of entities with an undetermined extension would be highly problematic (as it is the case for intuitionistic *lawless sequences*) and incompatible with the extensional conception of a set. So, we are led to the second alternative of the *simultaneous* choice. The idea is expressed by Bernays in the already mentioned essay:

> [Platonism] abstracts from the possibility of giving definitions of sets, sequences, and functions. These notions are used in a "quasi-combinatorial" sense, by which I mean: in the sense of analogy of the infinite to the finite …we imagine functions engendered by an infinity of independent determinations which assign to each integer an integer, and we reason about the totality of these functions. In the same way, one views a set of integers as the result of

> infinitely many independent acts deciding for each number whether it should be included or excluded. (Bernays 1964)

Now, the simultaneous choice of certain individuals is precisely what serves the purpose of *plurally* referring to such individuals. It follows that the arbitrary reference implicit in second-order quantification involves the same choice acts, whether second-order variables range singularly over sets or plurally over individuals.

In the plural interpretation, the locution "let *A* be arbitrary individuals" means "choose at will some individuals simultaneously and call them '*A*' ". In the set-theoretical interpretation, the locution "let *A* be an arbitrary set of individuals" means "choose at will some individuals simultaneously and call '*A*' their set". At this point, it is plain that sets are inessential. The alleged role of sets of collecting individuals turns out to be illusory: what selects the members of a set is not the set itself but the *act* of choosing them simultaneously. Thus, the arbitrary reference to certain individuals by no means presupposes the existence of their set; it merely presupposes the act of choosing them simultaneously. So, the *ontological* innocence of plural reference is vindicated.

We can conclude that the plural interpretation of second-order logic is less ontologically committal than the set-theoretical one: both involve the same acts of choice, but the plural interpretation does not involve any second-order entities (acts, unlike sets, being no entities). The doubts raised by Resnik position on this subject is clearly and Parsons are therefore superseded.

This conclusion does not entail, however, that second-order logic, in the plural interpretation, is no more problematic than first-order logic. Certainly it is, but not for ontological reasons. What is more problematic is the conception of *simultaneous* choice of (possibly infinitely) many individuals, compared with that of choice of a *single* individual involved in first-order logic. The question arises: given a suitable idealisation of *singular choice* (depending on the nature of the individuals we are dealing with), how can we idealise a *simultaneous choice*? Though he does not explicitly talk about choices, Black suggests that one can easily conceive the act of indicating several things at once:

> The notion of "plural" or simultaneous reference to several things at once is really not at all mysterious. Just as I can point to a single thing, I can point to two things at once, using two hands, if necessary; pointing to two things at once need be no more perplexing than touching two things at once. Of course it would be a mistake to think that the rules for "multiple pointing" follow automatically from the rules for pointing proper; but the requisite conventions are almost too obvious to need specification. The rules for "plural reference" are no harder to elaborate. (Black 1971)

Let us try to propose a suitable ideal picture of a *simultaneous choice*. Imagine that, instead of a unique agent, infinitely many agents are available. More precisely, imagine a *leader* agent at the head of a team of subagents, one for every individual of the universe of discourse. When the leader orders "choose!", each subagent shows ad libitum one of the signs 0, 1, say by lifting a shovel with the signs printed each in one of its faces. Relative to a simultaneous choice, an individual is *designated* if the corresponding agent shows 1. So, a simultaneous choice plays the role of the

characteristic function of the set of the designated individuals. In contrast, a *singular choice* simply consists in a choice of a single individual by the leader. Again such individual is said to be *designated* by the singular choice.

Now, if one accepts the ontology of sets of individuals, then he can regard a simultaneous choice as a device for arbitrarily referring to sets. But what is important is that, once this device has been introduced, sets, understood as genuine entities, become quite inessential for interpreting second-order monadic logic. In fact, second-order truth can be directly defined in terms of choices as follows.

Let ϕ be a second-order monadic formula whose free first-order variables are among $x_1, \ldots, x_m$ and free second-order variables among $X_1, \ldots, X_n$. Consider for each variable x_i a singular choice $x_i^*(i = 1, \ldots, m)$ and for each variable X_j a simultaneous choice $X_j^*(j = 1, \ldots, n)$. We will inductively define the truth value of ϕ relative to the choices $x_1^*, \ldots, x_m^*; X_1^*, \ldots, X_n^*$. We will state only the clauses for atomic formulas and for second-order quantifiers, the others being as usual:

(1) if $\phi \equiv X_j x_i$, it is true if the individual designated by choice x_i^* is designated by choice X_j^*;
(2) if $\phi \equiv \forall Y \psi$, it is true if, however a plural choice Y^* is performed, ψ is true relative to choices $x_1^*, \ldots, x_m^*, X_1^*, \ldots, X_n^*, Y^*$;
(3) if $\phi \equiv \exists Y \psi$, it is true if it is possible to perform a plural choice Y^* in such a way that ψ turns out to be true relative to choices $x_1^*, \ldots, x_m^*, X_1^*, \ldots, X_n^*, Y^*$.

Observe that, while Boolos' truth definition explains plural quantification in the object language by assuming plural quantification in the metalanguage, the present approach explains singular and plural quantification in the object language assuming the notion of quantification over choices in the metalanguage. Such explanation avoids any circularity and its importance rests on the fact that choices are not objects but *acts*. Any talk which reifies acts treating them as objects is to be paraphrasable, in principle, so to avoid any reification. In particular, quantification over acts is to be understood in a purely potential sense. Clause (2) does not quantify over a mysterious realm of all acts of choice; what it requires for the truth of ϕ is that, if the leader orders a simultaneous choice relative to variable Y, then, *independently* of what each subagent chooses, ψ turns out to be true. Besides, such independence is to be thought of as an *objective fact*, which obtains or not quite independently of the knowledge, even on the part of the leader, of which is the case. Similarly for the *possibility* involved at clause (3): the possibility that the subagents make their choices so as to verify ψ is an objective fact which obtains or not quite independently of their knowledge. This guarantees the validity of classical logic. Accordingly, one could take only one of the two quantifiers as primitive and define the other in terms of that. The truth of $\forall Y \psi$ can be understood as the *impossibility* of a simultaneous choice falsifying ψ; the truth of $\exists Y \psi$ as denying that ψ turns out to be false, *independently* of what the subagents choose. We prefer, however, to take both quantifiers as primitive, since none of them seems more elementary than the other; we believe that each of them can help to clarify the other.

So far our semantics has been concerned with *monadic* second-order logic. Boolos' treatment extends plural quantification to *full* second-order logic by tak-

ing the notion of ordered pair as primitive. Lewis has proposed a codification of a pairing function by combining plural reference with mereology. Within our framework, even ordered pairs can be introduced by means of simultaneous choices, as follows.

Call a *binary choice* an act consisting in the choice by every subagent of two (not necessarily distinct) individuals in a certain order. A binary choice is a *pairing choice* if, for all individuals x, y, a unique subagent chooses them orderwise. We will assume the possibility of a pairing choice (the notion of *possibility* being explained as above) and we will speak of (*ordered*) *pairs* understanding the reference to such a choice. In this way, our semantics extends to full (poliadic) second-order logic. As we have already observed, our notion of choice can be viewed as an extension to the plural case of that used by Hintikka in his game-theoretical semantics.

In fact, you could further stress the analogy with Hintikka's semantics by reformulating our semantics game theoretically. For, with every sentence of second-order logic you can associate a game played by two teams, the team of *verifiers* and that of *falsifiers*. Each team consists of a leader and of one player for every individual. A move of a team consists of a singular choice by the leader or of a simultaneous choice by his players. The game rules are then defined as in Hintikka's theory, the moves relative to second-order quantifiers being simultaneous choices. The truth of a sentence is defined as the existence of a *winning strategy* for the team of the verifiers. The problem of the ontological commitment to sets shifts to that of the ontological commitment to strategies. Of course, if these are understood as set-theoretical functions, no step forward has been taken. But a winning strategy can be understood, without any reification, in terms of the notions of possibility and independence explained above. To say that there is a winning strategy for a team means that this team *can* win *any* play, quite *independently* of the moves of the opposing team. To prefer the formulation *à la* Tarski or that *à la* Hintikka is a matter of taste. Both exploit the same primitives: *choice acts*, *possibility* and *independence*.

Acknowledgements I am grateful to my colleagues Pierdaniele Giaretta and Ernesto Napoli for a number of helpful comments. My best thanks to Aldo Antonelli for inviting me to write the paper.

References

Bernays, P. (1964). On platonism in mathematics. In P. Benacerraf & H. Putnam (Eds.), *Philosophy of mathematics* (pp. 274–286). Cambridge: Cambridge University Press.

Black, M. (1971). The elusiveness of sets. *Review of Metaphysics*, *24*, 614–636.

Boolos, G. (1985). Nominalist platonism. *Philosophical Review*, *94*, 327–344.

Bools, G. (1984). To be is to be a value of a variable. *Journal of Philosophy*, *81*, 430–449.

Frege, G. (1891). Function and concept. In P. Geach & M. Black (Eds.), *Tranlsations from the philosophical writings of Gottlob Frege* (pp. 137–156). Oxford: Blackwell.

Frege, G. (1893). *Grundgesetze der Arithmetik* (Vol. I). Jena: Pohle.

Frege, G. (1904). What is a function? In B. MacGuinness (Ed.), *Collected papers in mathematics logic and philosophy* (pp. 285–292). Oxford: Blackwell.

Gödel, K. (1984). Russell's mathematical logic. In P. Benacerraf & H. Putnam (Eds.), *Philosophy of mathematics* (pp. 211–232). Cambridge: Cambridge University Press.
Hintikka, J. (1997). *The principles of mathematics revisited*. Cambridge: Cambridge University Press.
Parsons, C. (1990). The structuralist view of mathematical objects. *Synthese*, *84*, 303–346.
Ramsey, F. (1925). The foundations of mathematics. *Proceedings of the London Mathematical Society*, *25*, 338–394.
Resnik, M. D. (1988). Second order logic still wild. *Journal of Philosophy*, *75*, 75–87.
Russell, B. (1908). Mathematical logic as based on the theory of types. In J. van Heijenoort (Ed.), *From Frege to Gödel* (pp. 150–182). Cambridge, MA: Harvard University Press.
Shapiro, S. (1997). *Philosophy of mathematics*. Oxford: Oxford University Press.

Chapter 13
The Priority of Arithmetical Truth over Arithmetical Provability

Abstract It is claimed that the idea of grounding truth on provability perverts the very nature of the intuitive notion of proof and that the possibility of grasping proof-conditions without presupposing some realist notion of truth is illusory.

13.1 Introduction

As is well known, Intuitionism identifies mathematical truth with provability. A basic philosophical tenet of Intuitionism is that mathematical entities exist only as mental constructions and that, in the absence of a reality determining mathematical truth, a mathematical proposition can be true only in virtue of a proof, whence the rejection of Tarski's semantics and classical logic. Therefore, Intuitionism regards the notion of intuitive proof as the fundamental semantic notion on which the meaning of a proposition is founded: to know the meaning of a proposition is exactly to know its *proof-conditions*, that is to grasp what counts as a proof of it. This position faces a major difficulty: for a working mathematician a proof is an argument which shows the truth of a proposition; so if propositions cannot be true independently of being proved, one wonders what a proof has to prove. I think that this obvious general objection to any proof-theoretical semantics has never received a satisfactory answer from the constructivist. For a formal language, the best answer is supplied, prima facie, by Heyting's inductive explanation of the proof-conditions of a compound sentence in function of the proof-conditions of its components. I will argue, however, that Heyting's strategy does not succeed in avoiding a realist notion of truth. I think that, in general, the idea of grounding truth on provability perverts the very nature of the intuitive notion of proof and that the possibility of grasping proof-conditions without presupposing some realist notion of truth is illusory. When discussing some of the difficulties of constructivism, this objection was expressed, though only in passing, by Gödel (1995, pp. 312–313):

> One may […] object that the meaning of a proposition about all integers, since it is impossible to verify it for all integers one by one, can consists only in the existence of a general proof. Therefore, in the case of an undecidable proposition about all integers, neither itself nor its negation is true. Hence neither expresses an objectively existing but unknown property of the integers. I am not in a position now to discuss the epistemological question as to whether

this opinion is at all consistent. It certainly looks as if one must *first* understand the meaning of a proposition *before* he can understand a proof of it, so that the meaning of "all" could not be defined in terms of the meaning of "proof". But independently of this epistemological investigation, I wish to point out that one may conjecture the truth of a universal proposition (for example that I shall be able to verify a certain property for *any* integer given to me) and at the same time conjecture that no general proof for this fact exists. It is easy to imagine situations in which both these conjectures would be very well founded.

A similar observation has been expressed by Hintikka (for instance in Hintikka (1997)). He criticises the intuitionistic standpoint as resting on a confusion between *determining* and *recognising* the truth of a proposition. Hintikka discusses the matter within the framework of his *game-theoretical semantics*. He defends the priority of truth over provability by emphasising that "to seek truth, one has to know what truth is" and stresses the fact that his semantical games are *truth-constituting* and are to be well distinguished from *truth-seeking* games, the latter being parasitic on the former. Furthermore, he stresses that semantical games have a constructive character, so that the notion of truth determined by means of them should be accepted by the intuitionists. If this were the case, since, for a standard first-order language, game-theoretical truth turns out to be equivalent to Tarski's classical truth, the intuitionistic acceptability of classical logic would follow. The rules governing the moves of semantical games are certainly intuitionistically unobjectionable. However, Hintikka's definition of truth also involves the highly problematic notion of a *winning strategy*, the truth of a proposition being defined, with reference to the corresponding game, as the existence of a winning strategy for the verifier. Hintikka treats this notion realistically, a winning strategy being thought of as something whose existence or non-existence is well determined by the rules of the game. On the other hand, such a winning strategy exists for an intuitionist only insofar as it is *known* (or at least *knowable*[1]), and fails to exist only insofar as the impossibility is *known* (or at least *knowable*) to find one. So an intuitionist can immediately reject Hintikka's argument (see Tennant 1998). I think, however, that Hintikka's argument is essentially correct and can be defended by holding that a suitable amount of realism should be accepted even by the most radical constructivist.

Here, I will try to defend the priority of truth over provability, restricting myself to the special case of first-order arithmetic.

The intuitionistic identification of truth with provability involves a modal notion, which can be understood in two essentially different ways. Accordingly, in Martino and Usberti (1994) a distinction was made between *orthodox* and *non-orthodox* Intuitionism. To begin with, I shall recall this distinction, which will be useful for the discussion at issue.

[1]For the modality involved in the word "knowable" see the distinction (discussed later) between orthodox and potential Intuitionism.

13.2 Orthodox Versus Non-orthodox Intuitionism

By *orthodox* Intuitionism, I mean the traditional Brouwerian Intuitionism. According to this, the equation *truth* = *provability* is intended in the sense that a mathematical proposition *becomes* true (or false) as soon as it (or its negation) has been proved. Until a proposition (or its negation) is proved, it has no truth value. According to this conception, the notion of possibility implicit in the word "provability" has a distinctively *epistemic* character. A proposition is provable insofar as we know how to prove it; there is no room for a *realist* notion of provability according to which a proposition would be provable or not independently of our knowledge.

The dependence of truth on the temporal element may seem somewhat strange to a working mathematician, who is used to regarding mathematical truth as objective and eternal. Some of today's constructivists, aiming to make Intuitionism more fitting to the common sense of the working mathematician, have tried to understand provability in the realist sense of the objective, knowledge-independent existence of a proof. According to this view, call it *non-orthodox Intuitionism*,[2] a proposition may be provable even if, as a matter of fact, nobody will ever prove it. An upholder of this view is Dag Prawitz (1987, 1998). He agrees with the intuitionistic tenet that identifies mathematical truth with provability, but contends that provability is to be understood as atemporal existence of a proof:

> …in pure mathematics there are phenomena that seem difficult to account for without such an objective notion of truth which does not refer to properties belonging to the speaker. We do not only assert sentences in mathematics, we also make conjectures and ask questions to ourselves. If we wonder whether there are infinitely many twin primes, we do not wonder whether this has been proved, but a verificationist must be allowed to wonder not only that, but also whether it can be proved. Similarly, he may conjecture that there are infinitely many twin primes, and normally he is then not making the conjecture that it will be proved that there are infinitely many twin primes, which is a conjecture about future history. From a verificationist point of view the natural way to take the conjecture is to understand it as saying that it is provable that there are infinitely many twin primes. This may also be expressed by saying that there exists a proof of the proposition that there are infinitely many twin primes, where 'exists' is to be taken in a tenseless sense, not as implying that a proof has already been constructed by us.
>
> I am thus arguing that even within a theory of meaning in terms of proofs (or verifications) we must make room for the possibility of entertaining ideas of provability or of abstract existence of proofs; it would be too narrow to construe our speech to be only about what is proved or about actual existence of proofs. Once we accept the notion of provability as legitimate, it is hardly controversial within verificationism that the truth of a proposition is to be identified with provability or existence of proofs (Prawitz 1998, pp. 47–48).

As observed by Dummett, Prawitz's view leads to a shift into a realist position, which justifies bivalence. For, if the truth of a proposition A consists in the existence of a proof of A and such a proof does or does not exist quite independently of our knowledge, it is natural to identify the falsity of A with the non-existence of a

[2] In Martino and Usberti (1994) this view was called "potential Intuitionism". Here, in order to avoid a possible confusion with the conception of potential infinite, which is espoused by both sorts of intuitionists, I prefer the locution "non-orthodox Intuitionism".

proof of it, so that A turns out to be determinately true or false independently of our knowledge. Prawitz seems to be aware of this objection and, in the sequel of the above-mentioned passage, replies:

> Although the idea of proofs existing independently of our hitting upon them certainly contains a flavour of realism, I do not think that it amounts to a full step to realism. I want to give two reasons for thinking so. Firstly, proofs as here understood are something that in principle can be known by us, and hence there is no talk about in principle unknowable proofs. Secondly, I do not see why the disjunction 'either there exists a proof of A or there does not exist a proof of A' must be taken in a classical way. Although we think of the proofs as having some kind of existence even before we find them, an intuitionist may still maintain that to assert the disjunction that either there is or there is not a proof of A requires that we find a verification either of the existence or of the non-existence of a proof of A. For an arbitrary A we do not know how to find such a verification, and we should then have no difficulty in resisting the thought that the disjunction in question is true (Prawitz 1998, p. 48).

Concerning the first reason, we observe that even classical proofs are knowable in principle, so that this feature of proofs does not characterise constructivism at all.

As to the second, the reason why the disjunction that there is or there is not a proof of A is to be understood classically lies precisely in the *realist* existence of proofs. If the non-orthodox intuitionist insisted that we are allowed to assert the disjunction in question only when we know how to verify the existence or how to verify the non-existence of a proof of A, he would prevent the possibility of asserting in his language his own thesis that a proof of A exists or not independently of any verification. He would clearly be incoherent if he denied the intelligibility of the classical reading of "there is or there is not a proof of A".

The matter is quite different for the orthodox intuitionist. Having rejected any knowledge-independent existence of proofs, he is forced to understand the existence of a proof of A as the knowledge of such a proof. Accordingly, the non-existence should be understood, at first sight, as a mental state in which no proof of A is known. But since such a state may be transitory, it fails to be a suitable candidate to establish the falsity of A, i.e. the truth of $\neg A$. That has led the orthodox intuitionist to interpreting the non-existence of a proof of A as the knowledge of the impossibility of proving A and to regarding such a knowledge as a proof of $\neg A$. It is important to realise that such impossibility cannot be understood, according to the orthodox intuitionist, as an objective (possibly unknown) state of affairs, in virtue of which a proof of A will never be found. Antirealism concerns not only the existence of mathematical objects but also modality. The impossibility of proving A is to be understood as the knowledge of how to reject any alleged proof of A. This justifies the rejection of bivalence, since, according to this conception, given any proposition A, we are not in a position in general, not even in principle, to prove A or to prove $\neg A$.

On the other hand, this motivation for a revision of classical logic is not available to the non-orthodox intuitionist. Having identified the truth of A with the objective (knowledge-independent) existence of a proof of A, he can well identify the falsity of A (i.e. the truth of $\neg A$) with the objective non-existence of a proof of A. No reason of non-intelligibility can force him to reject this interpretation, which is the most

natural, the one in accordance with the usual understanding of negation in natural language.

In particular, given any *decidable* arithmetical predicate $P(x)$, the non-orthodox intuitionist can reconstruct the classical meaning of $(x)P(x)$ by interpreting the truth of this as the *objective* existence, for every number n, of a proof of $P(n)$, the falsity (i.e. the truth of its negation) as the *objective* existence, for some n, of a proof of $\neg P(n)$. Since, for decidable sentences, classical truth and intuitionistic provability are equivalent, the interpretation in question of $(x)P(x)$ is equivalent to the classical one. More generally, the non-orthodox intuitionist can understand the classical Tarskian inductive definition of truth for the whole first-order arithmetic in terms of the objective existence of proofs of atomic sentences.

Prawitz could certainly reply that, though the realist interpretation is *intelligible* to the non-orthodox intuitionist, he is by no means compelled to adopt it. Indeed, the non-orthodox intuitionist, as a constructivist, might argue that the realist interpretation, even if intelligible, is unsatisfactory. A constructive interpretation must grant the knowability, in principle, of every mathematical truth; and this fundamental requirement is not fulfilled by the realist interpretation. That motivates the adoption of intuitionistic logic.

I think, however, that the requirement of knowability is inadequate to motivate a revision of classical logic. For, once the intelligibility of classical logic has been recognised, that requirement can be appropriately fulfilled by introducing an epistemic operator K, to be read as "it is knowable that…", where the modality is understood in the realist sense. We thus obtain *epistemic arithmetic*, in which, as is well known, intuitionistic arithmetic is interpretable (see Shapiro 1985). This interpretation meets the requirement of knowability, leaving the classical meaning of logical constants unaltered. For instance, the intuitionistic rejection of the principle of excluded middle is interpreted as the fact that sentences of the form $KA \vee K\neg A$ (where disjunction and negation retain their classical meaning) are not logically valid.

I conclude that the epistemic flavour of the non-orthodox intuitionist is not able to supply philosophical reasons supporting the intuitionistic revision of classical logic. For a more detailed discussion, we refer the reader to Martino and Usberti (1994).

Here, I want to argue that a more subtle form of realism lies behind the conception of mathematics of the Brouwerian Intuitionism, so that even the orthodox intuitionist has no right, at least as far as first-order arithmetic is concerned, to reject classical logic. Indeed, I think that this conclusion holds far beyond first-order arithmetic. But the intuitionistic reconstruction of higher mathematics, as the continuum and set theory, uses *undetermined universes* and even *undetermined objects* as *choice sequences*. So my claim for higher mathematics would need a discussion, which I do not wish to pursue here, of how such highly problematic entities affect the notion of truth.

13.3 The Constructive Notion of a Process

According to the constructive conception, the set N of natural numbers is generated step by step by iterating the successor operation. I think that the most perspicuous model of N, the best representative of the conception of *potential* infinite, is Hilbert's model of finite strings of strokes. I shall refer to that as the intended model of constructive arithmetic. In his famous paper "On the infinite" Hilbert (1926) observes that although the infinite does not exist in reality "it may still be the case that the infinite occupies a justified place *in our thinking* that it plays the role of an indispensable concept". Then, he describes the *pre-mathematical* concept of a finite string of signs as the most elementary notion of potential infinity, which the mathematician can rightfully exploit in a *finitary* reasoning:

> As a further precondition for using logical deduction and carrying out logical operations, something must be given in conception, viz., certain extralogical concrete objects which are intuited as directly experienced prior to all thinking. For logical deduction to be certain, we must be able to see every aspect of these objects, and their properties, differences, sequences, and contiguities must be given, together with the objects themselves, as something which cannot be reduced to something else and which requires no reduction. This is the basic philosophy which I find necessary, not just for mathematics, but for all scientific thinking, understanding and communicating. The subject matter of mathematics is, in accordance with this theory, the concrete symbols themselves whose structure is immediately clear and recognisable.

Hilbert's numerals can be inductively defined by clauses

N1 "|" is a numeral.
N2 If n is a numeral, then $n*$ "|" is a numeral (where $*$ is the concatenation symbol).

As is well known, inductive definitions are *pseudo-definitions*. In fact the above definition is circular, since the second clause presupposes the notion of numeral. But, according to Finitism, the above clauses are to be understood as a synthetic *description* of the intuitive process by which numerals are constructed step by step, starting with a single stroke and reiterating indefinitely the action of producing a new string by adjoining a new stroke to the last string already obtained. Grasping such a process is the very foundation of the conception of potential infinite.

Consider the language of first-order arithmetic. The quantifier-free sentences are *finitary* in Hilbert's sense: they are decidable and are regarded as unproblematic, since they can be interpreted as stating the result of a finite calculation. Hilbert recognises, however, a finitary meaning also to quantifier-free formulas with free variables: these are interpreted as hypothetical judgements asserting that every finitary sentence obtained by replacing the free variables by numerals is true. In this sense the universal closure of a quantifier-free formula has a finitary meaning, I shall call a sentence of this form a *universal finitary sentence* (for short, *u.f.-sentence*).[3] I shall refer to the fragment of first-order arithmetical language consisting of the quantifier-free formulas and their universal quantifications as the *finitary fragment* of number theory.

[3]Such sentences are usually called Π_1-sentences.

Constructivism denies an *intrinsic* truth value to u.f.-sentences; the intuitionistic meaning of such a sentence is determined by grasping the problem of searching for a *proof* that each of its instances is true. Similarly for the finitistic meaning, with the restriction to *finitary* proofs. What is the boundary of finitistic methods of proof is a highly controversial matter, but there is general agreement that finitary proofs include at least the ones formalisable in the system of *primitive recursive arithmetic*. As to the other quantified formulas of first-order arithmetic, they are, according to Hilbert, *ideal* elements, lacking any finitary meaning. The use of these in a formal system is to be justified by showing *finitistically* its consistency with the finitary fragment of the theory. Intuitionistically, they are to be reinterpreted in terms of proof-conditions, according to Heyting's explanation. Their classical meaning, in terms of truth conditions, is to be rejected, so the intuitionists claim, since it would rest on a realist conception of numbers, thought of as an actual infinity of entities, living in a platonic realm, existing quite independently of the human knowledge.

Classical mathematics has mostly been defended by stressing its utility, simplicity and beauty, as well as by questioning the cogency of the intuitionistic arguments against classical logic. Intuitionism, on the other hand, has often been appreciated as a legitimate and interesting philosophical position, though perhaps inadequate for a satisfactory reconstruction of mathematics. Sometimes, a moderate acceptance of classical logic has been proposed as a plausible extension of the intuitionistic conception, as a liberalisation of the too severe constraints imposed by traditional Intuitionism. In his "Constructivism liberalised", Velleman (1993), after discussing Dummett's thesis that, when dealing with quantification over an infinite domain, the appropriate logic is the intuitionistic one (see, for instance, Dummett 1994), proposes a more "liberal constructivism", which allows classical logic for first-order number theory:

> Constructive mathematics is sometimes described as being about the mental constructions that could be carried out by an "ideal mathematician" who can perform any finite computation but can never complete a computation involving an infinite number of steps. If we regard the impossibility of completing an infinite sequence of computational steps as merely "medical" then we are led to a more liberal conception of the "ideal mathematician", according to which infinite sequences of computations can be completed. Liberal constructivism could be thought of as being about the mental computations of this more capable ideal mathematician. Such an ideal mathematician could determine the truth value of any statement of first order number theory by simply checking all of the (infinitely many) cases involved. But note that even this ideal mathematician cannot check every case of a general statement about the real numbers, not because of a lack of time or computational speed, but simply because the question of what real numbers there are has not been adequately settled, so it is not clear precisely what computation must be done to check that some property holds for "all real numbers". Thus we are led to accept classical logic, and in particular the law of excluded middle, for quantificational over the natural numbers, but to reject it for quantification over the real numbers. Liberal constructivism adds only a small amount of realism to ordinary constructivism. Even many who today call themselves formalists seem to accept this amount of realism (Velleman 1993, p. 79).

I want to hold the stronger claim that, in order to accept classical logic for first-order number theory, there is no need to add any amount of realism to ordinary constructivism. In fact the amount of realism needed for the purpose lies already

hidden behind the intuitionistic idealisation of number theory. More precisely, I claim that

(i) Heyting's semantics, even restricted to the finitary fragment of number theory, involves a commitment to a certain form of realism, in virtue of which the whole classical first-order arithmetic is fully justified;
(ii) the intuitionistic proof-conditions of a sentence are much more problematic than its truth conditions and are intelligible only through the latter.

13.4 Computational Realism

Let $(x)P(x)$ be an u.f.-sentence. According to the constructivist idealisation, natural numbers are constructed step by step by an *ideal agent* in the course of an infinite process. For every number n, the agent is able to test $P(n)$ by means of the appropriate algorithm; and we can imagine that he carries out effectively the infinite process of testing one $P(n)$ at a time. Then, it is perfectly determined whether in the course of the process he will always find, for every $P(n)$, the value "*true*" or, for some, the value "*false*"; so the truth value of the sentence $(x)P(x)$ is perfectly determined.

This argument is trivial, but I find it very cogent. Neither the *actual* infinity of numbers nor the capacity of the agent to complete infinitely many computations is involved by this argument. What I claim to be well determined is a *fact* concerning the spatio-temporal activity of the agent.

An intuitionist might certainly object that I'm reasoning about the process in question as if it were a real process, while it is a purely *ideal* process, which does not take place outside of our mind, so that there is no reality which could determine the truth or falsity of a sentence concerning it. Such a sentence, the objection goes, can have a truth value only insofar as we succeed in *recognising* its truth or falsity.

Indeed, what leads an intuitionist to reject classical logic is not the mere thesis that mathematical entities are nothing but mental constructions, but rather his belief that our reasoning about mental constructions *has to take into account* that no reality can determine the truth or falsity of propositions about such constructions. My main claim is precisely that such a belief is wrong and in disagreement with the very intuitionistic perspective.

I must certainly agree that I refer to the ideal agent and his activity as if they were real; but that is in order. The mere *idea* of an agent free of empirical spatio-temporal limitations does not help me to decide, e.g., whether a very large number is prime or not. That could be decided by the ideal agent *if he really existed*. So the decidability of $P(n)$ for all n, accepted by a constructivist, refers to a counterfactual world where the agent exists and carries out his constructions. The fact that the ideal agent is a creation of our mind by no means prevents us from reasoning about him *as if* he were real. Moreover, we are *compelled* to do so, if we want to exploit our idealisation which, otherwise, would be of no use. The intuition of the *decidability in principle* of a recursive predicate P is based on our capacity of *imagining* the

process of testing the infinitely many $P(n)$'s as a real one, taking place in space and time. Therefore, if the device of the idealised mathematician is taken seriously, it commits the intuitionist to a form of *semantic realism*, since, in order to exploit what the idealised agent can do, he *must* refer to him as to an existing being. Thus, as far as the activity of the ideal agent is concerned, to deny the intelligibility of a realist semantics is inconsistent with the very same intuitionistic treatment of the theory of the ideal agent.

At this point, the intuitionist, even if he conceded to be committed to realism about the ideal mathematician, might object that Intuitionism is interested in what such an ideal mathematician can *know*. Now, his argument goes, there is an essential difference between testing $P(n)$, for any given n, and testing $(x)P(x)$. While the former is tested by means of a finite computation, in order to test the latter, by applying the computing procedure, the agent should be able to perform in a *finite* time all the infinitely many computations for all $P(n)$'s; and such ability goes far beyond the constructivist idealisation. Lacking that, the agent can decide the sentence only by searching for an abstract proof of it or for some false $P(n)$, but, in general, there is no guarantee that such research will be successful, whence the rejection, from the constructivist standpoint, of the general validity of the principle of excluded middle.

I reply that, in order to search for a proof of $(x)P(x)$, the ideal mathematician must know the *meaning* of $(x)P(x)$. And my point is that neither the ideal possibility of performing infinitely many computations in a finite time nor the abstract notion of a mathematical proof is involved in the *meaning* of $(x)P(x)$, such a meaning arising straightforwardly from the computational meaning of the instances $P(n)$'s. What $(x)P(x)$ says is that, in the course of the *endless* process of testing, one at a time, the infinitely many $P(n)$'s, the agent will always find the value "true". The realism of this interpretation, as a mere description of the outcomes produced along the process, is in agreement with the realist nature of the intuitionistic conception of an infinite process. Once the value "true" of each $P(n)$ has been defined as one of the two possible outcomes of the nth computation, performed by the ideal agent at stage n, it is straightforwardly determined whether, as a matter of fact, each computation yields the outcome "true" or some yields the outcome "false". The problem of knowing what is the case is not involved in the understanding of the *meaning* of $(x)P(x)$, but *presupposes* such a meaning. The intuitionistic objection that the alleged fact that every computation yields the outcome "true" can never obtain, because at no finite stage can all the infinitely many computations be performed, is very poor. Such a fact is to be understood as an infinite sequence of component facts: its obtaining is nothing but the obtaining, at each stage n, of the nth fact that the nth computation yields the outcome "true".

At this point, the intuitionist may invoke a Dummettian argument and hold that the alleged truth conditions, determined by the realist interpretation, are transcending the use of language. A mathematician, the argument goes, can *manifest* his understanding of the meaning of a proposition only through his capacity of deciding whether any proposed argument is a proof or not of that proposition. And what this capacity shows is merely the knowledge of the *proof-conditions* of the proposition. By no means can he manifest a knowledge of *truth conditions*. So, according to the

Wittgensteinian thesis that meaning is determined by use, the realist truth conditions would be extraneous to meaning.

My reply is that, first of all, there is no evidence that, in general, the knowledge of the proof-conditions of a proposition is manifestable. Until a proposition A has been decided, there seems to be no other way of manifesting the ability of recognising a proof of it than that of rejecting any alleged proof. So, until A is decided, there is no way of manifesting any ability to distinguish its meaning from that of any other undecided proposition. And since a proposition may remain forever undecided, there is no guarantee that such ability can ever be manifested.

Secondly, what is more important, I think that, manifestability aside, the ability of grasping the proof-conditions of a proposition without involving the notion of truth is illusory. When dealing with *formalised* proofs, such ability is granted by the mere knowledge of the axioms and the inference rules of the formal system. But intuitionistic semantics is grounded on an *informal* notion of proof. And it is hard to imagine what could provide the ability of recognising an *informal* proof of a proposition, if not the knowledge of its truth conditions. At first sight, the difficulty may seem to be superseded by Heyting's inductive explanation of the proof-conditions of a sentence in function of the proof-conditions of its components. Heyting's explanation certainly supplies a precious criterion for testing whether something is an intuitionistically acceptable proof. I shall argue, however, that, though Heyting carefully avoids any appeal to truth, one can hardly maintain that the intelligibility of Heyting's clauses does not involve an implicit grasp of realist truth conditions.

Let $P(x)$ be a decidable predicate again. According to Heyting's explanation, a proof of $(x)P(x)$ is a method of producing, for every number n, a proof of $P(n)$. At first glance, the clause seems to explain what a proof of the universal sentence is, provided it is known what is a proof of any of its instances, without any appeal to truth. Several of today's intuitionists have tried to improve Heyting's semantics by introducing a distinction between *canonical* and *non-canonical proofs*. They have observed that Heyting's original explanation of the proof-conditions of a compound sentence, in terms of the proof-conditions of its components, is circular. Since, e.g., $P(n)$ may be derived from $(x)P(x)$, a proof of the former may involve a proof of the latter. They have tried to remove the circularity by calling such a proof "non-canonical" and holding that to know the intuitionistic meaning of a proposition amounts to knowing what a "canonical" proof of it is. A canonical proof of a sentence should be a *direct* proof, which does not involve any concept of proof of a more complex sentence. A *general* proof of a sentence A is then defined as any *method* of finding a canonical proof of A. So Heyting's revised clauses should explain what a *canonical* proof of a sentence is in terms of the notions of the *canonical* proofs of its components. Now, a canonical proof of $P(n)$ consists in applying to $P(n)$ the decision procedure and obtaining the value "true". A *canonical* proof of $(x)P(x)$ is understood as any method of finding, for any given n, a canonical proof of $P(n)$. So, what such a method has to yield is nothing but the *knowledge* that the deciding procedure, applied to any $P(n)$, always gives the value "true". And since $(x)P(x)$ is *classically* true precisely if the decision procedure always gives the value "true", a canonical proof of $(x)P(x)$ cannot be anything else than an argument to show that the

classical truth conditions obtain. So, one cannot understand what such an argument has to show, if he does not understand the truth conditions of the universal proposition. If the understanding of the locution "the deciding procedure gives always the outcome 'true"' amounted, in turn, to understanding what a proof of $(x)P(x)$ is, Heyting's explanation would fail to explain anything. The deciding procedure for the predicate must be determined and understood before putting the problem of searching for a proof of the universal sentence. To give the output "true" for every input is a feature of the algorithm that holds, when it does, quite independently of the existence of such a proof. The task of the latter cannot be anything else than to bring to knowledge an *objective* state of affairs. But that is the task of any *classical* proof of $(x)P(x)$ as well. This does not mean, of course, that any classical proof of the proposition in question must be an intuitionistic proof. But the difference between an intuitionistic and a classical proof of our universal sentence lies in the difference between the intuitionistic and the classical notion of a method. And Heyting's explanation fails to give any account of the notion of a method: this is merely presupposed as a primitive notion, in terms of which even the notion of *canonical* proof is explained. Thus, the distinction between canonical and non-canonical proofs is of no help for avoiding any circularity. I conclude that all that is said by Heyting's clause about a proof (canonical or not) of $(x)P(x)$ is that, by means of it, we have to come to know that the Tarskian truth conditions for the universal sentence are satisfied.

Moreover, the intuitionistic claim that the proposition "the computing procedure, applied to whatever $P(n)$, will give always the value 'true"' can be true only in virtue of a proof of $(x)P(x)$ is in disagreement with the intuitionistic knowledge that the predicate $P(x)$ is decidable, expressed by asserting $(x)(P(x) \vee \neg P(x))$. By means of an intuitionistic proof of this, one recognises that the procedure, applied to any n, yields, in a finite time, the value "true" or the value "false". The intuitionist is therefore compelled to agree that the truth values of the various $P(n)$'s, *whatever* they are, depend solely on the computational procedure, so that what they are, in particular if they are all "true", by no means can depend upon a proof of $(x)P(x)$. Thus, the understanding of the truth conditions of $(x)P(x)$ is implicit in the intuitionistic knowledge that the predicate P(x) is decidable. The understanding of the proof-conditions of $(x)P(x)$, on the other hand, is a much more complicated matter, just because the agent cannot verify all the infinitely many instances, one at a time, by applying the computing procedure. A proof of the universal sentence consists of an *abstract* reflection on the decisional procedure, which may involve a great unsurveyable variety of considerations. In order to understand the proof-conditions, one needs to grasp some essential feature shared by all possible proofs of $(x)P(x)$. But such proofs do not share anything except the common purpose of showing that the testing procedure will verify all instances, i.e. that the truth conditions obtain.

The same considerations hold for the finitary doctrine relative to Hilbert's *real* part of mathematics, in particular to u.f.-sentences. In fact, according to Hilbert, such sentences, unlike the *ideal* ones, are to be interpreted *contentually* and, if true, are to be proved by means of a finitary reasoning. The main feature of such a reasoning is that it is grounded on the basic intuition of the generative process of finitary numbers, so that its conclusions should be certain beyond any doubt. Now, it is clear that the

very notion of reliable reasoning presupposes the pre-existing truth values of the conclusions: a reasoning is reliable insofar as the truth of its conclusion is granted. The u.f.-sentences of special interest for Hilbert's programme are the (arithmetical translations of the) metamathematical propositions expressing the consistency of a formal theory. Hilbert's proposal to search for finitary proofs of such propositions is motivated by the aim of *guaranteeing* the consistency, understood as an *objective* syntactical feature of a formal system. The programme rests on the view that, however uncertain and problematic the contentual intended interpretation of a mathematical theory may be, as is the case, in particular, whenever the *actual* infinite is involved, nevertheless, once a theory has been formalised, the metamathematical assertion of its syntactical consistency has a precise objective meaning, in virtue of which its truth or falsity is well determined. The constructive aspect of a consistency statement has nothing to do with the perverse idea according to which *to be* consistent would amount to the existence of a consistency *proof*. The majority of mathematicians, even if they believe, after the failure of Hilbert's programme, that the consistency of set theory is unprovable, nevertheless reasonably hope that set theory is consistent. The finitary flavour of a consistency statement is to be seen in its content which describes an elementary feature of the concrete *sign-figures* representative of formalisable proofs, i.e. the fact that the sentence "$0 = 1$" never occurs as an end-formula in any of such figures. These figures are constructed by the ideal agent, step by step, following a mechanical procedure, whether in the course of this process "$0 = 1$" occurs or not as the end-formula of some proof-figure is an objective fact, quite independent of any consistency proof.

Once the above realist interpretation has been accepted for the u.f.-sentences, one can easily extend it, by means of a straightforward inductive procedure, to all sentences of first-order arithmetic, in such a way that every arithmetical sentence describes an objective feature of some computational process. Take, e.g., a sentence of the form $(x)\exists y P(x, y)$, where $P(x, y)$ is a decidable predicate, and consider the mechanical process of listing all ordered pairs which satisfy the predicate. The formula describes an objective fact about this list, precisely the fact that every number occurs in the list as the first component of some ordered pair.

Thus, the finitistic conception of the *real* part of arithmetic and, a fortiori, the intuitionistic conception of numbers, entail the adoption of a sort of realism, call it *computational realism*, which provides an interpretation of first-order arithmetic in terms of the computational activity of the ideal agent. According to this interpretation, the arithmetical truths are exactly the classical ones. This conclusion does not entail, it should be noted, that any classical proof of an arithmetical sentence be justifiable according to computational realism, the possibility of such a justification depending upon the notions involved. A classical proof of a first-order sentence might use some higher-order notions not available in the perspective at issue, and it might involve, e.g., an essential use of the notion of *actual infinity*. But, as far as classical derivability from first-order Peano's axioms is concerned, computational realism can validate any classical proof, formalisable in first-order arithmetic. For, within the framework of computational realism, one can straightforwardly recognise that Peano's axioms are true and that the classical inference rules are truth-preserving; so,

within that framework, one can prove the soundness metatheorem for classical first-order arithmetic. On the other hand, the informal notion of arithmetical proof, correct according to computational realism, transcends Peano's first-order axiomatisation, nor can it be formalised in any extension of that. For, as soon as a computational realist recognises the truth of a decidable set of arithmetical axioms, he is able to recognise the truth of Gödel's undecidable sentence as well. This conclusion is in agreement with the fact that the computational framework allows any reasoning based on the intuition of a computational process. In fact, any axiomatisation of arithmetic allows an interpretation in some *non-standard* model, where the *finite* computability of arithmetical operations is lost. It follows that no axiomatisation can fully catch the intuition of a finite computation. This is the fundamental reason why constructive reasoning, as understood in the present paper, is incapable of any axiomatic characterisation.

Summing up, the so-called notion of *constructive mathematical truth*, grounded on some alleged truth-independent notion of proof, as received from traditional constructivism, is the outcome of the perverse attempt to reverse the natural priority of truth over provability. This attempt, contrary to its aim of focusing the central role of proof in mathematics, destroys the very nature of mathematical proof, understood as a compelling argument to bring to knowledge a certain objective state of affairs. It is not motivated by the rejection of the actual infinite in favour of the potential infinite. Such a rejection by no means leads from realism to antirealism, but rather from realism about mathematical objects, thought of as simultaneously existing in a static realm, to realism about infinite computational processes, thought of as carried on by an ideal agent in an ideal spatio-temporal world. The fact that such a world is merely imaginary does not prevent the working mathematician from reasoning about it *as if* it were real. On the contrary, the use of classical logic is compelled by this very idealisation, according to which computational processes are conceived of as well determined by their generating rules. To think otherwise is confusing the act of imagining with the *content* of the act: an imaginary computational process is imagined as well determined, while it is not imagined as imaginary. That is the reason why the fact that a process is imaginary cannot affect our reasoning about it.

References

Dummett, M. (1994). What is mathematics about? In A. George (Ed.), *Mathematics and mind* (pp. 12–26). Oxford: Oxford University Press.

Gödel, K. (1995). Some basic theorems on the foundations of mathematics and their implications. In S. Feferman (Ed.), *Kurt Gödel: Collected works* (Vol. III, pp. 304–323). Oxford: Oxford University Press.

Hilbert, D. (1926). On the infinite. In P. Benacerraf & H. Putnam (Eds.), *Philosophy of mathematics* (pp. 183–201). Cambridge: Cambridge University Press.

Hintikka, J. (1997). *The principles of mathematics revisited*. Cambridge: Cambridge University Press.

Martino, E., & Usberti, G. (1994). Temporal and atemporal truth in intuitionistic mathematics. *Topoi*, *13*, 83–92 (reprinted here as chapter 11).

Prawitz, D. (1987). Dummett on a theory of meaning and its impact on logic. In B. M. Taylor (Ed.), *Michael Dummett: Contributions to philosophy* (pp. 117–165). Dordrecht: Springer.

Prawitz, D. (1998). Truth and objectivity from a verificationist point of view. In H. G. Dales & G. Oliveri (Eds.), *Truth in mathematics* (pp. 41–50). Oxford: Clarendon Press.

Shapiro, S. (1985). Epistemic and intuitionist arithmetic. In S. Shapiro (Ed.), *Intensional mathematics* (pp. 11–46). Amsterdam: North Holland.

Tennant, N. (1998). Review of Hintikka J.: The principles of mathematics revisited. *Philosophia Mathematica*, *6*, 90–115.

Velleman, D. J. (1993). Constructivism liberalized. *Philosophical Review*, *102*, 50–84.

Chapter 14
The Impredicativity of the Intuitionistic Meaning of Logical Constants

with G. Usberti

Abstract Dummett's thesis that Heyting's explanation of the meaning of logical constants is circular is discussed in this chapter. We defend Dummett's position.

In his book *Elements of Intuitionism*, Dummett discusses at length the intuitionistic notion of proof and argues that Heyting's explanation of the meaning of the logical constants is circular:

> The principal reason for suspecting these explanations of incoherence is their apparently highly impredicative character [...] [T]he explanations require us, in determining whether or not a construction is a proof of a conditional or of a negation, to consider its effect when applied to an arbitrary proof of the antecedent or of the negated statement, so that we must, in some sense, be able to survey or grasp some totality of constructions which will include all possible proofs of a given statement. The question is whether such a set of explanations can be acquitted with the charge of vicious circularity (Dummett 1977, p. 390).

Dummett's answer is that Heyting's explanation, so as it stands, is certainly circular, because of the impossibility to grasp the general concept of proof of a given proposition. When Heyting says that a proof of $A \supset B$, for instance, is a method of transforming every proof of A into a proof of B, he cannot be referring—Dummett argues—to ordinary informal proofs. The reason is the following: an ordinary informal proof can use elimination rules, in particular *modus ponens*, so whatever we were previously ready to accept as being an informal proof of $A \supset B$, it would supply us with a method of transforming any proof of A into a proof of B: it would be sufficient to take an arbitrary proof of A, to annex to it the given proof of $A \supset B$ and to append a single application of *modus ponens*. In other words—Dummett concludes—Heyting's explanation of the meaning of $\supset$ does not restrict at all the class of entities we were previously disposed to consider as ordinary informal proofs; consequently, it has no explicative power, unless it is referred to a class of proofs smaller than the class of ordinary informal proofs. Then, he urges a suitable distinction between general proofs, also called demonstrations, and restricted proofs, also called canonical proofs. These should satisfy the requirement that a canonical proof of a proposition A does not involve any proof of a proposition of which A is a component, while a demonstration of A should be an effective method of finding a canonical proof of A. The meaning of the logical constants should be given only in terms of canonical

proofs. In particular, the meaning of implication would be explained by saying that a canonical proof of $A \supset B$ is a method of transforming any canonical proof of A into a canonical proof of B, with a predicative quantification on all canonical proofs of A.

At first sight, Dummett's project seems to be realised by Prawitz's definition of valid argument and Martin-Löf's type theory (which we will call ITT), where the distinction between canonical and non-canonical proofs is treated in a systematic way and the meaning of a proposition is given just by explaining what its canonical proofs are.

However, Dummett himself, just after stressing the importance of the notion of canonical proof as a means for avoiding the impredicativity in question, calls our attention to a hidden difficulty, which seems to make the notion of canonical proof useless. The explanations of implication, negation and universal quantification, even reformulated in terms of canonical proofs, involve the primitive general notion of method which, in turn, involves the general notion of proof. For instance, a method for transforming any proof of A into a proof of B, intuitionistically understood, is given only when it is recognised that it works as desired, i.e. when an intuitionistic proof is given that, by applying it to an arbitrary proof of A, one gets a proof of B. No restriction on such a proof is imposed by the condition that the method be applied only to canonical proofs of A in order to get canonical proofs of B. And since the antecedent of an implication may in turn be an implication, it follows that the restriction to canonical proofs of A (of the quantification occurring in the explanation of what a canonical proof of $A \supset B$ is) cannot serve the purpose of avoiding impredicativity.

We want to discuss the questions: (i) Is Dummett's criticism of Heyting's explanation cogent? (ii) Are the difficulties of Heyting's explanation superseded by Prawitz's or Martin-Löf's notions of canonical proof?

Certainly, Heyting's explanations are not able to explain to a non-intuitionist what an intuitionist proof is. Their intended meaning can be understood only if they are read intuitionistically, in particular only if the intuitionistic notion of method, with its implicit notion of evidence, is a priori understood. What they try to show is just how the intuitionistic meaning of the logical constants is reducible to the basic concept of effective method. The problem is whether, taking for granted an appropriate understanding of this primitive concept, Heyting's explanations are correct.

A crucial feature of the intuitive notion of proof is that it is an open notion: the totality of methods of proof acceptable as correct is indefinitely extensible. This is not a trouble in itself: it is a sign of the continual process of evolution of mathematics and it is in accordance with the impossibility, shown by the incompleteness theorem, of representing all correct reasonings about a given mathematical theory in a single formal system. What is problematic is whether it is correct to quantify over all methods of proof. According to Dummett, to give the meaning of a proposition by means of such a quantification makes the very meaning of propositions unstable:

> On the intuitionistic view, this evolution creates a special danger. If we look on the appeal to the full intuitionistic meaning of $\supset$, in proving a statement of the form $A \supset B$, as mediated by the invocation of a principle of the form '$A \supset$ there is a proof of A', an advance in our

> apprehension of the available modes of proof may lead us to weaken such principles, because restrictions on the means whereby A could be proved which formerly seemed reasonable no longer appear so. When this happens, some proof, involving a conditional $A \supset B$, that had formerly seemed acceptable, may be invalidated. Hence, because of the peculiarities of the intuitionistic interpretation of $\supset$, provability is not a stable property: we cannot think of an addition to our stock of methods of proof as merely allowing us to prove more than we could before, while all proofs we had already given remain intact, since such an addition may lead to a rejection of certain earlier proofs. The intuitionistic interpretation of $\supset$ does, indeed, give to the notion of proof a self-reflexive or impredicative character and to some degree weakens the conclusive and irreversible nature of mathematical results; mathematics becomes a subject whose results are fallible and liable to revision, like those of other sciences (Dummett 1977, pp. 401–402).

Dummett's argument may not seem to be very cogent. An intuitionist might try to defend the quantification over an undetermined domain of entities by arguing that the constructive interpretation of the universal quantifier is innocent. The impossibility, at any stage of knowledge, of surveying all possible proofs of a proposition A—he might argue—does not prevent the possibility of grasping the general concept of proof of A: knowledge of a concept by no means entails knowledge of all the objects that fall under it, and the quantification over such objects by no means presupposes that they are given all at once: it exploits only some very general features, which must be shared by any object, present or future, in order to fall under the concept in question. So, in order to possess a method of transforming any proof of A into a proof of B, one has to recognise that, whenever he will find a proof of A, he will be able to produce a proof of B; and this knowledge is to be reached from the analysis of the mere concepts of proof of A and of proof of B. If, because of what Dummett calls (with Wittgenstein) the "motley" of possible proofs of A, the general concept of proof of A is not able to guarantee the applicability of the method in question to any proof of A, one has no right to claim to have a proof of $A \supset B$. So, once a proposition has been proved, there is no danger that it will be disproved when new methods of proof are invented.

It is difficult to see why this argument is fallacious, but we think it is deceptive. The impossibility of surveying all possible proofs of A—it is argued—does not prevent the possibility of grasping the general concept of proof of A, i.e. of having a criterion to establish whether something is a proof of A. But what is necessary to have in order to have a criterion to establish whether something is a proof of A? Since a proof of A is by definition something that confers evidence to A, in order to be able to recognise it, we must know how something conferring evidence to A must be; and how it must be depends partly on how A is, partly on what we mean as evidence, that is on the standard of evidence we employ. Well, this standard of evidence may change through time. It is true that it is the fact of having a certain standard that permits us to establish whether something is or not a method of proof; but this does not prevent the possibility that the discovery of new methods of proof affects the standard itself. For example, before Gentzen's consistency proof of arithmetic, elementary number theory was reasonably held to formalise all the methods of proof concerning natural numbers intuitively acceptable from a finitist point of view. After the proof, induction up to ϵ_0, which is not formalisable in elementary number theory,

was added to the list of acceptable methods; that is to say, the standard of evidence has been modified so as to include as finitistically evident propositions proved by the new type of construction. Consequently, some constructions that before Gentzen's proof fell out of the domain of the possible proofs of a proposition, after the proof belong to that domain, that is to say, for some numerical proposition it is not possible to quantify over the totality of its possible proofs. Perhaps the following paradox may shed some light on the phenomenon.

Among the effective methods, let us call *first-level method* a method of computing a function from $\mathbb{N}$ (the set of natural numbers) to $\mathbb{N}$, and *second-level method* a method of mapping every first-level method into a natural number.

Let us associate with every second-level method F a first-level method F^* as follows:

1. $F^*(n) =_{def} F([n])$, where $[n]$ is the constant first-level method whose value is n.

F^* is to be identified with the following procedure: "In order to compute the output of F^* for the input n, apply F to $[n]$ and take the result as $F^*(n)$". So, given any first-level method f, it is decidable whether it is of the form F^*, for some F uniquely determined by f. We can then define the second-level method K by putting

2. $K(f) =_{def} \begin{cases} F(F^*) + 1 \text{ if } f \circ F^* \text{ for some } F \\ 0 \text{ otherwise} \end{cases}$

Here $\circ$ must be understood as *sameness* (intensional equality). K seems to be a well-defined effective method: given any first-level method f, check whether it is of the form F^*, for some F: in the affirmative case compute $F(f) + 1$ and take it as $K(f)$, in the negative case take 0 as $K(f)$. But we get the contradiction $K(K^*) = K(K^*) + 1$.

The paradox seems to show that the notion of second-level method, defined over all possible first-level methods, is contradictory. A natural way to escape the contradiction would be to restrict the domain of second-level methods to first-level methods whose constructions do not involve second-level methods. However, once a second-level method has been introduced, it generates new well-defined first-level methods not belonging to its domain. So we would be led to split the first-level methods into a ramified hierarchy.

We think that these considerations support Dummett's criticism to the explanation of implication: a proof of $A \supset B$ should be a method applicable only to proofs of A not involving proofs of more complex propositions. We conclude that Dummett is right in claiming that Heyting' s clauses are impredicative.

Let us turn to the questions: are Prawitz' definition of canonical argument and Martin-Löf's reformulation of Heyting's clauses still impredicative?

Here is how Prawitz sums up his notion of canonical proof:

> To be in possession of a canonical proof of $A \supset B$ […] we must be in possession of a proof of B from the hypothesis A […], which in turn means that we must be in possession of an argument for B from A[…] together with procedures associated with each line of the argument for which we recognize that when the hypothesis A is replaced by a proof of A

> […], the composition of all the procedures associated with the steps of the resulting argument constitute a method for finding a canonical proof of B […] (Prawitz 1987, pp. 160–161).

Since "what counts as a canonical proof of a compound sentence is […] determined in terms of proofs of the subsentences, which are in turn defined in terms of canonical proofs of the same sentences" (p. 161), the requirement of molecularity is satisfied. But since "we cannot assume that the immediate subproof of a canonical proof is again canonical" (*ibidem*), the complexity of the method in question is not bounded by the fact that the requirement of molecularity is satisfied; as a consequence, a proof of the fact that the method has the required properties will be necessary in general, and no bound is placed on the complexity of that proof, so that the problem of impredicativity is not avoided.

Moreover, the definition of validity of an open argument makes reference to the extensions of the set J of justifying procedures. The reason for this is that otherwise there might be an (open) argument for B from A as assumption, valid relative to a set J of justifying procedures but not to an extension J' of J, namely if J' contained new justifying procedures from which methods could be constructed for finding valid arguments for A not capable to be transformed (by the procedures in J) into methods for finding valid arguments for B. Although Prawitz's manoeuvre solves this difficulty, it renders the definition highly impredicative in character.

Let us pass to Martin-Löf. Martin-Löf's types are simple (i.e. not ramified); so, for instance, the functions from $\mathbb{N}$ to $\mathbb{N}$ constitute a single type, as well as the functions from $\mathbb{N}^{\mathbb{N}}$ to $\mathbb{N}$. However, an essential feature of the functions of ITT is that they are *extensional*, so that apparently they can escape the above paradox, which exploits intensional identity. Similarly, the proposition $A \supset B$ is a type whose elements are the extensional functions from proofs of A to proofs of B: so Heyting's quantification over all possible methods of carrying any proof of A into a proof of B becomes a quantification over all extensional functions from A to B, which, at first sight, does not seem to be paradoxical.

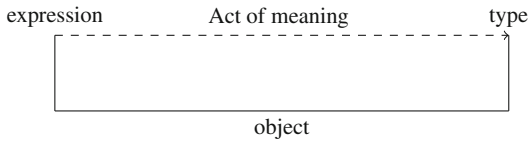
The problem, however, is how to understand extensional functions intuitionistically: from an intuitionistic point of view, a function is nothing but a method of computation. A first possible answer seems to be the following. One might draw inspiration from Frege's distinction between propositional functions and their extensions and try to hold that canonical functions, as mathematical objects, are not to be identified with the very methods by means of which they are calculated, but rather with the extensions of those methods. So coextensive methods would yield the same canonical function and the extensional equality of a type of functions would be authentic sameness. In this way, the quantification on a type would not involve any impredicative quantification over all possible methods of proof. This interpretation of Martin-Löf's type theory seems to be suggested by Beeson:

> $\lambda(f)$ is thought of as the extension of f [where f is the function *in intension*]. Note that $\lambda(f)$ is not the set of values of f, since it is not a set. But it is "the values of f", or "f with everything but the values abstracted away (Beeson 1985, p. 253).

We think, however, that this notion of extension would be appropriate from a platonistic point of view, but that it is unintelligible in the intuitionistic perspective.

Within the conception of potential infinity, there is no room for an entity made of the infinitely many values of a function. Such values cannot build up any object, since they do not exist all together; and their potential existence is nothing but the existence of a method for computing them. The attempt to single out the pure extension of a function by abstracting the method of computation away is hopeless: it would have the opposite effect of destroying their potential existence.

A second possible answer can be extracted from Martin-Löf (1984) and the unpublished Martin-Löf (1987a). A mathematical object, according to Martin-Löf, is nothing but a mathematical expression endowed with meaning,[1] as the following scheme suggests:



Here is how Martin-Löf explains his idea:

> The expression which stands for a mathematical object is the matter out of which it is made, and the type of the object is its form. In this sense, a mathematical object is a composite of matter and form. But one component has been forgotten, namely, the act of understanding through which the matter receives its form. It is through this act that the object comes into being (Martin-Löf 1987a, p. 57).

This view is in turn based on some general principles, some of which are stated below.

3. a. Although syntactical and semantical categories are distinct, a syntactical entity (an expression) can become a semantical entity (an object) through an act of knowledge. Indeed, it is only through this process that mathematical objects come into being.[2]
   b. An entity is not an object unless it is sorted into a semantical category, i.e. unless one knows (i) what its category is, and (ii) at which conditions two objects of that category are the same object.[3]
   c. Between objects and acts, there is an essential difference: while an object is necessarily an object of a certain category, an act does not belong to any category.

---

[1] "A mathematical object is the same as a meaningful mathematical expression" (Martin-Löf 1987a, p. 17).

[2] "The mathematical objects are the meanings of the mathematical expressions" (Martin-Löf 1987a, p. 49). "*A meaning is always the meaning of an expression*. We cannot say that the mathematical objects are meanings tout court, only that they are the meanings of the mathematical expressions" (*Ibid.*, pp. 60–61). "There is no other way of getting to the mathematical objects than by making sense of the mathematical symbols" (*Ibid.*, p. 17.).

[3] In this sense "[T]he type precedes the object" (Martin-Löf 1987a, p. 53).

d. A particular category of mathematical objects is the category of sets.[4] "To know a set, you must know the rules for forming or constructing the elements of the set as well as the rules for forming identical elements of the set. […] A set is defined by its constructors, that is, the primitive operations whose values are elements of the set[5]" (Martin-Löf 1987a, p. 90).

Let us consider for instance the functions from $\mathbb{N}$ to $\mathbb{N}$, whose type is conceived by Martin-Löf as a set. According to (3)(d), we must give the rules for forming the canonical elements of $\mathbb{N}^{\mathbb{N}}$ and the rules for forming identical canonical elements of the set. Here are the rules:

$$\frac{\begin{array}{c}(x:\mathbb{N})\\ b(x):\mathbb{N}\end{array}}{\lambda x b(x):\mathbb{N}^{\mathbb{N}}} \tag{14.1}$$

$$\frac{\begin{array}{c}(x:\mathbb{N})\\ b(x)=c(x):\mathbb{N}\end{array}}{\lambda x b(x)=\lambda x c(x):\mathbb{N}^{\mathbb{N}}} \tag{14.2}$$

In the conclusion of rule (14.1) the expression "$\lambda x b(x)$" is sorted into the type of functions from $\mathbb{N}$ to $\mathbb{N}$ through the act of understanding that confers evidence to the premise. Since the premise is a hypothetical judgement, what confers evidence to it is a hypothetical proof of $b(x):\mathbb{N}$ from the hypothesis $x:\mathbb{N}$[6]; it is through this proof that we become aware of the functional dependence of $b(n)$ on n, for every n.[7] In the conclusion of rule (14.2), two canonical elements of the type $\mathbb{N}^{\mathbb{N}}$ are declared to be identical provided they yield the same values for the same arguments.

What kind of object is $\lambda x b(x)$? Perhaps it is convenient to keep two possible ways to conceive it. According to the former the premise of the rule guarantees that we know a method m for associating to every number n a number $b(n)$, and "$\lambda x b(x)$" is nothing but a name of m; the canonical element of $\mathbb{N}^{\mathbb{N}}$ is therefore m itself. According to the latter, it is the very expression "$\lambda x b(x)$" that becomes a canonical element of the type $\mathbb{N}^{\mathbb{N}}$ when it is accompanied by the hypothetical proof of $b(x):\mathbb{N}$ from $x:\mathbb{N}$. The former seems to fit well with what Martin-Löf writes in Martin-Löf (1984), and

[4] As is clear from the quotation that follows in the text, every set is a category, but not *vice versa*. Within the terminological framework of type theory, "type" is systematically ambiguous between "category" and "set".

[5] Primitive elements are also, and more frequently, called "canonical".

[6] "The notion of hypothetical proof, […] which is a primitive notion, is explained by saying that it is a proof which, when supplemented by proofs of the hypotheses, becomes a proof of the thesis or consequent" (Martin-Löf 1985, p. 252).

[7] The essential characteristic of a proof of a judgement is just that it is an act of knowledge. It is this act that legitimates the introduction of the expression "$\lambda x b(x)$" and that transforms it into a true mathematical object. "[A] proof of a judgement is an act of knowing" (Martin-Löf 1987b, p. 417); "[T]he proof of a judgement is nothing but the act of knowing, or, perhaps better, the act of understanding or grasping" (*Ibidem*).

the latter seems to be the only plausible interpretation of Martin-Löf (1987a). So let us discuss the two interpretations separately.

Since methods are certainly not extensional entities, the same holds for the canonical elements of $\mathbb{N}^{\mathbb{N}}$: they are essentially the intensional functions from $\mathbb{N}$ to $\mathbb{N}$. This is perfectly compatible with the extensionality expressed by the rule (14.2). This is a mere stipulation, by which we decide to call "equality" the relation of coextensiveness; it is perfectly legitimate and motivated by the fact that in mathematics we are not interested in distinguishing coextensive functions. But it has nothing to do with the nature of the objects in question, which are what they are quite independently of any stipulation. In other words, to say that the equality of type $\mathbb{N}^{\mathbb{N}}$ is the extensional one does not mean that two canonical elements are *the same* provided they are coextensive: it means only that two coextensive elements, as mathematical objects, play the same role.[8] Thus the quantification over the type $\mathbb{N}^{\mathbb{N}}$ amounts to the quantification over all first-order methods. Not so, however, for the type $\mathbb{N}^{\mathbb{N}^{\mathbb{N}}}$, on which the condition of extensionality imposes an effective constraint. In fact a canonical element of this type must satisfy the condition of *respecting equality*, i.e. of mapping equal canonical elements of $\mathbb{N}^{\mathbb{N}}$ into the same natural number, so that not every second-level method is a canonical element of $\mathbb{N}^{\mathbb{N}^{\mathbb{N}}}$. Nevertheless, the notion of canonical element of $\mathbb{N}^{\mathbb{N}^{\mathbb{N}}}$ presupposes the general notion of second-level method, for, in order to introduce such a canonical element, one can take *any* second-level method and try to show that it respects extensionality; so *any* second-level method is a legitimate candidate for the test of respecting extensionality. In particular, the above paradoxical method K is a good candidate for the test; moreover, just because it is paradoxical, it passes the test in virtue of the *ex falso quodlibet* principle and the paradox follows. We conclude that, on the first of the two interpretations we have distinguished above, the impredicativity of the intuitionistic meaning of the logical constants is not superseded.

Let us pass to the second interpretation. According to it, meaning does not preexist to the sign, but it is the sign itself, provided it is associated with the act of understanding it. From this point of view, reference to methods is still present, indeed, but they are no longer conceived as *objects*, but as acts: it is the hypothetical proof of the premise of (14.1), as an act, that constitutes the method of computation. Since, according to principle (3)(c), an act does not belong to any category, reference to methods does not affect the category of the objects we know through them; the only objects we are dealing with are linguistic expressions. As a matter of fact, this seems not to solve but to make more evident the problem of impredicativity: if a function is conceived as a name plus an act of understanding, when we have two different canonical names we have two distinct functions, even if they associate the same values with the same arguments. However, an answer[9] to this difficulty can perhaps be found in the way identity is dealt with within ITT. The idea is that identity

[8] This point deserves a closer analysis. We will come back to it after having discussed the second interpretation.

[9] This answer has never been explicitly formulated by Martin-Löf in his writings; our discussion is then merely speculative.

conditions between objects cannot be stated in an absolute way, but are necessarily restricted to the type which the objects belong to; when this condition is respected, however, identity must be interpreted as *sameness*. Well, in the case of type $\mathbb{N}^{\mathbb{N}}$ the identity condition is, as we know, that two functions are the *same function* if they associate the same values with the same arguments; consequently, two expressions that, within the category of names, are different may be the *same object* within the category of the functions from $\mathbb{N}$ to $\mathbb{N}$, into which they are sorted by the acts of understanding them. In this way, we arrive at a purely extensional view of functions, and the paradox seems to have been avoided.

However, we hold that this solution is illusory. It is an undeniable matter of fact that in mathematics functions are individuated through methods of computation rather than through names.[10] So let us consider two methods m_1 and m_2 such that they associate the same values to the same arguments; even accepting the idea that they, as acts, do not belong to any semantical category, there is a clear intuitive sense according to which we can assert that m_1 and m_2 are different, if they are different.[11] Let us suppose that they are in fact different; it is then legitimate to denote them by two different names that, when they will be understood as names of the two different methods, become two different mathematical objects. Of course, as we noticed when discussing the first interpretation, this is perfectly compatible with the extensionality expressed by the rule (14.2), but the identity sign in the conclusion of (14.2) could not plausibly be interpreted as denoting sameness in such a situation: we have generated two different canonical functions with the same extension, and functions will be non-extensional entities again.

References

Beeson, M. J. (1985). *Foundations of constructive mathematics*. Berlin: Springer.

Dummett, M. (1977). *Elements of intuitionism*. Oxford: Clarendon Press.

Martin-Löf, P. (1984). *Intuitionistic type theory*. Naples: Bibliopolis.

Martin-Löf, P. (1985). On the meaning and justification of logical laws. In E. Bernardi & R. Pagli (Eds.), *Atti degli Incontri di Logica Matematica* (Vol. II, pp. 291–340). Siena: Università di Siena.

[10] This fact conflicts with principle (3)(a), and is the first of the reasons why we hold that the solution does not work. The third assertion quoted in footnote 2 seems to the antirealist to be convincing because he interprets it as an objection to platonism; but if we interpret it as an attempt to argue that names come before methods of computation in the process of individuating functions, this appearance vanishes altogether.

[11] This fact conflicts with principle (3)(c), and is the second of the reasons why we hold that the solution does not work. In particular, we do not see any non-ad hoc reason to prevent asking whether two acts of knowledge are identical or not, or at least to deny that there is some notion of identity with reference to which the question becomes legitimate.

Martin-Löf, P. (1987a). Philosophical implications of type theory. Lecture at the Institute of Philosophy, University of Florence

Martin-Löf, P. (1987b). Truth of a proposition, evidence of a judgement, validity of a proof. *Synthese*, *73*, 407–420.

Prawitz, D. (1987). Dummett on a theory of meaning and its impact on logic. In B. M. Taylor (Ed.), *Michael Dummett: Contributions to philosophy* (pp. 117–165). Dordrecht: Springer.

Chapter 15
The Intuitionistic Meaning of Logical Constants and Fallible Models

Abstract In this chapter, the problem of the failure of completeness of first-order predicate logic in an intuitionistic metamathematics is discussed and the philosophical significance of fallible models is analysed.

15.1 Introduction

I'll be concerned with the problem of intuitionistic first-order predicate logic IPC within an intuitionistic metamathematics.

By a *natural* interpretation of IPC, we mean what in the literature is often called an *intuitive* interpretation or, in Dummett's terminology, an *internal* interpretation (Dummett 1977)

Given Heyting's explanation of logical constants, natural models are the natural counterparts of classical models. More explicitly, a natural model is obtained by interpreting individual variables and constants in an inhabited domain D, by assigning, by means of an interpreting function I, to every atomic sentence A (in the language extended by constants for all members of D) a meaningful proposition $I(A)$ (what amounts to supplying proof-conditions for all atomic sentences). The proof-conditions for the compound sentences and for absurdity ($\bot$) (which counts as a primitive logical constant) are then supplied by Heyting's clauses.

Here is a formulation of Heyting's proof-conditions (Heyting 1956):

(1) For A atomic $\neq \bot$, a proof of A is a proof of $I(A)$;
(2) a proof of $\forall x A(x)$ is a method of proving *all* $A(d)$ ($d \in D$) [likewise for $A \wedge B$];
(3) a proof of $\exists x A(x)$ is a method of *some* $A(d)$ ($d \in D$)[likewise for $A \vee B$];
(4) a proof of $A \rightarrow B$ is a method of transforming every proof of A into a proof of B;
(5) Nothing is a proof of $\bot$.

Negation is defined by $\neg A =_{df.} A \rightarrow \bot$.

As it is well known, the completeness of IPC for natural semantics (within an intuitionistic metamathematics) is equivalent to a certain form of Markov's principle. Since the latter is far from being intuitionistically evident, this result provides an

argument against completeness. On the other hand, since Markov's principle seems to be compatible with the general principles of Intuitionism, this argument for incompleteness is not a conclusive one.

What is cogent, however, is the failure of *strong* completeness (i.e. of $\Gamma \vDash A \Rightarrow \Gamma \vdash A$, where Γ is any set of sentences and A any sentence). A very simple proof of that, which does not exploit anything beyond standard Intuitionism, was suggested to me by Troelstra. Here is the proof:

Proof Assume, by reduction, strong completeness. Let M be an arbitrary natural model and A an arbitrary sentence. Consider the set of sentences $\Gamma = \{\bot : M \vDash A \vee \neg A\}$. Since $\neg\neg M \vDash A \vee \neg A$, we have $\neg\neg(\bot \in \Gamma)$ and hence $\Gamma \vDash \bot$. By strong completeness $\Gamma \vdash \bot$ so that $\bot \in \Gamma$ and therefore $M \vDash A \vee \neg A$. It follows that $A \vee \neg A$ is logically valid and hence derivable, which is absurd.

This proof rejects strong completeness in virtue of the mere intended meaning of logical constants. It does not provide, however, any counterexample to strong completeness. I think that no such counterexample is achievable without exploiting deeper features of Intuitionism. In a recent paper, Charles McCarty (1991) gives a counterexample to strong completeness under the assumption of Church's thesis. But I am not very sympathetic for the intuitionistic version of Church's thesis: it imposes to the intuitive notion of proof too strong limitations of a mechanicalistic nature, which seem to me inappropriate. I prefer to use lawless sequences and exploit the indeterminateness of the universe of constructions. I have here a counterexample to strong completeness using a lawless sequence.

Let *HA* be Heyting's arithmetic extended with a new 1-place predicate symbol P and a new individual constant c. With reference to a lawless sequence α, define the sets of sentences:

$\Gamma = \{P(\underline{n})$: the least $m > n$ such that $\alpha(\underline{m}) = 0$ exists and is odd$\} \cup \{\neg P(\underline{n})$: the least $m > n$ such that $\alpha(\underline{m}) = 0$ exists and is even$\}$,
$\Delta = \{c > \underline{n} : n \in \omega\}$,
$\Sigma = HA \cup \Gamma \cup \Delta$.

Σ has no models.

Proof By way of contradiction, let M be a model of Σ. From the induction schema (extended to P), we get $M \vDash \neg\neg\forall x < c(P(x) \vee \neg P(x))$. On the other hand, if $M \vDash \forall x < c(P(x) \vee \neg P(x))$, then $M \vDash P(\underline{n}) \vee \neg P(\underline{n})$ for all $n \in \omega$ and it follows that $\forall x \exists y > x \alpha(y) = 0$, against lawlessness. Hence $M \vDash \neg\forall x < c(P(x) \vee \neg P(x))$, which is absurd. Thus $\Sigma \vDash \bot$. On the Other hand, every finite subset of Σ is interpretable in the standard model of ω, and therefore not $\Sigma \vdash \bot$.

Incompleteness for natural semantics shows that the formal inference rules are inadequate to capturing the intended meaning of logical constants, as expressed by Heyting's explanation.

The nature of this gap can be better understood in the light of the completeness results with respect to the so-called *fallible* models. As it is well known, Veldman

(1976), de Swart (1976) and others of the Nijmegen school reached in the seventies intuitionistic completeness proofs of IPC with respect to a modified type of Kripke and Beth models. The main feature of such models consists in allowing $\bot$ to be true at some node. For this reason, they are called sometimes fallible models. The simplest completeness proof for single formulas with respect to fallible models was found by Friedman, rearranged by Troelstra and published in the book *Constructivism in Mathematics* by Troelstra and Van Dalen (1988). In that proof it is constructed a universal fallible Beth model, in which are true exactly all derivable sentences of IPC.

Though Beth models are of no immediate significance for natural semantics, it is known that with every lawless path of a Beth model it is possible to associate a natural model in such a way that truth in the Beth model is equivalent to truth in all its paths. So, in virtue of the theory of lawless sequences, Beth validity turns out to be equivalent to natural validity.

By means of this connection between Beth models and natural models, one can translate the notion of fallible Beth model into a notion of *fallible natural* model, which appears as a direct generalisation of natural model. Sticking to this latter notion, by a *fallible model tout court* I shall mean a model defined as a natural model, except that Heyting's clause for $\bot$ is given up and the interpreting function I assigns to $\bot$ any proposition $I(\bot)$ *implying (the interpretations of) all atomic sentences*.

Soundness for fallible models is trivially verified. As to strong completeness, by generalising Friedman's completeness proof for single formulas, one can easily prove the following theorem:

Theorem 15.1 *Let* Γ *be any set of sentences. For every lawless* $0-1$*-sequence* α *there is a fallible model* M_α *such that the sentences derivable from* Γ *are just those true in all* M_α*'s.*

Observe that, taking for Γ the set of our counterexample to natural strong completeness, we have that not all M_α's are contradictory, while no M_α is non-contradictory.

It is a general feature of all M_α's (relative to any Γ) that no one of them is non-contradictory.

I'll try to draw some philosophical implications of this theorem.

First of all, as already observed by Dummett, completeness for fallible semantics shows that the only logical constant responsible of natural incompleteness is negation (or absurdity). In fact, for the $\bot$-free fragment of IPC, fallible models are natural models and strong completeness holds. Natural incompleteness of the whole IPC shows that the formal system of inference rules is inadequate to catch Heyting's clause for $\bot$; i.e.: nothing is a proof of $\bot$. Indeed, this is the only Heyting's clause not respected by fallible models. The only rule governing $\bot$ is the elimination rule, i.e. *ex falso quodlibet*, which expresses just the peculiarity of absurdity preserved by fallible models. The lack of introduction rules for $\bot$, on the other hand, is not itself an inference rule, so that no rule can guarantee the unprovability of $\bot$.

Perhaps it is worth noting that the formal system of *classical* predicate logic CPC is as well inadequate to catch the intended falsity of $\bot$. Indeed also classical logic

is sound for fallible models. The reason why this fact does not prevent classical completeness is that, classically, validity for natural models is sufficient to assure validity for all fallible models. In other words, both classical and intuitionistic logics are inadequate to capture the intended meaning of $\bot$, though only intuitionistically such inadequacy affects the notion of validity.

But, apart from its role in analysing the source of natural incompleteness, is fallible semantics in itself of any intuitive significance for the intuitionistic conception of truth?

First of all, observe that in fallible models, so as they are defined, $\bot$ not only has a non-standard interpretation, but, strictly speaking, is not even a logical constant, its interpretation varying from model to model arbitrarily (subject only to the clause of implying all atomic sentences). In order to remedy for this shortcoming, we can restrict the class of fallible models by requiring $\bot$ to be exactly *the conjunction* of all non-logical atomic sentences. More precisely, define a *positive model* as a natural model, except that Heyting's clause for $\bot$ is to be replaced by the following:

(5′) A proof of $\bot$ is a method of proving (the interpretations of) all non-logical atomic sentences.

The constance of $\bot$ is so recovered, in the sense that, once fixed the language, the interpretation of $\bot$ is determined by those of non-logical symbols. It should be noted, however, that, in a positive model, the meaning of a sentence containing $\bot$ is not merely determined by the meaning of *its own* symbols, the meanings of *all* non-logical atomic sentences of the language being involved by $\bot$. This behaviour is similar to that of equality: interpreting it as indiscernibility makes its meaning depending on all non-logical predicates of the language.

Theorem 15.1, so as it stands, fails for positive semantics. For, if Γ is the set of universal quantifications of all non-logic atomic formulas, then $\Gamma \models \bot$, but not $\Gamma \vdash \bot$.

The theorem holds, however, with a slight modification:

Theorem 15.2 *Given a language L for IPC, extend it to* $L' = L \cup \{P\}$*, where P is a new predicate letter (of arbitrary degree). Let* Γ *be any set of L-sentences. For every lawless* $0-1$*-sequence* α *there is a positive* L'*-model* M_α *such that the L-sentences derivable from* Γ *are just those true in all* M_α*'s.*

So, strong completeness is restored and $\bot$ is treated as an outright logical constant. Such a logical constant, of course, after having lost its main feature of unprovability, cannot anymore be regarded as the authentic absurdity. Nevertheless, I think that positive semantics is of a remarkable significance for natural semantics. For, the positive interpretation of $\bot$, as conjunction of all non-logical atomic sentences, not only does not involve any notion extraneous to Intuitionism, but it is explained in terms of the intended meaning of the other logical constants. Moreover, the recognition of soundness for fallible semantics does not seem to presuppose the usual notion of negation. This means that, as far as abstract deduction is concerned, Heyting's meaning of negation, understood as recognition of the impossibility to prove something, plays

no essential role in informal intuitionistic reasoning. In other words, soundness and strong completeness for positive semantics assure that, in order to draw the logical consequences of certain assumptions about an abstract structure, an intuitionist never needs to use the concept of unprovability. Therefore, from a logical point of view, Heyting's negation seems to be redundant and positive semantics may be regarded as a natural *negation-free* semantics for the whole IPC.

Not so for classical logic. The positive interpretation of $A \vee \neg A$ says that either A is true or it implies all non-logical atomic sentences. The classical justification of this principle seems to presuppose the standard classical notion of falsity in order to distinguish two cases according to whether A is true or false.

There is another approach to fallible semantics, which exploits some different insights: it preserves the intended meaning of $\bot$ and relativises the general notion of proof. The basic idea is suggested by the question: how may it happen that a working mathematician proves an absurdity? This happens in two recurring cases: when his reasonings are incorrect or when he is arguing under some false assumptions.

The first case is of no theoretical interest, since the highly idealised notion of proof, on which intuitionistic truth is based, is free of errors. In spite of its name, fallible semantics cannot be justified by invoking any fallibilistic conception of proof.

The second case, on the contrary, plays a very important role even in the activity of the idealised mathematician, whose hypothetical reasonings may well rest on some false assumptions. This observation suggests the attempt to generalise the notion of natural model by including, as an ingredient of a model, a *set of hypotheses*. Truth in a model should then be interpreted as informal provability under the given set of hypotheses and, in particular, the truth of $\bot$ as a *reductio ad absurdum* of the hypotheses.

The success of this idea depends on how the notion of proof of a proposition A under a set S of hypotheses is understood. The *orthodox* intuitionistic meaning of this notion is immediately given by combining Heyting's explanation of universal quantification and of implication: a proof of A under S is a method of transforming a method of proving every proposition of S into a proof of A. So understood, however, the notion of proof under hypotheses is of no help for strengthening natural validity and reaching completeness. For, if a proposition is provable *tout court*, it is provable *a fortiori* under any set of hypotheses. Validity so generalised would then be equivalent to natural validity.

It is possible, however, to modify the orthodox notion of truth under hypotheses into one that, without loss of intuitionistic intelligibility, serves the purpose of strong completeness.

Roughly speaking, according to the modified notion I am proposing, to prove A under S means that, when carrying on the constructions required from the proof condition of A, we are allowed to assume, step by step, some propositions of S. Unlike for the orthodox notion, no hypothetical method of proving all propositions of S is involved.

Precisely, let M be a natural model and S a set of (meaningful) propositions (not necessarily expressible in the language of IPC). Call *hypotheses* the members of S and *hypothetical model* the pair $\langle M, S\rangle$. Truth in $\langle M, S\rangle$ is defined in Heyting's style

by modifying the proof-conditions according to the following generalised inductive definition:

(1) for A atomic $\neq \bot$, a proof of $I(A)$ is a proof of A;
(2) a method of proving *all* $A(d)$ $(d \in D)$ is a proof of $\forall x A(x)$ [likewise for $A \wedge B$];
(3) a method of proving *some* $A(d)$ $(d \in D)$ is a proof of $\exists x A(x)$ [likewise for $A \vee B$];
(4) a method of transforming every proof of A into a proof of B is a proof of $A \rightarrow B$;
(5) if $H \in S$, a method of transforming every proof of H into a proof of A is a proof of A;
(6) nothing is a proof of A but in virtue of clauses (1)–(5).

According to this definition, it may well happen that a sentence true in M (i.e. provable *tout court*) fails to be true in $\langle M, S\rangle$ (i.e. provable under S). Example:

Let α be a lawless sequence and P a 1-place predicate letter. Interpret P on ω by taking $\alpha(\underline{n}) = 0$ as interpretation of $P(\underline{n})$. So $\neg\forall x P(x)$ is true. Now take as set of hypotheses $S = \{\alpha(\underline{n}) = 0 : n \in \omega\}$. Then $\forall x P x$ is trivially true under S so that, unless we know an m such that $\alpha(\underline{m}) \neq 0$, we cannot prove $\neg\forall x P(x)$.

Soundness for hypothetical semantics is straightforwardly verified: the inference rules are insensitive to hypotheses.

Now, a fallible model may be reinterpreted as a special hypothetical model. In fact, with every fallible model M we can associate the hypothetical model $\langle M', S'\rangle$ as follows: M' is the natural model obtained from M by restricting the interpreting function I of M to the non-logical atomic sentences; $S = \{0 \neq 0 : I(\bot)\}$. M and $\langle M', S'\rangle$ are easily seen to be equivalent.

It follows that Theorem 15.1 can be reformulated for hypothetical models and strong completeness holds.

In this way, we can preserve the intended meaning of all logical constants and exploit the intuition that a proof of $\bot$ is nothing but a *reductio ad absurdum*.

Summing up, I have indicated two possible intuitionistic accounts of fallible models, which try to exploit two different intuitions, respectively : the reinterpretability of negation in positive terms, and the relativisation of informal provability to a certain hidden set of assumptions.

In both positive and hypothetical models,

(i) truth is definable intuitionistically in Heyting's style in terms of the general intuitionistic notion of proof; and
(ii) the modifications, with respect to natural models, preserve, to a large extent, the intuitionistic intended meaning of logical constants.

I do not claim that these are the only possible intuitionistic accounts of fallible semantics. My point is that intuitionistically plausible reformulations of fallible models are possible and that, therefore, fallible models are of remarkable interest for the intuitionistic theory of meaning.

Theorem 15.1 (and its reformulations) shows that, using intuitionistic categories, one can construct a surprisingly wide variety of interpretations of predicate logic.

In fact, every set of sentences, closed under intuitionistic derivability, can be characterised as the set of sentences true in all members of a certain class of models. If, in particular, a set of sentences contains the universal closures of all instances of the excluded middle, its intuitionistic closure is a classical theory. Thus every classical theory is intuitionistically interpretable. So, in spite of the often claimed unintelligibility, from the intuitionistic point of view, of classical reasoning, there is a sense in which every classical reasoning is capable of intuitionistic meaningfullness. Namely, the classical logical consequences of any set Γ of sentences are just the intuitionistic logical consequences of Γ, relative to a suitably restricted class of fallible models. It seems therefore that, within the intuitionistic perspective, classical reasoning cannot rightfully be charged with unintelligibility, but rather with lack of generality.

Of course, an intuitionist may rightfully maintain, from his viewpoint, that classical mathematics is unintelligible, in the sense that the alleged intended structures, which certain classical theories would pretend to describe, are intuitionistically meaningless. For instance, an intuitionist may reject the classical theory of real numbers, for the reason that he does not know any privileged intuitionistic structure, of specific mathematical interest, satisfying all theorems of that theory. Of course, no one of the models provided by the strong completeness theorem is a good candidate for an intuitionistic counterpart of classical continuum (at least because all those models are potentially contradictory). Nevertheless, regarded as an abstract theory, whose aim is to study the features common to all structures satisfying certain axioms, the theory of classical continuum, as any axiomatic theory, is capable, in the sense explained, of intuitionistic reinterpretation.

In spite of Brouwer's aversion to formal logic, it is just the formal analysis of mathematical language, together with the abstract notion of mathematical structure, what confers to classical reasoning intuitionistic intelligibility.

References

De Swart, H. (1976). Another intuitionistic completeness proof. *Journal of Symbolic Logic*, *41*, 644–662.

Dummett, M. (1977). *Elements of intuitionism*. Oxford: Clarendon Press.

Heyting, A. (1956). *Intuitionism: An introduction*. Amsterdam: North-Holland.

McCarty, D. C. (1991). Incompleteness in intuitionistic metamathematics. *Notre Dame Journal of Formal Logic*, *32*, 323–358.

Troelstra, A., & Dalen, D. V. (1988). *Constructivism in mathematics* (Vol. II). Amsterdam: North-Holland.

Veldman, W. (1976). An intuitionistic completeness theorem for intuitionistic predicate logic. *Journal of Symbolic Logic*, *41*, 159–176.

Origin of the Essays

1. "Brouwer, Dummett and the bar theorem" (with Pierdaniele Giaretta), *Atti convegno nazionale di logica*, Montecatini Terme 1979, 541–558, Bibliopolis.
2. "Creative subject and bar theorem", *The L. E. J. Brouwer centenary Symposium*, North-Holland Publishing Company, 1982, 311–318.
3. "Natural intuitionistic semantics and generalized Beth semantics", *Atti incontri di Logica Matematica*, Univ. Siena, 1982, 373–376 (translated from Italian)
4. "Connection between the principle of inductive evidence and the bar theorem", *Mathematics Proceedings* A 86(3), 1983, 325–328, also published in *Indagationes Mathematicae*, vol. 45, fasc. 3, 1983.
5. "On the Brouwerian concept of negative continuity", *Journal of Philosophical Logic* 14, 1985, 379–398.
6. "Classical and intuitionistic semantical groundedness", *Atti incontri di Logica Matematica*, Univ. Siena, 1987, 91–98 (Translated from Italian).
7. "Brouwer's equivalence between virtual and inextensible order", *History and Philosophy of Logic* 9, 1988, 57–66.
8. "An intuitionistic notion of hypothetical truth for which strong completeness intuitionistically holds", *Teoria* VIII, 1988, 2, 131–144.
9. "Propositions and judgements in Martin-Löf" (with Gabriele Usberti), *Atti convegno Pontignano "Problemi fondazionali nella teoria del significato"*, 1991, 125–136.
10. "Negationless Intuitionism", *Journal of Philosophical Logic* 27, 1998, 165–177.
11. "Temporal and atemporal truth in intuitionistic mathematics", (with Gabriele Usberti), *Topoi* 13, 1994, 83–92.
12. "Arbitrary reference in mathematical reasoning", *Topoi* 20, 2001, 65–77.
13. "The priority of arithmetical truth over arithmetical provability", *Topoi*, 21, 2002, 55–63.
14. "The impredicativity of the intuitionistic meaning of logical constants" (with Gabriele Usberti), unpublished, 1988.
15. "The intuitionistic meaning of logical constants and fallible models", unpublished, 1992.

Index

A
Antirealism, 97, 98, 118, 136, 145
Arbitrary reference singular
  plural, 125–128, 130

B
Bar theorem, 1–3, 7–9, 12, 13, 21, 27, 28
Begriffsschrift, 77

C
Choice arbitrary
  plural, 129
  singular, 128–130
Choice sequences, 1, 13, 21, 24, 27, 32, 37, 38, 40, 41, 44, 95, 100, 137
Comprehension principle, 50, 95, 123, 127
Constructivism, 63, 93, 94, 133, 136, 139, 145
Creative subject, 13, 15–17, 25, 27–29, 31, 34–37, 40, 53, 61, 62, 82, 100

E
Epistemic mathematics, 105, 106

F
Forcing, 23, 70

I
Ideal agent, 119, 120, 140, 141, 144, 145
Imagination, 83, 114
Impredicativity, 100, 126, 148, 151, 154
Inductive evidence, 20, 21, 28
Infinite actual
  potential, 6, 135, 138, 145
Intensional identity, 38, 86–89, 151
Intuitionism orthodox
  potential (non-orthodox), 134–137

L
Lawless sequences, 37–41, 44, 45, 63, 65, 70, 87, 88, 90, 92, 94, 95, 127, 158, 159
Logical constants
  classical, 47, 81, 92, 94, 97, 100, 102, 104–106, 110, 113, 117, 118, 129, 133, 134, 136, 137, 139, 140, 145, 157, 159–161, 163
  intuitionistic, 81, 85, 88–90, 93, 97, 100, 102, 105, 106, 137

M
Markov principle, 24
Models Beth
  fallible, 157, 159, 160, 162, 163
  natural, 157, 159, 160, 162

N
Negative continuity, 31–33, 44, 45

O
Order virtual
  inextensible, 53–57, 60, 62

P
Proof canonical

informal intuitionistic, 161
Proof tree, 8

R
Realism computational
metaphysical, 82

S
Semantical completeness, 23–25, 50, 63, 85, 86, 88
Semantical groundedness, 47–49
Semantics game
natural intuitionistic, 63, 85
negationless, 50, 85, 86, 89, 90, 93, 94
Solipsism, 36
Strengthened liar paradox, 49

T
Truth classical
intuitionistic temporal
atemporal, 97, 99, 100, 106, 108, 165
Type theory ramified
simple, 123