ON LOGIC

Arthur Nieuwendijk

ON LOGIC

Inquiries into the Justification of Deduction

ILLC Dissertation Series DS-97-03

For further information about ILLC-publications, please contact

Institute for Logic, Language and Computation
Universiteit van Amsterdam
Plantage Muidergracht 24
1018 TV Amsterdam
phone: +31-20-5256090
fax: +31-20-5255101
e-mail: illc@fwi.uva.nl

ON LOGIC

Inquiries into the Justification of Deduction

Academisch Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit van Amsterdam,
op gezag van de Rector Magnificus
Prof. dr. J.J.M. Franse
in het openbaar te verdedigen in de
Aula der Universiteit
(Oude Lutherse Kerk, ingang Singel 411, hoek Spui)
op 8 april 1997 te 15.00 uur

door

Alexander Clemens Nieuwendijk

geboren te Hilversum

Promotor: Prof. dr. R. Bartsch

Co-promoter: Dr. F. Veltman

ISBN: 90-74795-68-4

Voor Enzo

Zijn is willen

Contents

Acknowledgments

Frank Veltman, Martin Stokhof, Dick de Jongh, Willem Groeneveld, Harry Stein.

Introduction

Views on logic

This dissertation is *on* logic. It examines three different ways of looking at logic. Three different ways of answering the same question: How is it possible to justify logical principles? It examines an *ontological* point of view in the fourth chapter, a *transcendental* point of view in the third one, and a *semantic* point of view in the first two. Despite this underlying thematic unity, the chapters differ with respect to their main subject.

Chapter IV, 'Semantics and Comparative Logic', examines the assumption underlying current western interpretations of the logic developed by the Indian school of philosophy called Navya-nyāya. Within this tradition the analysis of language as well as that of valid inference was assigned an important place. In western literature, it has become common practice to interpret and evaluate Navya-nyāya logic in the light of two-valued classical predicate logic. Normally, this results in a depreciation of the logical content of Navya-nyāya philosophy, or in an attempt to reconstruct this philosophy so that it fits the framework of classical predicate logic. Both are unfortunate, since it is unclear why classical predicate logic should be taken as the universal logical standard. In this chapter it is shown that taking classical logic as the prevailing logical norm brings with it a particular view on what meanings are. It is argued that this view on meaning does not fit the Navya-nyāya view on ontology. Taking this ontology as a starting point a different semantic view is developed, which, as it turns out, supports a non-classical logic, and which thus can be taken as explaining why Navya-nyāya logicians do not accept certain classical principles, among which the principle of the excluded middle is the most prominent one.

The same principle is also present in chapter II, called 'Logical revisionism and expressibility'. The subject of this chapter, however, could hardly be more different. This chapter, as are the remaining two, is grafted onto a specific feature of Michael Dummett's philosophy of logic. Chapter II is about one of two steps of a particular argument against classical logic. The blueprint of this argument is a recurrent theme in Dummett's writings, and it has become one of the hallmarks of modern anti-realism. The argument in its full form falls apart into two steps. The first one consists of an argument to the effect that the semantic principle of bivalence is unacceptable. The second one is to take us from the rejection of bivalence to the rejection of the logical law of the excluded middle. The first step is discussed extensively in chapter I. In chapter II, however, it is simply assumed that we have indeed good reasons for rejecting classical semantics, and our undivided attention is directed to the claim that a rejection of the principle of bivalence necessitates a revision of classical logic.

Against this claim there is an obvious and well-known counterexample provided for by super-valuations. This semantic theory shows that there is no technical obstacle to devising a semantics in which the principle of bivalence does not hold, but which, nevertheless, validates classical logic. Does this then put an end to the anti-realist's ambition to argue against the justifiability of classical logic? Not necessarily so. It is shown that classical logic can be validated by a partial, or non-classical semantics only at the expense of losing the means of expressing certain conceptual distinctions, most notably the distinction between direct and indirect evidence. Given, furthermore, that this distinction is crucial from the point of view of semantic revisionism, this offers a solid ground for rejecting the super-truth approach as a counterexample to logical revisionism.

In 'Use and logical revisionism', chapter III of this thesis, yet another way of determining the correct logic is examined. The basic idea is simple, beautiful and highly appealing. It says that logic is justifiable on the basis of an analysis of the concept of proof. The analysis is expected to explicate what is the least we need to be able to *do*, if we are to obtain anything at all that can be called a logical proof. Under this view a proof is seen primarily as the result of an *activity*, more precisely, the activity of *using language* in a particular way. The general idea is that we can use expressions, and most notably the logical constants, in many ways, but that some of these uses are special in the sense that if we were forbidden to use the constants in that way, we could not obtain a logical proof at all. Clearly, then, these uses are to be conceived of as constituting the conditions without which no proof is possible.

For an investigation into these 'minimal uses' of the logical constants to be successful there must be something like a clear conception of proof from which it is possible to determine a unique logic. But in the exposition of Dummett's 'transcendental deduction' of the possibility conditions of proof, it becomes clear that this assumption is untenable. For it turns out that the answer to the question which logic can so be justified depends heavily on external and quite trivial assumptions regarding the concept of proof.

Chapter II and chapter III both raise the question whether it is possible to argue for logical revisionism without being forced to step outside the context within which we are supposed to develop the argument, and which also is believed to give the argument its strength. This context is given by the constraints that a semantic revisionist imposes on any inquiry into problems of logic and of metaphysics. Chapter I, 'Semantic revisionism, anti-realism, and logic', discusses these constraints. They are presented as constituting a particular view on the subject of logic as it is developed in opposition to the traditional view according to which questions concerning the justification of logical principles can only be answered by means of metaphysical inquiries. In contrast to this, semantic revisionism claims that the only way to answer questions of logic and of metaphysics is by means of a meaning-theoretical investigation.

In chapter I we will not be concerned with the construction of a particular model of meaning, but only with the conditions that any such model must meet if it is to be adequate. In particular we will examine what the consequences are of the assumption that meaning is prior to both logic and metaphysics, both for our conception of what it is that a model of meaning has to account for, as well as for the form it must take if it is to live up to these expectations. Finally, given the form that a model of meaning must take, we will see that if we take classical semantics as our starting point, we cannot fulfill all the tasks that a model of meaning is expected to fulfill.

This rejection of classical semantics depends, in the end, on the acceptance of semantic revisionism. Dummett averts to semantic revisionism, because, as he believes, the problem of the justification of logic *cannot* be solved if we look at it from the traditional point of view. The interesting question is: Why *must* the traditional approach fail? In Dummett's writings one will search in vain for an answer. One will not even find an explicit formulation of the question. Let me make a suggestion. The traditional approach fails, because if we set out to determine what acceptable logical rules are in ontological terms, then, in the course of doing this, there will come a moment at which we will find ourselves forced to change or extend our original perspective. This, of course, does not explain why the ontological perspective necessarily leads to a theoretical inability. It only points at a direction that such an explanation could take. But despite this, by slightly generalizing it we can come to the following working hypothesis: A perspective on a philosophical problem has a foundationalist force if we are able to solve the problem completely within the limits that constitute the perspective itself.

Seen in the light of this working hypothesis, chapters II to IV all contribute to an evaluation of semantic revisionism. With respect to the argument developed in chapter II, the question is whether the need for the distinction between direct and indirect evidence can be argued for by reference to the use of statements only, that is, within the limits constitutive of semantic revisionism itself. In the case at hand, this is not an insurmountable problem. Yet, this does not conclusively establish the claim that a rejection of classical semantics necessitates a rejection of classical logic. Other counterexamples are still possible, and it cannot be said in advance whether it will always be possible to defend the revisionist consequences of anti-realism along similar lines. That will all depend on whether the relevant conceptual distinctions can be argued for within the context of semantic revisionism. And there is no reason at all to suppose that this can be done in all cases.

Seen in the same hypothetical light, the findings of chapter III have more far-reaching consequences. From the fact that we cannot uniquely determine a logic from a conception of proof, which is stated exclusively in terms of use, it follows that, in the course of determining what logical principles are acceptable, we are forced at some moment to leave the original perspective. And this is as good a reason for rejecting semantic revisionism, as is Dummett's reason for rejecting Frege's platonism.

In all this, chapter IV occupies a special position. It could be taken as a counterexample to the claim that the ontological perspective necessarily leads to theoretical inability as far as the justifiability of logical principles is concerned. But this would do no justice to the ambitions of a semantic revisionist. For he wants to determine *the* correct logical rules, the rules that can be accepted no matter how we think about reality. In chapter IV classical predicate logic is rejected as a logical standard on the basis of an examination of the Navya-nyāya views on ontology. And these are just given. According to a semantic revisionist this says no more than that classical logic ought to be rejected if the Navya-nyāya views on ontology are correct. And, he continues, the problem is that we do not know how to determine whether or not these views are correct. Indeed, what chapter IV does show is that it is not impossible to determine what logical rules are acceptable given a particular view on reality. A semantic revisionist wants to do more, but looked upon in the light of the first three chapters of this thesis, it is at least doubtful whether it is possible to do more. That is, it is doubtful whether the idea of a logical rule that is valid irrespective of how we view reality can be given any content at all.

I Semantic Revisionism, Anti–realism, and Logic

1 How to Justify Logic?

1.1 The Objectivity of Meaning

Suppose that one is led to an interest in the question: How, if at all, is it possible to criticize or question fundamental laws that are generally accepted?[1] Then how would one go on to answer it? This chapter examines Michael Dummett's attempt to answer this perplexing question. As is well known, for Dummett this question coincides with another one: Is it possible to argue persuasively for *anti-realism*?

An anti-realist denies that the logical law of the excluded middle has universal validity. This, he believes, is a direct consequence of the fact that there are statements of which we will never be in a position to know whether they or their negations are true. A *realist*, on the other hand, holds that the truth or falsity of a sentence does not in any way depend on our knowledge, or possible knowledge of what the truth-value of this sentence is. Therefore, he says, the law of the excluded middle is universal valid.

The problem with arguments like these is that they convince only those who are already believers. A realist will never be persuaded by the anti-realist's argument, just because what the anti-realist believes to be crucial —the dependence of the truth-value of a sentence on our ability to known this truth-value— is believed by the realist to be totally irrelevant, while the anti-realist's argument does not establish the relevance of this thesis, but simply assumes it. Does this mean that we cannot hope for a solution of the dispute between realists and anti-realists? It does not. But it does show that in order to answer

[1] Cf. Dummett 1992c, p. xix.

the *prima facie* logical question whether or not the law of the excluded middle is a universal valid law we need an answer to a question of a completely different nature, namely the question whether or not the truth or falsity of a sentence depends in any way on the abilities we possess for coming to know what this particular truth-value is.

The general point is that the realism/anti-realism debate is not *directly* concerned with the status of logical principles. Perhaps more surprisingly, it is not about questions of existence, either. This was already perceived by Kreisel, who once remarked that

> the essential issue between platonist and constructivist philosophies of mathematics concerns not so much the existence of mathematical objects as the *objectivity* of mathematical truth.[2]

It is this claim, which has been met with the approval of both Michael Dummett and Crispin Wright, another influential advocate of the anti-realist point of view on meaning and truth,[3] that serves as the starting point of the current investigation into Dummettian anti-realism. To be more precise, I will expound the thought that a Dummettian anti-realist rejects realism on the ground that realism, due to its characteristic assumptions, cannot account for the *objectivity of meaning*.

To support this claim I will take Dummett's criticism of Frege's explanation of the objectivity of understanding as a starting point. Dummett's criticism is not directed towards the general set up of Frege's account, but concerns only one, though decisive aspect of it: the thesis that thoughts can be conceived of as being independent of the means we have for expressing them. Frege's account of objectivity will be presented as consisting of two parts: one part comprising the general criteria that any explanation of meaning has to meet if it is to account for the objectivity of meaning, and another part consisting of Frege's elaboration on these criteria. The main point to be made is that acceptance of the first part of Frege's theory does not commit one to acceptance of the second part. Even more so, in Dummett's criticism we will read the claim that due to his assumption of thoughts being independent of the means we have for expressing them, Frege cannot meet his own criteria. In opposition to Frege Dummett endorses the claim that 'meaning is use', which, as it turns out, is nothing but the succinct

[2]Kreisel 1965.

[3]See, for example, Wright 1980, p. 5, Dummett 1992a, p. 508.

expression of the assumption that thoughts can only be conceived of as being dependent on the means we have for expressing them. It is in this context also that we come to see both the requirement of *manifestation*, and that of *recoverability* as mere reformulations of Frege's criteria for the objectivity of meaning. According to the former a theoretical explanation of meaning must be able to explicate for each sentence of the language, what knowledge of the meaning of that sentence consists in. The latter says that a meaning-theory must explain how the meanings it assigns can be learned by the speakers of the language from the use made of it by other speakers.

1.2 Semantic Revisionism and Anti-realism

The realism/anti-realism debate is about the correct model of meaning, and according to an anti-realist, a realist semantics is untenable, because it cannot meet the requirements that a semantic theory in general ought to meet. But if we are interested in the justification of logic, then why should we be concerned with meaning? To see this, we must turn to a characteristic presupposition of Dummett's overall strategy: the conviction that questions concerning the justification of logical principles can only be answered by developing a new view on the subject of logic in opposition to the traditional view according to which such questions can only be answered by means of metaphysical or ontological inquiries. According to the point of view explored by Dummett we do not any longer consider metaphysics to be prior to logic, but, on the contrary, we take meaning to be prior to both logic and metaphysics.

So, where traditionally it was thought that logic could be justified by examining what the correct metaphysical views are, we now concentrate on the question: What is the correct model of meaning? This question, Dummett believes, can be answered solely on the basis of what we already know about the meanings of the statements of our language. This knowledge is complete, because our statements have the meaning they have solely in virtue of the way we use them. Hence, to be able to use a statement correctly just is to know the meaning of that statement. This is not to say that we 'command a clear view of the working of our language.'[4] But it does say that in order to gain a

[4]The phrase, of course, is Wittgenstein's. Dummett uses it for his own purposes. See Dummett 1991b, p. 13.

complete understanding we have to do no more nor less than to scrutinize our practices, aiming, eventually, at a systematic description of them by constructing a model of meaning. The correctness of a model of meaning will not depend on any metaphysical presupposition, but only on the working of our language.

In this chapter, the idea that there is no need to go beyond the use of language in order to determine the correct model of meaning is interpreted as a *norm* constituting a particular *way of looking* at certain problems in metaphysics and logic. As a norm it states that a metaphysical position, or some set of logical principles, or some particular explanation of meaning is acceptable only if it can be justified without appealing to anything that does not directly concern the use of language. This induces a distinction between semantic revisionism and anti-realism. *Semantic revisionism* is characterized by putting forward a new perspective on logic, metaphysics and meaning. This perspective is essentially normative in that it determines the criteria for evaluating philosophical theories. *Anti-realism*, on the other hand, is not itself a perspective, but a position within the framework of semantic revisionism.

To get an idea of the interplay between semantic revisionism and anti-realism, we will start with examining Dummett's interpretation of the terms 'realism' and 'anti-realism', with special attention to the role of logic, and to the relation of this particular conception towards traditional, ontologically oriented realism/anti-realism debates.

2 Semantic Revisionism

2.1 The Primacy of the Theory of Meaning

Frege's greatness, Dummett says, consists in his having perceived that the theory of meaning is the fundamental part of philosophy which underlies all others. By this he means that Frege conceives of the theory of meaning as the only part of philosophy whose results do not depend upon those of any other part.[5] This 'revolutionary perspective' induces a change in the way philosophical questions are to be understood. Especially,

[5] Dummett 1992a, p. 669.

> the fundamental question of metaphysics, namely the resolution of the dispute between realism and idealism, comes to be seen as a dispute over the general form which a theory of meaning should take.[6]

The progress we make by understanding a metaphysical dispute as a dispute over 'the general form which a meaning-theory should take'[7] is that this will make a resolution of persistent problems possible. For traditional metaphysics, as Dummett sees it, the major problem is that we do not know how to solve the disputes that mark the history of Western philosophy:

> The moves and counter-moves are already familiar, having been made repeatedly by philosophers on either side throughout the centuries. The arguments of one side evoke a response in certain of the spectators of the contest, those of the other side sway others of them; but we have no criterion to decide the victors. No knock-out blow has been delivered. The decision must be given on points, and we do not know how to award points.[8]

This metaphysical deadlock, Dummett suggests, evolved out of a complete lack of a clear conception of the content of the opposing theses: we have no way to explain what accepting either of them amounts to.[9] Clearly, such a conception is a prerequisite for solving the metaphyis-cal dispute in question. But what exactly is meant by this? Here is a simple example. Consider the formal assertion 'Fa'. Asked for the content of this assertion one could proceed as follows. 'F' stands for a certain set of individuals, 'a' stands for a particular individual, and 'Fa' means that the individual a belongs to the set of individuals that have the property F, in other words, 'Fa' means that a is F. The simple statement 'Fa' is, of course, far from being a metaphysical assertion. But we need not resort to complex statements in order to observe that what tells us something about the content of a statement is not the final explanation 'Fa' means that a is F, but the way in which this explanation is brought about, the principles that underly it. In other

[6] Dummett 1992a, p. 671.

[7] I use 'meaning-theory' instead of 'theory of meaning' for reasons that will become clear below. The quotation mentions 'theory of meaning' due to the fact that the distinction between these two notions is not made explicitly before Dummett 1991b.

[8] Dummett 1991b, p. 12.

[9] *ibid.*

words, the content of statement is determined by the principles that govern its interpretation.

This only elaborates on Dummett's claim that a metaphysical doctrine originates from a 'particular way of understanding our thoughts about reality.'[10] It follows that in order to evaluate a metaphysical doctrine we need to explicate that particular way of understanding of which it is a manifestation. It is thought to consist of a set of principles that govern the way we interpret particular sentences. Together these constitute what is called the underlying *model of meaning.*

To settle a metaphysical dispute we should thus start out with explicating the models of meanings that underly the competing doctrines. A systematic representation of a model of meaning is called a *meaning-theory.* It is to be distinguished from the *theory of meaning.* The former gives a 'complete specification of the meanings of all words and expressions of one particular language,' while the theory of meaning 'enquires into the general principles upon which a meaning-theory is to be constructed.'[11] And these principles form the basis on which a specific meaning-theory is to be evaluated. Or, as Dummett frequently expresses it, the theory of meaning determines the general *form* that a meaning-theory should have if it is to be adequate.

What is the progress made by the transformation of the question as to the correct metaphysical doctrine into the question as to the correct model of meaning? For Dummett this is obvious: We cannot hope to find an answer to the former question, but, in contrast, the latter one can be decided. Why? Well, because the theory of meaning is the only part of philosophy whose results do not depend upon those of any other part.

This is puzzling. Not so much the claim that the theory of meaning forms the basic of philosophy, although, of course, one could have second thoughts about that, but the claim that because the theory of meaning has this status, an adequate standard for evaluating meaning-theories can be developed. The underlying thought is the following. In a genuine philosophical dispute the opposing positions must share some common ground, which is given by their agreement on what the disagreement is about. What is agreed upon is reflected in the terms in which the opposing positions are stated, it is the perspective into

[10] Dummett 1991b, p. 15
[11] *ibid.*, p. 22.

which the problem at hand is looked upon. A quite effective method for settling the disagreement is to derive the criteria for evaluating the opposing positions solely from the shared context, without which the controversy looses its point. What Dummett is claiming is that we cannot derive such criteria from the context alone if we do not take the meaning-theoretical perspective. In any other case we must, in order to settle the dispute, go beyond the perspective that gives the original dispute its point to a perspective, that does not at all contribute to the solution of the problem, but, instead, induces a new one.

Whether or not the meaning-theoretical perspective is indeed the exception that Dummett claims it to be is something that is yet to be seen. This makes semantic revisionism, that is the claim that we ought to take the meaning-theoretical perspective, highly programmatic. This programme is distinctively foundationalist in its outlook. There are basically two different ways of interpreting this foundationalist character of Dummett's position. The first is to take the belief that the theory of meaning is independent of, and basic to all other parts of philosophy as more or less the outcome of a philosophical investigation. Under this interpretation the belief is to be rejected if one fails to explicate how all the rest of philosophy depends on the theory of meaning. But Dummett's foundationalism is of a different kind. It is not some belief supported by some 'known facts', but the expression of a way of doing philosophy. The difference between these two forms of foundationalism is the way in which they can be refuted. Conceived of as a 'statement of a matter of fact' it is rather clear how to criticize it. But one cannot so easily refute a perspective, a way of looking. The only way to dismiss it is by arguing that the promise it makes cannot possibly be fulfilled on its own terms.

The upshot of these considerations is that Dummett's semantic revisionism is foundationalist, not in the sense of laying down some principles from which all other are said to be derivable, but by imposing a constraint on the way in which disputed principles are to be evaluated. To state, therefore, that the theory of meaning determines whether or not a meaning-theory is correct independently of any other part of philosophy, is not to state a matter of fact, but a *criterion*. Hence, a meaning-theory is correct, or acceptable, if and only if it can be so justified, that is, if it can be justified in meaning-theoretical terms only.

2.2 Metaphysics from the Point of View of Meaning; an Example

Semantic revisionism originates from the assumption that meaning is prior to both logic and metaphysics. But how are we to understand a metaphysical dispute as a dispute over the general form which a meaning-theory should take? Semantic revisionism is analytic in the sense that it claims that philosophical problems are best approached through the analysis of language. But the more far-reaching requirement is that this analysis should not in any way depend on traditional categories, like qualities, observable and theoretical objects, or events. For, as Dummett observes, the philosophical position traditionally considered to be anti-realistic offer the realist 'an easy victory' as soon as they fall prey to an appeal to traditional categories in their analysis of language.[12] The conclusion Dummett draws from this is that the only way for an anti-realist to achieve a philosophical respectable position is to eliminate all reference to traditional categories and redefine his position in semantic terms only, in other words, to redefine his position in the context of semantic revisionism.

Let me illustrate how the anti-realist comes to be so easily defeated by the realist by discussing Dummett's analysis of phenomenalism. The form of anti-realism Dummett aims to characterize takes the form of a species of reductionism, because its opposition to realism concerns the question 'whether or not entities of a particular kind are among the ultimate constituents of reality.'[13] Phenomenalism fits this description very well in that it denies the existence of material objects by claiming that all statements about such objects are reducible to ones about sense-data. For matters of convenience, we will call the class of statements about material objects the disputed class, and the class of statements about sense-data the reductive class.[14] The general idea is that statements of the reductive class express the existence of evidence for or against the statements of the disputed class.[15] Now phenomenalism assumes that the sense-datum language is logically prior to, and independently of the language we use to talk about material objects. And it is claimed that a sentence about a material object is true only if its translation in the sense-datum language is true.

---

[12] Dummett 1978b, p. 165.
[13] *ibid.*, p. 145.
[14] *ibid.*, p. 155.
[15] *ibid.*, p. 156.

The distinctive characteristic of phenomenalism also forms its weak spot. Opponents have rejected it on the ground that it is impossible to isolate the class of sense-datum statements, or because it fails to make clear how sense-datum statements acquire their meaning. A phenomenalist is vulnerable for this kind of realist objections, Dummett suggests, as long as he invokes notions like 'sense-data' in the development of his position. One way to characterize the difference between phenomenalism and realism about material objects is to point out that a phenomenalist would explain the meaning of material-object sentence in a different way than a realist would. But, as Dummett remarks, in order to maintain this difference it is not at all necessary to explain the meaning of material-object sentences by reference to sense-datum sentences, nor is it needed to explain the meaning of the latter in terms of sense-data. What makes the difference is not this, but the phenomenalist's insistence that the meanings of material-object statements be tied to the existence of observational evidence. And the idea of a reductive class, of a translation of sentences of the disputed class in sentences of the reductive class, and of the existence of sense data in terms of which the meanings of the sentences in the reductive class is explained, constitute a particular elaboration of this general point. The anti-realist can give this up, without his position collapsing into that of the realist.

What Dummett purports to argue for, then, is that even if a phenomenalist decides to drop the assumption of the existence of a reductive class, it does not follow that he assigns the same meanings to material-objects sentences as does a realist and that, consequently, he might still oppose the realist view. In particular, in contrast to a realist, the phenomenalist might deny that every a material-object statement is either true or false. In that case, it might very well be that

> some deductive inferences which would be valid on a realist view would fail on a phenomenalist view, and that this would necessarily result in a difference as to which material-object statements we have adequate indirect evidence for, and hence in which ones were believed.[16]

We thus reach a point at which we conceive of the dispute between phenomenalism and realism as a dispute over the validity of certain deductive inferences. And this, Dummett claims, is not unique for this particular dispute. Indeed, any traditional controversy between realists and anti-realists can so be reconstructed, since

[16]Dummett 1978b, p. 160.

> all disputes over realism entail a disagreement about the criteria for the validity of deductive arguments containing statements of the disputed class.[17]

We now have come close to the completion of the reinterpretation of a metaphysical dispute as a dispute over the correct model of meaning. Only one step is missing, the one that takes us from logic to meaning. In the last quotation it is stated that a disagreement on logic *is entailed by* a dispute over realism. And, according to Dummett, the only reason for this is that a dispute over realism is, in the end, a dispute over the correct model of meaning. The crucial idea here is that a disagreement over the correct model of meaning necessarily involves a disagreement over the validity of some logical principles. This is not at all a trivial assumption, one also which I believe to be false. It would take me too far astray to defend this claim here, but I will return to it in chapter II, where it will be the leading question.

Although this assumption is most interesting from the point of view of logic, and also quite essential for the intelligibility of Dummett's logical revisionism, it is not needed for the reconstruction of a metaphysical debate as a dispute over the correct model of meaning. To return to our example, when Dummett characterizes phenomenalism in terms of the denial that every material-object statement is either true or false, this already is sufficient to substantiate the basic thought that

> The conflict between realism and anti-realism is a conflict about the kind of meaning possessed by statements of the disputed class.[18]

It even rather nicely illustrates Dummett's more detailed ideas about the structure of the realism/anti-realism debate in general:

> Realism I characterize as the belief that statements of the disputed class possess an objective truth-value, independently of our means of knowing it: they are true or false in virtue of a reality existing independently of us. (...) The anti-realist, on the contrary, insists that the meanings of these statements are tied directly to what we count as evidence for them.[19]

Notwithstanding the fact that this quotation does mention a reality existing independently of us in the characterization of realism, the basic

[17] Dummett 1978b, p. 164–5.
[18] *ibid.*, p. 155.
[19] *ibid.*, p. 146.

idea it expresses is that, from the perspective of semantic revisionism realism is characterized by the thesis that every sentence is either true or false, while anti-realism distinguishes itself by the thesis that a sentence is true if and only if there is evidence for its truth, and false if and only if there is evidence for its falsifity.

2.3 Semantic Anti-realism

While it is essential under the traditional interpretation of phenomenalism that it denies the existence of material-objects, Dummett's reinterpretation dismisses the question as to the existence of material-objects as being totally irrelevant. In this sense, Dummettian anti-realism can be said to oppose not only traditional realism, but also traditional anti-realism. One could say that Dummett seeks to preserve the structure of the opposition between realism and anti-realism within a radically different framework. This new form of anti-realism is, one could say, a kind of isomorphic image of the traditional kind. In order to distinguish it from the latter I will name it *semantic anti-realism.*

We must be careful not to draw any ontological consequences from the semantic anti-realists theses. Take, for example, the rejection of Frege's assumption of a 'third realm' of independently existing objects, which we will discuss below. The semantic anti-realist comes to this rejection on the grounds that an appeal to those entities obstructs an account of the objectivity of meaning. From this we cannot conclude that he denies that there are independently existing objects. Nor, for that matter, that he affirms this. The only thing we can conclude is that on the anti-realist's score ontological issues are eliminated completely from the meaning-theoretical discourse. But are they thereby also eliminated from metaphysics?

Given that metaphysical disputes are now all conceived of as meaning-theoretical disputes, the answer appears to be: Yes. But giving it a second thought, I think we can do better than this. What is rejected under the new perspective is not so much ontology as such, but the assumption that we can characterize what it is to speak meaningfully only by reference to the objects that are spoken about. This is not so much a rejection of the notion of an object, as it is a challenge to reinterpret it by characterizing objects in terms of what it is to speak meaningfully about them.

I come to this, not with the pretension to meet this challenge, but because I believe it to be useful to spell out the implications of, what could be called, Dummett's semantic turn in metaphysics. It is only when we have a complete overview of the tasks it sets that we can begin to consider the possibility of extending Dummett's philosophical intuitions into a well-articulated philosophical method. If possible, it certainly will not be easy to accomplish this. The task can be compared with that which faced Kant, when he came to the conclusion that it is impossible to find a place in the framework current in his days for the notion of knowledge as ordinarily understood.[20] This brought him to revise the metaphysical framework by reinterpreting the notion of an object solely in terms of what it is to have representations and experience.

For Dummett it is not knowledge, but meaning that is problematic. His plea for a semantic turn is motivated for a large part by his belief that it is impossible to give a satisfactory account of meaning within a traditional metaphysical framework. This is a recurrent theme in the writings of Dummett: we cannot give an adequate account of meaning as long as we characterize entities as being independent of the meanings of the expressions we use in talking about those entities.

I will not here pursue any further the comparison between Kant's critical writings and what are the, up to now, scarcely recognized pillars of Dummett's semantic turn in metaphysics.[21] Instead, I wish to

[20] Cf. Bencivenga 1987, p. 48.

[21] I do believe that such a comparison could be rather fruitful. Reading Dummett one frequently encounters not only problems that have a distinct critical ring, but also the way of solving these problems that Dummett elaborates on, or hints at, do remind one of Kant's work. It would be interesting to see how far one could go in understanding Dummett's position as an attempt to reconstruct Kant's critical philosophy on the basis of Frege's philosophy of language. Kant, for example, sought to delineate those concepts of the understanding that have no empirical contents whatsoever, the so called *reine Verstandsbegriffe*. Dummett, on the other hand, seeks to locate those logical rules that have no extra-semantical content. These rules are referred to as being assumptionless, or as being *purely logical*. (See Dummett 1991b, p. 272, where Dummett stresses the importance of an adequate conception of the purely logical.) In order to determine the purely logical rules Dummett appeals to the idea of a proof-theoretic justification procedure. This, in its turn, might be compared with Kant's idea of a transcendental analysis. Furthermore, since the theory of meaning reflects an insight into the limits of our understanding Dummett's investigations can be called critical, in precisely the same way as Kant's studies are. The difference being that where Kant sought to explicate the limits of knowledge, Dummett seeks to determine the limits of understanding.

return to the original question: How is it possible to criticize generally accepted logical laws?

2.4 The Notion of a Semantic Theory

The basic reason for criticizing a generally accepted logical law is that it cannot be justified properly. But, what counts as a *proper* justification? The present subsection outlines the answer that a semantic revisionist is about to give. We start with the assumption that in framing the notion of validity we need a *semantic theory*, that is, 'a theory that displays the mechanism by which a sentence is determined as true or otherwise in accordance with its composition.'[22] This is not to say that a semantic theory specifies the truth-conditions of sentences, but, more generally, that feature of a sentence on which its truth-value depends. That could be a truth-condition, but also, for example, a classification of constructions into those that do, and those that do not prove the sentence. So, more precisely, a semantic theory is required to 'exhibit the way in which the semantic value of a sentence is determined by the semantic values of its components, and to give the general condition for a sentence to be true, in terms of its *semantic value*.'[23]

The notion of a semantic value is closely related to Frege's notion of reference. Frege explained it mainly in terms of the relation which a proper name has to its bearer. But, according to Dummett, the notion of reference so interpreted is far too specific to contribute to a description of the general form that a meaning-theory should take. To distinguish Frege's particular interpretation from the functional aspect of the notion of reference in the context of a meaning-theory in general, Dummett introduces the notion of the *semantic roles* of expressions. These are given by 'the properties we must ascribe to constituent expressions of each given type in order to determine each sentence as being true or otherwise.'[24] Semantic theories that serve the purposes of logic normally state the semantic role of each expression as 'its having something of an appropriate type associated with it.'[25] That 'something' is called the *semantic value* of an expression.

[22]For example, Dummett 1991b, p. 61.
[23]*ibid.* My italics.
[24]Dummett 1981, p. 149. Slightly paraphrased.
[25]*ibid.*

So far for the clarification of semantic value, or reference, as I will continue to use both notions interchangeably. A theory of reference, then, stipulates an association between expressions of the language and their semantic values. A theory of reference, or semantic theory, is needed to frame the notion of validity. According to Dummett, the reason for this is that 'we know in advance that what is required of a form of argument, for it to be valid, that it be truth preserving, that it carry true premises into a true conclusion.'[26] Hence, the one notion that really matters to logic is that of truth under an interpretation, and it is precisely this notion that is given shape by a semantic theory.

The selection of the semantic theory is all decisive, since 'it is this choice which has repercussions both on which logical laws we should hold to be valid and on which metaphysical views we ought to favour.'[27] Since a semantic theory uniquely determines a logical theory, it follows that the latter is acceptable if the former is.[28]

A semantic theory is acceptable —we now come to the crucial thesis— if it can be extended to a 'coherent and workable meaning-theory.'[29] This criterion derives directly from the definition of a semantic theory. Hence, a demonstration that some semantic theory cannot be so extended boils down to a demonstration that the theory is not a semantic theory at all. So, a logical theory is acceptable if it can be validated by a semantic theory.

The question as to the criterion for the correctness of a logical rule is thus reduced to the question as to our conception of a semantic

[26]Dummett 1991b, pp. 20, 40. It is rather doubtful that we indeed know this in advance. Recent developments (dynamic semantics, update semantics) have brought forward semantic characterizations of validity that do not take truth as their central notion. In the light of this, the burden is on the anti-realist to point out why he knows in advance that these characterizations are mistaken. I will not go into this matter.

[27]Dummett 1991b, p. 148.

[28]In the next chapter we will see that the converse does not hold, that is, a logic does not uniquely determine a semantic theory. For semantic revisionism this has the immediate and undesired consequence that we cannot determine the model of meaning that underlies a metaphysical doctrine by characterizing it in terms of the canons of deductive inference that it licenses to be valid. Hence, if a semantic revisionist is to retain his revisionist pretensions with respect to metaphysics, he needs to review critically his semantic reduction of metaphysical doctrines as expounded in section 2.2. I will return to this matter at the end of the next chapter.

[29]Dummett 1991b, p. 303.

theory, which, in turn, is said to depend on our conception of a workable meaning-theory.[30] A semantic theory is a part, or, as Dummett used to call it, the core of a meaning-theory,[31] but, clearly, both do not coincide. We will now turn to the questions that are left unanswered by a semantic theory.

A meaning-theory, as Dummett conceives of it, aims at a systematic representation of a speaker's linguistic knowledge.[32] Its function is to present 'an analysis of the complex skill which constitutes mastery of a language.'[33] This practical ability is analyzed by describing what someone who speaks the language may be said to know. In this sense, a meaning-theory offers a theoretical representation of a practical ability.[34] It is not intended as a psychological hypothesis, though. No doubt there will be certain psychological mechanisms without which we, as human beings, would not have the ability to speak a language. But a meaning-theory does not describe these. The correct meaning-theory will be such that knowledge of it can be ascribed to anyone, or anything —human being, robot, or martian— who could be said to speak the language, independently of their internal mechanisms.[35]

Being an account of what someone has to *know* if he is to be a competent speaker of the language, we require of a meaning-theory that the account it provides enables us to explain what the *objectivity* of that knowledge consists in. A meaning-theory describes what a speaker knows if he knows the meaning of an expression, that is, if he understands it. But the account it provides will not be satisfactory if it fails

[30]Note that, eventually, the question as to standards for correctness of logical rule is modified into a kind of conceptual analysis. I have found this quite typical for the semantic revisionism endorsed by Dummett.

[31]Dummett 1975.

[32]Dummett 1991b, p. 103.

[33]Dummett 1976, p. 70.

[34]In Dummett 1993, pp. 159–60. this is explained as follows: 'the idea was that we can describe the articulation of this highly complex practical ability —the ability to speak a particular language— by representing it as a possible object of propositional knowledge, while acknowledging that it is in fact not propositional or theoretical, but practical knowledge.' But, Dummett admits, as an explanation of what it is to *understand* a language, this will not do, because the ability to speak a language is not a straightforward practical ability, for it is necessary to have it in order to know what it is to do. This is to be taken as a qualification on the thesis he endorsed before, the thesis that a meaning-theory is a theory of understanding. (Cf. Dummett 1975, p. 99.)

[35]*ibid.*

to explain how people who understand the language can communicate with each other by means of it. Hence, the account of what a subject's understanding of an expression consists in should be supplemented by an account of how mutual understanding is possible.

3 Semantic Revisionism and Objectivity

3.1 Accessibility and Comparability

In communication, Frege believed, people convey thoughts. They do this by bringing about changes in the external world which are meant to be perceived by others, so as to give them a chance to grasp a thought and to take it to be true.[36] Clearly, Frege makes a distinction between mutual understanding, which obtains whenever the hearer associates the same thought with the sentence uttered, as the speaker does,[37] and the recognition of the truth-value of that thought.

To make someone take the thought expressed by a given sentence as true or otherwise, is not the only possible goal of linguistic communication. We can also use language to ask a question, or to give a command. So, for communication to be successful it is not enough that the hearer grasps the thought expressed by the speaker by uttering a sentence: the hearer must also 'grasp' the way in which a particular expression is used. This aspect of meaning is called the force (*Kraft*) of a linguistic utterance. For our purposes, however, we need not go into this feature of Frege's philosophy of language, for it is only by considering the possibility of mutual understanding —in the restricted sense of grasping the same thought— that Frege develops his conception of objectivity.

Frege never actually doubted that speakers of the same language can, and on many occasions do understand each other. To explain this we have to perform two tasks.[38] In the first place, we need to explicate what, in general, it is for two people to understand each other. For Frege this raises the question how it is possible that the same thought can be grasped by different people. To avoid psychologism, Frege answers this question this by analyzing the concept of thought itself, and not the psychological mechanism that accompanies the grasping of a thought.

[36] Frege 1918, p. 77.
[37] *ibid.*, p. 65.
[38] Taylor 1992, p. 65.

The question thus becomes: What features do we have to ascribe to thoughts so that different people, at different spatio-temporal locations can grasp the same thought? I will refer to the requirement that people can grasp the same thought as the *accessibility condition*.

It is necessary, but not sufficient. Any account that meets only this condition is able to explain how communication is possible. But that it is possible does not imply that it is actual. You and I may be able to grasp the same thoughts, but, as far as communication by means of language is concerned, what is of equal importance is whether you grasp the same thought as I do, when I utter a particular sentence. So, the second task to be performed is to explain how it is possible to determine whether or not this is indeed the case.

This point can be illustrated by drawing on a kind of argument that is due to Putnam.[39] Imagine two speakers, one of which uses 'skating' to refer to skating, while the other uses 'skating' to refer to swimming in ice-cold water, and 'swimming in ice-cold water' to refer to skating. Both use 'warm' to mean warm. The one says: 'I always have difficulties in staying warm when I am skating.' The other replies: 'I know what you mean, I have the same problem.' Notwithstanding what the second speaker believes, it is clear to us that he does not know what the other means. That is to say, he did not grasp the very same thought the first speaker intended to express by uttering the sentence 'I always have difficulties in staying warm when I am skating.' It is not that he cannot grasp that particular thought, for he would have grasped it, if the other had said 'I always have difficulties in staying warm when I am swimming in ice-cold water.' But in that case too, he would have grasped a thought different from the one the speaker intended to express. So, again, they would not have understood each other, even though, in that particular situation neither one would have any reason to believe that they did not.

If accessibility is not sufficient for mutual understanding, then what is? Think of someone speaking Urdu to me. Since I do not know that language, I will not understand a word of what he is saying, which, in the present context, says that I do not know what thoughts he intends to express by his utterances. But exactly the same can happen when someone speaks English to me, since, as we saw, we must reckon with the possibility that the speaker systematically associates different

[39] *Locus classicus*: Putnam 1980b.

thoughts with the sentences he utters, than I do. Nevertheless, I would say in the latter, but not in the former case that the speaker and I speak the same language. It follows that attaching the same senses to the same expressions is not necessary for speaking the same language.[40]

It is difficult, if not impossible, to delineate sharply the border between speaking the same language and speaking a different one. But that, I think, is not a reason for rejecting this distinction, for at the extremes we find significant and appealing differences. No speaker of a language, not even if it is his mother-tongue, is expected to understand every possible sentence of that language. Here I use 'understanding' not as meaning 'attaching the intended thought to the sentence uttered', but more basically as 'attaching one thought or other to the sentence uttered.' Confronted with an utterance of which one does not know, right at the spot, what thought to associate with, the situation appears to be very similar to one in which someone addresses you in Urdu. But the essential difference is that, if I speak the same language as my conversant does, I have the means of coming to know the thought he intended to express. Being spoken to in Urdu, in contrast, I just would not know what to do.

The means of finding out what someone else means by what he says comprise more than just the ability to engage in conversation with each other. Take the 'skating-swimming' example mentioned above. Imagine that the knowledge of English of our speakers does not go beyond the expressions that are quoted. Then how could they possibly find out that contrary to their belief, they do not understand each other? Not by just speaking to each other. That much is clear. But now consider the following situation. Having agreed to go skating one of our speakers has put on several layers of clothing, while the other has put on only a thin tracksuit. Once on the ice the first speaker wonders: had not the other said to have difficulties in staying warm when he was skating? Asked for it, the other replies: 'Yes, indeed, but I am not skating now.' We see that in many cases two speakers may come to the conclusion that they did not understand each other in the way they thought they

[40] Or, if we want to hold on to an explanation of what it is to speak the same language in terms of associating the same sense to the same expressions, we would have to conclude that there is not an essential difference between the Urdu and the English case: in both cases the disputants speak different languages. Generalizing this, we would probably come to the conclusion that the notion of a common language is quite empty, and that, really, there are only idiolects.

did by means of what they *do* under specific circumstances.

When two people speak the same language, I would say, both must be able, on the basis of what the other says, to form expectations about what the other will do under which circumstances. But this is not enough. Suppose that someone utters a sentence in Urdu to me, and suppose that it means ‘According to the tradition of my family we use several fine herbs for baking bread’, while I do not know this, then what would stop me from forming any expectation I want to? Nothing. But still, we do not speak the same language. Why not? Because I lack the means to express how the expectations I have concerning the way he will act under specific circumstances are connected to the utterances he has made. If, on the basis of his utterance, I believe that he cannot swim, then I will expect him not to jump into the water voluntarily, and it will frighten me to see him doing just that. And I will be unable to parry his subsequent amazement by reference to what he has said. This is not just because I cannot reproduce the utterance. For even if I can do that, it will only increase his amazement to hear me explaining my fear by saying: ‘According to the tradition of my family we use several fine herbs for baking bread’.

It is not necessary for two people to speak the same language that they have the same understanding, while conversely, two people will not be able to come to understand each other if they do not speak the same language.[41] Two people have the same understanding if they grasp the same thought, but to speak the same language they must be able to compare the thoughts that they attach to the sentences of their language.[42] This latter requirement will be called the *comparability condition.* If it is not possible to compare the thought associated with some sentence by some speaker with the thought that another speaker associates with the same sentence, it could never be claimed for sure that both speakers actually understand each other, even if it is conceded that, at least in principle, they can grasp the same thoughts.

[41] I concentrate here on mutual understanding through the use of language, ignoring, but not denying, the possibility that people may come to understand each other in other ways.

[42] Cf. Dummett 1992a, p. 584–5, where it is claimed that ‘differences in sense must be detectable and resoluble.’ Frege once remarked that variations of sense may be tolerated ‘as long as the thing meant remains the same.’ (Frege 1892, note to p. 27.) This suggests that the comparability of sense presupposes the comparability of reference.

It follows that any account of mutual understanding necessarily has to meet the comparability condition. Both conditions together are sufficient. Since mutual understanding is specified as grasping the same thought, it follows that both conditions should be taken as constraints on any account of thoughts, and as such as constituting our conception of what is for a thought to be objective.[43]

Frege nowhere explicitly makes the distinction made here between the general criteria that any explanation of meaning has to meet if it is to account for the objectivity of meaning, and the explanation of meaning itself as constrained by these conditions. However, when he does address the question as to the objectivity of meaning directly, or indirectly via a rejection of a subjectivist account of meaning, he almost always mentions the requirement of being publicly accessible[44] as well as that of being comparable,[45] and this indicates that Frege was well aware of the importance of these two conditions.[46]

3.2 Frege's Analysis of Thoughts

Does Frege's conception of thoughts satisfy both the accessibility and the comparability condition? With a view to the former Frege distinguishes thoughts from subjective impressions, while he seeks to meet the latter by distinguishing what is objective and real (*wirklich*) from what is objective but not real.

A close reading of Frege's remarks about impressions reveals three distinguishing properties. The first of these is that impressions cannot be perceived; 'I see a green field, I thus have a visual impression of the green. I have it, but I do not see it.'[47] We cannot be acquainted with

[43]Why is accessibility not enough to establish the objectivity of meaning? Because what is at stake is not the objectivity of meaning conceived of in isolation of particular languages: what is at stake is the objectivity of the *understanding I have* of a particular expression.

[44]Frege 1892, p. 29–30, note p. 32, Frege 1894, p. 317, Frege 1895, p. 74, Frege 1918, p. 69.

[45]Frege 1892, p. 30, Frege 1894, p. 317, Frege 1918, p. 67. It should be noted, though, that in earlier works like Frege 1884 objectivity is discussed almost exclusively in terms of 'independence'.

[46]Dummett, for example, also holds these two conditions to be fundamental to Frege's conception of objectivity. See Dummett 1992a, p. 154, and Dummett 1976, p. 135.

[47]Frege 1918, p. 67.

impressions otherwise than by *having* them. And to say that I have an impression is the same as saying that it belongs to the content of my consciousness at a particular moment. Because of this, and given that a consciousness is always somebody's consciousness, it follows that impressions are completely dependent on the existence of consciousness beings, or, citing Frege, 'a sensation [i.e., impression] is impossible without a sentient being.'[48] Impressions thus need a bearer, without which they cannot exist. This is the second distinguishing characteristic.

The relation between an impression and its bearer is such that the former cannot exist without the latter. This can mean that an impression can exist as long as it belongs to the content of somebody's consciousness, but not necessarily the same consciousness at all times. But we can also read it as saying that an impression exists as long as it belongs to the consciousness of one and only one particular bearer. Frege clearly has the latter reading in mind, so as a third distinguishing feature we record that impressions belong essentially to the content of a particular consciousness: 'It is so much the essence of any one of my ideas to be a content of my consciousness, that any idea someone else has is, just as such, different from mine.'[49] It follows that I cannot have the impressions that somebody else has, and given also that sense impressions are completely dependent on the existence of consciousness, it follows that the impressions I have depend for their existence on the existence of my consciousness. Since I take my consciousness to be of finite existence, this conclusively establishes the finiteness of impressions. Notice that the second characteristic implies nothing more than that impressions are finite if the existence of consciousness is finite (not the existence of individual consciousness beings, but the existence of consciousness as such).

Thoughts are publicly accessible, because they are not subjective: they do not need a bearer, and do not belong essentially to the content of a particular consciousness. Thoughts are, in one word, completely independent of the subjects that have access to them. This distin-

---

[48]Frege 1918, p. 67.

[49]*ibid.* For my exposition of Frege's analysis of impressions I roughly followed Carl 1994. Carl, however, claims that the second characteristic is implied by the first, and the third by the second (*opus cit.*, p. 81). I cannot agree with this. From my exposition it will be clear that I take the third feature not as a consequence of the second, but as a restriction on it. I also fail to see how the second could follow from the first.

guishes thoughts from sense impressions, but it does not yet discern them from material objects, since these are thought of as being independent in the same way. For Frege this means that thoughts as well as material objects are objective, but not that thoughts are material objects. Hence, he distinguishes material objects that are objective and real from thoughts that are objective and nonreal. This distinction is based on differences in the conditions that have to be satisfied for us to have knowledge of either kind of entities. We cannot have knowledge of what is real otherwise than by means of sense impressions. However, perceptual knowledge involves more than having sense impressions alone:

> But do I not then see that this flower has five petals? We can say this, but if we do, the word 'see' is not being used in the sense of having a mere visual experience: what we mean by it is bound up with thinking and judging. Newton did not discover the laws of gravitation because his senses were especially acute.[50]

Any knowledge of what is real requires sense impressions, but also something else: thinking a thought. It follows that a thought cannot be real, because, otherwise, our knowledge of a thought would require sense impressions, but also something else. Hence, to assume that a thought is real would make the notion of a thought redundant, or it would lead us into an infinite regress.

Summing up we find that Frege distinguishes the subjective from the objective in terms of the notion of dependence, while the objective is differentiated in terms of the epistemological relation in which we may stand to objective entities: material objects are perceived, while thoughts are *grasped*.

Digression: Objectivity in Time

Contrary to what Frege appears to hold,[51] the analysis of thoughts by contrasting them with impressions and physical things by itself does *not* support the claim that thoughts are timeless and immutable. Frege finds no difficulties in the claim that different persons may perceive the

[50] Frege 1969, p. 149.
[51] See, for instance, Frege 1918, p. 76.

same physical object.[52] The realm of the physical is discerned from that of thoughts by being real. According to Frege's terminology that means that physical objects are part of a causal order, and thus that they have an origin and are not timeless.[53] Hence, on Frege's own account, timelessness is not a necessary condition for being publicly accessible. What is necessary is that thoughts do not have a mental origin, not that they have no origin at all.

I stress this point so that we avoid the mistake of believing that a difference of opinion regarding the timelessness of meanings is also a difference of opinion regarding the objectivity of meaning. It is not, or, at least, it need not be: one can accept Frege's views on the objectivity of meaning, without accepting the claim that thoughts are timeless. The timelessness of meaning has thus to be argued for independently of arguments establishing the objectivity of meaning.[54]

Frege nowhere explicitly argues for the timelessness of thoughts. Perhaps he just could not think of it to be otherwise, a prejudgment that is more understandable, when one takes into account Frege's preoccupation with mathematical statements, which, if true, he held to be 'timelessly true, true independently of whether anyone takes them to be true.'[55] But when he discusses examples like the Pythagorean theorem, Frege seems more concerned with the status of their truth-value, than with the status of the thoughts expressed by them. And it is quite conceivable that the claim that thoughts are timeless is not a premise, but a consequence of the claim that mathematical statements, if true, are eternally so, in conjunction with the claim that thoughts are bearers of truth-values.

What the case may be, it is indisputable that Frege took the timelessness of thoughts as a starting point and, through analysis, tried to

[52]We cannot conclude from this that perceptions of physical objects meet the two conditions. Perception is not, like 'grasping' some extraordinary capacity. In contrast to the former, the latter does enable one to acquire knowledge without mediation of the senses. Hence, knowledge of the external world has to be extracted from sense impressions. To guarantee the objectivity of knowledge of the external world, Frege appeals to 'something non-sensible'. It is unclear what purpose Frege's conception of an object serves.

[53]Frege 1918, p. 76.

[54]These questions are not totally unconnected. At least this seems certain: if one does not accept the objectivity of meaning, there is hardly any point in arguing for its being timeless.

[55]Frege 1918, p. 69.

establish that apparently deviating phenomena in natural language like tense and indexicality could nevertheless be explained in terms of eternal thoughts. In more recent semantic theories, however, precisely this aspect of Frege's theory has been disputed. In Perry 1979, for example, it is argued that, in many cases we cannot dispense with indexicality if we want to explain belief. And in Barwise and Perry 1983 we even find that indexicality is taken as a basic feature of meaningfulness. It goes without saying that in this approach meaning is not explained by reference to eternal, unchanging entities, but in terms of constructions made out of physical, hence, transient things.

Grasping a Thought

In order to explain the possibility of mutual understanding we have to account for the objectivity of a subject's knowledge of meaning. For Frege the objectivity of this knowledge is based on both the independence of the object of that knowledge, and the epistemological relation in which we stand towards these objects. It is clear that the former accounts for the accessibility of meanings. However, the bare fact that the entities known are publicly accessible does not imply that a subject's knowledge of such an entity is also publicly accessible. And it goes without saying that if we are to compare the understandings different speakers have, these understandings must be publicly accessible. It is therefore that we need an answer to the question: What is it to grasp a thought? For it is the notion of 'to grasp' that must bring us from the accessibility of thoughts to the accessibility of a subject's understanding of a statement.

A brief discussion of perception will help us to get some grip on the problems involved. The assumption that material objects are independent, that is, that they do not cease to exist as soon as they are no longer perceived is normally taken as a necessary ingredient of an account of the objectivity of perceptual knowledge. The assumption is appealing, for if we would take the object known to be a constitutent of the knowledge a subject has, it would be impossible to overcome the subjective character of this knowledge. On the other hand, it must be conceded that although we cannot be wrong about the perceptual knowledge we have, we can be wrong about the object perceived. Hence, if we are to explain perceptual error, we cannot take perceptual knowledge to be given by some direct perception of material objects. Instead we must stipulate that we have perceptual knowledge of material objects via something else, sense impressions, for example. These, however, are essentially private and the difficulty is to explain how the knowledge we acquire via these private impressions is, nevertheless, objective. And it is just not enough to say that the objects perceived exist independently of the acts of perceiving.

Frege himself suggests that there is some parallel between perceptual knowledge and knowledge of thoughts.[56] The problem of grasping a thought, Frege argues, is not more complicated than to explain how we can perceive a physical object. In perceiving a thing, all we can be sure of is that we have visual impressions. But these belong to the individual and cannot be identified with the thing perceived. So, Frege concludes, for the external world to open up for us, something must be added, and this something must be non-sensible. In the same way, Frege suggests, 'something non-sensible, even without the co-operation of sense impressions, could also lead us out of the inner world and enable us to grasp thoughts.'[57] We already saw that Frege suggests that in the case of perception this 'non-sensible something' is a thought.[58] But if so, it is quite mysterious what this something could be in the case of grasping a thought. If knowledge of material objects requires sense impressions, as well as thinking a thought, it follows that the objectivity of perceptual knowledge depends on the objectivity of understanding. And that makes an account of what it is to grasp a thought all the more urgent.

What troubles the explanation of the objectivity of perceptual knowledge is that, although the real is independent of our psychological constitution, our knowledge of the real is not. This transposes to knowledge of thoughts. That sense impressions are not needed for us in order to grasp a thought is not to say that psychological mechanisms do not play any role at all. Frege appeared to have been led by the belief that the need to take into account psychological considerations originates from the failure to differentiate the real from what is not real. But he was wrong. That need is induced by the assumption that the objectiv-

[56]As the following discussion will show, the comparison of 'grasping' with 'perceiving' is most unfortunate from Frege's point of view. This has also been acknowledged by Notturno 1985, p. 68, and Haack 1978, p. 239-40.

[57]Frege 1918, p. 75.

[58]See the quotation on page 28.

ity of knowledge can be established via the objectivity of the objects known, in conjunction with the belief that objectivity coincides with independent existence.

Putnam comes to a similar conclusion when he remarks that grasping a Fregean thought is still an individual psychological act, and to know the meaning of a word 'just a matter of being in a psychological state'.[59] However, from this alone one cannot conclude that on Frege's account a subject's understanding is inaccessible, and, hence, that the account ends in a form of psychologism it was designed to refute. That this conclusion, nevertheless, appears as being inevitable is due to Frege's rather strict conception of a psychological state, one that Putnam associates with *methodological solipsism*.[60] Methodological solipsism is characterised by the assumption that we would not call anything a psychological state if it presupposes the existence of any individual other than the subject to whom that state is ascribed. As Putnam observes, 'Frege's argument against psychologism is only an argument against identifying concepts with mental particulars, not with mental entities in general.'[61] But could we not, as an alternative resort to a less strict conception of a psychological state? Putnam suggests a weaker form of mentalism which differentiates between mental states as tokens and mental states as types. This strategy will leave intact the subjective character of impressions (they are tokens), while circumventing the question as to the accessibility of thoughts, by allowing tokens to be identical at the level of types. We will then say that someone grasps the sense S of expression E if he is in the psychological state of 'knowing the sense of E to be S', and we explain that John and Mark associate the same sense S with E if both are in the state of 'knowing the sense of E to be S.'

The merit of this approach is that it is compatible with the Fregean framework in almost every respect: we assume independently existing thoughts, which are strictly distinguished from grasping subjects. We also acknowledge that to grasp a thought is to be in a psychological state which is utterly private. The only difference being that, in addition, we assume that individual states are of a certain type which can be characterized, partly at least, in terms of a relation between senses

[59] Putnam 1980a, p. 218.
[60] *ibid.*, p. 220.
[61] *ibid.*, p. 222.

and expressions, both of which are existing independently of the grasping subject. The rejection of Frege's methodological solipsim has the desired consequences: we can now effortlessly meet both the accessibility and the comparability condition. But, as Putnam points out, it also has a serious drawback. By a simple argument we can now establish that two subjects who are in the same state necessarily associate the same senses with the same expressions. In conjuction with the Fregean assumption that sense determines reference, it follows that two subjects in the same psychological state necessarily assign the same references to the same expressions. But this, Putnam argues, is wrong at least as far as natural language is concerned.

In one version his argument runs as follows.[62] Suppose that, unless one has some very specialized knowledge, one cannot tell the difference between molybdenum pots and pans and aluminum pots and pans. Imagine, furthermore, a world that is like ours in all respects except this one: in that world, call it Twin Earth, whatever is made of aluminum in our world is made of molybdenum, and whatever is made of molybdenum in our world is made of aluminum on Twin Earth. The English speaking inhabitants of Twin Earth, furthermore, use 'aluminum' as the name of molybdenum and 'molybdenum' as the name of aluminum. If, now, we take two average speakers of English, one inhabiting Earth, and one inhabiting Twin Earth, such that neither one has the knowledge necessary to be able to distinguish aluminum from molybdenum, then, Putnam claims, there may be no difference at all in their psychological states when they use the expression 'aluminum pan'. Yet, the extension of this expression is different in both worlds: on Earth the extension of 'aluminum pan' is the set of aluminum pans, while on Twin Earth it is the set of molybdenum pans. So, Putnam concludes, the psychological state of the speaker, even if we conceive of it as a type instead of as a token, does not determine the referent of an expression.

I end this paragraph with the following summarizing conclusion. If, like Frege, we embrace methodological solipsim there is little hope that the stipulation of a third realm will enable us to escape subjectivist scepticism. If, on the other hand, we take a psychological state in its wider sense, we become vulnerable to Putnam's argument, which threatens the Fregean conception of sense and reference. For a semantic

[62]Putnam 1984, p. 123. Cf. Putnam 1980a.

revisionist this states a dilemma, for not only does he require of any meaning-theory that it accounts for the objectivity of understanding, but also that it takes Frege's distinction between sense and reference as a starting point.[63] In the next section we will see how, in order to overcome this dilemma, a semantic revisionist takes issue with its first horn.

4 Understanding and Use

The notion of a third realm that Frege introduces in the course of his analysis of sense is quite commonly interpreted as being akin to the platonist conception of a realm of ideas and forms. But, contrary to what one is inclined to believe, there is no established consensus on what the correct interpretation of Frege's philosophy is. Sluga, for example, has quite persistently argued against Dummett's conception of Frege as a platonist, by pointing out, what he takes to be, strong Kantian influences in Frege's writings. I have no intentions to take sides in this debate.[64] I only want to remark that either standpoint may taken to support my characterization of Frege's account of objectivity as being object-oriented, since the disagreement primarily concerns Frege's conception of an object, and not the general form of his account of objectivity. From my point of view, the interesting question is not which of the two interpretations presents the most plausible exegesis of Frege's writings, but, firstly, what interpretation of the objectivity of

[63]See §§ 2.4 and 4.1 of this chapter.

[64]Although, admittedly, I believe that a good case can be made for a non-platonistic interpretation. Frege explicitly states that by 'objectivity' he does not understand something independent from reason. According to Dummett this should be read as saying that the objective cannot be apprehended save by, or by reference to, rational thought, and not as saying that its existence depends upon our thinking about it. (Dummett 1991c, p. 125.) But Dummett seems to overlook another possible interpretation according to which the objective is such that if we were to think about it, we could not think about it otherwise. What distinguishes this interpretation from the one that Dummett provides is that it eliminates all reference to the existence of the objects *an sich*. This certainly seems to cohere with Frege's remarks that, contrary to first thoughts, we can only be certain of our subjective experiences, and that, furthermore, we can only enter the outer world if our sense impressions (Vorstellungen) are accompanied with something non-sensible, i.e., thoughts. This implies that we may never be sure that the way we know the world does coincide with the way the world really is, that is, that all we know is how the world is *für sich*.

understanding would be yielded by either reconstruction, and, secondly, what their respective merits and demerits are. The limitations set by this chapter preclude me from paying any more attention to this.

To concentrate on the ontological status of thoughts is a natural thing to do once we accept Frege's conception of thoughts as being logically independent from the means we have for expressing them. Frege believed that it is possible, at least in principle, to grasp a thought otherwise than as expressed linguistically, although he conceded that for us, human beings, such may not be possible.[65] The problem we encountered is, to speak with Dummett, that 'Frege's account of sense does not show *how* that is possible, that is, how a sense can be grasped otherwise than as the sense *of* an expression.'[66] In the preceding subsection we found that there are basically two ways of extending Frege's analysis of sense to an explanation of understanding. One extension fails to be a satisfactory account of the objectivity of understanding, while the other makes it depend, ultimately, not on the features of what is understood, but on the features of our way of knowing it.[67] As an account of the objectivity of understanding the first one cannot be accepted in general, while from the point of view of semantic revisionism the second one is also unacceptable.

It thus appears that it is *impossible* to account for the objectivity of understanding in a way that is acceptable for a semantic revisionist, if we assume thoughts to be independent from the means we have for expressing them. So, the proper conclusion to draw for a semantic revisionist is not that we should ground the independence of thoughts differently, but that we should give up that independence itself and take seriously the fact that senses are *of* expressions.

As I understand Dummett's writings, the slogan 'meaning is use' is meant to call on to precisely this shift of perspective.[68] It urges one to conceive of a thought as a '*feature* of the utterance, accruing to it in consequence of a variety of facts: the sentence of which the utterance is a token; the context of the utterance, and the language considered

[65] Frege 1969, p. 288.

[66] Dummett 1993, pp. 10, 25. Cf. Dummett 1991a, p. 260 ff.

[67] Since understanding is a cognitive notion, I do not distinguish Putnam's solution from a neo-Kantian one.

[68] That the slogan 'meaning is use' is not only programmatic will be the subject of chapter III.

as a conventional practice.'[69]

The appeal to use is not completely alien to Frege's ideas, for Frege never denied the importance of the use of language for human beings to grasp thoughts.[70] But the question remains how this can clarify what it is to grasp a thought. At one place, Frege explicitly states that he believes it to be necessary for us that a thought of which we are conscious is connected with some sentence or other.[71] What this suggests, in effect, is that thinking (grasping a thought) coincides with the ability to use the sentence connected with that thought. As a consequence we stipulate that our ability to grasp thoughts can be explained by explaining our ability to use the sentences of our language.

The appeal to the dictum 'meaning is use' is thus seen as being inspired by Frege's writings. But that is not the same as holding that

> Frege's thesis that sense is objective is (...) implicitly an anticipation (in respect of that aspect of meaning which constitutes sense) of Wittgenstein's doctrine that meaning is use (or of one of the family of doctrines so expressed)[72]

This, I believe, is to misconceive the import of the appeal to use. It is not merely an elaboration on what is already implicit in Frege's writings, but, rather, it comes down to a rejection of the pattern underlying his analysis of understanding, that is, it involves a dismissal of the object-oriented approach to objectivity, which rejection has its origin in the idea that

> the objectivity of sense is sufficiently guaranteed by its being expressed within the common language: it was not necessary for Frege, in order to safeguard that objectivity, to view it as having an existence independent of even the means of expressing it.[73]

In one sense the appeal to use is thus highly programmatic: it sets the task of accounting for the objectivity of understanding without assuming senses to be logically independent entities. We must now examine what consequences this has for the general form that a meaning-theory ought to take.

[69] Dummett 1991a, p. 261
[70] See, for example, Frege 1969, pp. 143, 189.
[71] Frege 1969, p. 288.
[72] Dummett 1976, p. 135. See also Dummett 1992b, p. 94.
[73] Dummett 1992a, p. 680. Cf. Dummett 1993, p. 25.

4.1 The Form of a Meaning-theory

From the point of view of semantic revisionism, Frege is to be criticised for holding that the account of what it is to grasp a sense belongs to epistemology and is not a part of a meaning-theory. We can state the same by saying that a semantic revisionist has a different conception of the task of a meaning-theory than Frege would have had: he requires of a meaning-theory not only that it says *what* is known by someone who speaks a language, but also that it explains what knowing this consists in. A semantic revisionist transforms the *epistemological* question what it is to grasp a sense into the *meaning-theoretical* question what it is to take some word as expressing that sense, as is illustrated by the following quotation:

> A grasp of the concept of a sheep may be being thought of in either two ways. Either it is thought of as prior to his understanding of the English word 'sheep', or it is thought of as attained precisely by gaining an understanding of that word. In the former case, the claim of the meaning-theory to be in any way explanatory rests on the possibility of giving an explanation of what it is to grasp a concept which is independent of taking some word to express that concept; for if we could explain what it is to grasp the concept expressed by the word 'sheep' only by stating what it is to take *some* word (not necessarily that one) as expressing that concept, we should not need any explanation of what it is for someone who already grasps that concept to associate it with the particular English word 'sheep'.[74]

The appeal to use, it was said, is motivated by the demand that we should take seriously the thought that senses are dependent on the means we have for expressing them. These means do not only comprise the words we have, but also, and perhaps especially, the ways in which these words are used. This naturally gives rise to the thought that a word expresses a particular sense, because it is used in a special way. We can thus identify the dictum 'meaning is use' as the stipulation that the way an expression is used constitutes its having the sense that it happens to have.[75] Note that, since the way a word is used is a purely accidental property of it, this is in complete accordance with

[74] Dummett 1991b, p. 111.

[75] This is quite reminiscent of Frege context-principle, according to which a word has meaning only in the context of a sentence, or, differently put, a word has meaning only if it is *used* in the construction of larger units.

the idea that senses do not intrinsically belong to the expressions that have them.

Assuming that the way a word is used constitutes the sense that it has, we can now say that someone grasps the sense of an expression if he knows how to use that expression. Placed in this light, not only our explanation of what it is to understand an expression changes, but we must also indicate what the accessibility and comparability of knowledge of meaning —the two conditions that constrain any account of the objectivity of a subject's understanding of an expression— consists in by reference to use.

If the appeal to use is to ensure the objectivity of understanding, it must be possible to attain knowledge of the meaning of an expression from use only. That is to say, this knowledge may not contain any feature that cannot be learned or recovered from the way it is used by others. This *recoverability* condition, as I will call it, is the shape the accessibility condition takes if meaning is to be explained in terms of use.

For knowledge of meaning to be comparable it must be possible, at least in principle, to determine whether or not two people have the same understanding of some expression. If the way people use an expression is to tell us anything about whether or not they have the same understanding of it, it has to be the case that, whatever it is they know when they understand an expression, this knowledge must manifest itself *fully* in the use made of the expression. For, if knowledge of meaning does not fully manifest itself in use, two people could use some expression in the same way, without understanding it in the same way. Under the dictum 'meaning is use' the comparability condition appears as, what I will call, the requirement of *manifestation*.

The appeal to the use of language in order to explain what possession of knowledge of meaning consists in may be successful, because, unlike other practical abilities, the ability to use language has as an essential ingredient the ability to appraise one's own and other's uses.[76] If this would not be the case, that is, if the possession of the ability to use language would not imply any ability of self-appraisal, nor any capacity to recognize if, or how, prevailing circumstances are apt for its exercise, it would not be possible to attribute knowledge of meaning on the basis of observable use of language. To hold that meaning is

[76] Cf. Wright 1993a, p. 247.

use is one thing, to say that knowledge of meaning can be attributed on the basis of observed use of language is quite another thing. The latter implies the possibility of knowledge ascription on the basis of observed use, the former does not. It is possible to hold that meaning is use, to observe that someone uses a certain expression correctly, without thereby being obliged to hold that he knows the meaning of that expression.

We have seen the reasons a semantic revisionist has for requiring of a meaning-theory that it specifies, not only what it is that someone knows who understands an expression, but also in what knowledge of the sense of an expression consists, i.e., what counts as a *manifestation* of that knowledge.[77] However, the resulting requirements of manifestation and recoverability constrain the theory of sense, while, as was pointed out earlier, the logic we ought to favour depends on the choice of a semantic theory, or theory of reference. Since we are interested in logical revisionism, we must now examine how constraints on the theory of sense have any bearings on the selection of a semantic theory.

How does the task of constructing a theory of sense fits in into the more comprehensive task of constructing a meaning-theory? Let me begin by pointing out that the anti-realist's criticism of Frege's conception of sense should not be taken as relating in any way to the general structure of Frege's philosophy of language. On the contrary, for, as Dummett puts it, 'we would not have the least idea how such a meaning-theory might be constructed, if we were not familiar with the distinction, introduced by Frege, between sense and force.'[78] Hence, the appeal to use should be seen as an attempt to remedy for what the anti-realist believes to be the shortcomings of Frege's theory without distorting the structure of the theory in general.[79] Precisely what this

[77]Dummett 1976, p. 70.

[78]*ibid.*, p. 72.

[79]A critical reader may feel disappointed, because the assumption that a meaning-theory should take the shape of Frege's meaning-theory is not in any way questioned here. Of course, he is right; this is a major assumption of Dummett's semantic revisionism. It is, for example, only in this context that a simple statement like 'To know the meaning of a sentence is to know what has to be the case for it to be true' is to be taken as determining a meaning-theory almost completely. It seems more than justified to ask why we should accept this claim as to the form of a meaning-theory. The reason that I do not touch upon it is that, in this chapter, I am concerned, not with the cogency of Dummettian semantic revisionism in general, but only with the role the notion of objectivity plays therein. And it is only against

structure is, need not concern us here. We confine us to the theory of reference, conceived of as a programme for a theory of sense.[80]

Earlier it was said that a semantic theory is required to exhibit the way in which the semantic value of a sentence is determined by the semantic values of its components, and to give the general condition for a sentence to be true, in terms of its semantic value. Thus conceived of, a semantic theory will yield a notion of truth under an interpretation. For the purposes of logic, this is quite sufficient,[81] but for the purposes of a meaning-theory it is not, since its task is to explain how we come to interpret some expression in one, rather than another way: a meaning-theory 'is concerned only with a single interpretation of language, the correct or intended one: so its fundamental notion is that of *truth* simpliciter.'[82] It is for this reason that the theory of reference has to be supplemented by a theory of sense, since the assignment of specific interpretations to the words of the language is accomplished entirely by the theory of sense.[83]

The intended interpretation of an expression is given by the sense that is associated with it. A meaning-theory is required to explain what constitutes this association. This cannot be done by specifying

---

the background of the assumptions concerning the structure of the theory that that role can be made visible.

However, not to remain completely silent on this point, let me remark that the form of a meaning-theory, as Dummett conceives of it, really is not a single assumption, but a cluster of claims, and could be criticized in equally many ways. This helps one to develop a systematic perspective on Dummett's writings: The distinction between sense and reference, for example, has been challenged by Kripke and Quine, among others. Hence, Dummett's extensive and critical discussions of these authors. Compositionality, to name something else, is also an important feature of a meaning-theory, but, as a principle, it is rejected by the position labeled 'holism'. No surprise then, that we find holism as a frequently recurring theme.

There are lots of other examples to give. But the point I want to make is that we cannot accuse Dummett of taking the form of the theory for granted, and that a substantial part of his writings can be read as arguing that a theory of meaning should take that particular form. Whether or not his arguments are convincing is a subject quite different from the one examined in this chapter.

[80]Dummett 1991b, p. 140.

[81]It is, because Dummett holds that 'we know in advance that what is required of a form of argument, for it to be valid, is that it be truth-preserving, that it carry true premises into a true conclusion.' (Dummett 1991b, p. 20. Cf. Dummett 1976, p. 115.)

[82]Dummett 1991b, p. 20. See also p. 61.

[83]*ibid.*, p. 86.

the semantic values that the expressions of the language actually have, because the semantic value of an expression is not an ingredient in its meaning: to know the semantic value of an expression is to know more than is needed for understanding it.

Now, it might be said that the specification of the reference *says* what the reference is to be, and, by saying it in a particular manner, *shows* what the sense is to be. And that, therefore, the task of constructing a theory of sense coincides with that of constructing a theory of reference. The objection is mistaken in believing that a meaning-theory is required to do no more than showing (to someone who understands the metalanguage in which it is formulated) what the senses of the words of the object-language are. But, as has been stated repeatedly, this is not enough: a mere statement of what the sense of a word is, does not in any way explain what constitutes taking the word as having that sense.[84]

Hence, the task of constructing a theory of sense does not consist only in giving a recursive characterization of what the sense of expressions of each logical type is to be. Recursive definitions of sense are known from intensional semantics. In Montague grammar, for example, the 'sense' of an expression is represented as a function from possible worlds to references (entities, truth-values, sets of entities, etc.). This, of course, is completely in line with the conception of the sense of an expression as the way in which its reference is given. But, nevertheless, from the current point of view, it fails as a theory of sense: it states *what* the sense of expressions of each logical type is to be, but it does not in any way explain *how* sense is associated with the expression that expresses it, that is, because, ultimately, it conceives of sense as being independent of the means we have for expressing it.

So far for considerations questioning the need for a theory of sense as distinguished clearly from a theory of reference. We now ask the opposite question: Can we not dispense with the theory of reference? No, because we need the notion of semantic value in order to explain that of sense. The sense of an expression, on Dummett's view, is 'something which, in view of how the world is, determines whatever we have taken to be the semantic value of the expression.'[85] This conceptual dependence of the notion of sense on that of reference makes that the

[84]Dummett 1991b, p. 149.
[85]Dummett 1981, p. 246.

theory of reference serves as a foundation of the theory of sense, and that it determines, within strict limits, the form that a theory of sense can take. The latter point can be illustrated by the following. Let us assume that the theory of reference stipulates an association between names and objects. The theory of sense must supply an answer to the question what constitutes the association between a given expression and its semantic value. If we, as Frege did, explain sense by reference to truth-conditions, this question becomes: 'What feature of our treating certain conditions as those required for the truth of sentences containing the name 'a' constitutes our taking 'a' to stand for the object u?'[86] Hence, the choices made in the semantic theory determine the task set for the theory of sense.

It is because a semantic theory constrains the form that a theory of sense based on it can take, that the former can be criticized on the basis of an evaluation of the latter. That is to say, a semantic theory will be subjected to criticism if it can be argued that no theory of sense having the form determined by the semantic theory in question can explain adequately what the intended interpretation consists in. A semantic theory is correct, on the other hand, if it can be extended with a theory of sense that specifies, for each expression of the language, what its meaning is. And to do that we not only need the notion of a semantic value, we also need to know what constitutes the association between a given expression and its reference.[87]

To use Dummett's terminology, a semantic theory is correct if it can be extended to a workable meaning-theory. A meaning-theory is *workable* if, firstly, it 'enable us to explain in a non-circular manner what a speaker's grasp of the sense of any expression is,' and, secondly, it 'accords with our practice, to the greatest extent that is possible.'[88] Ignoring the second one, we see that the first condition makes the workability of a meaning-theory, and thus, indirectly, the cogency of a semantic theory, dependent on the adequacy of the theory of sense. That it should not be circular here means that the explanation we are after should not invoke 'a capacity on the part of the speaker which cannot intelligibly be ascribed to him in advance of his knowing the language.'[89] It

[86] Dummett 1981, p. 158.
[87] *ibid.*
[88] Dummett 1991b, p. 340.
[89] *ibid.*, p. 341.

should not escape us that in this explanation quite explicitly a speaker's grasp of the sense of an expression —a piece of linguistic *knowledge*— is connected with a *capacity*. And the reason an anti-realist has for requiring that linguistic knowledge be equated with a practical ability derives from, as we saw, his dissatisfaction with Frege's account of the objectivity of understanding. We thus return to the claim we started out with, the claim that the anti-realist's conception of the form that a meaning-theory ought to take is, for a substantial part, determined by his ideas about how the objectivity of understanding is to be accounted for. How this supports a rejection of classical semantics will occupy us in the next subsection.

To end this subsection, let me remark that, as I see it, the merit of my reconstruction of the realism/anti-realism debate is, at least, that it provides a rationale for the requirement that knowledge of meaning must be capable of manifestation in a speaker's behaviour. In Dummett's writings I have searched in vain for an explicit defense of this demand. Prawitz has a more pronounced view on this matter. But he takes the manifestation, and recoverability conditions as assumptions that justify the appeal to the dictum 'meaning is use.'[90] The contrast between his views and mine could not be greater, since I take both conditions not as justifying, but as being justified by the appeal to use, which, in its turn, has shown to arise from a conception of how the objectivity of understanding is to be accounted for.

4.2 Truth-conditions and Understanding

A semantic theory according to which the semantic value of a sentence consists in its truth-value cannot be extended to a workable theory of meaning. That is the claim we will discuss in this subsection. It is argued that a truth-conditional theory of meaning cannot meet the requirement of recoverability, nor that of manifestability. Having established earlier that these are objectivity conditions, and given that a truth-conditional account is characteristic of a realist view on meaning, it follows that a realist cannot account for the objectivity of understanding.

A meaning-theory is required to account for the knowledge that a speaker of a language possesses. The constraint imposed on it is

[90]Prawitz 1977.

that the representation of what a speaker knows if he understands an expression is to be such that it enable us to connect that knowledge with what the speaker says and does. If recoverability and manifestation are to be met, the connection between them must be very strong: the knowledge a speaker possesses of the meaning of an expression should not contain any feature that could not be recovered from the way that expression is used by the community of speakers, while, at the same time, that use must fully express the knowledge that speakers have which use the expression in that way. The connection between linguistic knowledge and the ability to use the language meet these if only if we can *equate* knowledge and ability. That is, if it can be established that someone understands an expression if and only if he has a certain capacity.

Let us consider a meaning-theory that characterizes the meaning of a sentence in terms of its truth-conditions. On this view, knowledge of meaning will be explained as knowledge of truth-conditions: someone is said to know the meaning of a sentence, if he knows what has to be the case for it to be true.[91] To meet the objectivity conditions, we must now point out which practical ability we can equate a speaker's knowledge of truth-conditions with.

To state the truth conditions of some sentences, Dummett says, is just to state its semantic content in other words.[92] And there is no objection to ascribing knowledge of meaning to someone who has this *ability*. The knowledge a speaker has if he is able to state the truth conditions of a sentence is called *explicit* knowledge. But not all linguistic knowledge can be explicit: only the knowledge of the meanings of sentences introduced to a speaker 'by purely verbal explanations in the course of his progressive acquisition of the language' can.[93] Verbal explanations presuppose a prior understanding of the expressions in terms of which the explanation is given. A meaning-theory is required to account for the meanings of all expressions of the language. And it could not escape circularity if it assumed that the understanding of each expression of the language consisted in the ability to give a verbal explanation of it.

It follows that, if understanding consists in a knowledge of the truth

[91] *Locus classicus*: *Tractatus*, 4.024.
[92] Dummett 1976, p. 80.
[93] *ibid.*

conditions, at least some part of a speaker's linguistic knowledge must be *implicit*, that is, not verbalizable. And if it is not the ability to give a verbal explanation, i.e., to state the truth conditions, then with what practical ability is this kind of knowledge to be equated? As the only suitable candidate appears the capacity 'to evince recognition of the truth of a sentence when and only when the relevant condition is fulfilled.'[94] One may object that recognizing that the condition for the truth of a sentence obtains is not really doing anything. Or that one may fail to recognize some condition as obtaining due to psychological reasons (divertedness, inattentiveness, etc.), and that, hence, this cannot be taken as indicating that one has not the capacity to do so. We circumvent these problems by equating a speaker's knowledge of the truth condition of a sentence with his 'mastery of a procedure for deciding it, that is, his ability, under suitable prompting, to carry out the procedure and display, at the end of it, his recognition that the condition does, or does not, obtain.'[95]

Sentences for which there is such an effective procedure are called *decidable*. Unfortunately, natural language contains many sentences which are not effectively decidable. Examples given are: quantification over an infinite or unsurveyable domain (e.g. over all future times), subjunctive conditionals, or expressions explainable only by means of it and references to regions of space-time inaccessible to us in principle.[96] And for an *undecidable* sentence, we cannot equate a speaker's knowledge of its truth condition with his capacity to recognize whether or not this condition does in fact obtain. The reason for this is not that we will never be in a position to recognize that the conditionfor its truth is satisfied, or that it is not. The reason is that, by hypothesis,

> Either the condition is one which may obtain in some cases in which we are incapable of recognizing the fact, or it is one which may fail to obtain on some cases in which we are incapable of recognizing that fact, or both (...)[97]

In other words, an undecidable sentence is such that there is no guarantee that we recognize the fact that the condition for its truth obtains,

---

[94] Dummett 1976, p. 81.
[95] Cf. Dummett 1977, p. 374.
[96] Dummett 1976, p. 81.
[97] *ibid.*, p. 82.

or that it does not, every time we are in a position to do so. It follows that a knowledge of the meaning of a sentence cannot, in general, be equated with a capacity to recognize the conditions for the truth of the sentence. Given, furthermore, that there is only one way for a person to manifest his implicit knowledge of meaning, namely, by his readiness to accept a statement as true whenever the conditions for its truth obtains and he is in a position to recognize it as obtaining,[98] it follows that there is no capacity that can be equated with implicit knowledge of truth conditions, and hence, that a truth conditional account of meaning cannot possibly meet the constraints that the theory of meaning imposes on any meaning-theory.

5 Concluding Remarks

We just saw why, *given* semantic revisionism, realism must be rejected. The question that has not been addressed directly but that yet appears to be more basic is: Why should we take the point of view specified by semantic revisionism, to begin with? Following Dummett, the best reason for averting to semantic revisionism is that it will make us succeed where otherwise we would fail. Why? Well, because there will be no need to appeal to anything else but our use of language in determining what the correct logical rules are. This is a cryptic way of expressing two ideas. The first of these is that is there is an intimate connection between our ability to solve philosophical problems and the way we are looking at them. The second is that a problem is solvable if in the course of solving it we need not change our point of view from which we are looking at it originally. Admittedly, this is still rather cryptic, but it is helpful in getting some grip on the underlying structure of Dummett's overall strategy. The first idea can be traced back to Dummett's claim that the traditional, ontological-oriented approach towards logic and meaning makes it impossible for us to answer questions concerning the acceptability of logical principles. The only way to overcome this inability, Dummett claims, is to take a different look at the same problem. Hence his semantic revisionism.

The second idea is not so easily illustrated. The reason is that Dummett nowhere seriously explains why the ontological-oriented approach

[98]Cf. McDowell 1981, p. 227. According to McDowell this restriction reflects an assumption concerning the internal articulation of a theory of meaning.

must fail. The only thing he says in this connection is that this is apparent from the history of philosophy. However, the shortcoming of philosophy so far could be just a contingent fact, and as long as we cannot exclude this possibility, there is hardly any reason for averting to semantic revisionism. So, for a semantic revisionist it is important to have at least some notion of what makes it necessary that we cannot decide on the correct logical rules as long as we take the ontological-oriented point of view. Here the second idea comes in: this approach fails, because if we set out to determine what acceptable logical rules are in ontological terms, then, in the course of doing this, there will come a moment at which we will find ourselves forced to change or extend our original perspective.

This is a working hypothesis. It does not, of course, explain why the ontological perspective necessarily leads to theoretical impotence. It only points at a direction that such an explanation could take. It applies very well, though, to our discussion of Frege's inability to account for the objectivity of meaning. Starting from his ontological analysis of meaning, Frege cannot account for the objectivity of meaning, without turning to epistemology, where he has to face a problem very similar to the one he sought to solve by stipulating a third realm of meanings. Following the second idea we could say that Frege's ontological approach fails, *because* it forces us in the course of our investigation to change our perspective.

We thus see how the inability to account for the objectivity of meaning gives a reason for rejecting Frege's general perspective on meaning. But even so, this is not yet a positive argument in favour of semantic revisionism. The only positive argument I can think of just consists in an actual justification of logic on the basis of the use of language alone. The proof of the pudding is, so to speak, the eating of it. But whether Dummett, or anybody else, will ever be able to go beyond the programmatic stage that semantic revisionism is still in, is more than doubtful as will become increasingly clear in chapters II and III.

II Logical Revisionism and Expressibility

1 Bivalence and the Law of the Excluded Middle

A difference in theory must be reflected in a difference in practice, a difference in the conception of meaning in a difference of use.[1] Since Dummett presents this as a general principle, it ought to be applicable to the anti-realist thesis that the realist conception of meaning is mistaken. What consequences does a rejection of this conception of meaning have for the way language is, or at least, ought to be used? This chapter concentrates on the claim that every departure of realism entails a deviation from classical logic. And, since the acceptance of the semantic principle of bivalence is taken to be characteristic of realism, this claim can more succinctly be formulated as the thesis that bivalence is a necessary condition for classical logic.

As has been noted by Wright,[2] there is *prima facie* evidence that this claim cannot be sustained. Supervaluations, for example, present a counterexample: they do not respect the principle of bivalence, while they validate every classical tautology.[3] Does the anti-realist have a reply to that? Wright does not bother to provide us even with a clue. But so much seems certain, if the anti-realist has one, it will be based on an analysis of the conditions that semantic models underlying supervaluations are subjected to.

The crucial question is this: are these conditions acceptable in the light of the basic tenets of anti-realism? With this in view, the chapter shows that it is possible to devise a partial, that is, a non-classical

[1]Dummett 1987b, p. 225.

[2]Wright 1987.

[3]Cf. Rasmussen and Ravnkilde 1982, among others. Dummett, too, seems to accept it as a counterexample, see Dummett 1987b.

semantics that validates classical logic only at the expense of expressive power. More precisely, the chapter shows that under the super-truth approach we cannot express the distinction between direct and indirect evidence.

As such this is not a sufficient reason for rejecting the super-truth approach as a counterexample to the claim that a dismissal of classical semantics necessitates a revision of classical logic. It is only a first step to this. The second step consists in an argument to the effect that the distinction between direct and indirect evidence is essential for anti-realism. In this chapter the latter step receives comparatively little attention. We will deal briefly with it in section 3.1, mainly to show that a good case can be made for it. To establish it conclusively much more need to be said, but that would carry us too far away from the main goal of this chapter. Hence, as far as anti-realism is concerned, the claims defended in this chapter are conditional. Under the assumption that the distinction between direct and indirect evidence is essential for anti-realism, it is shown that supervaluations do not provide an acceptable counterexample to the claim that a rejection of bivalence necessitates a rejection of the law of the excluded middle. In addition it is shown that if we want to be able to express this distinction in full we cannot be satisfied with the conditions that are characteristic of the semantic models that validate intuitionist logic. From which it follows that, contrary to what has become common practice, intuitionist logic should not be taken as the prime example of a logic that is acceptable from the point of view of anti-realism.

The analysis of the conditions imposed on non-classical semantic models that validate classical logic enters into the notion of bivalence itself. A semantic theory respects the principle of bivalence if it assigns to every sentence of the language exactly one of two truth-values 'true' and 'false'. 'Exactly one' means: at most one and at least one. At most one truth-value amounts to the principle of non-contradiction. But it is not this aspect of bivalence that is contested by the anti-realists, but the other one, the thesis that every sentence has at least one truth-value, in other words, the assumption that there are no truth-value gaps. The super-truth approach does not respect bivalence, because it allows for such gaps. This, however, is not exclusive for supervaluations, there are lots of semantic theories that exhibit the same feature. As a group they are referred to as *partial* semantics. Yet, most of those theories do *not* validate classical logic. So, what is special about supervaluations?

What makes that it, notwithstanding its partial nature, does validate classical logic?

In order to shed some light on this, three different semantic theories will be compared: supervaluations, Kripke semantics for intuitionist logic, and semantics for Nelson's logic of constructible falsity, Nelson semantics for short. Partiality is their common denominator, but the three differ with respect to the logic they validate. The first one validates classical logic, the second one, as the name indicates, intuitionist logic, while the third validates a logic which is neither classical nor intuitionist. How can these differences be accounted for?

2 Partial Semantics

One cannot contrast different theories unless there is some fixed standard of comparison available. Given a logical language $\mathcal{L}$, a model for $\mathcal{L}$ provides an interpretation of its atomic sentences. On the basis of such an interpretation, and assuming that $\mathcal{L}$ is compositional, we can determine the interpretation of the complex sentences of $\mathcal{L}$ by means of the truth, and falsity conditions provided by a semantic theory. Semantic theories may differ with respect to these conditions, while they agree about the interpretation of the atomic sentences of the language. That is, despite their differences, they may give the same answer to question: how are we to interpret atomic sentences in a partial setting? In order to answer this question we will develop a notion of a partial model that is general enough to serve as a basis for, respectively, intuitionist logic, classical logic, and constructible falsity. In a subsequent section I will give a precise content to the concepts direct and indirect evidence, by interpreting the expressions 'φ is indirectly verified', and 'φ is indirectly falsified' in terms of these models.

2.1 An Informal Introduction

The notion of a partial model is, of course, well known, at least among logicians. However, since the main goal of this chapter is to contribute to the development of an argument to the defense of a philosophical thesis, I will take some time to explicate the ideas that partial models as defined in subsection 2.2 are supposed to substantiate. Starting from scratch, the question is: How are we to interpret atomic sentences? The principal idea underlying partial semantics is that their interpretation is not absolute, but relative to the information we may possess about the relevant situations. That information will, in general be incomplete, it will contain gaps, but gaps that may subsequently be filled when we learn more about the situation.

An example. Imagine you are driving your motorbike, when suddenly it stops. You are inexperienced, but eager to maintain your motor yourself, so enthusiastically you start taking your bike into pieces, to find out, finally, that you were out of gas. What has happened? Of course, you immediately realized that the malfunctioning of only some part of the machine must have caused the trouble. But, you had not the slightest idea which part. You had only questions : Could it be the carburetor? Could it be the sparking plugs? And so on. The information state you are in at that moment will be represented by a 'checklist': The carburetors work properly, the ignition is in order, etc.; all propositional sentences none of which you know to be true, or to be false. Then you started to eliminate possibilities. Everytime you inspected a part, you gained new information: The carburetors worked indeed properly, the ignition was indeed in order, and so on, until you hit the jackpot. And every time you acquired new information, the information state you were in changed.

An example like this may give the impression that information states represent the various stages in the actual process of acquiring knowledge. But that is not the case. An information state does not represent a psychological, but (part of) a factual situation. In the models we are considering informational states are thought to be ordered in a particular way, depending on the amount of information they contain. Surely, a knowing subject may acquire knowledge in the same way as described by the model, but he need not. One of the requirements on the order of information states is that if a certain state contains the information that, say, the sparks are perfectly in order, then every subsequent state will contain at least that piece of information. In other words, it is assumed that information is *persistent*. Obviously, this would be too strong a requirement on a representation of the actual process of acquiring knowledge, since in real life we happen to forget most of the information that is presented to us.

A state of information represents the facts one is acquainted with. A fact is conceived of as a constituent of the world. Given this conception of an information state, there are two kinds of development that have to be distinguished. In the first place, there is the growth of the knowledge a person has of the world. And, in the second place, there is the development of the world itself. Suppose, for example, that you have some shares you want to sell, and you wonder whether the time has come to do so. Since you want a maximal profit from your investments you inform yourself about the mood of the stock market. But while you are reading the signs in order to make up your mind, that mood might change, and with it the exchange value of your shares. The point is that in order to acquire more detailed information about how the world is, you start with the information you have about how the world is, to end up with a knowledge state which, at its best, tells you how the world *was*.

There is little reason to believe that, at least in this respect, the mood of the stock market is essentially different from other phenomena that make up the world. But to say that we can only learn how the world was, and not how it is, is not to deny that factual situations are fixed in one way or the other. We take them to be fixed in a way similar to how we assume that what has happenend in the past is fixed.[4]

Returning to the motorbike example, I want to point out that it exhibits a feature that we will encounter below as an optional requirement on the ordering of states of information. Since a motorbike has only finitely many functional parts nothing withholds us from assuming that there exists information states in which all the questions of our checklist have received a definite answer. Such an information state is said to be *complete*. That is not to say that in a complete state of information all the questions have received the correct answer; it is nothing more than a possible factual situation. We must also distinguish between the existence of such a state and its attainability. We might, for example, not be able to attain a complete state of information —whether correct or incorrect— about some factual situation in the past, simply because we lack the means, or sources, to determine whether some of our conjectures about that situation are either true or false. Or, a complete state of information may just contain too much information. Too much, that is, to acquire in a finite amount of time.

Another point is this. We imagined an inexperienced driver. But

[4]In this lies the key for extending representations of growth of information about fixed factual situations to representations of growth of information about evolving factual situations. We represent the development of a factual situation as a chain of successive situations, and treat each of these as a fixed one.

now imagine a more experienced driver, familiar with the particular sounds his motor makes under different circumstances. In the same situation, no doubt the first thing to occur to him would have been: 'It must be some defect in the fuel supply'. The very same situation, then, provides different information for different persons in the sense that for the inexperienced driver all possibilities were open, while for the experienced driver they are not. Still, we cannot say that the experienced driver *knows* that he has run on empty as soon as the engine stops. The best we can say is that he knows something about the way in which his knowledge will develop. This kind of information will be called *second level* information and it is to be distinguished from *direct*, or *first level* information on the basis of which a sentence can be assigned a definite truth-value.

On the basis of the information we have, some sentences are assigned definite truth-values and we expect other sentences to receive a particular truth-value when our knowledge will be extended. But where do we get our information from? Let us consider again the notion of a complete state. Naturally, we assume such a state to tell us 'how things really are.' But, as said, the abstract notion of a complete state does not guarantee that; a state in which all atomic sentences have the wrong truth-value is as complete as one in which they all get the right one. An interesting question is the following. Suppose that a complete state is attainable. Will it then always be possible to attain a complete state that is also correct? In large that depends on where we assume we get our information from. Surely, our sources can be unfaithful. But unfaithful to what? There seems only one answer possible: unfaithful to the facts.[5] So we can read, for example, that our subjects 'cannot but get their data from what is in fact the case.'[6] If this is true, then, of course, it will always be possible to attain a complete and correct state of information.

Since we do not doubt the reliability of our means of extracting information from what is in fact the case, a state of information can be identified with that part of the world our subject has knowledge

[5] This brings us to the problem of belief revision, a topic that I will not touch at all. I only want to mention that theories that model belief revision are complementary to theories that model growth of information given that the facts are taken to be the ultimate source of information.

[6] Veltman 1985, p. 159.

of. Still, there are at least two questions that should be asked. The first concerns the ontological status of what is in fact the case. Is it conceived of as dependent, or as independent from our means of knowing it? The other one is this. What can be known *immediately* from what is the case? I get the information that, say, the vase is red immediately from it being the case that in fact the vase is red. But do I get the information that the vase is not blue also directly from it being the case that in fact the vase is not blue? And if so, does that commit me to the existence of negative facts? In a similar vein we might ask whether there are logically simple facts next to logical complex ones. For it seems that only if we admit logically complex facts it is possible to hold that we obtain our data from the facts, and that we can have direct information that the vase is red or blue. Below I will return to these questions.[7] Now I want to turn to the more formal elaboration of the notion of partiality.

2.2 Partial Models

States of information are partial because they, characteristically, may not support the assignment of determinate truth-values (either 'true' or 'false') to *all* atomic sentences. It follows that, in general, 'true' and 'false' are not complementary predicates, and hence that the informational content of a state of information cannot be thought to be determined solely by the sentences that are true in it, or that are false in it. We need to know both. For this reason, we associate each state of information s with a valuation V_s, a partial function from $\mathcal{P}$, the set of atomic sentences of $\mathcal{L}$, into $\{1, 0\}$. States and valuations are not to be identified, for we want to be able to speak of two different states that nevertheless verify exactly the same sentences.

Let me add that, at this stage, the predicates 'is true' and 'is false' must be taken as purely formal qualifications. As we will see, different interpretations of these predicates, and of their relation to each other give rise to different semantic theories.

The set of atomic sentences that are verified by a valuation V_s will be called V_s^+, while V_s^- denotes the set of all atomic sentences that are false under V_s. That is, $V_s^+ = \{p \in \mathcal{P} \mid V_s(p) = 1\}$, and $V_s^- = \{p \in \mathcal{P} \mid V_s(p) = 0\}$. Since valuations are partial, V_s^+ and V_s^- together need not

[7] See § 4.2 of this chapter.

exhaust $\mathcal{P}$.

I will limit myself to a language for propositional logic, consisting of a set of atomic sentences together with all the complex ones that be formed from them by means of the operators '$\neg$', '$\vee$', '$\wedge$', and '$\rightarrow$', to which a new logical constant, '$\boxplus$' is added.

1. Definition. (Logical language) *The logical language $\mathcal{L}$ is the smallest set of sentences such that $p \in \mathcal{L}$, for p atomic, and further, for any formulas φ and ψ, if $\varphi \in \mathcal{L}$, then $\neg\varphi \in \mathcal{L}$, and $\boxplus\varphi \in \mathcal{L}$, if $\varphi \in \mathcal{L}$ and $\psi \in \mathcal{L}$, then $\varphi \wedge \psi \in \mathcal{L}$, $\varphi \vee \psi \in \mathcal{L}$, and $\varphi \rightarrow \psi \in \mathcal{L}$.*

The constant '$\boxplus$' deviates from the others in that it has a meaning that is defined independently of the different semantic theories we will consider, while the meanings of the other logical constants may vary from theory to theory. The meaning of a sentence of the form $\boxplus\varphi$' is discussed in subsection 3.2.

What is of central interest in partial semantics is not so much the partial states themselves, but the ways in which those states can resolve their partiality. Given a valuation V we will say that any valuation V' *extends* V iff every atomic sentence that has a definite truth-value under V has that same value under V', while atomic sentences that are undefined under V may be assigned a definite value under V'. In other words, if a valuation V_v extends a valuation V_s, V_v can only differ from V_s with respect to the atomic sentences that are undefined under V_s.

2. Definition. (Extension) *Let V_s and V_v be two valuations.* V_v extends V_s *iff* $V_s \subseteq V_v$.

A valuation specifies which atomic sentences are made true in a state, which are made false, and which are left undefined in it. As such, one could say, it gives content to the state it is associated with.

A process of resolving partiality is represented by a model which is constructed out of states and valuations associated with those states. The following definition lays down the requirements that such a model has to meet.

3. Definition. (Partial model) *A* partial model *for a language $\mathcal{L}$ is a triple $\langle S, \leqslant, V\rangle$ such that:*

1. *S is a non-empty set of states partially ordered by $\leqslant$;*
2. *V is a function that assigns to every $s \in S$ a partial function V_s from the set of atomic sentences of $\mathcal{L}$ into $\{1, 0\}$;*
3. *For all $s, s' \in S$: if $s \leqslant s'$ then $V_{s'}$ extends V_s.*

If a model is to represent the successive stages in the way partiality can be resolved, we must ensure that if one state succeeds another the valuation associated with the former extends that associated with the latter. Due to the third requirement we can be sure that this will always be the case. Notice, furthermore, that this implies that an atomic sentence that is true (false) in a certain state, will remain true (false) in every extension of it. That is, it boils down to the requirement of persistency for atomic sentences.

As said before, we use states to represent the successive moments in the process of resolving partiality. But it would be incorrect to look upon a state as a stage in that process. Rather, a state represents one of, perhaps, many possibilities at a certain stage. Consider, for example, a state s in which the two atomic sentences p and q are undefined. Suppose that, according to our model, s can be extended to a state s' that verifies p, but not q, and to a state s'' that verifies q, leaving p undefined. According to our definition of extending, neither one of s' and s'' can extend into the other, but s can extend into both. Hence, s' and s'' represent two mutual exclusive possibilities in the next stage in the process of resolving s.[8]

We are now in a position to answer a question this subsection set out with: How are we to interpret atomic sentences in a partial model? In the light of what has been said above, the following answer lies readily to hand. Given a partial model $\langle S, \leqslant, V\rangle$, the interpretation of an atomic sentence p in M is:

$$I_M(p) = \langle\{s \in S \mid V_s(p) = 1\}, \{s \in S \mid V_s(p) = 0\}\rangle$$

I will omit the subscript whenever possible. The interpretation of p is thus determined as the ordered pair consisting of, respectively, the set of states in which p is true under V, and the set of states in which p is false under V. Occasionally, the former will be referred to as $I^+(p)$, and the latter as $I^-(p)$. An interpretation I will be called $I^+(p)$-dependent if, for all atomic sentences p, $I^-(p)$ can be defined in terms of $I^+(p)$.

We may impose additional restrictions on the partial models we consider in interpreting the language. One of them is, what I will call, *completeability*. It has already been mentioned above. We will now give it a precise content.

[8]The distinction between a state and a stage is thus brought out by determining the latter as the set of states that can be reached in a certain number of steps.

A valuation is *total* if it assigns the value 'true' or 'false' to every atomic sentence in the language:

4. Definition. (Total valuation) *Let $\mathcal{P}$ be the set of atomic sentences of language $\mathcal{L}$. A valuation V_s is* total *w.r.t. $\mathcal{L}$ iff $V_s^+ \cup V_s^- = \mathcal{P}$.*

The state s with which V_s is associated is, accordingly, called a total, or complete state.

5. Definition. (Path) *Let S be a set of partial states, and $s \in S$. A* path *through s, P, is a maximally linearly ordered subset of S with $s \in P$*

6. Definition. (Strong completeability) *A path is* strongly completeable *if it contains a total state. A partial state s is strongly completeable iff every path through s is strongly completeable. A model is strongly completeable iff every $s \in S$ is strongly completeable.*

Another condition that we will frequently encounter below is called *cofinality.*

7. Definition. (Cofinality) *A model is* cofinal *iff for all state $s \in S$ and all atomic sentences $p \in P$: if $\forall s' \geqslant s \exists s'' : s'' \Vdash^+ p$, then $s \Vdash^+ p$.*

Strictly speaking this is *positive* cofinality. The definition for *negative* cofinality is obtained by replacing all occurrences of '$\Vdash^+$' by occurrences of '$\Vdash^-$'.

The leading thesis of this chapter is that, in a partial setting, we cannot express the distinction between direct and indirect evidence if we are to validate classical logic as well. Cofinality and completeability will play an important role in establishing this: it will be argued that, on the one hand, cofinality and completeability are necessary for validating classical logic in terms of supermodels, while on the other hand in any model that satisfies these two conditions the expressions 'there is direct evidence for φ' and 'there is indirect evidence for φ' are logically equivalent.

3 Direct and Indirect Evidence

The best reason to distinguish direct from indirect evidence is contained in the idea that a state not only contains information about which statements are true, false, or undetermined in that state, but also, what has been called above, second level information: information about what are possible ways for that state to be extended. And whatever second level information a state contains is what is called here the *indirect evidence* it provides.

Probably the simplest way to illustrate the difference between direct and indirect evidence is this. Suppose there are two boxes, called A and B. One of them contains a coin. But you do not know which. The other one is empty. All this you know on the basis of the evidence available to you in what I will call the initial state of information, state 1. Next you open box B. It is empty. The state of information you are in now will be referred two as state 2b. Being in this state you conclude that the coin *must* be in box A. The question is whether you can with equal right conclude that the coin *is* in box A? To answer this question affirmatively is to deny that there is any qualitative difference between knowledge state 2b and a state you are in when, having opened box A, you have seen the coin , touched it, bit on it, etc.. Call this state state 2a.

What reason could there be to hold that the states 2a and 2b are to be distinguished? We will now briefly evaluate different reasons that could be given. The first is that state 2a is to be distinguished from state 2b, because being in state 2a we can say not only that the coin is in box A, but also that it is a dutch five guilder coin, that it is brand new, and all other kind of details we cannot know of being in state 2b.

If this differentiates direct from indirect evidence, it does so only by providing a negative criterion: to possess indirect evidence is to lack certain information that would be present if only we had direct evidence at our disposal. But does this criterion indeed suffice to distinguish direct from indirect evidence in all cases? Suppose now, to the contrary, that you first opened box A. You would then conclude that box B must be empty. But what would opening box B and seeing its emptiness add to your stock of evidence? What details would be revealed to you that you could not have known otherwise? None, I believe, for the thought that the emptiness of box B might be significantly different from the emptiness of, say, some box C, has no attraction at all. It follows then that our negative criterion lacks the required generality.

In a last attempt to the rescue of the negative criterion it might be put forward that by opening box B we do not learn any details about the emptiness of box B, but we do learn something about its interior. The attempt, however, is fruitless. We can easily modify our example in such a way that already in the initial state we have a detailed knowledge of the interior of both boxes. The only viable conclusion, then, is that we cannot, in general, distinguish direct from indirect evidence, if possession of the latter is characterized as the lack of certain information that we would possess if direct evidence were available.

There is another train of thoughts, equally wrong. It might be said that knowledge on the basis of indirect evidence is not as *certain* as knowledge on the basis of direct evidence. Being in state 2a, opening box B does add to my knowledge, because as long as I have not seen with my own eyes that it is empty indeed, there is always the possibility that it is not empty. What is wrong about this? Under this conception an expression like 'it must be the case that φ' is interpreted as 'it is most probably the case that φ'. But that interpretation cannot be reconciled with the principle of persistency. In the initial state it is known that one of the boxes is empty. That knowledge should be preserved even after you have come to know that the coin is in box A. It is hard to see how the evidence available to you in the initial state has deflated, in state 2a, to evidence for a statement to the effect only that *probably* one of the two boxes is empty.

It is wrong to interpret 'must be true' as 'it is most probably true', because, for the subject being in state 2a, it is not a real possibility at all that box B turns out to be not empty: the exclusion of that possibility can, in an important sense, even be said to be constitutive for the state of information he is in. What does that mean? That our subject, still in state 2a, just will not accept it if he opens box B and find it not empty. Would he then go as far as to deny the facts? No, there are other alternatives. First, he might deny that the box he opened is box B, and believe that it is some other box with which box B has been replaced when he was inspecting box A. Or, he might come to the conclusion that it is not true after all that one of the two boxes is empty and *revise* his state of information accordingly.[9] Thus, if the subject takes the evidence available to him seriously, he will not deny what he observes, but he will deny that what he observes is evidence for

[9] A third possibility is that the subject concludes that the situation his knowledge is about has changed. I do not mention it explicitly, because under the conception endorsed here that situation is taken to be fixed. But this simplification has no bearings on the current issue: the subject would take his observation as evidence for the fact that box B is not empty *now*, but he will deny that 'box B now' is part of the situation his knowledge is about.

the fact that box B is not empty; for him, there cannot be such a fact. He can only accept it as such by revising, that is, changing, the state of information he is in. It follows then that the exclusion of the possibility of certain states of affairs is constitutive for a state of information in the sense that it is impossible to allow for those possibilities while being in the same state of knowledge.

3.1 Inference as Indirect Evidence

Our discussion of the distinction between direct and indirect evidence in the preceding subsection serves a double purpose. On the one hand it gives a rationale to the more formal interpretation of the notion of indirect evidence as it will be introduced later. At the same time it is a starting point for a discussion of the role and content of the notion of indirect evidence as it frequently recurs in Dummett's writings on anti-realism. In this subsection, we will briefly review the way in which the distinction between direct and indirect evidence is embedded in Dummett's views on meaning.

In the preceding subsection, to recall, indirect evidence has been explained as second level information that characterizes a knowledge state by constraining the way in which this state can be, or could have been, extended. As such it appeals to the distinction we are accustomed to make between how things are in themselves and how they appear to us. For second level information tells us of certain statements that these have a particular truth-value, even though we may never *actually* be in the position to directly verify it. This is not to say that second level information transcends the abilities we have for establishing the truth-value of a statement. It only says that if we had been at a different place at the same or at a different time, or at the same place at a different time, then, using our normal abilities, we would have been able to establish directly what we already knew, albeit indirectly. To speak with Dummett, it admits 'a notion of truth explained in counterfactual terms as attaching to what we *should have observed* to be so had we had, or taken, the opportunity to do so.'[10]

This notion of truth, and thereby our conception of how things are independent of any observations we make, Dummett believes, does not have any ontological significance, but is induced completely by our

[10]Dummett 1991b, p. 220.

own linguistic practices, and especially by the existence of deductive reasoning:

> If we spoke a language devoid of all inferential practices whatever, then there would be nothing to debar us from equating the truth of a statement with its having been directly verified. (...) The introduction of deductive practice into the practice of speakers of such a language would immediately compel them to frame a broader conception of truth.[11]

Dummett is not implying that the admission of inferential procedures necessitates a distinction between the notion of truth and that of what may justifiably asserted. But, he holds, the practice of deductive reasoning does yield the need of

> a conception of what can be *indirectly* established as assertible, even though it has not been *directly* so established, just because that is what, in any interesting case, inference does, namely to establish its conclusion indirectly.[12]

An inference presents indirect evidence for the statement that is its conclusion. But what counts as *direct* evidence for that statement? And: In what respect does it differ from indirect evidence?

> For any given sentence, there will always be something which we regard as being the most *direct* means of recognizing it as having one or other truth-value; not in the sense of that means which involves the least expenditure of effort, nor the most practicable or the most certain; but that which corresponds, step by step, with the way in which the sense of an expression is determined from those of its components.[13]

So, while there are many indirect means of establishing a statement indirectly, there is only one, unique way of establishing it directly, since this means is determined by the senses of the constituent expressions as well as the by internal structure of the statement. In contrast to the indirect means, furthermore, knowledge of the direct means is constitutive of our understanding of the statement:

[11] Dummett 1994, p. 59.
[12] *ibid.*
[13] Dummett 1992a, p.237.

> The possibility of establishing the statement directly must be envisaged by anyone who grasps the meaning of the statement, construed on this model: the possibility of establishing it by indirect means need not be (...)[14]

So, to know the meaning of a sentence is to know what one has to do in order to determine the truth-value of the sentence directly, that is, in a way that is determined by the structure of the sentence. This does not imply that it is always possible to determine the truth-value of a sentence by direct means alone, nor is it said that the direct means provide the simplest or shortest way of establishing the truth-value of a sentence. As an example of the latter, consider the sentence 'It is raining, or it is not raining.' Perhaps the simplest way to verify this sentence is to recognize it as a tautology, but that is not the most direct means. For our grasp of the sense of a disjunction tells us that, in order to establish its truth-value, we ought to go out to see for ourselves whether one of the two disjuncts is true. Following the direct route we would never be in a position to know that 'It is raining, or it is not raining,' without knowing whether or not it is in fact raining.

Sometimes it is not even possible to determine the truth-value of a statement on the basis of the direct means alone. Take, for example, a universally quantified number-theoretic statement. According to Dummett's analysis, my understanding of the quantifier consists in the knowledge that a universally quantified sentence is true if each of its instances is true, and false otherwise.[15] Hence, in order to determine the truth-value of a universally quantified number-theoretic statement I will just have to go through all the natural numbers and determine whether or not the predicate applies. But although it is the direct means, it is not a means available to us, simply because we do not have the ability to check infinitely many numbers in a finite amount of time. It follows that, if we are to establish its truth-value at all, it must be established by indirect means, namely, by proving it.

Recapitulating we find that, on Dummett's account, the distinction between direct and indirect evidence is induced by our linguistic practices, more in particular, by the fact that deductive reasoning is among these practices. We also saw that for the explanation of understanding only the direct means are important: to know the meaning of

[14]Dummett 1978a, p. 313.
[15]Dummett 1992a, p. 634.

a statement is to know what the direct means are for establishing its truth-value. The merit of this account, Dummett believes, is that it enables us to explain how genuine epistemic advance by means of deductive inference is possible.[16] Consider once again a universally quantified number-theoretic statement. On the basis of our understanding of it we know what to do in order to establish it, but we are not capable of doing it. Hence, we do not actually know whether or not it holds. If, now, we are presented with a proof of it, this will tell us something we could not possibly have known solely on the basis of our original understanding of the statement in question. And in this sense, the proof provides us with genuinely new information.

But, the question is raised, given this account of the fruitfulness of inference, will it still be possible to explain why inference is compelling? For, if the knowledge that a particular proof establishes a certain statement is not in any way constitutive for our understanding of this statement, then why should that proof be taken as evidence for the truth of it? Or, turning the tables, what would stop us from taking just any inference we like as evidence for the truth of the statement in question? The solution, Dummett claims, lies in the observation that although direct and indirect evidence do not coincide, they are not totally unconnected.[17] How we are to conceive of this connectedness can be read of from the following criterion for the validity of a proof. A proof is *valid*

> if it provides an effective method for transforming a sufficiently detailed current observation that verifies a statement of a certain form into a verification of a related statement of a different form.[18]

Let me explain this by means of an application of Euler's solution to the problem of the bridges at Königsberg. Euler proved that in order to cross every bridge at Königsberg, one must cross at least one bridge at least twice. It is quite conceivable that, due to a moment of inattention, one has verified that during a connected course someone has crossed every bridge in Königsberg, without thereby having verified that at least one bridge is crossed at least twice. Hence, Euler's proof does *not*

[16]Cf. Dummett 1992c, p. 313, Dummett 1992a, p. 635, Dummett 1991b, p. 195 ff..

[17]In Dummett's explanation of this connectedness the notion of a conservative extension plays an important role. See chapter III for a detailed discussion thereof.

[18]Dummett 1994, p. 60.

show that any verification of the premise is also a verification of the conclusion;

> What it establishes, rather, is that any sufficiently detailled observations that serve to verify the former can be so arranged as simulteously to verify the latter.[19]

In other words, we accept a deduction as indirect evidence for a statement, because we take it as showing that if we are in a situation in which we can establish the premise(s) directly, then that situation can be transformed into a situation in which the conclusion can be verified directly.

Another example. Consider a situation in which we have eaten all the apples and pears we had, but that before eating them we had counted the apples and pears separately and found that there were eight apples and five pears:

> Though they have now been eaten an argument (...) convinces us that, if we had counted all fruit together, we should have found that there were thirteen of them: thirteen, that is, according to the criterion we had, namely counting, before we were ever introduced to the procedure of addition.[20]

Since the fruit has been eaten, it can only be established indirectly that there were thirteen pieces of fruit altogether, by adding eight to five. We accept this as indirect evidence, because we take it that if we were in a position to establish directly that there are eight apples and five pears, it would also be possible for us to establish directly that that there are thirteen pieces of fruit. Hence, the procedure of addition is acceptable as an indirect means of establishing a statement, because things could not possibly turn out to be otherwise if we would limit ourselves to the direct means of establishing it.

From what has been put forward in this paragraph it will be clear that the distinction between direct and indirect evidence does play an important role in Dummett's views on meaning and logic, and, moreover, that his conception of indirect evidence is rather similar to our analysis of indirect evidence as second level information. Much more can and need to be said about this issue. For the purposes of this chapter, however, we need not pursue the matter any further.

---

[19]Dummett 1994, p. 59.

[20]Dummett 1991b, p. 219. Cf. pp. 176–78, and Dummett 1978a, p. 300.

3.2 The Interpretation of 'Indirect Evidence'

Up to now we have only considered the distinction between *direct* and *indirect* evidence. We payed no attention at all to the distinction between *evidence for the truth* of a sentence, and *evidence for its falsity.* But if we do take this into consideration, as we will in the sequel, we must accordingly distinguish the distinction between direct and indirect verification, from that between direct and indirect falsification. And, beforehand, nothing justifies the belief that a reason to accept one distinction is not necessarily a reason to accept the other as well.

Next to the question whether there are indeed good reasons to discern direct from indirect evidence, we must ask: What conditions does a logic have to meet in order to be able to express the conceptual distinction between direct and indirect verification on the one hand, and direct and indirect falsification on the other? Questions of expressibility normally arise with respect to the properties of the frames underlying a logic. But in the present case no such property seems to be in play. I will therefore take resource to a different, though simple, strategy in tackling this question. The main idea is to find suitable formal 'translations' for locutions like 'state s verifies φ directly', and 'state s verifies φ indirectly'. We must be careful to ensure that these translations do not depend on any logic, although, of course, they cannot be independent from the language in which they are formulated. This seriously limits us to languages that are interpretable in a partial framework. For the present purposes, however, that is all that is needed.

The question thus becomes what happens when a logic is extended by adding an operator '$\boxplus$' such that '$s \Vdash^+ \boxplus\varphi$' is to be read as '$s$ verifies φ indirectly.' I will say that the distinction between direct and indirect evidence is expressible in that logic only if $\nVdash \boxplus\varphi \rightarrow \varphi$. In the same spirit it can be determined whether or not the distinction between direct and indirect falsification is expressible. There is no reason to expect that a difference in verification can be expressed iff the difference in falsification can. In fact, it easy to demonstrate that supervalutions (SV) cannot express either one, that Kripke semantics for intuitionist logic (KI) can only express the positive distinction, and that in Nelson semantics (NS) both can be expressed.

The first question is how the operator '$\boxplus$' is to be interpreted. Let us compare some alternatives. To begin with, we stay close to the informal characterization that figured in the previous subsection. There it was said that a sentence is indirectly verified (falsified) if it is impossible that it will turn out to be false (true). Using the terminology introduced above we might be consider the following formulations:

direct verification (primitive)	$s \Vdash^{+} \varphi$
indirect verification	$\neg\exists s' \geqslant s(s' \Vdash^{-} \varphi)$
direct falsification (primitive)	$s \Vdash^{-} \varphi$
indirect falsification	$\neg\exists s' \geqslant s(s' \Vdash^{+} \varphi)$

But this will bring us into trouble as soon as it is allowed that for some state s in which a sentence φ is undefined, there might be no extension in which φ is defined, one way or the other. In such a case s would indirectly verify, as well as indirectly falsify φ, and that would be a contradiction.

Since this situation might occur only in case sentences can be 'forever undecided' there is an obvious way to escape it, namely by demanding completeability. As an alternative we could change the characterizations mentioned above as follows:

indirect verification	$\forall s' \geqslant s \exists s'' \geqslant s'(s'' \Vdash^{+} \varphi)$
indirect falsification	$\forall s' \geqslant s \exists s'' \geqslant s'(s'' \Vdash^{-} \varphi)$

This definition, which will be named the '$\forall\exists$-definition', does not demand that every path through a state s must contain a state s' that verifies φ if s is to be said to verify φ indirectly. This can be seen from the following model. To describe it I make use of the notion of an *immediate successor*. Let s and s' be two different states. s' is called an *immediate successor* of s iff (1) $s' \geqslant s$, and (2) There is no state v such that $s' \geqslant v$, and $v \geqslant s$. The model is constructed as follows. Let s_1 be the initial state, and let an atomic sentence p be undefined in it. For every state s_n, furthermore, it holds that if n is odd it has two immediate successors s_{n+1} and s_{n+2}. Of these, the 'even state', state s_{n+1}, has no successors and does verify p. The odd state leaves p undefined. I will name a model of this form a 'ridge-model.'

According to the $\forall\exists$-definition state s_1 indirectly verifies the sentence p, since any extension of it has an extension that verifies p directly. For consider an arbitrary extension s_n of s_1. If n is even s_n is itself the extension that verifies p. And if n is odd, it is extended by s_{n+1} which verifies p. At the same time it cannot be said that every path through s_1 contains a state that verifies p, since we have the path

$P = \{s_n \mid n \text{ is odd }\}$ such that for every $s \in P$: $s \nVdash p$.

The problem with this definition is that it lacks generality, since it is only in special cases that we can be sure that we will never reach a state in which a sentence φ is falsified, even though it may never be verified.[21] Therefore, I prefer a slightly stronger interpretation, according to which a state s indirectly verifies a statement φ only if every extension of that state will eventually evolve into a state that verifies φ. Since this boils down to the demand that every path through s contains a state that verifies φ, we can define this interpretation by means of the notion of a *bar*:

8. Definition. (Bar) *Let S be a set of states with $s \in S$. A* bar for *s, B, is a subset of S such that for each path P through s: $P \cap B \neq \emptyset$*

9. Definition. (Future bar) *A bar B for a state s is called a* future bar *iff for all $v \in B$, $s < v$*

10. Definition. (Direct and indirect verification) *A state s verifies a sentence φ of $\mathcal{L}$* directly *iff $v' \Vdash^+ \varphi$. A state s verifies a sentence φ of* $\mathcal{L}$ indirectly *iff there is a bar B for s such that for every $v \in B$: $v \Vdash^+ \varphi$*

The definition for direct and indirect falsification is obtained by replacing every occurrence of '$\Vdash^+$' by '$\Vdash^-$'.

From this definition it is immediately clear that direct evidence implies indirect evidence. A state s directly verifies a statement φ only if it verifies φ itself. Since, trivially, every path through s contains s, $\{s\}$ is a bar for s. Hence, s also verifies φ indirectly.

To have any appeal definition 10 requires persistency. We would not speak of a state as providing direct evidence for a statement if in subsequent states that statement may become undefined again, or even false. A parallel thought applies to the definition of indirect falsification. Without persistency it would even be possible that a state s directly verifies some statement φ, while it also, although indirectly, verifies $\neg\varphi$. This would not be a contradiction in the strict sense of the word, but something very close to it.

The notion of a future bar is needed if we want to distinguish states that directly verify a statement from states that *only* indirectly verifies that statement. To that end we could modify defintion 10 by replacing

[21] I would say that we can be sure of that only if we assume that there are no contingencies. Compare § 4.1 of this chapter.

every occurrence of the notion 'bar' by that of 'future bar'. Given persistency we would still have that direct evidence implies indirect evidence: A state s directly verifies a statement φ only if it verifies φ itself. By persistency it follows that every state s' extending s verifies φ. Hence, the set $\{v \mid s \leqslant v\} \sim \{s\}$ is a future bar for s such that for every element v of it: $v \Vdash^+ \varphi$.

The bar-definition is stronger than the $\forall\exists$-definition. If there is a φ^+–bar for a state s then every extension of it has an extension that verifies φ. The converse does not hold, as is apparent from the model described above. But if we limit ourselves to finite models both definitions become equivalent. Let it be that for some s in a finite model it holds that every extension of it has an extension that verifies φ. Suppose there is no φ^+–bar for s. Then there is a path P through s such that for no $s' \in P$: $s' \Vdash^+ \varphi$. Let v be the final state of P. Surely v is an extension of s, but since v does not have any extensions and given that $v \nVdash \varphi$, v is an extension of s that does not have an extension that verifies φ. Contradiction.

We do not have to limit ourselves to finite models in order to be able to use the bar-definition and the $\forall\exists$-definition interchangeably. Another condition that will do the same job is *strong completeability*. Let s be a strongly completeable state such that $\forall s' \geqslant s \exists s'' \geqslant s'(s'' \Vdash^+ \varphi)$ holds. Suppose that there is no φ^+–bar for s. Then there is path P through s such that for no $s' \in P$: $s' \Vdash^+ \varphi$. Since s is strongly completeable it follows that there must be a $s' \in P$ such that $s' \Vdash^- \varphi$. Clearly, s' extends s, but, given persistency, it is also clear that there is no extension of s' that verifies φ. This is a contradiction.

3.3 Partial Models for Intuitionist and Classical Logic

For a first application of definition 10, we will now briefly compare partial models for intuitionist logic with partial models for classical logic. In order to avoid complications that have to do with the notion of falsification, we will confine ourselves to the distinction between direct and indirect verification. That is, we will allow only the notion of verification in the explanation of the meanings of the logical constants. The models that will be so constructed are called *positive* models:

11. DEFINITION. (Positive partial model) *A* positive partial model *for a language $\mathcal{L}$ is a triple $\langle S, \leqslant, V^+\rangle$ such that*

1. *S is a non-empty set of states partially ordered by $\leqslant$;*

2. *V^+ is a function that assigns to every $s \in S$ a partial function V_s from the set of atomic sentences of $\mathcal{L}$ into $\{1\}$;*
3. *For all $s, s' \in S$: if $s \leqslant s'$ then $V_{s'}$ extends V_s.*

The following conditions constitute what is knwon as *Kripke's semantics for intuitionist logic* (KI). Any partial model that satisfies these conditions validates all intuitionistically valid formulas:

1. $u \Vdash^+ \varphi \rightarrow \psi$ iff $\forall v \geqslant u$, if $v \Vdash^+ \varphi$, then $v \Vdash^+ \psi$.
2. $u \Vdash^+ \varphi \wedge \psi$ iff $u \Vdash^+ \varphi$ and $u \Vdash^+ \psi$
3. $u \Vdash^+ \varphi \vee \psi$ iff $u \Vdash^+ \varphi$ or $u \Vdash^+ \psi$
4. $u \Vdash^+ \neg\varphi$ iff $\forall v \geqslant u$: $v \nVdash^+ \varphi$.

Using definition 10 we can easily verify that on the basis of this semantics we can express the distinction between direct and indirect verification. To establish this, all we have to do is to construct a counterexample to the statement:

$$\Vdash^+_{KI} \boxplus \varphi \rightarrow \varphi$$

Consider the model consisting of two states u and v, with $u \leq v$. Let $I_p = \{v\}$. $\{v\}$ is p^+-bar for u. Hence, we have $u \Vdash^+ \boxplus p$, but we do not have $u \Vdash^+ p$.

What conditions does a partial model have to satisfy in order to validate all classically valid formulas in $\mathcal{L}$? Observe that given the four conditions mentioned above we already have that the truth-conditions are classical at *endpoints*. An endpoint is just a state that has no successors. Assuming that u is an endpoint, we can rewrite our four conditions as follows:

1. $u \Vdash^+ \varphi \rightarrow \psi$ iff, if $u \Vdash^+ \varphi$, then $u \Vdash^+ \psi$.
2. $u \Vdash^+ \varphi \wedge \psi$ iff $u \Vdash^+ \varphi$ and $u \Vdash^+ \psi$
3. $u \Vdash^+ \varphi \vee \psi$ iff $u \Vdash^+ \varphi$ or $u \Vdash^+ \psi$
4. $u \Vdash^+ \neg\varphi$ iff $u \nVdash^+ \varphi$.

These are just the classical truth conditions. It follows that at every endpoint, all classical tautologies are valid. This suggest that we obtain a partial model that validates classical logic if we impose the following condition on our models: if, for an arbitrary state v, it holds that every extension of it will eventually extend into a state in which some formula φ is valid, then state v must validate φ, too. This amounts to requiring (positive) *cofinality*.[22] The underlying idea is that in order to

[22]van Benthem 1986. It is defined on page 58.

determine the validity of a formula we should not only take into account the truth-values of its atoms right now, but, if some of the atoms are undefined, also the truth-values that those atoms will acquire when the model is extended. Cofinality systematically connects the truth-value of an atom at an instance and its truth-values at possible extensions.

However, as it turns out, cofinality for atomic sentences just is not enough to ensure cofinality for all formulas. Let us take a closer look. Given cofinality for atomic sentences we might try to proof for KI that cofinality holds for all sentences by induction. Problematic is the case of disjunction. Let it be that $\forall s' \geqslant s \exists s'' \geqslant s'(s'' \Vdash^+ \varphi \vee \psi)$. In order to apply the induction hypothesis, we have to appeal to the truth definition for disjunction, which tells us that $s'' \Vdash^+ \varphi$, or $s'' \Vdash^+ \psi$. But since it need not be the case that for *every* s'', $s'' \Vdash^+ \varphi$, or that for *every* $s'' \Vdash^+ \psi$, we cannot apply the induction hypothesis to derive $s \Vdash^+ \varphi$, or $s \Vdash^+ \psi$. And, hence, we cannot establish that $s \Vdash^+ \varphi \vee \psi$, even though every extension of it has an extension at which it does hold.

In this we find a motivation for the following alternative truth definition for disjunction:[23]

$$3' \qquad u \Vdash^+ \varphi \vee \psi \text{ iff } \forall v \geqslant u \exists w \geqslant v(w \Vdash^+ \varphi \text{ or } w \Vdash^+ \psi)$$

Application of this definition tells us that for every extension v of s'' there is an extension v' such that $v' \Vdash^+ \varphi$, or $v' \Vdash^+ \psi$. From this it follows immediately that every extension of s has an extension that verifies φ or ψ. And hence that $s \Vdash^+ \varphi \vee \psi$.

When we replace condition 3 of KI by condition $3'$, and, in addition, require cofinality for atomic sentences, we obtain a *Kripke semantics for classical logic* (KC). Although in KC every classical tautology can be shown to hold, KC does not respect the principle of bivalence. The idea underlying KC is that for a state s to verify a statement φ it is enough if there is no extension of s that verifies $\neg\varphi$. Since this condition is imposed on models for intuitionistic logic, it follows that every state in every model verifies all classical tautologies, for even in intuitionist logic it is impossible to verify the negation of a classical tautology. It is, of course, not impossible to verify the negation of a contingency. And even if, for some state s and a contingency φ, in every extension of s, φ eventually receives a definite truth-value, it need not be the same

[23] Adopted in van Benthem 1986. As an alternative $\varphi \vee \psi$ can be defined as: $\neg(\neg\varphi \wedge \psi)$.

truth-value in every extension. Hence, we cannot apply cofinality in order to assign a definite truth-value to φ in s. Note that for the same reason it was impossible to derive cofinality for all formulas from the requirement of cofinality for atomic sentences, without changing the rule for V.

That KC does not respect the principle of bivalence makes it a counterexample to the anti-realist's claim that bivalence is necessary for classical logic. However, the appeal to cofinality has a serious drawback: it has the consequence that the distinction between direct and indirect verification cannot any longer be expressed. Let B be a φ^+–bar for s. Then every path through s contains some s' such that $s' \Vdash^+ \varphi$. By cofinality we then have: $s \Vdash^+ \varphi$.

4 Negation

Negation stands in need of a separate treatement, because, more dominantly then in the case of the other logical constants, we have to face several alternatives for its interpretation, while it is far from obvious that we should prefer the one above the other. There are two different ways of approaching negation, depending on whether or not we take the concept of falsification as primitive. In this section I will confine myself to the intuitionist interpretation of negation, which, characteristically, appeals to verification conditions only. However, as we will see when discussing supervaluations in the next section, a semantic theory that imposes verification, as well as falsification conditions need not necessarily accept falsification as primitive.

The usual way of treating negation is by discussing the rules that govern the introduction and elimination of the symbol '$\neg$'. In contrast to this, but in line with the foregoing, I will approach negation primarily from the perspective of semantics. To that end I introduce a new logical constant, '$\perp$', with the following meaning:

$$\forall s,\ s \nVdash^+ \perp$$

This being given, we obtain intuitionist negation by defining $\neg\varphi$ as follows:

$$\neg\varphi =_{\mathrm{df}} \varphi \rightarrow \perp \tag{II.1}$$

This yields the following interpretation for '$\neg\varphi$': $s \Vdash^+ \neg\varphi$ iff $\forall s' \geqslant s$, if $s' \Vdash^+ \varphi$, then $s' \Vdash^+ \perp$. Since we have that for no s: $s \Vdash^+ \perp$, it follows that $s \Vdash^+ \neg\varphi$ iff $\forall s' \geqslant s$: $s \nVdash^+ \varphi$.

When applied to mathematical reasoning this boils down to the condition that a statement of the form $\neg\varphi$ is assertable in a state s iff s contains the information that a proof of φ is impossible. The interpretation of negation as the absence of a proof is already implicit in Brouwer's criticism of the acceptance of the law of the excluded middle as a generally valid principle. It lacks the alleged generality because in many cases checking the correctness of a mathematical statement does involve the construction of infinitely many mathematical objects. And in such cases it cannot be determined a priori whether or not an object that proves the statement is among all the possible constructions.[24]

What is not clear is whether this interpretation of negation is necessary for the form of criticism formulated by Brouwer. What is essential is that we cannot survey infinitely many possible constructions, and not that a negation is assertable only if it turns out that none of these constructions proves the statement at hand.[25] But, one could reply, how can we conceive of a mathematical object that proves a statement to be false, or a negation to be true? Questions like these, because they are hard to answer, might divert our attention from our original question: What reasons can there be for not treating falsification on a par with verification, even if we confine ourselves to mathematical reasoning? I will return to it below.

4.1 Double Negation and Indirect Evidence

Under the current definition of negation, as is well known, φ and $\neg\neg\varphi$ do not express the same meaning. The former is assertable if and only if a proof of it actually has been constructed, while the latter is assertable in case it is impossible to attain a state in which a proof of φ is impossible. This difference corresponds almost completely to the distinction between direct and indirect evidence I have introduced above. Almost, because the assertability of $\neg\neg\varphi$ does not guarantee

[24]'Want het geheel der uit de vraagstelling te ontwikkelen systemen is *aftelbaar onaf*, dus niet a priori methodisch te onderzoeken ten opzichte van de aanwezigheid of afwezigheid van een de vraag beslissend systeem.' (Brouwer 1908, p. 156.)

[25]One might even go as far and ask whether the assumption that mathematical objects are mental constructions is essential to Brouwer's criticism of the law of excluded middle. It is certainly sufficient. But even if we assume that mathematical objects are not constructions, but just exist in some way or other, it could still be argued for that we might never reach a state in which we have discovered the proof object.

that we will, eventually, be able to construct a proof of φ, while under the bar definition I have give above, we can be said to have indirect evidence for φ in some state s only if every way of extending s will, in the end, result in a state s' in which φ is directly verified.[26]

The difference is clear, but need not be exaggerated. For one thing, the bar definition does not require that the directly verifying state actually obtains; it may take an infinitely long process to get there, which may be asked to much for a human being. It remains, of course, that if we could go on infinitely, we would always get there. For this reason $\boxplus\varphi$ is stronger than $\neg\neg\varphi$ in the sense that the former implies the latter but not vice versa.

By imposing certain restrictions on the models we are considering we might obtain an interpretation of $\neg\neg\varphi$ that coincides with that of $\boxplus\varphi$. There are several options, none of which affect the logic. That is, from the standpoint of intuitionist logic there can be no preference for one class of models, rather than another. The first thing we could do is to limit ourselves to finite models. Since intuitionist logic has the finite model property it is complete with respect to the class of finite models, while, as we have seen above, in finite models the bar-definition is equivalent to the $\forall\exists$-definition.

Alternatively, we could require strong completeability. A model is strongly completeable if and only if, for all $s \in S$, every path through s contains a total state. In § 2.2 of this chapter the notion of a total state has been defined in terms of positive and negative valuations. Since, presently, we are only considering positive models, we must slightly adjust our general definition. In a positive model a state s will be called *total* iff $s \Vdash^+ p$ or $s \Vdash^+ \neg p$ for all atomic sentences p.

We must distinguish strong completeability from weak completeability. A model is called *weakly completeable* if and only if for all $s \in S$ and all formulas φ of $\mathcal{L}$, every path through s contains a state s' in which φ is either true or false. The difference is that in the case of weak completeability every sentences will have a definite truth-value at some instance, but, in contrast to strong completeability, it need not be the case that there is an instance at which every sentence has a definite truth-value.

Let us consider strong completeability first. To require it would not make any difference to the logic. Let M be a counter model to $\Delta \Vdash^+ \varphi$.

[26] See the 'ridge-model' described on page 67.

Then, by the finite model property, there is a finite counter model M^*. Since every finite model is strongly completeable, it follows that M^* is completeable. Hence, intuitionist logic is complete with respect to the class of completeable models.

This argument makes essentially use of the fact that intuitionist logic has the finite model property. It establishes that the expressions '$\boxplus\varphi$' and '$\neg\neg\varphi$' have the same meaning, if we limited ourselves to strongly completeable models. It is, however, not necessary to exclude certain models in order to ensure that $\boxplus\varphi$ and $\neg\neg\varphi$ express the same. That is, we can make the slightly stronger claim that, if, in some model for intuitionist logic '$\boxplus\varphi$' and '$\neg\neg\varphi$' do not have the same meaning, we can extend this model so that they will become equivalent, without affecting the logic:

1. PROPOSITION. *Every model $M = \langle S, V^+, \geqslant\rangle$ for intuitionist logic such that $\not\Vdash^+_M \boxplus\varphi \leftrightarrow \neg\neg\varphi$ can be extended to a model M^* such that:*

1. $\Vdash^+_{M^*} \boxplus\varphi \leftrightarrow \neg\neg\varphi$
2. $s \Vdash^+_M \varphi \Leftrightarrow s \Vdash^+_{M^*} \varphi$ *for all states $s \in S$ and formula $\varphi \in \mathcal{L}$.*

Such an extended model M^* wil be called a *completed model*. The following defintion lays down how it can be obtained:

12. DEFINITION. (Completed model) *Let $KI = \langle S, \leqslant, V\rangle$ be a Kripke-model for IL. The* completed model *KI^* is the model obtained from KI by adding to each path P of KI a state s_p such that:*

1. *For each $s \in P$: $s_p > s$;*
2. *For all atomic formula $\varphi \in \mathcal{L}$: $s_p \Vdash^+ \varphi$ iff $\forall s \in P \exists s' \geqslant s$: $s' \Vdash^+ \varphi$.*

A completed model is a Kripke model. We can see that completion preserves persistency as follows. Suppose that for some $s \in P$: $s \Vdash^+ \varphi$. Then, by persistency, $\forall s' \in P$: $s' \geqslant s \Rightarrow s' \Vdash^+ \varphi$. At the same time we have for all $s'' \leqslant s$: $\exists v \geqslant s''$: $v \Vdash^+ \varphi$, namely s itself. Hence, by definition 12: $s_p \Vdash^+ \varphi$.

If KI is a Kripke model, and KI^* its completion, we can easily prove by induction on the complexity of formulas that, for all $s \in S$: $s \Vdash^+_{KI} \varphi \Leftrightarrow s \Vdash^+_{KI^*} \varphi$. It is also not difficult to show that for any completed model KI^* it holds that $\Vdash^+_{KI^*} \boxplus\varphi \leftrightarrow \neg\neg\varphi$.

Every completed model is strongly completeable, in the sense that every path contains a state in which either φ or $\neg\varphi$ is forced. For, if a path did not already contain such a state, it will contain one as soon as a state s_p is added to it in accordance with definition 12. Suppose that $s_p \nVdash^+ \varphi$. Then $\exists s \in P$: $\neg\exists s' \geqslant s(s' \Vdash^+ \varphi)$. But then $s \Vdash^+ \neg\varphi$. By persistency: $s_p \Vdash^+ \neg\varphi$.

As a corollary we find that for all $s \in S$: for all strongly completeable models M $s \Vdash^+_M \varphi$ iff for all weakly completeable models M' $s \Vdash^+_{M'} \varphi$. From right to left because strongly completeable model is also weakly completeable. From left to right, since the completion of a weakly completeable model is strongly completeable.[27]

We have seen that every Kripke model M is equivalent to a Kripke model M' such that $s \Vdash^+_M \neg\neg\varphi$ iff $s \Vdash^+_{M'} \boxplus\varphi$, while $\Gamma \nVdash^+_M \psi$ iff $\Gamma \nVdash^+_{M'} \psi$. Informally, this means that the intuitionist interpretation of negation may be a sufficient condition for intuitionist logic, but not a necessary one: we can retain the logic even if we decide to interpret '$s \Vdash^+ \neg\neg\varphi$' as 'in state s there is indirect evidence for φ.' The proper conclusion to draw, then, is that intuitionist logic lacks the expressive power to express the difference between the intended interpretation of $\neg\neg\varphi$ and the meaning given to $\boxplus\varphi$.[28]

Let me end this subsection by pointing out that under the definition of negation we have been using so far we cannot express the difference between direct and indirect verification of a statement of the form '$\neg\varphi$'. In other words, we have: $\Vdash^+ \boxplus\neg\varphi \leftrightarrow \neg\varphi$. Suppose that for some s there is bar B such that for every $s' \in B$: $s' \Vdash^+ \neg\varphi$. This can only mean that s' has no extension in which φ is true. Consequently, s has no extension in which φ is true, hence $s \Vdash^+ \neg\varphi$.

What this says is that the notions of 'being false' and 'impossible to become true' coincide. This is also quite typical for intuitionist logic in which $\neg\neg\neg\varphi$ and $\neg\varphi$ are equivalent. This seems quite appropriate when we limit ourselves to mathematical reasoning, but less so when we consider everyday reasoning. So, the natural question is: Is there anything in which mathematical deduction differs from ordinary reasoning

[27] Although it makes no difference for the logic which of the two we adhere to, it could make a difference philosophically, since the basic realistic thought that 'there is something in virtue of which our sentences are true or false', seems to demand strong completeability.

[28] We may suspect far-reaching consequences for Dummett's thesis that two pictures of reality differ from each other only if they have different practical, i.e., logical consequences. What we see here is that due to a lack of expressive power a logic may not be able to discriminate between contrasting semantic interpretations. And I would say that, for example, the picture of the construction of mathematical reality as being a completeable process is very different from one that is essentially not completeable

that can account for this?

Let us take a closer look. Given the intuitionist notion of negation, the interpretation of a formula φ is determined completely by the set of states in which φ is true. In other words, the interpretation I is I^+–dependent. And this is acceptable, I claim, only in case there are no contingencies, but only contradictions and necessary truths. And the reason then that intuitionist logic fits mathematical reasoning so well is that an intuitionist, too, holds that mathematical reality contains no contingencies. Consequently, all mathematical statements are either necessarily true, or lead to a contradiction. And this justifies the strong interpretation of '$\neg\varphi$ is true' as 'φ is a contradiction.' A statement is a contradiction if it carries the information that the world is differently from what it says.[29] From the constructive point of view, this means that φ is a contradiction if it is impossible to construct a reality in accordance with what φ says. And the negation of a contradiction just says that such a construction is impossible, which does imply that such a construction will never be established.

If our picture of reality does include contingencies, statements of the form $\neg\varphi$ cannot in general be so interpreted. If φ is a contingency we can never be sure that, if φ has not been realized so far, that it cannot be realized at all. So $\neg\varphi$ carries different information depending on whether φ is a contingency or not.[30]

4.2 Primitive Falsification

What happens if we treat falsification on a par with verification, that is, if we not only list truth conditions, but falsity conditions as well? This can, but need not, affect the logic. Much depends on additional requirements. And this means that an appeal to falsity conditions does not imply that the semantic theory in question takes falsification as a primitive notion.

Why should we take falsification as primitive? One reason might be that outside the realm of mathematics we cannot but reckon with contingencies, for which the intuitionist interpretation of negation seems rather inadequate. But could we not also derive a strong argument

[29] Cf. Landman 1986, p. 37.

[30] The notion of a contingency is more general than that of a contradiction. The latter is a special case —a border case to borrow Wittgenstein's terminology— of the former.

against primitive falsification, by asking what acceptance of it commits us to? To see whether primitve falsification has unacceptable consequences we will now inquire whether it is possible to render the notion of direct falsification philosophically intelligible.

The notion of a direct falsification does appear to be problematic when we consider perception as the paradigmatic source of *direct* evidence. Assuming that we can only perceive the presence of things, negative entities[31] are nowhere to be met in experience. It follows that, if the absence or lack of a property can be known at all, it can be known only indirectly, i.e., by inference from positive entities which can be known directly. In Taylor 1952 this assumption is rejected as a 'curious prejudice' and it is put forward that we can perceive positive as well as negative entities.[32] Taylor presents us with a picture of two circles, one of which containing a dot, while the other one is empty. He then claims that, just as in the case of the circle with dot we cannot cite any other evidence for our knowledge of the fact that it contains a dot than simply the fact that it does, we have no other reasons to believe that the other circle does not contain a dot than the fact that it is empty.

Undoubtedly, Taylor is right in claiming that our knowledge that the circle does not contain a dot *follows* from the knowledge that it is empty. But as such the example fails to demonstrate that the former knowledge can be obtained directly, for to know that it contains no dot, and to know that it is empty can certainly not be equated. What remains is that we might have direct knowledge from the fact that the circle is empty. But it is not clear that this can be taken as a support for the claim that negative entities may be directly perceived.

Let us, for the sake of the argument, assume that we cannot have direct knowledge of negative entities. Is this sufficient for rejecting the notion of a direct falsification as unintelligible? No, not if we can make any sense out of the idea that a statement is false if its truth is incompatible with that of other statements that are already known to be true. We will then say that it can be directly known that a statement φ is false, because it can be directly known that facts obtain that are incompatible with the occurrence of a fact that would make φ true.

[31] I use the notion 'entity' as a general term comprising things, facts, and whatever else can be thought of as a constituent of the sensible reality.

[32] Taylor 1952, p. 444.

The appeal to incompatibility requires an adjustment of our conception of a fact, if facts are conceived of as the ultimate constituents of reality. If the world, at every moment, is conceived as the totality of all facts, then nothing can count as a fact unless it is actually the case. But to say that some fact f is incompatible with some other fact g is to imply that f does not obtain if g does. This presupposes that we can conceive of f even if it does not obtain, in other words, if f is looked upon as a *possible* fact. This means that a fact can no longer be seen as a constituent of the actual reality, but only as a constituent of a possible reality, a possible world, that is, as a constituent of our representation of the actual world.

Does this suffice to avoid a commitment to negative facts? Not completely, for we still have to answer the following question: How do we know that, say, the fact that something is yellow all over is incompatible with the fact that it is green all over? This will definitely bring us into trouble if we assume that a belief, or a statement, if it is true, is made true by the facts. The only way out is to reject the claim that *every* true belief or statement is made true by the facts. In particularly beliefs concerning the incompatibility of two facts are themselves not made true by the facts. But if it are not the facts, then what makes such beliefs true?

The best way of approaching this question is by arguing that it is the wrong question. Starting point of such an argument would be that incompatibility statements explicate the conditions that a representation of reality has to meet if it is to be intelligible to us. That is, they constitute the possibility of such representations, but are not themselves part of them. I will not pursue this matter here, and say only a few words to support the claim that reality is underdetermined with respect to incompatibility claims.

As physics tells us our perception of a spot having a certain colour depends on the frequency of the light waves that are reflected by that particular spot. It is also known that there is not one, but a range of frequencies that we experience as, say, red. Now imagine a creature that experiences a difference as clear as that between red and orange for us between the first half of the range of 'red–frequencies,' and the second half. He even has different names for them: 'rred', and 'redd'. For him, clearly the fact that something is rred is incompatible with the fact that it is redd, while for us it is not: the representation of something that is both rred and redd is perfectly intelligible. This little thought experiment merely illustrates that the actual occurrence of a spot reflecting light waves having a certain frequency does not in itself determine any experience of it as an experience of the fact that it is red. And if it cannot do this, it certainly cannot determine which facts are incompatible.

Leaving this for what it is, let us return to the general ideas underlying the notion of incompatibility. The first is that facts are now located at the level of representation. The second is that not every combination of facts, as a representation of (a part of) reality, makes sense to us, and that whether or not some combination does, is not itself a fact. The latter has the implication that we must give up the thought that *every* statement, if it is true, is made true by the facts. We can find a more formal elaboration of these ideas in Veltman 1981, and Landman 1986, where we can find conditions a state of information has to meet if it is to grow into a complete representation of reality. We may add that a state of information presents a representation of reality that is intelligible to us only if it can be extended to a complete representation. A set of facts R is said to be complete if for every possible fact f it holds that either $f \in R$ or there is some fact $g \in R$ such that g and f are incompatible. Note that a complete representation need not be a correct representation. Note further that it is not assumed that for every fact there is a unique fact, or set of facts, it is incompatible with.

Due to this latter feature the approach under discussion avoids a commitment to negative entities,[33] while it admits of an acceptable explanation of the notion of direct falsification: a statement φ is directly falsified in a state of information s if s contains some fact g that is incompatible with the fact f that would make φ true.

Does this approach makes an appeal to negative entities redundant in all cases? It works fine for the classical cases like 'The vase is not red.' But in such cases it is clear that we can find at least one fact the occurrence of which is incompatible with the fact that it is red. But now consider the statement 'Water has no colour.' With Taylor we might ask what positive fact, or facts there could be that are incompatible with the fact that water has a colour. Perhaps, in this case, we could say that in contrast to the vase being green, that water is colourless

[33]This can also be seen as a rationale for representing information structures as lattices that are not Boolean algebras; in the latter but not in the former every element has a unique complement.

is a necessary property, which means, in the current context, that we would not accept anything as an intelligible representation of reality, if it would represent water as having a colour. Hence it is not that I hold that statement to be true on the basis of the information I presently have at my disposal, but it is because this is true, that I can distinguish between intelligible and unintelligible representations of reality.

An opponent need not take the trouble of pointing at the weakness of this reply. He may simply present an example of a contingent fact for which it is equally unclear that if it does not occur there must be some incompatible fact. Russell once told how he could not bring Wittgenstein to accept that the statement 'There is a rhinoceros in this study' is false. We can now see why. The first thing to realize is that it is highly unlikely that we will find any positive fact that is incompatible with the fact that there was a rhinoceros in Russell's study. And if, in addition to this, we reject negative facts, the absurd consequence seems to be that, indeed, there is no reason at all to believe that it is false that there was a rhinoceros in his study.

This is not a peculiar property of Russell's study, it holds for mine, too. And, admittedly, I see no quick way out that does not imply a commitment to negative facts. It does not help a bit to say, for example, that it is incompatible with the fact that I do not keep animals at the institute. In the first place not, because in a sense it reiterates the original statement by generalizing it, and in the second place not because the same kind of problem will reappear as soon as we are asked what fact is incompatible with the fact that I do keep animals at the institute.

Our tentative conclusion must be that by appealing to the notion of incompatibility we have not succeeded completely in avoiding negative facts. However, the examples discussed in the last paragraphs are equally difficult to deal with on the traditional interpretation of the notion of fact. Hence, it may very well be that there are independent reasons on the basis of which a commitment to negative entities cannot be avoided. But if that is so, then a rejection of the idea of a direct falsification, because acceptance of it would imply a commitment to negative entities can easily be refuted. On the other hand, if, eventually, we will succeed in explaining direct falsification in terms of incompatibility, there can also nothing be to prevent us from taking direct falsification serious. Either way, commitment to negative entities just does not provide a good reason to reject direct falsification.

4.3 Supervaluations

In § 3.3 of this chapter we saw how we can obtain a Kripke semantics for classical logic by imposing cofinality for all formulas on a Kripke semantics for intuitionist logic. We have considered only positive models for the positive fragment of the language, but the addition of negation, defined as above, is unproblematic. That is, no other requirements, and no other modifications are needed to isolate the class of partial models that is characterized by the set of all the classical tautologies in $\mathcal{L}$.

There is another way to delineate the class of models that are so characterized. It is known as the *super-truth* approach (SV).[34] Under this approach we confine ourselves to completeable models. A completeable model, to recall, is model of which every path contains a total state. We now define super-truth as follows:

13. Definition. (Super-truth) *Let* $M = \langle S, V, \leqslant \rangle$ *be a model. Then we have super-truth iff, for all* φ,

1. $\Vdash^+_M \varphi$ *iff* $s \Vdash^+_M \varphi$ *for all total states* $s \in S$
2. $\Vdash^-_M \varphi$ *iff* $s \Vdash^-_M \varphi$ *for all total states* $s \in S$.

The principle of super-truth boils down to the requirement of cofinality for all formulas. The difference between the super-truth approach and Kripke semantics for classical logic is that, in contrast to the latter, the former does not define negation in terms of implication and falsum, but, instead, distinguishes between truth and falsity conditions

For a Kripke semantics for classical logic, the definition of negation in terms of implication and falsum is essential, because it induces, what I have called, I^+– dependency.[35] Following the super-truth approach this is established by imposing *completeability*, the requirement that every path contains a total state. In fact, completeability and I^+– dependency are equivalent. For suppose that there is a path P that does not contain a total state. Then, for some p there is not a $s \in P$ such that either $s \Vdash^+ p$ or $s \Vdash^- p$. Hence, $I^- \neq S \sim I^+$.Conversely, suppose that I^+– dependency does not hold, then there is a state s and an atomic sentences p such that: $\neg\exists s' \geqslant s(s' \Vdash^+ p)$, while $s \notin I^-$. Hence, there is a path P through s such that for all $s' \in P$ neither $s' \Vdash^+ p$, nor $s' \Vdash^- p$.

---

[34]See Fine 1975.

[35]See page 57.

Given completeability it can be demonstrated that definition 13 is the only truth definition that respects the principles of persistency, fidelity, and resolution. By *fidelity* Fine refers to the requirement that in complete states the distribution of truth-values is classical for all formulas and not just for the atomic sentences.[36]

To require *resolution* is to require that if a sentence φ is undetermined in some state u, then there is a state $v \geqslant u$ that verifies φ, and there is a state $w \geqslant u$ that falsifies φ. Resolution can be seen as a reformulation of the requirement of cofinality. The only difference being that now we do not only consider what is verified in total states, but also what is falsified in these states. Just as in the case of Kripke's semantics for classical logic it is not immediately obvious how we are to interpret the logical constants if we are to ensure resolution for all formulas given resolution for the atomic sentences. The following truth conditions will do the job:

1. $s \Vdash^{+} \neg\varphi$ iff $s \Vdash^{-} \varphi$
   $s \Vdash^{-} \neg\varphi$ iff $s \Vdash^{+} \varphi$
2. $s \Vdash^{+} \varphi \wedge \psi$ iff $s \Vdash^{+} \varphi$ and $s \Vdash^{+} \psi$
   $s \Vdash^{-} \varphi \wedge \psi$ iff $\forall v \geqslant s \exists w \geqslant v (w \Vdash^{-} \varphi$ or $w \Vdash^{-} \psi)$
3. $s \Vdash^{+} \varphi \vee \psi$ iff $\forall v \geqslant s \exists w \geqslant v (w \Vdash^{+} \varphi$ or $w \Vdash^{+} \psi)$
   $s \Vdash^{-} \varphi \vee \psi$ iff $s \Vdash^{-} \varphi$ and $s \Vdash^{-} \psi$.
4. $s \Vdash^{+} \varphi \rightarrow \psi$ iff for all $v \geqslant s$, if $v \Vdash^{+} \varphi$, then $v \Vdash^{+} \psi$
   $s \Vdash^{-} \varphi \rightarrow \psi$ iff $s \Vdash^{+} \varphi$ and $s \Vdash^{-} \psi$.

Using these truth conditions we can now proof by induction on the complexity of formulas that resolution holds for all formulas given resolution for atomic sentences.

It should be noticed that the super-truth approach yields exactly the same truth conditions as does Kripke semantics for classical logic. This, of course, was to be expected if only because super-truth just boils down to the requirement of cofinality for all formulas, given fidelity, completeability and persistency. In part, this establishes the thought that both semantic theories are *equivalent*. I will call two semantic theories $\mathcal{T}$ and $\mathcal{T}^*$ equivalent if it holds that, for any model $\langle S, \geqslant, V_s \rangle$, and any $s \in S$:

[36]Strictly speaking, the notion of fidelity extends the notion of completeability. To list it separately is in a sense redundant, because, given that bivalence holds for atomic sentences, we can derive that it also holds for complex formula by means of the truth-conditions.

1. $s \Vdash_T^+ \varphi$ iff $s \Vdash_{T^*}^+ \varphi$
2. $s \Vdash_T^- \varphi$ iff $s \Vdash^- T^* \varphi$

As far as the equivalence of KC and SV is concerned, what is left to be shown is that they also have the same falsity conditions. How can this be done, since KC does not explicitly lay down any falsity conditions at all?

Given that SV yields the following truth-condition for negation:

$$s \Vdash^+ \neg\varphi \text{ iff } s \Vdash^- \varphi,$$

and given that in KC we have $\neg\varphi =_{\mathrm{df}} \varphi \to \bot$ we would establish that SV and KC have the same falsity conditions by showing that:

$$s \Vdash_{KC}^+ \varphi \to \bot \text{ iff } s \Vdash_{SV}^- \varphi$$

The, in this respect, crucial observation has already been made: in both semantic theories it holds, for any atomic p that $I^-(p) = \{s \in S \mid \neg\exists s' \geqslant s : s' \in I^+(p)\}$. By induction we find that this also holds for complex formulas, thus proving their equivalence.

We see that, despite the differences in their appearance, both KC and SV embody the same concept of negation. One conclusion to draw from this is that under the super-truth approach the concept of falsification is not primitive, but can be expressed by a statement that appeals only to the concepts of verification, implication and falsum. I will return to this below and relate it to the inability to express the difference between direct and indirect falsification, which, of course, is inherent to SV as it is to KC.

4.4 A Summing Up

Having found that the super-truth approach embodies a concept of negation that, despite the appeal to falsity conditions, is essentially the same as that underlying definition II.1 (page 72) we are now in a position to make the following observations.

The first is that, given definition II.1, the class of cofinal models is characterized by the formula $\neg\neg\varphi \to \varphi$. In other words, let $\mathcal{C}$ be the class of cofinal models. We have:

$$\Vdash_M^+ \neg\neg\varphi \to \varphi \Leftrightarrow M \in \mathcal{C}$$

From left to right: consider some model M such that $M \notin \mathcal{C}$. Then there is a state s_0 in S such that, for some formula φ, it holds that

$\forall v \geqslant s_0 \exists u \geqslant v(u \Vdash_M^+ \varphi)$, while $s_0 \nVdash^+ \varphi$. Now suppose that $\Vdash_M^+ \neg\neg\varphi \to \varphi$. Given definition II.1 we have that $s \Vdash^+ \neg\neg\varphi$ iff $s \Vdash_M^+ (\varphi \to \bot) \to \bot$ iff every successor of s' of s has a successor s'' that validates φ, i.e., iff $\forall s' \geqslant s \exists s'' \geqslant s'(s'' \Vdash^+ \varphi)$. Hence we have $s_0 \Vdash_M^+ \neg\neg\varphi$, and thus: $s_0 \Vdash^+ \varphi$. Contradiction. Hence not $s_0 \Vdash_M^+ \neg\neg\varphi \to \varphi$. And, consequently, not $\Vdash_M^+ \neg\neg\varphi \to \varphi$.

Conversely, suppose that M does not validate $\neg\neg\varphi \to \varphi$. Given the meaning of '$\neg\neg\varphi$' under the current definition of negation, it follows that there must be some s, such that $\forall s' \geqslant s \exists s'' \geqslant s'(s'' \Vdash^+ \varphi)$, and $s \nVdash^+ \varphi$. Hence $M \notin \mathcal{C}$.

The following observation is analogous to the previous one. It connects the law of excluded middle to completeability. Let the meaning of '$\vee$' be determined by:

$$s \Vdash^+ \varphi \vee \psi \text{ iff } \forall s' \geqslant s \exists s'' \geqslant s'(s'' \Vdash^+ \varphi \text{ or } s'' \Vdash^+ \psi)$$

Let $\mathcal{C}'$ be the class of completeable models. We have:

$$\Vdash_M^+ \varphi \vee \neg\varphi \Leftrightarrow M \in \mathcal{C}'$$

From right to left is trivial. From left to right almost: suppose $M \notin \mathcal{C}'$. Then there is a state s such that, for some formula φ, it holds that $\exists s' \geqslant s \forall s'' \geqslant s'(s'' \nVdash_M^+ \varphi \text{ and } s'' \nVdash_M^+ \neg\varphi)$. Hence, $s \nVdash_M^+ \varphi \vee \neg\varphi$. And thus $\nVdash_M^+ \varphi \vee \neg\varphi$.

These two observations generalize the remarks made when discussing Kripke semantics for classical logic and the super-truth approach. The latter supports the claim that completeability is a necessary condition for classical logic. When discussing the meaning of double negation in intuitionist logic we already saw that it is not sufficient. The former advocates the assertion that cofinality is necessary for classical logic.

5 Constructible Falsity

The following semantics is the propositional fragment of the semantics presented in Thomason 1969. There, Nelson's logic for constructible falsity is proven to be complete with respect to the full semantics. It has one important feature in which it differs from the semantic theories discussed so far. To speak with Thomason:

> The falsity, for instance, of a formula such as $P(a)$ at a stage of construction is not defined in terms of the failure of the individual named

> by a to have the property corresponding to P; rather, this falsity is conceived of as a feature which is discovered directly, as the truth of $P(a)$ is discovered.[37]

Although Thomason takes a first-order sentence as example, the point should be clear; the present semantics intends to take falsification as serious as it does take verification. That results in the following truth and falsity conditions for our language:

1. $s \Vdash^{+} p$ iff $V_s(p) = 1$
   $s \Vdash^{-} p$ iff $V_s(p) = 0$, for atomic p
2. $s \Vdash^{+} \neg\varphi$ iff $s \Vdash^{-} \varphi$
   $s \Vdash^{-} \neg\varphi$ iff $s \Vdash^{+} \varphi$
3. $s \Vdash^{+} \varphi \wedge \psi$ iff $s \Vdash^{+} \varphi$ and $s \Vdash^{+} \psi$
   $s \Vdash^{-} \varphi \wedge \psi$ iff $s \Vdash^{-} \varphi$ or $s \Vdash^{-} \psi$
4. $s \Vdash^{+} \varphi \vee \psi$ iff $s \Vdash^{+} \varphi$ or $s \Vdash^{+} \psi$
   $s \Vdash^{-} \varphi \vee \psi$ iff $s \Vdash^{-} \varphi$ and $s \Vdash^{-} \psi$
5. $s \Vdash^{+} \varphi \rightarrow \psi$ iff for all $u \geqslant s$, if $u \Vdash^{+} \varphi$ then $u \Vdash^{+} \psi$
   $s \Vdash^{-} \varphi \rightarrow \psi$ iff $s \Vdash^{+} \varphi$ and $s \Vdash^{-} \psi$

What immediately attracts attention is that, except for negation, the truth-conditions as laid down by Nelson's semantics (NS) are exactly the same as those given by Kripke's semantics for intuitionist logic. That means that, trivially, we have: $s \Vdash^{+}_{KI} \varphi \Leftrightarrow s \Vdash^{+}_{NS} \varphi$, for the positive fragment of $\mathcal{L}$. The question then arises whether these two theories are equivalent? They are not, not even if we limit ourselves to the positive fragment of the language. Suppose they are. Then it would hold that:

$$s \Vdash^{+}_{KI} \varphi \rightarrow \bot \Leftrightarrow \Vdash^{-}_{NS} \varphi$$

But since in NS we have that $s \Vdash^{-}_{NS} \varphi$ iff $s \Vdash^{+}_{NS} \neg\varphi$, that would amount to:

$$s \Vdash^{+}_{KI} \varphi \rightarrow \bot \Leftrightarrow \Vdash^{+}_{NS} \neg\varphi$$

which is obviously not true. It holds from right to left, but the following is a counterexample to the converse. Consider a model consisting of only two states s and s' such that $s \leqslant s'$. Let $I^{-}(p) = \{s'\}$. Then we have: $s \Vdash^{+}_{KI} \neg p$, but $s \nVdash^{+}_{NS} p$.

In contrast to NS, in KI falsification is not primitive, or, what comes down to the same, the interpretation of a sentence p is dependent

[37] Thomason 1969, p. 251.

on $I^+(p)$. Now the thought may come up that, in analogy with imposing cofinality on KI–models, we might obtain intuitionist logic by imposing *negative cofinality* on NS–models. By negative cofinality I understand the condition:

$$\forall s' \geqslant s \exists s'' \geqslant s'(s'' \in I^-(p)) \Rightarrow s \in I^-(p)$$

This will not do, however, for the simple reason that I^+– dependency implies negative cofinality, but not the other way round. Suppose that I^+– dependency does not imply negative cofinality. Then there is a state s and a formula p such that $\forall s' \geqslant s \exists s'' \geqslant s'(s'' \in I^-(p))$, while $s \notin I^-(p)$. By I^+– dependency it follows that there must be a $u \geqslant s$ such that $u \in I^+(p)$. But then $\exists v \geqslant s$, $v \in I^+(p)$ and $v \in I^-(p)$.

A counterexample to the converse is given by the following model. Let $S = \{s, u, v\}$, and $s \leqslant u$, and $s \leqslant v$. Let $I^-(p) = \{v\}$. Negative cofinality holds, but while $\neg \exists s' \geqslant s(s' \in I^+(p))$, $s \notin I^-(p)$.

6 Expressive Power

In intuitionist logic, as we saw, we cannot express the distinction between direct and indirect falsification. The definition of negation induces I^+– dependency, which implies negative cofinality. Hence, if for some s there is future bar B such that for every $s' \in B$: $s' \Vdash^- \varphi$, then, by negative cofinality, $s \Vdash^- \varphi$. However, as we have seen, in intuitionist logic we can express the difference between direct and indirect verification, since KI–models do not satisfy the requirement of positive cofinality. On the supertruth approach models are required to satisfy I^+– dependency, or equivalently, completeability, and hence negative cofinality, and in addition to add positive cofinality. Consequently, in classical logic we cannot express the difference between direct and indirect verification, nor that between direct and indirect falsification. In contrast to these, NS–models satisfy neither negative, nor positive cofinality. It follows that in Nelson's logic for constructible falsity both distinctions are expressible.

It thus turns out that Nelson's logic exceeds intuitionist logic in expressive power, which, in its turn, exceeds that of classical logic. This is the informal content of the following claim. Let $M = \langle S, \leqslant, V \rangle$ be a partial model. Then there are valuations $\Vdash_c^+, \Vdash_i^+$, and $\Vdash_n^+$, and there are translations φ° and $\varphi^\star$ such that for any $s \in S$:

$$s \Vdash_c^+ \varphi \Leftrightarrow s \Vdash_i^+ \varphi^\circ \Leftrightarrow s \Vdash_n^+ \varphi^\star \tag{II.2}$$

Given V, a partial function from the atomic sentences of $\mathcal{L}$ into $\{1, 0\}$, define the following valuations:

$s \Vdash_c^+ \varphi$ iff there is a bar B_s for s such that $\forall s' \in B_s : V_{s'}(\varphi) = 1$
$s \Vdash_c^- \varphi$ iff there is a bar B_s for s such that $\forall s' \in B_s : V_{s'}(\varphi) = 0$
$s \Vdash_i^+ \varphi$ iff $V_s(\varphi) = 1$
$s \Vdash_i^- \varphi$ iff there is a bar B_s for s such that $\forall s' \in B_s : V_{s'}(\varphi) = 0$
$s \Vdash_n^+ \varphi$ iff $V_s(\varphi) = 1$
$s \Vdash_n^- \varphi$ iff $V_s(\varphi) = 0$

And the following translations:

$p^\circ = \neg\neg p$	$p^\star = p$
$(\neg\varphi)^\circ = \neg\varphi^\circ$	$(\neg\varphi)^\star = \varphi^\star \to \bot$
$(\varphi \wedge \psi)^\circ = \varphi^\circ \wedge \psi^\circ$	$(\varphi \wedge \psi)^\star = \varphi^\star \wedge \psi^\star$
$(\varphi \vee \psi)^\circ = \varphi^\circ \vee \psi^\circ$	$(\varphi \vee \psi)^\star = \varphi^\star \vee \psi^\star$
$(\varphi \to \psi)^\circ = \varphi^\circ \to \psi^\circ$	$(\varphi \to \psi)^\star = \varphi^\star \to \psi^\star$

Claim II.2 can now be proved by induction. As far as translation φ° is concerned that proof is well-known. Translation $\varphi^\star$ is trivial; it just maps every formula of intuitionist logic onto itself.

We may draw the following conclusions. Since $\Vdash_c^+ \varphi \leftrightarrow \varphi^\circ$, but $\nVdash_i^+ \varphi \leftrightarrow \varphi^\circ$, it follows that in intuitionist logic 'φ°' expresses something that cannot be expressed in classical logic. In the same vein, since $\Vdash_i^+ \neg\varphi \leftrightarrow (\varphi \to \bot)$, but $\nVdash_n^+ \neg\varphi \leftrightarrow (\varphi \to \bot)$, what is expressed in Nelson's logic by means of '$\neg\varphi$' cannot be expressed in intuitionistic logic.

7 Logical Revisionism Vindicated?

In this chapter we have taken issue with the claim that supervaluations, or, for that matter, any partial semantic theory equivalent to the theory of supervaluations, presents a counterexample to the anti-realist claim that a rejection of classical truth-conditional semantics implies a rejection of classical logic. Despite supervalutions an anti-realist can still hold on to the claim that, given the basic tenets of semantic revisionism, bivalence is a necessary condition for classical logic, because, from the point of view of semantic revisionism, the distinction between direct and indirect evidence cannot be dispensed with, while, at the same time, this distinction cannot be expressed in any language which is interpreted in terms of a partial semantics that validates classical logic.

The second part of this argument has been established in two steps. First, it was pointed out that the semantic restriction embodied by *cofinality* is necessary for a validiation of classical logic by reference to partial models. Then it was demonstrated that cofinality precludes the expression of the distinction between direct and indirect evidence. The latter step can succinctly be stated as follows. Let $IV = \{u \mid \forall u' \geqslant u \exists u''(u'' \Vdash^{+} \varphi)\}$, and $DV = \{u \mid u \Vdash^{+} \varphi\}$. IV represents the meaning of the expression 'being an indirect verification of φ. Consider an arbitrary $s \in IV$. Then, $\forall s' \geqslant s \exists s''(s'' \Vdash^{+} \varphi)$. By cofinality: $s \Vdash^{+} \varphi$. Hence $s \in DV$. A similar argument can be given for (in)direct falsification.

The comparison between Kripke semantics, supervaluations and Nelson's semantics has given rise to the following observations. Given a Kripke semantics for classical logic, or, equivalently, the super-truth approach, we cannot express either distinction. On the basis of a Kripke semantics for intuitionist logic we can express only the difference between direct and indirect verification. We also saw that as far as falsification is concerned, Kripke semantics for intuitionist logic recognizes only indirect falsification. As we saw, this is due to the specific definition of negation. This definition, which is characteristic for intuitionist logic, plays a crucial role in the reduction of the notion of direct falsification to that of verification. Only in Nelson's semantics, which maintains a basically different notion of negation such a reduction is not possible.

With respect to expressibility, Nelson's logic is the strongest of the three logics considered in the sense that in it one can express more conceptual distinctions than in the other two. These findings, taken together, establish the thesis that, in a partial setting, we cannot retain classical logic without a considerable loss of expressive power in the logical language. Assuming that this is unacceptable from the point of view of anti-realism, this provides sufficient reason for rejecting supervaluations as a counterexample to the anti-realist claim that a revision of classical logic is needed.

There is one more point that we should not overlook. On the basis of our findings in this chapter it is not only justified to hold that the super-truth approach does not present a counterexample against logical revisionism, but also that intuitionism cannot be the paradigm of a logic that is acceptable on anti-realist standards, since the conditions that have to be imposed on a semantic theory in order to validate it require that it be impossible to express the distinction between direct and indirect falsification in the object language. The only way to overcome this is to revise the concept of negation. Hence, it is precisely there that the revision of logic advocated by anti-realism should have its beginning.

I want to conclude this chapter by briefly evaluating the argument presented here in the light of the basic tenets of Dummettian anti-realism. For anti-realism it is not only the result that counts; even more important is *how* it is brought about. For it is the way in which the result is obtained that does or does not endow it with conclusiveness. In this chapter the question as to the correct logic is approached by first asking: What are the philosophically significant concepts to distinguish? The second question is: What are the conditions that have to be imposed on a semantic theory so that we can express these conceptual distinctions in the object language? The third and last question is: What is the logic validated by a semantics so constrained?

Of these three questions only the last is purely logical. In order to answer it we prove of a particular logic that it is complete with respect to the semantics that is obtained by answering the second question. The logics discussed in this chapter, notably classical logic, intuitionist logic and Nelson logic for constructible falsity, are known to be complete with respect to, respectively, the semantic theories KC, KI, and NS.[38] A completeness proof, however, contributes to the justification of logic only in the light of our answers to the first two questions. And this means that, following the route of this chapter, in the end, the justification of logic depends on the reasons we have for making certain conceptual distinctions.

For an anti-realist this is acceptable only if the need to make these distinctions can be argued for solely on the basis of an examination of how we use our language.[39] In the case of the distinction between direct and indirect evidence it is not to be expected that this will cause grave difficulties. It remains, however, that in this chapter we have been concentrating on supervaluations, and that, as yet, we cannot exclude

[38] See, for example, Troelstra and van Dalen 1988, and Thomason 1969.

[39] This is characteristic of, what I have called elsewhere, *semantic revisionism*. See chapter I.

the possibility of other counterexamples. If there are any, then in order to refute these along similar lines, other conceptual distinctions may be needed, but there is no reason to suppose that the need for just any distinction can be argued for within the limits set by semantic revisionism.

What this chapter establishes is that, despite supervaluations, anti-realism may still have logical revisionist consequences. This, however, is only a negative thesis in that it says that a particular logic —classical logic— is *not* acceptable, given semantic revisionism. But does semantic revisionism also give rise to a positive thesis? In other words, Is it possible to determine which logic *is* acceptable, given semantic revisionism? This will occupy us in the next chapter.

III Use and Logical Revisionism

1 Introduction

In this chapter I examine what bearings the dictum 'meaning is use' may have on the choice of a logical theory. I will confine myself to the role it plays in, what I call, *semantic anti-realism* as it is developed in the writings of Michael Dummett. If we take a *normative perspective* on logic to be a cluster of assumptions which, taken together, may give us reasons to prefer one logic over another, the leading question of this chapter may also be specified as: What normative perspective does a semantic anti-realist derive from the thesis that meaning is use?

Limiting myself to the impact this thesis has for the selection of a logical theory, I will not go into the arguments that support it. However, a short sketch of its background may be helpful in positioning the current enquiry in the anti-realist landscape. As I have argued elsewhere,[1] the anti-realist's appeal to use is quite properly looked upon as being rooted in his dissatisfaction with Frege's account of the objectivity of understanding. This objectivity, the anti-realist claims, can only be accounted for if it is possible to equate the linguistic knowledge that is ascribed to a speaker with some *practical ability.* And a realist account of knowledge of meaning, the argument continues, cannot, in general, meet this requirement, and ought therefore to be abandoned.

Placed in this light, the dictum 'meaning is use' is nothing but a succinct expression of the demand that linguistic knowledge must be systematically correlated with practical abilities, and as such it can be seen to generate a requirement imposed by the anti-realist on the form that a meaning-theory, i.e., an explanation of a speaker's understand-

[1] See chapter I.

ing, should take. But what consequences does this have for the selection of a logical theory? Quite commonly, the anti-realist rejection of classical semantics, when accepted, is taken up as an incentive to develop a meaning-theory starting with a semantic theory based on verificationist starting points.[2] And, as it turns out, it is rather unlikely that we will be able to validate classical logic in terms of such a semantic theory: the logic we can validate is, or comes close to, intuitionistic logic.

This, however, is not a good reason for rejecting classical logic. It is not, because the rejection of classical truth-conditional semantics[3] does not force one to give up classical logic as well, for the simple reason that acceptance of classical logic does not necessarily commit one to classical semantics.[4] It is, furthermore, not immediately obvious that a verificationist semantics suits all the anti-realist needs. And even if it does, what tells us that it is the *only* viable alternative? For as long as we have not established this conclusively, we ought to reckon with the possibility of another, as yet unknown, semantic theory that meets the anti-realist's requirements, but that does not validate intuitionistic, but some other non-classical (or perhaps even classical) logic.[5]

So, as it stands, the anti-realist's examination of the form that a meaning-theory ought to take fails to constitute a normative perspective on logic for two reasons. In the first place, it does not provide sufficient reasons for rejecting classical logic as being incorrect. In the second place, it fails to single out a *single* correct semantic theory, and, hence, it leaves open the possibility that there may be *several* correct logics. As far as logical revisionism is concerned, then, the analysis of the form of a meaning-theory has primarily a *negative* impact: it clearly fences off the semantic theories that are to be labeled as incorrect. This does not immediately give us a criterion for distinguishing the incorrect

[2]See, for example, Prawitz 1977, Wright 1993b, Tennant 1987

[3]The difference between classical and verificationist semantics is not that the former does, while the latter does not assign a central role to truth-conditions in the analysis of meaning. On the contrary, their difference concerns how truth-conditions are to be explained: in terms of truth functions, as is characteristic for classical semantics, or in terms of verification conditions, as is preferred by a verificationist.

[4]This is not to say that the anti-realist's argument against classical semantics could not be *extended* to an argument against classical logic. But, as I argue in chapter II, such an extension depends crucially on the assumption that we should distinguish between direct and indirect evidence.

[5]For an attempt in this direction, see Cozzo 1994.

logical theories from the correct ones, but it may be a first step in that direction.

The importance, for anti-realism, of a decisive, normative perspective on logic is indisputable. On Dummett's view, to mention one reason, a metaphysical position consists, essentially, in a canon of deductive inferences that are held to be valid.[6] This reduction of metaphysics to logic, he claims, is the first step to a resolution of metaphysical disputes, since, in contrast to these disputes themselves, differences of opinion concerning logic can be settled in a definite way. Which says, in effect, that Dummett holds that it is possible to determine what *the* correct logic is.

An important question is thus whether the dictum 'meaning is use', besides having this negative impact mentioned above, may also be taken to support a *positive* claim. In other words, is it possible, by appealing to use, to single out the correct logic *directly*, instead of indirectly, by merely pointing out which logics are not correct? Yes, Dummett claims, this can be done by means of, what he calls, *proof-theoretic justification procedures*. These are invoked in order to determine the logical principles that one may appeal to in the process of giving a semantic explanation of the way one understands the logical constants. What such a semantic explanation consists in, is determined by the analysis of the form that a meaning-theory may take. Having determined what a semantic explanation *is*, we now take one step back and ask how such an explanation is *possible*. And the thesis to be examined in the sequel is that in answering this question we will come to see the logical principles that are *unconditionally correct*; that are acceptable no matter how we interpret our thoughts about reality.

To conclude this introduction I want to comment briefly on the subject of this chapter. It is not my intention, nor my ambition to answer the question: What is the correct logic? Rather, my interest in the subject is motivated by the second-order question: How can one obtain an answer to the question 'What is the correct logic?'? The reconstruction presented below of the way in which Dummett derives a normative perspective on logic from an analysis of use, or, to be more specific, from an analysis of the general *form* that a proof can take, is only a first step in answering this meta-question. I speak of a *reconstruction* of Dummett's views and this is to be taken literally. It

[6]For a more detailed discussion of this point of view, see chapter I, § 2.2.

is my contention that the question whether this reconstruction accurately describes Dummett's views, or is merely inspired by these, is less important than the question whether it presents a conceivable and characteristic way of answering the question as to which the correct logical principles are. The main goal of this chapter is to show why the latter question ought to be answered affirmatively, and, in doing so, to add to the raw material on the basis of which we may once come to more general conclusions. Once, perhaps, but not in this chapter. Here I will concentrate on the concepts in terms of which Dummett's idea of a proof-theoretic justification procedure is stated. The explanation of these concepts will show itself to be instrumental in many respects. Among other things, it will enable us to see what particular interpretation of the dictum 'meaning is use' is presupposed by the belief that the correct logical laws can be singled out by means of proof-theoretic justification procedures.

2 Self-justifying Rules

Proof-theoretic justification procedures, PT-procedures for short, may be conceived of as a method for bringing to light all and only those logical rules that are *self-justifying*. A natural response to this would be: If the notion of a self-justifying rule is to make any sense at all, then why do we need some kind of a *justification* procedure in order to establish what they are?

The distinction needed here is that between being self-justifying and being self-evident: a rule can be self-justifying without being known as such.[7] And we do not come to know that some rule is self-justifying by inspecting the form it takes, or by some immediate insight. We need a method to *show* us the self-justifying rules. The distinction between being self-evident and being self-justifying points at the basic thought that 'being self-justifying' is not to be conceived as an *intrinsic* property of a logical rule. By this I mean that in considering a logical rule to be self-justifying, we do not ascribe a property to it that it possesses due to its very nature, but one that it possesses *in virtue of our general conception of proof.* This is the key idea: to conceive of the notion of a self-justifying rule as being correlated to our general notion

[7]Cf. Dummett 1991b, p. 251: 'It is essential to develop a characterization that will allow us to recognize a set of logical laws as self-justifying by their very form.'

of proof, in the sense that the self-justifying rules are all and only those that together constitute the essence of our conception of what a logical proof is. If we look at it in this way, it is only natural to say that it is not self-evident that a rule is self-justifying, because it is only through an analysis of our conception of proof that a rule can be shown to be self-justifying.

This explanation is neutral with respect to questions of logical foundationalism. Surely, one may take it as suggesting that any rule can be shown to be self-justifying, just by adopting the appropriate conception of proof, but the notion of a self-justifying rule does not force one to hold that we can indeed modify our conception of proof at will. And, of course, a Dummettian anti-realist would deny that we can.[8] But if the foundationalist character of anti-realism resides in the status an anti-realist assigns to a particular conception of proof, then what use does it have for him to invoke the notion of a self-justifying rule in his pursuit of determining *the* correct logical rules? An anti-realist is a foundationalist, but he is not a platonist, nor does he have any affinities with the Tractarian point of view. It is through the notion of a self-justifying rule that he dissociates himself from these positions. Logical principles do not *show themselves* to be valid from beyond the borders of our intellectual capacities, but *are shown* to be valid in the light of our ability to construct proofs. And the idea of a valid principle, a logical law, would vanish as soon as we would stop using our language in that particular way: both are inextricably connected with each other. In this we read the germ of the anti-realist's reply to logical scepticism: to be able to prove something is just to be able to *use* language in a particular way. As such, the possibility of constructing a proof is intimately connected to the possibility conditions of the use of language in general. And to explicate our general conception of proof is to explicate precisely those conditions to which the use of language is subjected to, that make the construction of a proof possible. Possibility conditions cannot be denied without denying the idea of proof

[8]From this, it follows that a logical rule, if it is self-justifying, is necessarily so, and, hence, that 'being self-justifying' is an essential property of some rules. On the other hand, it has been said that a logical rule is self-justifying only in virtue of a general conception of proof. Taking this together it appears that something may possess some property essentially, in virtue of something else. This, I would say, raises some interesting question concerning the connections between the notions 'essence' and 'independence'. I will not pursue these here.

itself. So, for the anti-realist there are only two possibilities: either one accepts the notion of a proof, or one rejects it, but there is no such thing as *another* conception of proof.

An anti-realist conceives of the PT-procedures as a means of pointing out what is the least we need to be able to do, if we are to obtain anything that we can call a logical proof: 'we cannot have a proof unless we have some *means* of proof,'[9] Dummett observes, and the PT-procedures are developed in order to single out precisely these means. Hence, for an anti-realist the question whether or not we should adhere to the conception of proof as explicated by means of PT-procedures makes no sense at all. The PT-procedures, he would say, show us not just some conception of proof, but the conditions without which no proof is possible.

Looked upon in this way, the question as to what the correct logical rules are is basically a *transcendental* one. The idea that the PT-procedures are best taken as a kind of transcendental deductions is never stated explicitly by Dummett, nor by any other defender of anti-realism, at least as far my knowledge goes. However, it gains force if we try to understand what Dummett does state explicitly. How, for example, are we to understand the remark that if the logic validated by PT-procedures is not '*the* right logic, at least it may serve as a medium by means of which to discuss other logics'?[10] To see how this may be taken as a support of the present point of view we must get a clear view on how Dummett conceives of the nature of a controversy over the validity of a logical law, of what is needed to solve it, and of how the PT-procedures contribute to their solution.

2.1 Disputes over the Validity of Logical Laws

Not every dispute over the validity of a logical law is problematic. It might very well be that someone just fails to see that some principle is valid given the acceptance of certain other rules. If this could not occur, there would no point in making exercises, nor in examining a student of logic. Such a disagreement can be resolved by showing how the law under dispute can be derived from other, already accepted rules. Nevertheless, in this case, too, we may speak of the justification of a

[9] Dummett 1991b, p. 245. My emphasis.
[10] *ibid.*, p. 300.

logical law. Dummett speaks of a justification of the first grade.[11]

A disagreement between, for example, a classical mathematician and an intuitionist cannot be solved in this way. Their quarrel is not over the validity of some particular derivation, but over which principles one may legitimately appeal to in the course of such a derivation. And, thereby, also over what counts as a demonstration of the *untenability* of an alleged logical principle. It is for this reason that we cannot take an intuitionist as accusing a platonist of committing logical fallacies. Such an accusation makes sense only against the background of an agreement on what the valid principles are and on what counts as a correct application of these. For it is only then that we will be able to make someone see that he has made a logical mistake, when he has made one. In the case of platonism and intuitionism such a common background is missing. How, for example, could an intuitionist convince a classically inclined logician that it is a logical mistake to appeal to the law of the excluded middle? Only by deriving a contradiction from an instance of this law, the latter would reply. But this is exactly what an intuitionist cannot do; he cannot refute $\neg A \vee A$, since he acknowledges the validity of $\neg\neg(\neg A \vee A)$. So, his argument must take a different form. He will, for example, show that the law of the excluded middle is unacceptable by arguing that accepting it means that 'we ought to have certain knowledge which we in fact do not possess.'[12] Whether or not this is a reasonable claim, the fact remains that it is not a logical claim, and as such for a platonist not recognizable as being an ordinary and for him acceptable way of criticizing logical deductions.

The intuitionist notion of a counterexample is thus much stronger than that of a platonist. For a platonist a counterexample to some general statement just consists in a proof of the negation of an instance of it. An intuitionist, in contrast, would already accept as a counterexample a demonstration that an instance of it cannot be proved unless some other problem is solved. Given that we have a uniform way of constructing similar 'counterexamples' for each unsolved problem of the same form, we can conclude that the general statement will never be intuitionistically provable.[13] But for a platonist this would never be a strong enough reason to refrain from using the statement in the course

[11]Dummett 1991b, p. 192.
[12]Troelstra and van Dalen 1988, pp. 8-16.
[13]See Dummett 1977, pp. 44–45.

of deductive inference.

An intuitionist, then, finds himself in a situation in which the very reasons he has for rejecting certain classical principles deprive him of the means to convince a platonist that he should give up these principles. The main problem is that for a platonist a counterexample coincides with a proof of the negation of the challenged principle, while for an intuitionist these two do not, in general, coincide. They disagree not only on what follows from what, but also on what counts as a refutation of a given inference. In such a case, Dummett points out, the controversy turns on the *meanings* of the logical constants.[14] This is completely in line with the observation we made that none of the parties involved can convince the other by purely logical means to give up his position: what the *logical* means are depends to a large extent on how the connectives are interpreted.

A disagreement on the meaning of the logical constants is considered by Dummett to be 'deep' in distinction to a 'superficial' one.[15] A disagreement is *superficial* or trivial if it is a mere verbal disagreement. In that case it can be resolved by introducing two new words that both parties understand alike, each of which carries one of the meanings of the word under dispute.[16] A disagreement on the meaning of some expression is *deep* or genuine when one of the parties cannot accept the meaning attached by the other to the expression under dispute. Hence, deep disagreements

> turn on different conceptions of what it is possible to mean, and hence, ultimately, on different conceptions of what meaning is.[17]

As an example, consider the classical analysis of the meaning of the standard logical constants in terms of the familiar truth-tables. To this, it might be objected that a different set of truth-tables would be more appropriate to represent their meanings. This would be a superficial disagreement. A deep disagreement would originate not by rejecting a particular truth-table, but by putting to doubt the very possibility of

[14] Dummett 1991b, p. 302.

[15] The terms 'superficial' and 'deep' appear for the first time in Dummett 1991b. The distinction itself is made earlier. See, for example, Dummett 1992a (first published in 1973), p. 668.

[16] Dummett 1991b, p. 193.

[17] *ibid.* For an extensive treatment of the anti-realist criticism of the realist account of what meaning is I refer to chapter I.

representing the meanings of the logical constants in terms of truth-tables. The conflict between a realist and an anti-realist as Dummett conceives of it goes deeper; the latter rejects the meaning attached to the logical constants by the former as unintelligible.[18]

The distinction between deep and superficial controversies is highly theory-relative: it follows naturally from the assumption that the acceptance of a logical principle as valid is invariably tied to a particular conception of what meaning is, and that, furthermore, not any such conception is coherent. It is only because of this that an anti-realist may conceive of the realist's understanding of the logical constants as a mere illusion, and not as 'change of subject,' to use a phrase of Quine's.[19] It cannot be, an anti-realist would say, because the alleged subject changed to is just not there.

In a *genuine* dispute over the validity of logical laws the issue is not *what* meanings are to be assigned to specific logical constants, but *how*, that is, in what terms, the meaning of any logical constant ought to be framed.[20] Clearly, before we may think of solving the dispute, we need to know what the opposing positions consist in, that is, we need an explicit statement of how both disputants specify the logical constants. That is where the notion of a *semantic theory* comes in, because the proper way to explain how one understands the logical constants is by means of a semantic theory.[21]

2.2 The Explanatory Value of a Semantic Theory

Dummett takes the controversy between an intuitionist and an adherent of classical logic to be the paradigm of a genuine dispute over the validity of logical principles. Both quarrel over which logical principles are valid. But this quarrel, Dummett claims, is secondary: it derives from a difference of opinion on how the meaning of the logical constants is to be accounted for. It follows that, in order to get a clear view on what really is at stake, we need to explicate, for both disputants, the particular conception of meaning that underlies their logical preference. We do so by formulating a semantic theory. So far, so good. But now we must confront the following problem: How do we know whether a

---

[18] Dummett 1991b, p. 12.
[19] Quine 1986, ch. 6.
[20] Dummett 1991b, p. 192.
[21] *ibid.*, p. 55. See also chapter I, § 2.4.

proposed semantic theory does indeed capture the way in which one of the disputants thinks that the meaning of the logical constants is to be framed?

The only thing that makes an observable difference is the canon of deductive inferences held to be valid by either disputant. So a minimal requirement is that every acceptable derivation can be shown to be valid with respect to the proposed semantic theory. This amounts to requiring *correctness*. However, since intuitionist logic is correct with respect to any semantics with respect to which classical logic is correct, it is not sufficient.

An additional requirement is called for. And although Dummett nowhere states it explicitly I read him as demanding that the semantic theory should have, what I will call, *explanatory value*; the semantic explanation it offers should be non-trivial. By this I mean the following. A semantic theory has explanatory value if it explicates the *semantic assumptions* on which the acceptance of some logical principles as valid depends. We just saw that the least we require of a semantic theory, if it is to capture the conception of meaning underlying a logical theory, is that we can show that all the relevant principles hold in the language so interpreted. But we cannot do this, unless we assume at least some logical principles to hold in the metalanguage. Now, if a semantic theory is to have any explanatory value at all, then the fact that we can show a logical principle to hold in the objectlanguage, given this semantic interpretation, should not depend completely on the assumption that this very same principle holds good in the metalanguage, too. A semantic theory is expected to shed some light on the reasons someone has for accepting some logical principle as valid, and to say that one does so because it is held to be valid in the metalanguage is not to explain anything at all.[22]

Given the nature of the method explored by Dummett, we must assume that not every logical rule has a non-trivial semantic explanation. The reason is this. Suppose that every rule does have a non-trivial

[22]An explanation that falls prey to this kind circularity is called by Dummett a *programmatic* interpretation. It is maximally sensitive to the logic of the metalanguage, that is, it validates all and only those logical laws that are valid in the metalanguage. Dummett does not consider it to be a semantic theory proper. An example of a semantic theory that is 'maximally *in*sensitive to the logic of the metalanguage,' is Kripke's semantics for intuitionist logic. (See Dummett 1991b, p. 55.)

semantic explanation, that is, that for every rule it holds that its validity depends, at least in part, on some semantic assumption. The basic thought is that the acceptance of a logical rule is to be evaluated by evaluating the semantic assumption on which acceptance of it depends. However, in order to determine, for some rule, the semantic assumption on which it depends, we need to frame a semantic theory and, and demonstrate, at least, that the rule in question holds in a language so interpreted. To do so, we must assume at least some principles to be valid in the metalanguage. Now suppose that we do not need to assume the rule in question. Then, for the rules we do need to assume it still would be the case that we can only be sure that it is legitimate to accept them as valid, if it is legitimate to make the semantic assumptions on which they depend. But in order to determine whether or not that is the case, we need to frame a semantic theory ..., and so on. Hence, we conclude that the assumption that every rule has a non-trivial semantic justification leads to a 'backwards' regress: the argument shows that, in that case, we will never reach a position in which we can *start* justifying our rules.

Returning to a quotation mentioned earlier,[23] we may now specify the '*medium* by means of which we can discuss other logics' as consisting of precisely those rules that do not have a non-trivial semantic explanation. The above argument purports to show only that there must be such rules, it does not say anything about how to find them. Clearly, trivial semantic explanations will not be of much help in this respect. So, it appears that the method we are looking for could not be called a semantic explanation at all. It will be one that precedes such explanations, that makes them, in a sense, possible.

I must now adjust the contention expressed earlier that what the logical means are, depends on how the logical constants are interpreted.[24] This is not to be taken as saying that every interpretation yields a completely disjoint set of logical means. We can see now that, at least according to the present point of view, this is not true: some means are available under any interpretation, and the task we face is to determine which these are.

The rules we are after are those on which there cannot be any disagreement, and that as such, make the disagreement over the validity of

[23]Page 98.
[24]See p. 100.

other rules possible. Since a disagreement over the validity of a logical rule is conceived of as being always, essentially, a disagreement over a conception of meaning, precisely those rules will be indisputable the justification of which does not in any way depend on a conception of meaning. Instead, we take them to be justified by a demonstration that we cannot abandon them without abandoning the idea of a logical proof altogether. At this level, PT-procedures take over the role of semantic explanations. They must show us the rules that can be accepted no matter how we interpret our thoughts about reality. So conceived of, the PT-procedures do not give us any reason to dismiss a particular rule as being incorrect. The only thing they can show is whether the acceptance of the rule can be justified in terms of our conception of proof only. If not, the proper conclusion to be drawn is that acceptance of it commits us to some *external*, i.e., semantic assumption. And the question whether or not the relevant semantic assumption is justifiable, is not a question that can be answered by means of PT-procedures.[25]

From another angle the PT-procedures appear, as we have seen, as a means for circumventing the obvious regress lurking in the usual attempts to justify logical rules. According Dummett, the origin of the problem has to be sought in the assumption that a logic can only be justified in terms of a semantic theory. A semantic justification, as we saw, always involves the appeal to at least some logical principles, which, in turn, are in need of a semantic justification, and so on. To escape this, Dummett postulates a metalanguage the logic of which is not motivated semantically, but on the basis of an analysis of the *use* of language. And precisely here we come across the importance that the dictum 'meaning is use' has for Dummettian revisionism: through an analysis of use it seeks to determine the limits of any semantic justification.

2.3 Conservative Extensions

Proof-theoretic justification procedures, it has been said, provide a method for determining all and only self-justifying rules. A rule is self-justifying if it is constitutive of our conception of a logical proof. In the preceding subsection it has been argued that from the very nature of Dummettian revisionism it follows that there must be at least some

[25] Cf. Currie 1994, p. 470.

rules that resist a non-trivial semantic explanation. We will now start to investigate how we may 'derive' these from our conception of what a logical proof is by introducing the notion of a conservative extension.

The idea that the validity of at least some logical rules is relative to a notion of proof can be traced back to a suggestion made by Belnap in his reply to a short note by Prior. In that note Prior introduces a connective, named *tonk*. Its meaning is given by the following two rules: From A infer A *tonk* B. And: from A *tonk* B infer B. It can be seen immediately that adding *tonk* to some logical theory will enable us to infer B from A for arbitrary A and B. A logical theory that validates the inference of any sentence from any other sentence fails to do what we expect a logic to do: to discriminate between valid and invalid inferences. We call such a theory inconsistent.

Prior's example raises the general question how to distinguish interpretations of logical connectives that do make sense, from those that do not. In his reply to Prior, Belnap observes that 'we are not defining our connectives *ab initio*, but rather in terms of an *antecedently given context of deducibility*, concerning which we have some definite notions.'[26] Now we may define a connective *tonk*. But that we are able to do this does not guarantee that it also *exists*: it does so only if the definition is *consistent* with our prior conception of deducibility in the light of which the connective is defined. This suggests that we may determine whether or not some given connective exists by first determining the notions that constitute our conception of deducibility, followed by a demonstration that the alleged definition is consistent with these. So, What are these notions? And: How are we to know whether or not some definition is consistent with these?

Belnap characterizes deducibility by the structural rules of Gentzen: weakening, contraction, permutation, and transitivity. And he states the demand for the consistency of the definition of a new connective, say *plonk*, as follows: adding *plonk* to a logical system must result in a *conservative extension* of it, that is, it must not be possible to derive any statements not involving *plonk*, that could not be derived before. We thus find at the heart of his proposal the claim that if the addition of a connective to a logical system not containing it yields a conservative extension of this system, the definition of the connective in question is consistent with Gentzen's structural rules.

[26] Belnap 1967, p. 133.

When we consider this carefully we find that Belnap's proposal is stronger than needed for excluding the connective *tonk*. We saw that adding *tonk* to a logical system will render it inconsistent. This, however, does not in any way depend on whether or not the system in question respects Gentzen's structural rules. So, we could have confined ourselves to the demand that the addition of a connective should not make the system inconsistent. Belnap's criterion is stronger in the sense that it does imply this, while the converse does not hold: we can conceive of an extension of some logical system that is consistent, but not conservative. Hence, this criterion rules out not only definitions, like that of *tonk*, that do make the resulting logic inconsistent, but also some that do not.

Why did Belnap believe that we should require more than just consistency? Recall the basic assumption that we define our connectives in terms of an antecedently given context of deducibility. This says, in effect, that we assume our notion of deducibility to determine a limited set of admissible definitions. If we are to appreciate Belnap proposal in full, we should take it as a device to single out precisely the set of definitions that are admissible, given Gentzen's structural rules. The distinction we have to make is that between the demand that a logical system be consistent, in the sense that it allows us to discriminate between valid and invalid forms of inference, and the demand that the definitions of the connectives that constitute the system be consistent with an antecedently given notion of deducibility. Belnap's insight was that in order to meet the second one, it is necessary, but not sufficient to meet the first one. Looked upon from this perspective, Prior's *tonk* presents an interesting testcase for Belnap's criterion. It passes it, but that should not blind us for the more important fact that it excludes other definitions, which, at least *prima facie* are far less harmful, as well. Of these the most salient one is the classical rule for double negation elimination: from $\neg\neg A$ derive A. Adding negation subjected to the classical rule to the negation-free fragment of intuitionist logic yields a non-conservative extension. We will now be able to derive several classical laws that do not involve negation but are intuitionistically invalid such as Peirce's law: $((A \to B) \to A) \to A$. This law presents a clear example of a principle in which there is no occurrence of negation, which is not valid in our original fragment, but which, nevertheless, becomes valid as soon as we add a classically interpreted negation operator to the language.

The importance of the notion of a conservative extension is, to summarize, that it may be used in order to give substance to the thought that only a limited number of logical operations are admissible given an antecedently given notion of deducibility. In the next subsection we will examine how the notion of a conservative extension is taken up by Dummett in order to make the quite general requirement that there must be a kind of consonance, or harmony, between the two aspects of the (assertoric) use of expressions more precise.

2.4 Consonance

On the view we are pursuing, any logical rule is responsible to our general conception of proof. This conception of proof has been characterized as a transcendental one, and as such an invariable context in which an invariable set of logical rules can be shown to be self-justifiying. It has also been said that the possibility of constructing a proof is thought to be intimately connected to the possibility of using language in general. The use of language in proof is, after all, just *a* special case of language use. It will, therefore, not come as a surprise that, from the anti-realist point of view, the requirements that a logical principle has to meet if it is to be self-justifying are nothing but a special case of the general requirements that any assertoric use of language has to meet if it is to be acceptable. So, what will occupy us now is, firstly, what requirements we have to impose on the use of expressions in general, and, secondly, how these are 'translated' into constraints on the use of the logical connectives.

To say that meaning is use is not to say, at least not according to Dummett, that meaning is determined by actually use. Instead, Dummett seeks to derive a criterion for meaningfulness from the requirements we have to assume that our uses of language comply with, given that these uses are *successful*. In other words, Dummett holds that use of language cannot be successful unless this use is constrained in some way or other. For Dummett, then, the dictum 'meaning is use' says: 'meaning is correct use', that is, meaning is use in accordance with the constraints that make successful language use possible.

Dummett does not tell us very much about what the goals of language use are. Of these he only says that they are '*internal*: there is no form of description of what a language is required to do —to communicate thoughts, for example— that would represent it as something in principle achievable without the use of language.'[27] Important is not so much which goals the use of language serves, but the claim that we may fail to realize these ends (fully), and that 'the possibility of failure arises primarily because of the multiplicity of principles governing our linguistic practice.'[28] These principles fall apart into two categories: one consisting of the conventions governing the occasions on which an assertion is appropriately made, the other consisting of the 'conventions governing both the responses of the hearer and what the speaker commits himself to by making the utterance.'[29] So we find that, according to Dummett, we should distinguish between two aspects of the (assertoric) use of an expression: the *conditions for* its utterance, and the *consequences of* it.

The question whether or not linguistic practice has these two aspects is closely related to the question whether an established practice could be subject to revision. As a possible defect Dummett mentions the inconsistency of some of the principles by which assertions are warranted. This is a defect in as much as it undermines the reliability we expect such principles to possess. If we concede this, Dummett continues, we must also recognize that such an inconsistency may go unnoticed and hence, 'that the mere fact that it is established affords no ground for assuming a linguistic practice free from defect.'[30]

Assuming that a linguistic practice may be inconsistent in this sense, and, therefore, in need of revision, the question arises whether there are other defects, and thus other reasons for criticizing an established practice. There are, Dummett claims, for the two complementary features of any linguistic practice ought to be in *consonance*, and there is no automatic mechanism to ensure that they will.[31] To require this is to deny that either one can be determined independently of the other. For Dummett it is 'plain that we may legitimately demand a certain consonance,'[32] he calls it 'difficult to make precise, but intu-

[27]Dummett 1991b, p. 210. Cf. Dummett 1981, pp. 29–30.

[28]Dummett 1991b, p. 210.

[29]Dummett 1992c, p. 221.

[30]Dummett 1991b, p. 214. See also Dummett 1992a, p. 358.

[31]See, for example, Dummett 1991b, p. 215, Dummett 1992a, p. 397. For reasons of clarity, I will reserve the notion 'harmony' for the more technical specification of what it is for two logical rules to be in consonance. Dummett does not maintain this terminological discipline.

[32]Dummett 1992a, p. 397.

itively compelling.'[33] But in none of this we find a convincing reason for requiring consonance, nor does the remark that 'the use of language has a *point*, as, indeed, a game has a point, and, if the necessary harmony between the different components is lacking that point is destroyed or impaired.'[34] For this only says that the lack of consonance is a defect because it will stop us from realizing the goals of our use of language. But this, of course, tells us only what a defect is, it does not say anything *why* we cannot attain our linguistic goals if our practice lacks consonance.

We saw, to summarize, that it is the multiplicity of different features of the use of sentences and the demand that these be in consonance, that makes it possible to criticise an existing linguistic practice.[35] But we are still completely in the dark about the legitimacy of the demand of consonance. Probably the best reason for it is that we cannot give it up without abandoning the idea of a compositional meaning-theory, because consonance '*within the language as a whole* is a prerequisite for the possibility of a compositional meaning-theory.'[36]

The argument runs as follows.[37] In a compositional meaning-theory the meaning of a sentence 'must be explicable in a way that presupposes only a restricted range of other sentences—sentences with a lower degree of complexity, in a generalized sense of 'complexity'.[38] It follows that the content of a sentence cannot be taken as fixed by the totality of means by which its truth could be established, since these will include deductive arguments of an unbounded degree of complexity. Hence, we must distinguish between *direct*, or *canonical*, and *indirect* grounds for the truth of a sentence.[39] The derivation of a statement is called

[33]Dummett 1991b, p. 215.

[34]Dummett 1981, p. 30.

[35]Dummett 1992c, p. 223. Cf. Dummett 1992a, p. 397

[36]Dummett 1991b, p. 247.

[37]See Dummett 1991b, p. 229–30 and Dummett 1992c, p. 222.

[38]Dummett 1991b, p. 229. We need a generalized sense of complexity, because we cannot, in general, assume that the knowledge of the meaning of a sentence presupposes knowledge only of the meanings of sentences that are of a lower logical, i.e., syntactical, complexity. Take, for example the sentence, 'First-order predicate logic is incomplete.' The sentence itself has a quite simple logical structure. To understand it, however, one need to be familiar with, among other things, definitions, each of which may contain sentences of a higher complexity. See also Dummett 1987a, p. 249.

[39]See also chapter II, in particular § 3.1.

canonical if it establishes the conclusion, step by step in accordance with its logical structure. As such it is said to constitute the content, or meaning, of its conclusion. All other ways of establishing the sentence stand in need of a justification: 'The validity of such arguments must flow from the meanings of the logical constants, or of non-logical expressions occurring essentially in them. It has to be shown that the argument is valid in virtue of the meanings of those expressions, as independently given.'[40] Now, if we do not assume, for any sentence of the language, that the canonical means of establishing a statement as true should match, and be matched by, the consequences which accepting that statement as true is canonically treated as having, then it might very well be that a complex derivation establishes a conclusion on the basis of some premises, which, taken on their own, would not, given the meaning of the conclusion, suffice to establish it. Hence, we would not be able to show that the validity of the derivation flows from the meaning of the conclusion.

Let me conclude this rather abstract argument by recapulating what has been pointed out so far. The possibility of a revision of a logical practice has been shown to be given with the demand of a consonance between the two features of our linguistic practice. A subsequent digression brought to light that this demand itself is inherent to the idea of a compositional meaning-theory. Having established this we now return to the notion of consonance. Let us first consider an example of how the lack of consonance may give us a reason to revise an established practice. A rather clear case, Dummett believes, is presented by pejorative expressions. As an example he discusses the term 'Boche':

> The condition for applying the term to someone is that he is of German nationality; the consequences of its application are that is he barbarous and more prone to cruelty than other Europeans. We should envisage the connections in both directions as sufficiently tight as to be involved in the very meaning of the word: neither could be severed to alter its meaning. Someone who rejects the word does so because he does not want to permit a transition from the grounds for applying the term to the consequences of doing so.[41]

The moral Dummett draws from this is that since the grounds for accepting a statement in which the expression 'Boche' occurs as true

---

[40]Dummett 1991b, p. 229–30.
[41]Dummett 1992a, p. 454.

are not in consonance with the consequences that follow from this, we should revise the expression 'Boche', or exclude it from our practice altogether.

This certainly gives us an impression of the general idea, although in this specific case I find the conclusion drawn from it highly implausible. Let me make two comments. Firstly, I think it is just not correct to believe that being a German is a sufficient reason for someone to be called a Boche. For example, I would never have called my grandmother a Boche just because she was of German nationality. As it appears to me we will call someone a Boche only if he is of German nationality, *and* displays a particular behavior. If this is correct it might quite well be the case that, *pace* Dummett, our use of the term 'Boche' is in consonance, and as such does not stand in need of revision.

But if, secondly, we do accept Dummett's criterion for the correct application of the term 'Boche', it is still doubtful whether, in this case, the lack of consonance is a good reason for revising the term. It appears to me that in the argument Dummett confuses the evaluation of the uses of language as a means to attain certain goals with the evaluation of these goals themselves. We saw that a lack of consonance is considered by Dummett to be a defect, because it keeps us from realizing the ends of language use. No matter how improper one may believe it to be, the fact is that sometimes language is used for no other reason than to offend, insult, or hurt someone else. So, the only relevant question here is whether a lack of consonance does stop one from attaining such a goal. And if we look at the example in this way, we find that the answer must be: No. What is more, the example points out that it is precisely because of a lack of consonance that we succeed in offending or insulting the hearer. And as such it even presents a counterexample to the claim that we can attain the ends of the use of language only if the different aspects of our linguistic practice are in consonance.

Clearly, Dummett would find this unacceptable, since, as we saw, he conceives of consonance as a necessary condition for the realization of the goals of language use. If we are to keep to this thought we have no other choice than to reject Dummett's criterion for the application of the term 'Boche'. But then, as the first comment made clear, we have no longer any reason to believe that our use of the term is not in consonance, and, hence, no reason to revise it. This does not affect the cogency of the demand of consonance itself, it only says that the example is rather ill-chosen.

But despite this criticism it is worthwhile to look once more at the example. Ignoring the term 'Boche', we may read it as saying that, if X is an expression added to some language, and X may be inferred from A, and C from X, the use of the sentence X can be legitimately objected to, if, in its absence, we are not willing to recognize the validity of the inference of C from A.[42] And this suggests, in effect, that the demand of consonance, which, so far, we have taken at face value, is closely related to the notion of a *conservative extension*. How exactly this connection is to be conceived of will occupy us in the next section.

3 Proof-theoretic Justification Procedures

3.1 Consonance, Harmony, and Conservativity

The preceding section ended with the observation that the example Dummett provides of a case in which consonance is lacking, really is an example of a failure of conservativity. This makes one wonder how precisely these two notions, i.e, consonance and conservativity, are related. Could it even be that it is possible to explain the informal notion of consonance in terms of the more formal requirement of conservativity? In answering this question I limit myself to the logical connectives. So, the question I will start this section with is: Is it possible to specify for the logical constants what it is for the different aspects of their use to be in consonance, in terms of the notion of a conservative extension?[43]

The gist of Dummett's idea's on this matter can be found in the following quotation. It will serve as the starting point of the present inquiry:

> In the case of a logical constant, we may regard the introduction rule governing it as giving the conditions for the assertion of a statement of which it is the main operator, and the elimination rules as giving the consequences of such a statement: the demand for harmony between

[42] Dummett 1992a, p. 397.

[43] The answer to this question will consist, for a large part, in making precise the concepts employed in Dummett 1991b. Although most of these may appear to be derived directly from general proof theory, and especially from the work of Dag Prawitz (most notably Prawitz 1965), the definitions below have no other aim than to capture formally the informal formulations Dummett uses. For this reasons the adopted terminology may deviate from any standard or usage currently adopted in general proof theory.

> them is then expressible as the requirement that the addition of the constant to a language produces a conservative extension.[44]

What this says, actually, is that it is not consonance, but *harmony* that can be expressed in terms of conservativity. Harmony and consonance, however, are closely related. As the term is used here, harmony strictly refers to the specific requirement that is obtained by applying the general demand of consonance to a particular class of expressions: the logical constants. Let us see what this specific requirement looks like. Consonance has been characterized as the demand that, in a way, the *consequences of* asserting a sentence are determined by the *conditions for* asserting it, and vice versa. In the case of a logical constant **c**, the former are specified by the elimination rule that governs **c**, and the latter by its introduction rule. These rules can be said to be in consonance 'if there is a deductive path from the premises of the introduction rule to the conclusion of the elimination rule without invoking the rules governing the constant **c**.'[45] This can be made more precise as follows. First, the notion of a *local peak* is defined:

1. DEFINITION. (Local peak) *A* local peak *for a constant* **c** *is any part of deductive inference where a* **c***-introduction rule is immediately followed by a* **c***-elimination rule.*[46]

The following presents an example of a local peak:[47]

$$\dfrac{\dfrac{\begin{matrix}[A]\\ \vdots \\ B\end{matrix}}{A \to B}\,I_{\to} \qquad A}{B}\,E_{\to}$$

In deriving the conclusion B the implication is first introduced, and then, in the next step, eliminated. Quite obviously, this is a rather

---

[44] Dummett 1992a, p. 455.

[45] Dummett 1991b, p. 248.

[46] *ibid.* In Prawitz 1965, the *locus classicus*, it is called a local maximum.

[47] I will use square brackets to mark off assumptions from premises. Any formula occurrence that is not the consequence of an introduction, or elimination rule, and that is not withdrawn in the course of the derivation is called a *premise*. Any other formula occurrence that is not the consequence of an introduction, or elimination rule is called an *assumption*.

roundabout way of deriving B. It is legitimate to introduce the implication, because B can be derived under the assumption of A. It is, furthermore, justified to eliminate the implication, if A is available, not just as an assumption, but as a premise. Taking this together, B could equally well be derived directly from the premise A:

$$\begin{array}{c} A \\ \vdots \\ B \end{array}$$

In other words, if some formula can be derived by means of an inference in which the introduction rule of '$\rightarrow$' is followed immediately by its elimination rule, then it is possible to derive this formula without appealing to the implication at all, by 'eliminating' or 'levelling' the local peak. If, for a logical constant, this holds, the rules that govern its use are said to be in intrinsic harmony.[48] So, we have:

2. DEFINITION. (Intrinsic harmony) *The rules for* **c** *are in* intrinsic harmony *if and only if any local peak for* **c** *can be levelled.*[49]

All this gives reason to rephrase the question we set out with as the question whether the demand for intrinsic harmony between the rules that govern a logical constant is indeed expressible as 'the requirement that the addition of the constant to a language produces a conservative extension.' Clearly, it would be if both notions are logically related such that each one implies the other. A careful analysis, however, shows that the alleged logical relation between both notions does not exist; intrinsic harmony and conservativity are, in fact, logically independent. The argument to this end sets out with a definition of a conservative extension:

3. DEFINITION. (Conservative extension) *A language* $\mathcal{L}'$ *is a* conservative extension *of a language* $\mathcal{L}$ *if and only if for any formula* $\varphi \in \mathcal{L}$ *and any set of formulas* $\Delta \subseteq \mathcal{L}$: $\Delta \Vdash_{\mathrm{L'}} \varphi \Leftrightarrow \Delta \Vdash_{\mathrm{L}} \varphi$.

[48]Dummett speaks of *intrinsic* harmony, and not of harmony simpliciter in order to be able to distinguish it from the demand of *total* harmony. The reasons for this will become clear below.

[49]Dummett 1991b, p. 248. Cf. p. 250: 'we may thus provisionally identify harmony between the introduction and elimination rules for a given logical constant with the possibility of carrying out this procedure, which we have called the levelling of local peaks.'

Quite often, Dummett speaks of *total harmony* instead of conservativity. The former notion is defined in terms of the latter:[50]

4. DEFINITION. (Total harmony) *A set of rules governing a logical constant* **c** *is in* total harmony *with respect to some language* $\mathcal{L}$ *if the language* $\mathcal{L}'$ *that is formed by adding* **c** *to the vocabulary of* $\mathcal{L}$ *is a conservative extension of* $\mathcal{L}$.

The actual argument consists of two steps. The first of these is a demonstration of the fact that intrinsic harmony is not sufficient for conservativity. The second step consists in an argument to the effect that it is not necessary either. We start with refuting the thought that intrinsic harmony is sufficient for total harmony, or conservativity, i.e., the claim that for an arbitrary constant **c**, if the rules for **c** are in intrinsic harmony, then the language $\mathcal{L} \cup \{\mathbf{c}\}$ is a conservative extension of $\mathcal{L}$. As a counterexample to this Dummett describes a language $\mathcal{L}$ comprising only two constants, '$\wedge$' and '$+$', where '$\wedge$' has the meaning it usually has, and the meaning of '$+$' is given by the following three rules:

$$\frac{A}{A+B}\,I_{+} \qquad\qquad \frac{B}{A+B}\,I_{+} \qquad\qquad \frac{A+B \quad \dfrac{[A]}{C} \quad \dfrac{[B]}{C}}{C}\,E_{+}$$

In words, '$+$' has the same meaning as the normal disjunction, except that a formula of the form $A+B$ may be replaced by some formula C only if C can be derived *in one step* from both A and B. In $\mathcal{L}$ it is not possible to derive $(A \wedge B)+(A \wedge C)$ from $A \wedge (B+C)$. But as soon as '$\vee$' is added to $\mathcal{L}$ it can be so derived. And this means that adding the normal, unrestricted disjunction to $\mathcal{L}$ yields a non-conservative extension of it, while, clearly, the introduction and elimination rules for normal disjunction are in intrinsic harmony.

Contrary to first thoughts as Dummett, too, seems to have them,[51] conservativity does not imply intrinsic harmony either. Consider the logical constant '$\Box$' the use of which is governed by two sets of rules, each one consisting of an introduction and an elimination rule. It depends on the syntactical structure of the formula A which of these sets should be appealed to in the course of a derivation when an occurrence of '$\Box$' is to be introduced or eliminated:

[50] Dummett 1991b, p. 250.

[51] See, for example, Dummett 1991b, p. 247.

$$\frac{A}{\Box A}\, I_{\Box} \qquad \frac{\Box A}{A}\, E_{\Box} \qquad \text{if '}\Box\text{' does not occur in } A$$

$$\frac{A}{\Box\Box A}\, I'_{\Box} \qquad \frac{\Box A}{A}\, E'_{\Box} \qquad \text{if '}\Box\text{' does occur in } A$$

Let $\mathcal{L}$ be a logical language and $\mathcal{L}'$ the language obtained from $\mathcal{L}$ by adding '$\Box$' to it. Then $\mathcal{L}'$ is a conservative extension of $\mathcal{L}$. Note first that in $\mathcal{L}'$ we have $A \vdash \Box A$ and $\Box A \vdash A$. Consider a derivation of some conclusion C in which there is no occurrence of '$\Box$'. Then, if '$\Box$' occurs in the derivation of C we obtain a derivation of C in which '$\Box$' does not occur by replacing every occurrence of formula of the form of '$\Box A$' by formula of the form 'A'.

Although the addition of '$\Box$' to a logical language yields a conservative extension, the rules that govern the use of '$\Box$' are not in intrinsic harmony. We have the following counterexample. Let A be atomic. The following local peak cannot be levelled:

$$\dfrac{\dfrac{\Box A}{\Box\Box\Box A}}{\Box\Box A}$$

This concludes my argument to the effect that intrinsic harmony and conservativity are logically independent. But if this is so, then why would Dummett explain intrinsic harmony in terms of conservativity? Or, stated differently, why would he demand conservativity rather than intrinsic harmony? It is justifiable to demand intrinsic harmony in as far as it is nothing but the formal counterpart of the requirement of consonance, which, in subsection 2.4, has been found to be motivated on the basis of the analysis of the assertoric use of language. But since intrinsic harmony and conservativity are logically independent this informal analysis does not in any way justify the demand of conservativity. Yet, even in the absence of an acceptable motivation, there can be no misunderstanding about it that it is conservativity that is crucial. For, as Dummett puts it, the *point* of the requirement of harmony is that, for every logical constant, its addition to the fragment of the language containing only the other logical constants should produce a conservative extension of that fragment.[52]

[52]Dummett 1991b, p. 290.

The situation we are facing now, is the following. It is clear that Dummett holds that the fact that the addition of some logical constant to some given language not containing it yields a conservative extension is more important than is the mere fact that the rules governing the use of that constant are in intrinsic harmony. Yet, it is only the latter requirement that can be motivated on the basis of Dummett's general analysis of the constraints that make successful (assertoric) use of language possible. So, the question is: What reasons does a Dummettian anti-realist have for requiring conservativity rather than intrinsic harmony? This will occupy us in subsection 3.3.

In the meantime, we will continue to examine the connection between conservativity and intrinsic harmony. From the fact that intrinsic harmony and conservativity are logically independent we cannot conclude that they are not related at all. Dummett, being aware of the fact that conservativity does not follow from intrinsic harmony directly, suggests that the former follows from the latter *conditionally.* To be more precise, his *conjecture* is that intrinsic harmony implies total harmony in a *stable* context. What this conjecture amounts to will be explained in the next subsection.

3.2 Dummett's Conjecture Explained

Dummett's counterexample to the claim that intrinsic harmony implies conservativity presents, as we saw,[53] a rather unusual connective '+' and the argument continues to show that if a perfectly natural connective like '$\vee$' is added to a language consisting only of '$\wedge$' and '+', this will result in a non-conservative extension of the original language. Having observed this, there are two ways to proceed. It can be taken as a reason for rejecting '$\vee$' as a possible extension of our language, or it can be taken as a reason to revise the original language. The first option makes the acceptability of a new connective depended on what connectives are already contained by the original language. If it contains both '$\wedge$' and '+', '$\vee$' is not acceptable, but if it contains only '$\wedge$', '$\vee$' is acceptable. This is far from the spirit of the PT-procedures, which are developed in order to determine the correct, or self-justifying rules independently of the context in which these rules are functioning.

Consequently, there is no alternative but the second option: non-

[53]See page 115.

conservativity urges us to scrutinize the original language. Hence, whether or not $\mathcal{L}'$ extends a language $\mathcal{L}$ conservatively tells us something about $\mathcal{L}$, too, and not only about $\mathcal{L}'$: if $\mathcal{L}'$ is a conservative extension of $\mathcal{L}$, then $\mathcal{L}$ is in order; if not, $\mathcal{L}$ is not in order. So, from the fact that adding '$\vee$' to a language $\mathcal{L}_{\{\wedge,+\}}$ containing only '$\wedge$' and '$+$' yields a non-conservative extension we learn that something is wrong with $\mathcal{L}_{\{\wedge,+\}}$. Now, in this simple case, it is most likely that the thought forces itself upon us that all our problems originate with the connective '$+$'; in contrast to the 'perfectly natural' connective '$\vee$', '$+$' appears as the result of a highly artificial restriction on the form that an elimination rule is allowed to take. Following this line of thought, we find that our interpretation of the result of an application of the requirement of conservativity will only then yield the desired result if it is completed with certain intuitions concerning the form of a logical rule. We will return to this in § 4.2 below.

Taking conservativity as yielding information about the original language and not about the extension also implies that the mere fact that the addition of a connective to some language yields a conservative extension does not in any way guarantee that the rules that govern the added connective are in order. Consider, for example, the language $\mathcal{L}_{\{\wedge,\vee\}}$. Adding the connective '$+$' to it yields a conservative extension, but this does not make '$+$' an acceptable, or even a plausible connective.

How are we to accommodate the intuition that something is wrong with a connective like '$+$'? Since the rules that govern '$+$' are in intrinsic harmony, this notion is of no use to us in understanding why '$+$' is not acceptable. Conservativity will not do, either. As we just saw, a semantic anti-realist has no choice but to take (non-)conservativity as qualifying the original language only. And, having observed, by means of conservativity, that the original language is not in order, he is as yet unable to say in what respect it is disfunctioning. *A fortiori*, he has no clear idea of what it is for a language to be in order.

A stronger procedure is called for, one that enables us to single out any set of logical rules containing the rules for '$+$' as incorrect, in any context. In order to develop such a procedure, Dummett takes up a suggestion by Gentzen, according to which an introduction rule fixes the meaning of the logical constant in question, and its elimination rule is taken to be no more than a consequence of this definition.[54]

If, on the basis of Gentzen's suggestion, we want to develop a procedure that is applicable to any logical law, we should take the notions of an introduction, and elimination rule not to narrow. Let us say that any rule the consequence of which has the constant **c** as its principal operator is an introduction rule for **c**. In the same fashion, we say that any rule is an elimination rule for **c** if one of its premises is required to have **c** as its principal operator.[55] Surely, this is only a first approach to a quite difficult task: to give a suitable definition of what it is to be an introduction or elimination rule. A shortcoming of the characterization just mentioned is that it ignores the logical complexity of the formulas involved as, respectively, the premise(s) and conclusion of the rule. Given this 'definition', for example, the distributive law qualifies as an elimination rule for conjunction:

$$\frac{A \wedge (B \vee C)}{(A \wedge B) \vee (A \wedge C)}$$

To call this rule an elimination rule is undesirable, since the logical complexity of the conclusion is higher than that of the premise of the rule. Surely, Dummett is aware of the difficulties involved in giving a suitable definition of what it is to be an introduction or elimination rule. In fact, he explicitly mentions two conditions a rule has to satisfy if it is to qualify as an introduction rule. The first one states that 'in

[54]Dummett 1991b, p. 251. The suggestion can be found in Gentzen 1934, p. 189. See also Prawitz 1965, p. 33. Genzten's idea has not remained unchallenged, as is explicitly recognized in Sundholm 1994. In the writings of Martin-Löf, for example, the possibility of defining the meaning of the logical constants in terms of the introduction rules is rejected. As an alternative, Martin-Löf develops a system that takes the elimination rules as meaning giving. This controversy, though interesting in itself, will not receive here any further attention, because Dummett's notion of a PT-procedure does not depend on the assumption that the meaning of a logical connective is determined by its introduction rule. As we will see below, it is characteristic of these procedures that they are indifferent to whether we take the introduction, or the elimination rules as endowing meaning.

[55]See Dummett 1991b, p. 256. If an elimination rule has more than one premise we distinguish between its *major* premise and its *minor* premise(s). The major premise is that premise the principal operator of which is being eliminated. Any other premise is called a minor premise.

any application of it, the conclusion will be of higher logical complexity than any of the premises and than any discharged hypothesis.'[56] The second one imposes a condition on the occurrences of free variables.[57] Of these two the first one, perhaps, could be called 'intuitively compelling', the second one certainly not. Now, it is not my intention to evaluate these conditions here by inquiring whether they are necessary or sufficient, nor will I go into the question of how to motivate them. The point I want to stress is that the PT-procedures presuppose a clear and workable notion of what it is to be an introduction or elimination rule, and that, apart from the fact that it is exceedingly difficult to make this notion explicit, it is not at all clear that it can be determined solely on the basis an analysis of use as Dummett has presented it.

For the moment we must be satisfied with this observation and direct our attention to precisely that aspect of the PT-procedures where the importance of a presupposed notion of an introduction can be seen most clearly. Once again we turn to the distributive law, asking this time how we are to *justify* this rule, given that the meaning of the connectives is fixed by their introduction rules. First, we examine how the premise of the distributive law can be obtained using only introduction rules:

$$\dfrac{A \qquad \dfrac{B}{B \vee C}\,{\scriptstyle I_\vee}}{A \wedge (B \vee C)}\,{\scriptstyle I_\wedge}$$

Then, we inquire whether or not the conclusion of the distributive law can be derived, given a legitimate derivation of the premise of the law, using the same assumptions and only introduction rules. Evidently, it can:

$$\dfrac{\dfrac{A \quad B}{A \wedge B}\,{\scriptstyle I_\wedge}}{(A \wedge B) \vee (A \wedge C)}\,{\scriptstyle I_\vee}$$

This says that, given the introduction rules for disjunction and conjunction, we can dispense with the distributive law, because, if we have

[56] Dummett 1991b, p. 258.

[57] '(...) if one or more of the premises of an application of the rule contain a free variable, but the conclusion does not, that free variable should occur in any of the hypotheses on which the conclusion depends.' Dummett 1991b, p. 259–60.

a legitimate derivation of its premise, we can derive its conclusion without appealing to it. It is in this sense that the distributive law, as an elimination rule for conjunction can be said to be a consequence of the meanings of disjunction and conjunction as fixed by their respective introduction rules.

Some comments are in order. Note first that for this procedure to be successful it must always be possible to obtain a *legitimate* derivation of the premise of the elimination rule. A derivation is legitimate if and only if it appeals to no other rules than those that fix the meanings of the logical constants in question. In the case at hand, we stipulated that the meanings of the constants is given by the introduction rules. Hence, a legitimate, or *canonical* derivation should appeal exclusively to introduction rules.[58] We can only be sure that a canonical derivation exists, if it holds that, if a statement having **c** as its principal operator is derivable at all, it can be established by a derivation the last step of which consists in an application of the introduction rule for **c**.[59]

A further comment is that if the rule to be justified is an elimination rule the major premise of which contains at most two atomic subformulas, the justification procedure described here comes down to the levelling of local peaks. For example:

$$\frac{A \wedge B}{A}\,{\scriptstyle E_\wedge} \qquad\qquad \cfrac{\cfrac{A \quad B}{A \wedge B}\,{\scriptstyle I_\wedge}}{A}\,{\scriptstyle E_\wedge} \qquad\qquad \begin{array}{c} A \\ \vdots \\ A \end{array}$$

The procedure described above is referred to by Dummett as the *upwards* justification procedure. It is characterized by the stipulation that the meaning of a logical connective is given by its introduction rules. We now define what it means for a logical rule to be justifiable in this way. In general, the set of rules which is taken to fix the meaning of the logical constants is called the *base* of a justification procedure. A derivation is called *canonical* if and only if it makes use of no other rules than those comprised by the base. This brings us to the following definition:

5. Definition. (Upward justification procedure) *A rule of inference R is justifiable with respect to a base consisting of introduction rules*

---

[58]We already have come across the notion of canonical derivation. See page 109.

[59]This is what Dummett calls 'the fundamental assumption.' It certainly is not without difficulties, but this is not the place to discuss these.

only, if and only if it holds that if there exists a canonical derivation of the premises of R, there exists a canonical derivation of the conclusion of R.

Next to an upward, we also have a *downward* justification procedure, which does not take the introduction, but the elimination rules as its base:

6. DEFINITION. (Downward justification procedure) *A rule of inference R with premise P and consequence C is justifiable with respect to a base consisting of elimination rules only if and only if for any consequence C'_n of C it holds that, if there is a canonical derivation of C'_n from C, then there is a canonical derivation of C'_n from P.*

Once again, I turn to the distributive law, but this time as an example of an introduction rule for '$\vee$'. It is justifiable by means of the downward procedure as follows. We first look at the consequences that can be drawn from the conclusion of the distributive law:

$$\dfrac{(A \wedge B) \vee (A \wedge C) \qquad \begin{matrix}[A \wedge B] \\ \vdots \\ D\end{matrix} \qquad \begin{matrix}[A \wedge C] \\ \vdots \\ D\end{matrix}}{D}$$

If D is to be the conclusion of a canonical derivation, it must be equivalent to A, or to some formula D' that is derivable from A using only elimination rules. It can easily be verified that in both cases we can obtain it from the premise of the distributive law using only elimination rules.

We now come to the notion we were after: stability.

7. DEFINITION. (Stability) *Let $\mathcal{L}$ be a logical language, and I the set of introduction rules governing the logical constants of $\mathcal{L}$. Let E be the set of elimination rules that are justifiable by means of an upwards justification procedure having I as base. Let I' be the set of introduction rules justifiable by means of a downwards justification procedure with E as its base. Then $\mathcal{L}$ is* stable *if and only if $I = I'$.*

It is inessential that, in this definition, we start with the set of introduction rules as the base of an upwards justification procedure. If a language is stable, we could equally well have started with the set of elimination rules as the base of a downward justification procedure;

the outcome would be exactly the same. When a language is not stable in this sense, then, of course, we cannot reverse the order without modifying the outcome.

The two notions of intrinsic harmony, and total harmony or conservativity are brought together in Dummett's conjecture, according to which intrinsic harmony implies total harmony, i.e., conservativity, in a context where stability prevails:

1. CONJECTURE. (Dummett's conjecture) *Let $\mathcal{L}$ be a logical language and* **c** *a logical constant such that E_c is in intrinsic harmony with I_c. Let $\mathcal{L}'$ be the language obtained from $\mathcal{L}$ by adding* **c** *to it. Then $\mathcal{L}'$ is a conservative extension of $\mathcal{L}$ if $\mathcal{L}$ is stable.*[60]

If correct, this hypothesis does indeed help us to understand why the addition of normal disjunction to the language $\mathcal{L}_{\{\wedge,+\}}$ containing only the two constants '$\wedge$' and '$+$' does not yield a conservative extension, even though the rules that govern disjunction are in intrinsic harmony. If we take a close look, we find that $\mathcal{L}_{\{\wedge,+\}}$ does not qualify as a stable context. Hence, given Dummett's conjecture, we cannot expect that adding disjunction to it yields a conservative extension.

Contrary to what Dummett wants us to believe, it is far from obvious that this conjecture does indeed hold. But before we put any effort into proving it, or disproving it, as the case may be, we would do well to consider whether total harmony, i.e., conservativity, provides an adequate criterion for determining whether or not a logical rule is acceptable. It will take the rest of this chapter to argue that it does not.

3.3 Conservativity and Use

At the end of subsection 3.1 we left unanswered the question what reasons a semantic anti-realist has for demanding conservativity rather than intrinsic harmony. To answer this question we need to make explicit what particular interpretation of the dictum 'meaning is use' must be presupposed by the belief that the correct logical laws can be singled out by means proof-theoretic justification procedures.

We begin by noting that, according to Dummett, harmony between logical rules 'can be demanded only when those rules are held *completely*

[60]Dummett 1991b, p. 290.

to determine the meanings of the logical constants.'[61] Connecting this with the observation that conservativity is the point of demanding harmony, and assuming that Dummett is after a distinguishing characterization of what it is for rules governing the use of some expression to exhaust the meaning of that expression completely, the thought lies ready to hand that Dummett believes conservativity to be this characteristic. This raises two questions. The first one is: What is the origin of the thought that rules governing the use of an expression are acceptable if and only if they completely determine its meaning? The second question is: Is conservativity, as it has been defined earlier, the suitable criterion for distinguishing rules that do exhaust the meaning of an expression from rules that do not? The latter question will be taken up in the next subsection, while the current one will be devoted to the former.

That we should assume that the meanings of the logical constants are completely determined by the rules that govern their use is quite understandable, if not inescapable, in the light of the dictum 'meaning is use', as it is understood by Dummett. And the assumption that there is nothing more to the meaning of a logical constant than the use made of it in accordance with the rules that govern it, has its full expression in the demand of conservativity. To see this, consider a language containing only one connective: '$\rightarrow$', governed by the familiar rules. In this language we cannot derive Peirce's law: $((A \rightarrow B) \rightarrow A) \rightarrow A$. But we can, as soon as we add classical negation. Assuming, firstly, that, in general, what counts as a ground for accepting some statement as true depends on the meaning of that statement,[62] and, secondly, that what counts as a valid derivation of a statement that has some logical constant **c** as its principal operator depends on the meaning of **c**, we must come to the conclusion that in a language containing both '$\rightarrow$' and classical negation, the rules that govern the use of '$\rightarrow$' do not completely determine its meaning. Adding classical negation to the language changes the way in which we use (some) expressions with '$\rightarrow$' as principal operator. It follows immediately that we cannot hold both

[61] *ibid.*, p. 286. My emphasis.

[62] This does not commit us to verificationism, since we need not take this relation between meaning and grounds for verification to be the relation of identity. But, I think we should agree with Dummett that there must be at least some connection between the meaning of a statement and the grounds there are for accepting it as true.

that the meaning of a connective is determined completely by the rules that govern its use, and that, in the present case, we can extend the language coherently by adding classical negation to it. For, evidently, the new rules for the negation sign bring about some new aspects of the meaning of '$\rightarrow$' by licensing particular uses of it that were not, and still are not, justifiable only on the basis of the rules that govern '$\rightarrow$'.[63]

To say that only those rules for the use of an expression are acceptable that determine its meaning completely is to say that the use made of it in accordance with these rules should not underdetermine, nor overdetermine its meaning. This is precisely why intrinsic harmony, as a criterion, is not sufficient, as the following quotation indicates:

> The fact that the consequences we conventionally draw from [a complex statement with a given principal operator] are in harmony with these acknowledged grounds shows only that we draw no consequences its meaning does not entitle us to draw. It does not show that we fully exploit that meaning, that we are accustomed to draw all those consequence we should be entitled to draw.[64]

The point is that if an elimination rule can be shown to be in intrinsic harmony with an introduction rule this only tells us that applying this rule will not take us beyond 'the most' we are allowed to do given the introduction rule. But that does not guarantee that in adopting that elimination rule we use the constant in question to the full extent that its meaning allows us to. And in that case, the rules that govern the use

[63]There is an interesting parallel to be drawn between Belnap and Dummett in this respect (See also page 106). Both believe that conservativity is an adequate criterion for the acceptability of a logical rule, while at the same time both fail to give a convincing argument for its adequacy. In both cases the argument structure is the same. First, a criterion for the correctness of a logical rule is introduced that has an immediate plausability (Belnap's consistency), or that can be argued for on the basis of some already accepted assumptions (Dummett's intrinsic harmony). Then conservativity is introduced as the condition that has to be met if the logical rules are to comply with the criterion for their acceptability. For both cases it holds that conservativity is *stronger* than the criterion is supposed to explain. So, what actually happens is that conservativity *replaces*, respectively, consistency and intrinsic harmony as an acceptability criterion. And here lies the problem, since the reasons we may have for requiring that logical rules be consistent, or in intrinsic harmony, only partly support the stronger claim that the demand of conservativity has to be met.

[64]Dummett 1991b, p. 287.

of the constant would *under*determine its meaning, and would, thus, not determine it completely, as the dictum 'meaning is use' requires.

It is only if the rules that govern a constant are stable that we can be sure that these determine not only the least we are allowed to do with it, but also the most, and, hence, that these rules, indeed, completely determine the meaning of the constant at hand. This observation certainly sheds some light on the goal of the PT-procedures: they are meant to delineate all and only those rules that completely determine the meanings of the logical connectives. It will be clear that we cannot expect any procedure to achieve this, unless it is assumed that the meanings of the connectives are given independently of the rules that govern the way in which they are *actually* used. That is to say, the dictum 'meaning is use' is, in the present context, not to be interpreted as saying that the meaning of an expression is given by the way in which we are actually using it, for in that case our use of an expression would by definition exhaust its meaning.

The contrast to be drawn here is that between the interpretation of the dictum 'meaning is use' as 'meaning is actually use', and as 'use ought to determine the meaning completely'.[65] Thus formulated the latter has, in contrast to the former, a strong normative impact. If two people use one and the same expression differently, there is, under the second interpretation, no reason to assume that for one person that expression has a different meaning than it has for the other. On the contrary. Such a case will give reason to believe that one of them, or perhaps even either one, uses the expression in question incorrectly by using it in accordance to rules that do not completely determine its meaning.

Even though meaning and use are not in this way to be identified, it does not follow that Dummett allows for the conception of the meaning of an expression independently of any notion of what it is to use that expression. At first sight, this may appear to be in flagrant contradiction to the claim just stated that we must assume that the meanings of the connectives are given independently of the rules that govern the way in which they are actually used. But the contradiction is only apparent. And seeing why brings us to the heart of what I take to be Dummett's conception of the dictum 'meaning is use'. For Dummett,

[65] Dummett ascribes the former interpretation to Wittgenstein. He strongly opposes it. See, for example, Dummett 1992d.

the claim that the meaning of an expression is determined by the rules that govern its use does not imply that we actually use that expression in accordance with these rules. It only says that if the rules that govern our actual use coincide with the rules that are constitutive for its meaning, then our use of it is correct. And we can find out whether we use an expression correctly by checking whether or not the rules that govern our actual use are 'maximal' in the sense that they exhaust the meaning of the expression in question. Whether or not a rule is maximal is not something that we in any way are able to decide. On the other hand, a maximal rule, i.e., a rule that is taken to be constitutive of meaning, does not commit us to the existence of an extra-linguistic reality, either. So, the position Dummett seeks to develop is opposed to radical conventionalism, but it is also opposed to realism in that it holds that a maximal rule that governs the use of an expression is not determined by the use we just happen to make of that expression, nor that it is determined by something that is totally independent from the activity of language use. Instead, maximal rules are thought to be given with the possibility of speaking a language in general. Using the same terminology, we may characterize the PT-procedures as a means for determining the maximal rules, thus supporting the view that they are to be taken as the tools of a transcendental analysis.

All this may give the impression that a definite conception of meaning and of use is already available. This is not the case, at least not according to Dummett. The problem is, he believes, that as yet we do not even have an adequate idea of how to construct these conceptions. And it is this problem he wants to tackle. It has two closely connected aspects. On the one hand, the question is how we can form an adequate conception of meaning. Elsewhere I have argued at length that it is a cornerstone of Dummettian anti-realism that a conception of meaning is adequate only if it can be explained completely in terms of use.[66] In this chapter, we have focused on the other aspect, the assumption that rules of use are correct only if they completely determine the meaning of the expression that they govern. And the question that must now be taken issue with is whether the demand of conservativity, as it is defined above, adequately expresses this assumption, at least as far as the logical connectives are concerned.

[66]See chapter I.

3.4 Normalization

How is it possible to distinguish rules that do exhaust the meaning of an expression from rules that do not? Can this be done on the basis of the demand of conservativity? Let us once again return to the language $\mathcal{L}_{\{\wedge,+\}} \cup \{\vee\}$, which is a non-conservative extension of $\mathcal{L}_{\{\wedge,+\}}$. Using Dummett's conjecture the fact that adding '$\vee$' to $\mathcal{L}_{\{\wedge,+\}}$ yields a non-conservative extension can be accounted for, since $\mathcal{L}_{\{\wedge,+\}}$ is not stable.

In his informal analysis of the reason that prevents conservativity to occur, however, Dummett does not make use of the notion of stability. Adding normal disjunction to a language containing '$\wedge$' and '$+$' as its only connectives does not result in a conservative extension, Dummett states, because '*normalization* is impossible in a logic containing all three connectives.'[67] Roughly, a logic is normalizable if, whenever there is a derivation of some formula A from a set of premises Γ in that logic, there is a *normal* derivation of A from Γ. And a derivation is normal if it proceeds without any detours, that is, if there is not some formula with principal operator **c** which occurs both as the consequence of the introduction rule for **c** and as the major premise of the elimination rule for **c**.

Dummett's appeal to normalization is understandable in as much as this notion is closely connected to the notion of intrinsic harmony. Intrinsic harmony has been defined in terms of the notion of a local peak, which is nothing but a special case of a detour: a local peak is a detour in which one and the same occurrence of a formula with **c** as its principal operator is both the consequence of the introduction rule for **c** and the major premise of the elimination rule for **c**. In a similar vein normalization can be seen as a generalization over intrinsic harmony.

The observation that a language containing '$\wedge$', '$\vee$', and '$+$' is not normalizable is correct. Consider the following derivation of the distributive law for '$\wedge$' and '$+$'.

$$\dfrac{\dfrac{\dfrac{A \wedge (B+C)}{B+C} \quad \dfrac{[B]}{B \vee C\ *} \quad \dfrac{[C]}{B \vee C}}{B \vee C\ *} \qquad \begin{matrix}[B]\\ \Pi_1\\ \vdots\\ (A\wedge B)+(A\wedge C)\end{matrix} \qquad \begin{matrix}[C]\\ \Pi_2\\ \vdots\\ (A\wedge B)+(A\wedge C)\end{matrix}}{(A\wedge B)+(A\wedge C)}$$

[67]Dummett 1991b, p. 290. My emphasis.

The following indicates why normalization is impossible. Notice first that the derivation does contain a detour: we first introduce a disjunction, which is eliminated two steps later. Following Prawitz 1965 the elimination of such a detour proceeds in two steps. First, if necessary, we rearrange the deductive steps in such a way that the application of the introduction rule is followed immediately by the application of the elimination rule (that is, maximal segments of length greater than 1 are reduced to maximal segments of length 1). We then replace this local peak by its reduction. In that way we eliminate, step by step, every detour, thus obtaining a normal deduction.

The problem for the derivation of $(A \wedge B) + (A \wedge C)$ above is the following. The starred occurrences of $B \vee C$ form a maximal segment of length 2. We want to reduce it to a maximal segment of length 1. Following Prawitz, this is done by converting the above derivation into:

$$\dfrac{\dfrac{A \wedge (B + C)}{B + C} \qquad \begin{array}{c} \dfrac{[B]}{B \vee C} \\ \vdots \\ (A \wedge B) + (A \wedge C) \end{array} \qquad \begin{array}{c} \dfrac{[C]}{B \vee C} \\ \vdots \\ (A \wedge B) + (A \wedge C) \end{array}}{(A \wedge B) + (A \wedge C)} \; E_{+}$$

But the result is not a valid deduction; the minor premises of the $+$-elimination rule are not directly derived from the assumptions B and C. Hence, E_{+} cannot be applied to obtain $(A \wedge B) + (A \wedge C)$ by eliminating $B + C$. We conclude that the conversion preserves validity, only if we 'unrestrict' the disjunction and appeal to $E_{\vee}$ instead of E_{+}. If we want to hold on to the restricted disjunction, we cannot lower the length of the starred segment, which says, in effect, that we cannot obtain a normal form for the deduction of F from $A + (B \wedge C)$.[68]

Now, what is the connection between normalizability and conservativity? Does the fact that the resulting language is not normalizable explain why this language is a non-conservative extension of the original language? As an answer to the first question, we can see that normalization is sufficient for conservativity, as the following argument

[68] The conclusion to be drawn from this is that normalization is impossible if we confine ourselves to the normal procedures. However, solely on the basis of what is said we cannot exclude the possibility that normalization is possible in a different, non-standard way. That, in this case, this may seem highly implausible is not enough reason to dismiss this possibility outright.

conclusively establishes: 'Normalizability implies that, for each logical constant **c**, the full language is a conservative extension of that obtained from omitting **c** from its vocabulary. For, if we have a proof whose final sequent does not contain **c**, any sentence containing **c** must first have introduced by an introduction rule, and then eliminated by an elimination rule; hence, by normalization, we can obtain a proof not involving that sentence.'[69]

Normalizability, however, is not a necessary condition for conservativity. We can see this as follows. Let $\mathcal{L}$ be a logical language which contains only two connectives: '$\vee$' and '$\wedge$'. Now we add '$+$' to it. The resulting language $\mathcal{L}'$ is a conservative extension of $\mathcal{L}$. It is, however, not normalizable; we still cannot get a normal proof of the distributive law for $\wedge$ and $+$. And because of this, our answer to the second question must be that a failure of normalizability does not explain why conservativity fails.

Taking both observations together we must conclude that normalization is a stronger demand than is conservativity and this gives rise to the thought that, apparently, requiring mere conservativity will not do for the purposes at hand. Is there any rationale for believing this? Indeed, there is. Throughout this subsection we have assumed that if we would have a criterion for conservativity, we would thereby have a means for determining whether or not the rules we take as governing some logical constant are self-justifying. And, clearly, these means should enable us to decide this solely on the basis of the interpretation of the constant in question, and independently of any other considerations. It is precisely at this point that mere conservativity falls short, because whether or not adding a given constant **c** to some logical language $\mathcal{L}$ yields a conservative extension depends not only on the interpretation of **c**, but also on which connectives $\mathcal{L}$ already contains. Consider, for example, the difference between extending a language that contains only '$\wedge$' by first adding '$\vee$' to it and then '$+$', with extending that same language by first adding '$+$' to it and then '$\vee$'. While the resulting languages will be identical in both cases, it is only in the second case that we will be confronted with a failure of conservativity, and hence with a reason to revise the interpretation of one or more of the constants in question.

To eliminate any possible dependency on the order in which the

[69]Dummett 1991b, p. 250.

constants are introduced, we say that the rules that govern a constant are self-justifying if they are for any possible order of introduction. In other words, we require that, relative to some given language, every extension of every fragment of it is a conservative extension. I will call a language for which this holds *conservatively closed*:

8. DEFINITION. (Conservatively closed language) *Let $\mathcal{L}_\Gamma$ be a language comprising the non-empty set of logical connectives Γ. Then $\mathcal{L}_\Gamma$ is said to be* conservatively closed *if and only if for every language $\mathcal{L}_\Delta$ with $\Delta \subseteq \Gamma$ it holds that every language $\mathcal{L}_{\Delta'}$ such that $\Delta \subseteq \Delta' \subseteq \Gamma$ is a conservative extension of $\mathcal{L}_\Delta$.*

And it is not difficult to see that a language is conservatively closed if the normalization theorem holds for it, since, if it holds for the language as a whole, it holds for every proper part of it. Hence, every extension of every proper part is normalizable, and thus, a conservative extension.

4 Concluding Remarks

4.1 Dummett's Conjecture Revised

In the final section of this chapter I will use its main results to develop a critical perspective on the PT-procedures as they originate from Dummettian anti-realism. A first question is: How are we to evaluate Dummett's conjecture in the light of the analysis discussed in the preceding section? Recall that this conjecture stipulates a condition which together with intrinsic harmony would ensure conservativity. But having established that mere conservativity does not provide a workable criterion for the delineation of the set of correct, or self-justifying rules, Dummett's conjecture, if true, answers the wrong question. We do not need to know what guarantees conservativity, but what ensures normalizability. Could we not rephrase Dummett's conjecture in this spirit? That would yield a proposition like this: Intrinsic harmony implies normalizability in a context where stability prevails. Unfortunately, it is false, as we already saw above; a logic containing '$\wedge$' and '$\vee$' as its only connectives is stable, but adding '$+$' to it yields a logic that is not normalizable. The problem, of course, is that the resulting logic is itself not stable.

In this particular case the question is: How can we exclude a connective like '+'? This question was raised earlier,[70] but this time we are in a position to provide a direct answer: we require, of each logical constant that is presented to us as a candidate for extending our language, that the rules that govern its use are stable. It is immediately obvious that '+' does not pass this test, and this we could take as a sufficient reason for the claim that the rules that govern '+' do not qualify as self-justifying.

The thought underlying this move is that we take stability as a criterion for being self-justifying. And, following Dummett's programme, there is also good reason for doing so. Analyzing the general conception of use, or more particularly, the conditions under which successful language use is possible Dummett comes to the conclusion that the different aspects of the (assertoric) use of expressions are to be in *consonance*. The term 'intrinsic harmony' has been introduced to refer to the requirement that results from applying the general demand of consonance to the specific case of the logical connectives. We also found that a rather essential element in Dummett's conception of use is that the rules that govern the use of some expression *completely* determine its meaning. More formally, this boils down to requiring stability. As a condition, stability is stronger than intrinsic harmony, so we need not mention the latter separately. We thus find that stability, as a criterion for being self-justifying follows quite naturally from Dummett's interpretation of the dictum 'meaning is use'.

Something is missing from this account. What is the connection, if there is any, between stability and normalizability? My answer will not be definite. I will present it as a hypothesis, the truth of which, I believe, is of greater importance to a semantic anti-realist, than is the truth of Dummett's conjecture. Let me start my by raising a question: Is the language $\mathcal{L}_{\{\wedge,\vee,+\}}$ in order? On the basis of our findings in section 3 different answers could be given, depending on how we construct the language in question. One answer could be the following. We view $\mathcal{L}_{\{\wedge,\vee,+\}}$ as the result of adding '$\vee$' to $\mathcal{L}_{\{\wedge,+\}}$. As such, it is *non-conservative*, and from that we conclude that the original language $\mathcal{L}_{\{\wedge,+\}}$ is not in order.[71] If the original language is not in order, there is little reason to expect that any extension of it, like $\mathcal{L}_{\{\wedge,\vee,+\}}$, is in order,

[70] See p. 118.

[71] Why we conclude this and not that the rules governing '$\vee$' are not in order is discussed on page 117.

even though we cannot tell exactly in what respect it is not in order.

The second answer is this. We construct $\mathcal{L}_{\{\wedge,\vee,+\}}$ as the result of adding '+' to $\mathcal{L}_{\{\wedge,\vee\}}$. As such it is a conservative extension of the original language. But now we have another reason to dismiss $\mathcal{L}_{\{\wedge,\vee,+\}}$: the rules that govern '+' are not in order since they are not *stable*. Hence, no language containing '+' is in order.

The third and last answer is the shortest: The language $\mathcal{L}_{\{\wedge,\vee,+\}}$ is in order if and only if it is *normalizable*. In a way these three answers represent the successive stages we have gone through in section 3. Overviewing these different answers we find that in answering the question whether or not some language is in order we eliminate all considerations pertaining to the way in which that language is evolved. All that counts is the language as it is, and it is of no importance at all that the very same language can come into being in rather different ways. We have gone a long way to reach this conclusion, which may not be Dummett's. In fact, there is no explicit evidence I know of that it is. And perhaps the fact that Dummett does not clearly distinguishes between the object of evaluation viewed as the process of evolving, and as the result of a process of development, and the inherent tension between these two views, is just one of the reasons that make it so very hard to get a clear sight on the ideas underlying the PT-procedures. But whatever the case may be, it is clear that the requirement that for any connective the rules that govern it be stable, or the demand that the language as a whole be normalizable, does not take into consideration any aspect relating to the way in which the language was constructed.[72]

2. CONJECTURE. *If $\mathcal{L}$ is a language such that for every connective* **c** *of $\mathcal{L}$ it holds that the rules governing* **c** *are stable, then $\mathcal{L}$ is normalizable.*

This, I claim, rather than Dummett's conjecture, is the hypothesis semantic anti-realism inevitably is driven to. I will not attempt to prove it here, as Dummett did not took the trouble of proving his conjecture. From a strictly logical point of view this may be quite unsatisfactory. But the goals aimed at in this chapter are not of such a logical nature. They are far more modest in that they are concerned only with an evaluation of the method explored by Dummett. And

[72]The difference between process and result, more specifically that between proof as act and proof as object, deserves more attention than I can give it here. This distinction was first applied in Sundholm 1983. See also Sundholm 1993.

we need not answer the questions raised by it in order to be able to evaluate the method, or basic features of it. Apart from being a logical claim, conjecture 2 says that Dummett's conjecture, if correct, answers the wrong question. And looked upon from the evaluative point of view this is a significant result.

4.2 The Notion of a Logical Rule

There is yet another reason for not pursuing the questions raised by Dummett's conjecture any further. The claim that it is not Dummett's conjecture, but conjecture 2 that adequately formulates the problem an anti-realist has to face only amends a particular aspect of the Dummettian programme, leaving unimpeded its general outlook. That is, it leaves intact the belief that the goals the programme sets itself can be reached. But are we not too lenient in this? Let us once again turn to the notion of stability. It was found to be the most plausible candidate as the criterion for being a self-justifying rule. But consider the following, very odd language with two connectives '$\vee$' and '$\wedge$' governed by the following rules:

$$\frac{(A \wedge B) \vee (A \wedge C)}{A \wedge (B \vee C)}\, I_\wedge \qquad\qquad \frac{A \wedge (B \vee C)}{(A \wedge B) \vee (A \wedge C)}\, E_\wedge$$

$$\frac{A \wedge (B \vee C)}{(A \wedge B) \vee (A \wedge C)}\, I_\vee \qquad\qquad \frac{(A \wedge B) \vee (A \wedge C)}{A \wedge (B \vee C)}\, E_\vee$$

Actually, the connectives of this language are governed by only one rule: the distributive law, alternately used as introduction and as elimination rule. The introduction and elimination rules of this language are stable, yet it is hard to imagine anyone who would accept these rules as self-justifying. An anti-realist would certainly not accept them as such, since it is clear that they do not in any way determine the meaning of '$\wedge$' and '$\vee$' completely. Hence, there seems to be no other possibility than to conclude that stability is not the criterion we were looking for.

But if stability is not strong enough, then what condition is? Instead of searching for again a stronger condition, I think we would do better to inquire whether it is possible to find such a condition at all. Recall the observation that the PT-procedures presuppose a clear and workable notion of what it is to be an introduction or elimination rule.[73] And if this applies to the PT-procedures in general, this surely also applies to stability, a notion in terms of which the PT-procedures are explained. That is to say, stability as a requirement will only enable us to to single out the set of self-justifying rules given a suitable notion of what it is to be an introduction or elimination rule. This is not to say that the demand of stability does not have any impact at all on our conception of what it is to be an introduction or elimination rule. It does, because it has consequences for our conception of what it is to be a logical rule in general. But, and that is the cause of our problem here, it does not determine what is to be an introduction or elimination rule *completely*.

A natural response would be to look for those features which, together with stability, do completely determine the notion of an introduction and elimination rule. Undoubtedly, the logical complexity of the formula obtained by applying an introduction or elimination rule would play a role in this, and probably other features as well. But in doing this we would sidestep what I belief to be the crucial question: Are the introduction and elimination rules indeed as fundamental as Dummett believes them to be?

Dummett's programme is anti-realist in the sense that it rejects the thought that the meaning of the expressions of our language is determined by some language-external reality. Dummett's alternative to realism builds on the thought that our conception of the meaning of an expression cannot and must not exceed our abilities to use that expression. As a consequence, Dummett focuses entirely on the rules that govern the use of an expression in an attempt to develop simultaneously an adequate conception of meaning and of what it is to be a correct rule. In this chapter we have concentrated on the latter task, limiting ourselves to a particular class of expressions: the logical connectives. It is also clear that for Dummett, the possibility of a systematic meaning-theory coincides with the possibility of a compositional meaning-theory. A meaning-theory is compositional if the meaning it assigns to a complex expression is completely determined by the meanings of its constituent expressions. Assuming that meaning is use, compositionality demands that the use of complex expressions be determined completely by the use of its constituent expressions. This says, in effect, that for Dummett, the basic rules are precisely those that govern the use of the

[73] See page 119.

basic, or atomic, expressions. We saw, furthermore, that, according to Dummett, the use of an expression is governed by the rules that lay down the conditions for the utterance of the expression as well as by those that determine the consequences of uttering it. Applied to the logical connectives this yields the claim that the use, and thereby the meaning, of a logical connective is determined by its introduction and elimination rules.

The question to be raised is whether the claim that the meaning of an expression is determined by the rules that govern its use implies that the meaning of an expression, say a logical connective, is determined by its introduction and elimination rules. In other words: Is it correct to identify the introduction and elimination rules of a connective with the rules that govern its use? Throughout this chapter we have taken this assumption for granted, and we have done so because it is implicit to the Dummettian programme. We just saw why. Compositionality demands that the rules that govern complex expressions are, in a sense, derivative from the rules that govern the use of the basic expressions, like logical connectives. As a consequence, these latter rules are taken to be 'primitive', and it is by notions like stability that Dummett seeks to delimit the notion of such primitive rules. What is important to note, however, is that this approach cannot take into consideration rules that are not specific for some particular expression, like the *structural* rules. Structural rules are rules of inference that do not involve any logical constant. Belnap's 'context of deducibility', as we came across it in subsection 2.3, was stated in terms of structural rules weakening, contraction, permutation, and transitivity. Recent work on this topic has shown that what rules for the use of a logical connective, i.e., introduction and elimination rules, are possible and needed, is in fact determined by the structural rules that are accepted.[74] At the time of writing this, I believe this to be uncontroversial. But if we really take it seriously, it has far-reaching consequences for the Dummettian programme. The PT-procedures are based on criteria that the set of introduction and elimination rules of a language has to meet if it

[74]Since structural rules are rules for manipulating the context in which logical connectives occur, and assuming that it are the structural rules that determine the logical space for the interpretation of the logical constants, it appears that a strong case can be made for the claim that logical connectives are *essentially indexical*. Although I believe this claim to be very promising and of great interest, I will not pursue it here.

is to be self-justifying. But given that the 'interpretation space' of the logical constants is determined by the structural rules, this means that the PT-procedures single out the rules for the use of a connective that are correct only in the light of an antecedently given set of structural rules.

There is no set of structural rules that appears to be more natural, or more compelling than any other. For Dummett's programme this has at least the following consequences. In the first place, it should be investigated if the analysis of use as it has been described and discussed above could be applied so as to determine the set of correct or self-justifying structural rules. I will not pursue this matter here since it would take me far beyond the subject of this chapter. I only want to remark that, as I see it, it is far from obvious that it could be done successfully. The reason is that Dummett's analysis of use is always an analysis of the use of certain class of expressions, whereas the structural rules do not apply to expressions at all, but to the contexts in which these expressions occur.

The second consequence is, to conclude with, that as long as such an investigation has not been conducted satisfactorily, that is, as matters stand now, we have to face the fact that the proof-theoretic justification procedures as conceived of by Dummett necessarily underdetermine the question as to which the correct logical rules are.

IV Semantics and Comparative Logic[1]

1 Towards a Formal Interpretation of Navyanyāya logic

1.1 Logic and Language

Current Western-oriented interpretations of Navyanyāya logic have in common that they all, in some way or other, call upon technical devices derived from Western classical logic in order to clarify the basic concepts of Navyanyāya logic. Underlying this method there is, of course, the assumption that it is possible to formalize Navyanyāya logic and that such a formalization can add to an illumination and understanding of the most important features of Navyanyāya logic in general, and to an insight in the techniques for logical analysis as developed by the Navyanyāya logicians in particular. This assumption, I urge, ought to be distinguished carefully from the assumption that the logical framework within which a formalization of Navyanyāya logic should take place is already known and available. I insist on this distinction for, although I do not dispute the possibility and, eventually, the fruitfulness of the use of devices taken from Western formal logic for a clarification of Navyanyāya logic, I believe that the question as to which logical framework is most suitable for this purpose is still to be decided on.

It seems to me that the failure of recognizing the distinction between these two very different assumptions has given rise to a misunderstanding of some of the basic tenets of Navyanyāya logic, rather than to a clarification of the same. It should be noted, however, that what I consider to be an uncritical attitude towards deviant logics is believed by others to have a firm philosophical foundation. I am referring to a

[1]First published in the *Journal of Indian Philosophy* **20**: 377–418, 1992. Reprinted with kind permission from Kluwer Academic Publishers.

Western tradition known under the name 'Ideal Language Philosophy', according to which there is one and only one logic, where 'logic' is to be understood as extensional first order predicate logic (FPL). This tradition, the motivations of which have a strong nominalistic flavor, takes arguments formulated in some language or other as its sole concern. The relation of 'logical consequence' is looked upon as a relation between sentences, and is given a semantic explication in terms of truth in models, a syntactic explication in terms of proof or deducibility, and both kinds of analyses are tied together by a completeness theorem. Very basic to this approach is the delimitation of a set of certain linguistic items, the so-called 'logical constants', consisting of 'and', 'or', 'not', 'implies', 'some', 'every', and, occasionally, 'identity'. In the tradition of Ideal Language Philosophy it is claimed that logic is nothing but the study of the properties of these logical constants and, furthermore, that anything that cannot be defined in terms of these is outside the scope of logic.

Obviously, an adherent of Ideal Language Philosophy will not accept my distinction of the two assumptions mentioned above. For him they naturally merge into one. The question as to the formalization of Navyanyāya logic, he would say, simply boils down to the question as to which extent Navyanyāya logic can be translated into the language of first order predicate logic. The task to which he appoints himself thus, is, as I see it, really twofold. At first he will try to locate or, if necessary, generate within the language of FPL those expressions into which certain technical devices employed by the Navyanyāya logicians are to be translated. He then proceeds to subject the translated concepts to the calculus of classical logic in order to see whether the results thereof are accepted by the Navyanyāya logicians. In doing so, he will find out that this is not always the case, and so he will have to face a serious problem. Keeping with his view he has no other alternative than to declare those elements of the Navyanyāya theory of inference that obstruct the alleged acceptance of certain classical principles to be logically 'unjustified' or 'superfluous'. As I will point out later in more detail, this is exactly what happens whenever the Navyanyāya discernment of three kinds of inferences, or the Navyanyāya scheme of inference is analyzed within the framework of FPL.

1.2 Extending the Scope of Logic

In current Western logic it is more and more acknowledged that there are a lot of phenomena apparently exhibiting patterns of systematic relationships which just cannot be analyzed adequately within the framework of FPL. This has given rise to a new, and in some respects radically different, approach to logic. Underlying this new approach is, what some have called, the 'common sense view of logic'. It seeks to establish a conception of 'logic' by entering first into the ordinary usage of the words 'logic', 'inference', and 'meaning'. The usage of these words, it is stated, defines a natural subject matter that is broader than logic presently studied. Next to this it is noticed that in ordinary usage the meaning of the word 'logic' is not unequivocal, that is to say, it is acknowledged that speaking of, for example, 'the logic of believing', or 'the logic of perception', generally does not amount to the same. This has given cause to the assumption that the different concepts on the basis of which, for example, human beings are able to communicate with each other, to obtain scientific results, to experience the world in daily life and to organize these experiences in a systematic way, all have their own logic. Under this view inference is just one of the means available to human beings to organize and to gain knowledge about their world. Inference then, must be looked upon as having its own logic, too. An analysis of inference taking the 'common sense view of logic' as its starting point will try to explicate this logic in contrast with, and, at the same time, in connection with the logic of other concepts governing human life, as, for example, perception.

To speak about 'inference having a logic of its own' will sound like a pleonasm to the proponents of logical nominalism, for what else can logic be about if it is not about the rules and principles in terms of which the soundness and validity of certain inferences can be accounted for? Indirectly, this question brings up another, as it suggests, one that is to be thought of as being already answered; What is logic all about?

This chapter has its ultimate motivations, first of all in the conviction that the answer to this rather important question provided by logical nominalism is not the only one possible, but most of all in the conception that the way in which this issue is given its solution does have consequences for the properties of the logical calculus emanating from it. In short then, this chapter originated from the opinion that an answer to the question 'What is logic about?' does matter and it can be looked upon as an attempt to illustrate this point.

Studying Navyanyāya has given rise to the thought that, in contrast with FPL, Navyanyāya logic is not about language. To face this possibility, however, is not to say that an understanding of this logic can go without an understanding of the closely related notions of inference and meaning. But it does carry with it the claim that for an understanding of these notions we should not limit ourselves to language. This, in fact, is one of the basic claims of what has become known as *situation semantics*, and it vindicates the methodological choice for a comparison of Navyanyāya logic with situation semantics, rather than with FPL. It follows from the extended notion of meaning held in situation semantics that inference is looked upon not as the correct application of appropriate rules operating on syntactical structures, but as 'an activity that attempts to use facts about the world to extract additional information, information implicit in the facts.'[2] A sound inference, it is argued, does not need language at all, a conviction that is unmistakably on a par with the Navyanyāya conception of inference.

The framework offered by situation semantics can be seen as an attempt to develop the common sense view of logic. It seeks to develop accounts of information and inference that do not presuppose language. Speaking in a more general fashion, it is intended to give an over-all theory of meaning within which a theory of meaning for human languages can be developed as a special case. It is also intended to provide us with the means for representing mental states and to allow for an understanding of the meaning of these. Before outlining the starting points of situation semantics in more detail, I will concentrate on the notion of *jñāna*. I will provide an interpretation of this notion as a preliminary to a comparison of Navyanyāya logic and situation semantics with respect to the representation of mental states and the role assigned to these in the analysis of inference. Next to this I will focus my attention to the notion of *vyāpti*, in particular to the conditions on the grounds of which the acceptance of a vyāpti-relation is justified. Finally, I will turn to the inference scheme as maintained by the Navyanaiyāyikas, thereby proposing an interpretation for which I will call upon some insights and devices derived from a situation semantical framework.

[2]Barwise 1989, p. 39.

2 Representing the Mental

2.1 *Jñāna*

A main feature of the Navyanyāya analysis of human knowledge can be traced back at least as far as Vātsyāyana's commentary on the Nyāyasūtra's, where it is plainly stated that 'cognitions of several things appear one after another.' To support his claim Vātsyāyana appeals to a pre-theoretical intuition: 'it cannot be denied, since it is directly perceptible by each man for himself'.[3] These remarks state explicitly what unmistakable underlies the Navyanyāya approach to knowledge implicitly; human cognition is analyzed and described in terms of successive 'cognitions', called *jñāna* or *buddhi*.[4]

From an ontological point of view, a jñāna is subsumed under the category called *guṇa*. As such I take a jñāna to be a non-repeatable, momentary entity.[5] Being a guṇa a jñāna is a particular, a unique characteristic of exactly that entity that possesses it, for as long as it lasts. Although a jñāna shares these characteristics with other guṇa's, it also has a distinctive which it does not share at all; typical of a jñāna is its property of having no other locus than the substance *ātman*.[6] Since, to a certain extent, this substance can be looked upon as a constituent of the individual, internal reality, I will consider a jñāna as, what I will call, an 'epistemic event'. I want to state explicitly that the word 'epistemic' is meant to indicate nothing more than that a jñāna has ātman as its only substance, and that the word 'event' should be taken only as an expression of its non-repeatable, momentary character.[7]

The same intuition, perhaps, that reveals the successive occurrence of jñāna's or mental states, also tells us that the way in which they do so, generally, is not arbitrarily, but exhibits certain regularities. An awareness of there being smoke at a particular place, for example, is normally followed by the awareness of there being fire, too. The

[3]See *Nyāyabhāṣya* under *Nyāyasūtra* 3.2.56 - 3.2.58.

[4]*Tarkasaṃgraha* 34.

[5]Cf. Potter 1954, Potter 1957, and Matilal 1986.

[6]*Tarkasaṃgraha* 17.

[7]The reason for this remark is that I have not, no more than Matilal has, found any evidence for believing that the Navyanaiyāyikas made an explicit distinction between 'intern' and 'extern', or 'mental' and 'physical' as it is common practice in Western philosophy ever since Descartes (cf. Matilal 1986, p. 128). It is noteworthy, furthermore, that in situation semantics this distinction also is hard to seek.

Navyanyāya theory of inference is an attempt to account for just this kind of regularities. Stated in this way it seems obvious that, since the Navyanyāya theory of inference seeks to account for certain regularities in the successive appearances of jñāna's, the relation of entailment holding between the set of premises and the conclusion of a valid inference really is to be conceived as a relation between jñāna's. Although I do not think that this is so obvious as it might seem, the thought of it naturally gives rise to two related questions. If jñāna's are taken as the constituents of an inference, and if, consequently, the relation of entailment is construed as a relation between these *mental* entities, then it might be put forward that entailment is constituted by the psychological mechanisms being laborious in a cognitive agent. In short, it might be retorted that a psychological interpretation of Navyanyāya logic is justified.[8]

The other question connects to the possibility of objective knowledge. The conception of cognition as being analyzable in terms of *non-repeatable, momentary* entities brings with it that the 'epistemological given' cannot be seen apart from the knowing subject in which, nor from the moment at which it occurs. As a consequence, knowledge becomes a highly subjective matter. The question is, then, how it is possible to achieve the generality demanded for by philosophical analyses, in other words, how is it possible to gain insight in knowledge in general?[9]

Both questions, as I see it, have their origin in laying too much stress on the ontological status of jñāna's; the former question relates to the 'mental' character of a jñāna, while the latter focuses only on its being non-repeatable and momentary. This is not to say that both questions are unimportant, but only that they are rooted in a one- sided portrayal. I myself think that it is possible to interpret a jñāna so as to do justice to its ontological distinctives, while at the same the alleged generality can be preserved without resorting to psychological mechanisms. In order to spell this out more fully I will borrow some devices from situation semantics. As a preliminary then, a global discussion of this theory must be brought up.

[8]Such an interpretation is proposed in Mohanty 1985.

[9]Considerations along these lines persuade Mullatti to interpret a jñāna as a proposition in a Fregean sense of the word. The interpretation of 'jñāna' as I propose it is also meant to be a criticism of this conception.

2.2 Situation Semantics on Mental Representations

As I have stated earlier situation semantics aims at a general theory of meaning within which a theory of meaning for human languages can be developed as a special case. Consequently, the proposed conception of 'meaning' is not limited to language. Typical of situation semantics is a strong commitment to a form of realism, that is to say, to the claim that meanings do not reside in the head nor in some Platonic realm, but are constructions out of real things. The theory also allows for a representation of mental phenomena. Mental states are conceived as (real) states reflecting the way the world is. They thus contain information about the agent's environment and it is claimed that they can therefore be adequately characterized by the information they contain. Then, just as linguistic meanings are, the informational content of mental states is taken to be a construction out of real things.

Basic Assumptions of the Theory

(1) The *primitives* of the theory are individual particulars. Included are ordinary things but also situations (including events) and spatio-temporal locations. Properties (or 1-ary relations) and (n-ary) relations are also taken as primitives.[10]

Situation semantics is rather liberal in its choice of primitives. FPL, in contrast, does not take properties and relations as basic, but as sets of n-tuples of individuals. The admission of so many primitives has its roots in the observation that anything humans systematically use is an invariant across situations and as such can easily become, upon human reflection, objectified, and so be treated as a thing in its own right.[11] Thus, in situation semantics everything that is used can be objectified and talked about. Relations are freely and frequently used, and so relations too can be objectified and treated as objects of the theory. Relations are very important in situation semantics; they are said to be the 'glue' that hold things together, the primary facts that go to make up reality. This brings me to a second assumption:

[10] This assumption, and the other ones to be discussed in this subsection can be found listed in Barwise 1989, p. 81. For their representation I have taken some minor, mainly typographical, liberties.

[11] This view is also justified by referring to the ease with which we can nominalize in natural language.

(2) There is a single world W, the real world having no concrete alternatives, that determines the basic facts as to what particulars stand in what relations.

Each relation comes with a set of argument places, roles that can be filled to get a basic state of affairs. The relation of 'holding', for instance, comes with three argument roles to be filled by the spatio-temporal location l, the holder a, and the thing b hold by a at l. What becomes clear from this is that, generally, a n-ary relation comes with a set of $n + 1$ argument places one of which is to be filled by a spatio-temporal location.[12] This is only true on the whole, since situation semantics also allows for unlocated relations, as we shall see.

(3) Legitimate assertions are about the world or portions thereof. These portions are *situations*. To each particular situation s there corresponds a set of *facts*, the facts that hold in s. In general, the set of facts associated with a given situation may be a proper subset of all the facts of W.

A situation is conceived of as a part of reality that can be comprehended as a whole in its own right. This, however, is not to say that a situation can be comprehended as totally independent of other parts of the world. In fact, a very important feature of situations is that they always, in some way or other, interact with other things (including other situations), or, what amounts to the same, they have properties or relate to other things. It should be noted, furthermore, that, since a situation comprises a spatio-temporal location, it is viewed as a momentary, non-repeatable object.

Situations are taken to be the primary semantic objects. Since these are conceived of as being parts of bigger situations, i.e. they stand in the relation *is a part of* holding between situations, rise is given to an

[12]The conception of a relation as accompanied by roles to be filled does, to a certain extent, remind of the Navyanyāya view of relations as coming with a 'limitor' (avacchedaka) and a 'determiner' (nirūpaka), since both can be looked upon as naming argument places. It should be noted also that, in contrast, Navyanyāya does not allow for relations with three or more arguments; these are always reformulated in terms of (complexes of) binary relations. Navyanyāya, furthermore, does not assign a spatio-temporal as an additional role to a relation, although, as it seems to me, the term of the relation that can be conceived of as the *adhikaraṇa* does not rarely imply a fixation at a real spatial location.

important characteristic of situation semantics in which it rather deviates from FPL. This feature is commonly referred to by the notion of *partiality*. It should be mentioned, to be sure, that partiality can come to light for two different reasons. Partiality in, what I call, the *semantic* sense of the word sets in whenever a theory employs partial functions, functions that do not return a result for every argument in the domain. A partial semantics, then, is one making use of an interpretation function that does not assign to every statement one of the two definite truth-values 'true' or 'false'. Partiality in this sense, for example, is encountered in theories that seek to account for presuppositions semantically. In the *ontological* sense of the word partiality means that the theory under consideration takes partial objects as its primary semantic objects. So, situation semantics is partial in the ontological sense of the word. But, since taking recourse to partial objects always brings with it the employment of partial functions, situation semantics is partial in the first sense of the word, too. The converse, though, does not hold. That is to say, the use of partial functions does not need to go together with taking partial object into consideration.

Due to partiality situation semantics exhibits some features in which it radically differs from FPL, a logic that is based on a so-called 'total' semantics; i.e. a semantics on the basis of which every statement is assigned a definite truth-value and that does not take partial objects into account. Two hallmarks I want to touch upon here concern the law of the excluded middle and, what is sometimes called, the question as to the persistence of statements.

Turning to partial objects has as its direct consequence that the law of the excluded middle is not longer tenable. For, although it may be the case that every total world satisfies '$\mathrm{R}(a) \vee \neg\mathrm{R}(a)$', it does not follow that a given part of the world should do so, too. It is important to realize that under the view I am discussing the question whether a certain statement holds in the actual world, that is to say, whether it describes a state of affairs that holds in some real situation, is answered not by taking the total world into account, but *by reference to a given situation s*, the situation under consideration. This situation, of course, might be the whole world, but this certainly does not go without saying, as it does in FPL. Now, if the situation under consideration is not the total world but a part thereof, then it is to be acknowledged that this situation possibly does not provide us with sufficient information to decide whether a described state of affairs holds in s or not. A situation s, for instance, represented by, say, 'in s: at l: Sitting, John, chair; yes' does not provide the information on the basis of which it can be decided whether the state of affairs described by 'John reads the newspaper' holds in it or not.

The question as to the *persistence* of statements also is one that naturally arises within any semantical framework based on partial objects like situations. It amounts to the question whether a statement having some truth-value with respect to a given situation s will have the same truth-value when a bigger situation s' such that s is a part of s' is taken into consideration. If so, the statement is called persistent. An example of a statement that is not persistent is: 'All the workers have gone on strike'. This statement being true, perhaps, with respect to a particular factory, will turn into a false statement if other factories are also taken into account.

Notwithstanding the fact that situations are really non-repeatable, momentary entities, it is also acknowledged that different situations can exhibit common features or, so-called, invariants across situations. This is captured by allowing for the assignment of properties to situations:

(4) Just as there are properties and relations among other particulars, there are properties of, and relations between situations and other particulars. A property of a situation is called a *type of situation*. Since these properties are not extensional it is not supposed that two distinct types of situations will be types of distinct real situations. Neither it is supposed that every type of situation is the type of some real situation.

A type of situation is obtained by abstracting from a number of constituents of a situation. It is used in order to capture the invariants across situations. Among these invariants, it is acknowledged, there are not just objects and relations, but also congeries of these. Take, for example, a situation s in which John reads the newspaper. By abstracting from all the argument places the type of situation S is obtained, a type that can be expressed verbally by 'someone reads something at some spatio-temporal location'. It is, to be sure, also possible to obtain other types of situations, for example S' expressed by 'someone reads a newspaper at some spatio- temporal location'. It is said that the situation s is of type S or, what amounts to the same, that s has the property of being of type S, or that s exemplifies S. A type of situation can also be viewed as an abstract object reflecting the internal structure of any situation that is of that type. The notion of a type of situation plays a key role in the conception of meaning put forward in situation semantics. It is claimed that meaning resides in systematic relations of a special sort between different types of situations. These relations are called constraints and it is in terms of these that the meaning of the primary semantic objects is determined.

In situation semantics 'being meaningful' amounts to 'containing information'. A situation s can contain information about another situation s' only if there is a systematic relation C that holds between situations sharing some configuration of uniformities with s and situations that share some other configuration of uniformities with s'.[13] That is to say, a situation is meaningful, only if it exemplifies a type of situation that is systematically related to another type of situation.

The theory of constraints, i.e., systematic relations between types of situations, as it is developed within situation semantics cannot be seen apart from the intention to formulate a thoroughgoing realistic theory of meaning. In keeping with this aspiration the theory claims that constraints, and consequently meanings, are actually nothing but regularities in the world. The fact, furthermore, that such regularities obtain is itself a fact among other facts: facts like these are (parts of) situations.

Being a fact among other facts, a constraint is basically the state of affairs that two types of situations stand in the primitive relation of *involvement* at the universal location l_u. In recent developments, however, this relation is conceived of as being unlocated, a conception I will adopt, too.[14] Symbolically I will represent a simple constraint C as: $C :=<<\text{involves}, S, S' >\text{yes}>$, or simply as: $C := S \Rightarrow S'$.

Linguistic as well as non-linguistic events are said to be meaningful if they have the property of conveying information about, or 'involving', other parts of reality obtaining in the world. The information conveyed by events is accessible for an agent by means of an *attunement* to the

[13] Barwise and Perry 1983, p. 12.

[14] In the context of situation semantics 'unlocated' means primarily 'universally located' and is, thus, not to be confused with the Navyanyāya notion of 'aprasiddha' for this notion, rendered as 'unlocatable', amounts to 'being nowhere locatable'. It should be noted, furthermore, that situation semantics also allows for constraints that do not hold universally, but only 'regionally'. I will not discuss these, since vyāpti-relations holding only regionally are, as far as I know, not acknowledged in Navyanyāya logic.

constraints that obtain in the world. That is to say, an attunement to the systematic relation between the types of situations S and S' is what allows an agent to pick up the information that is contained in a situation s, given that s is of type S. One of the examples occasionally used resorts to the types of situations S and S', where $S :=<\text{at } l < \text{being smoky}, a > \text{yes} >$, and $S' :=<\text{at } l <\text{being firy}, a > \text{yes}>$, and a constraint C having S and S' as its constituents: $C := S \Rightarrow S'$. It is said that on the basis of an attunement to C an agent is allowed to extract the information that there is fire in some situation s' from a given situation s being of type S, that is, a situation in which there is smoke.

What has become clear from all this is that systematic relations between types of situations, constraints, are what allows one situation to contain information about another situation and, hence, to be meaningful. These constraints are also called *type-meanings*. An attunement to such constraints is what allows an agent to soundly infer from the one thing being the case to the other thing being the case. In other words, type-meaning is what allows an event of a particular type to have *situation-meaning*. So, a given situation s in which there is smoke has a particular situation-meaning that it is of type S, and, given an attunement to the constraint C mentioned above, that type of situation is known to involve there being a situation s' of another type S', a situation in which there is fire. Succinctly, then, this can be expressed by saying that 'smoke "means" fire'.

Given these basic assumptions on the basis of which a general theory of meaning is formulated, the meaning of linguistic items can be developed as a special case taking recourse to the next assumption:

(5) The chief semantic value to be associated with a declarative sentence P is its meaning, a relation $U_p \Rightarrow D_p$ between the type U_p of the situation where P is assertively uttered and the type D_p of the situation thereby described. A legitimate assertive utterance u of type U_p is true if the situation s_u that u is about is of type D_p. The information content of the utterance u is that s_u is of type D_p.

Since I am primarily concerned with the relation between logic and semantics, and since in situation semantics linguistic meaning is conceived of only as a special case of a general semantics the basic assumptions of which are outlined above, I will not discuss assumption (5). Instead, I want to turn my attention to the question as to the representation of the mental as it is put forward in situation semantics.

Faithful to its realistic outlook, the theory commits itself to the existence of real mental states. In fact, mental states are conceived as just being events among the various other states of affairs and events in the world. Mental states, furthermore, are taken to be representational, that is to say, as representing the way the world is. Now, since the world consists of situations, mental states are taken to be mental representations of situations and, consequently, it is claimed that they can be described by the same means as those employed in the analysis of situations.

Just as other events, mental states are meaningful, for they are systematically linked to other situations, external and internal. One could also say that mental states have an informational content, thereby recalling that the basic intuition about the informational content I_s of an event s is that it is information about something besides s. In connection to the question regarding the informational content it is to be noted that mental states differ from other situations in one important aspect, for the former are representations, while the latter are not. As a distinctive, representations have the property of being usable in different situations getting at quite different contents in these circumstances.[15]

Since mental states are taken to be representational, they are classified in terms of types of situations, and hence in terms of (constructions out of) real things. So, types of situations play a double role, since they are used to classify what is seen as well as how it is seen. In this way the theory is loyal not only to its realistic background, but also to common sense. For it is said that by this method justice is done to the common sense picture of the mental that there are different ways of seeing the same thing. The type of situation S classifying the cognitive state an individual a is in at a location l is connected to the world by a setting, a list of individuals that fill in the argument roles of S, thereby relating the mental state to the particular object that is cognized, and entering the circumstances that, to a certain extent, contribute to the informational content of that particular mental state.

[15]This in contrast to a view of representations as having an intrinsic content, independent of circumstances.

2.3 Jñāna's and Situations

In Navyanyāya literature jñāna's are discussed by taking recourse to certain verbal expressions, usually sentences, which are 'marked' by means of the word 'iti' immediately following after such an expression. In contexts like these the word 'iti', which has many meanings, is usually translated into quotation marks. So, what we have are jñāna's on the one hand, and 'quoted' sentences by which jñāna's are talked about on the other. This being the case it will not come as a surprise that in FPL-based approaches to Navyanyāya logic the notion of jñāna is interpreted by establishing some direct relation between a jñāna and the sentence by which it is talked about. Hence, since FPL is solely concerned with sentences and their meanings, a jñāna is not rarely taken to be, in some way or other, the meaning of the sentence associated with it.[16] But what reasons are there for this conception? Are jñāna's meanings of sentences? I myself am not at all convinced that they are and to express this opinion I will outline an interpretation according to which they are not.

In its essence my suggestion comes down to the idea of taking jñāna's not as entities that *are* meanings, but as entities that *have* meanings. From this it is readily seen why I took recourse to situation semantics as a methodological choice. For situation semantics not only adheres itself to a form of realism that in many respects turns out to be on a par with the backgrounds of Navyanyāya logic, but it also provides the means which suit my purposes well. The major shift of focus I propose, then, is to take jñāna's to be entities that, just as linguistic entities, have meanings, and to assume that the meanings of both kinds of entities are constructions out of real things. It should be noted that under this conception the means by which the meaning of a linguistic expression is analyzed coincide with those by which the meaning of a jñāna is estimated. This, however, is not to say that jñāna's are meanings of sentences.

At first sight, perhaps, it may sound odd to seek for the meaning of a jñāna. But it should be realized that I take 'meaning' in a very special sense of the word. For I conceive of meaning as providing a relation between two events obtaining in the actual world. That is to

[16]It is, for example, not uncommon to take a jñāna as a proposition in the Fregean sense of the word.

say, I retain a relational account of 'meaning'.

In Navyanyāya to have a cognition always means to have a cognition of something. Subsequently, a jñāna is characterized by the property of being related to some object or other (saviṣayakatva).[17] In the language of Navyanyāya, this aspect of a jñāna is captured by saying that, roughly, the cognized object possesses the property *viṣayatā*, a so-called relational abstractum meaning something like 'being the object of [x]', a property that is determined (nirūpita) by the cognition by which that very object is cognized. Or, the other way round, it is said that the cognition has the property *viṣayitā*, perhaps to be rendered as 'having [x] as its object', which is determined by the particular object that is cognized. The cognition of an object taken as the presence in a certain cognitive agent at a fixed moment of an epistemic event being related to the object cognized by *viṣayatā* or, depending on the line of approach, *viṣayitā*, presents that object together with one or more distinguishing attributes.[18] In Navyanyāya, then, the cognition of an object is conceived of as being of a qualitative nature, a conception that is reflected in the way a jñāna is analyzed. For its description breaks up into a qualificandum (viśeṣya), a qualifier (viśeṣaṇa or prakāra), and a relation between both (viśeṣya-viśeṣaṇa-saṃbandha).

In order to be able to determine the value of the 'pseudo-variables' just mentioned with respect to a given jñāna recourse is taken to the assumption that it is possible to associate the jñāna under consideration with a linguistic expression.[19] In keeping with my view, the 'criterion of association' is not that the jñāna is the meaning of the expression, but that the jñāna and the expression associated with it have the same meaning. That is to say, a jñāna can be associated with a certain expression if and only if that expression relates to exactly the same object from which the jñāna derives its content. The relation between an expression φ and its denotation is established by stating that the denotation (śakya) has the property of 'being the denotation of' (śakyatā),

[17] Compare Matilal 1968, §§2, 3.

[18] That is to say, the cognition relates to an object by being the determiner of the complex property viśiṣṭa-viṣayatā (qualified object-ness), a property that is made up of the viṣayatā's resident in the several entities of which the cognized object is composed.

[19] It has to be noted that only those jñāna's which have a relational structure are allowed for to be associated with a verbal expression. jñāna's having a non-relation structure, however, fall outside the scope of this chapter, so I will not discuss these.

a relational abstractum that is thought of as being determined by the expression φ taken as an entity in its own right. From a methodological point of view it can be remarked, leaving aside all kinds of technicalities, that the relation between a cognition and its object is fixed along lines very similar to those followed to capture the relation between an expression and its denotation. In both cases, furthermore, the object possessing the property *viṣayatā* or *śakyatā* is taken to be a real object obtaining in the actual world. Taking this as a starting point, the general idea underlying my 'criterion of association' could be formulated in terms more familiar to the language used in Navyanyāya, perhaps, by saying that an expression φ can be associated with a jñāna σ if and only if φ determines a *śakyatā* resident in an object, which object also possesses the property *viṣayatā* being determined by σ.

The syntactical structure of the associated expression contributes to the analysis of a jñāna, for, to a certain extent, it is taken as a guideline for determining which entities are qualifiers of which other entities. This, however, is not to say that the meaning of the sentence is derived solely from its syntactical analysis. As I see it, in Navyanyāya the meaning of an indicative sentence is the object described by the sentence, and this object has, in fact, a much more complicated structure then one would suspect by taking only the syntactical structure into consideration. So, to explicate the meaning of a sentence fully, the object referred to has to be submitted to an ontological analysis. The expression '[This] mountain possesses fire' (parvato vahnimān), for example, describes a complex object that ontologically falls apart into the particulars 'mountain' and 'fire', the generics 'mountain-ness' and 'fire-ness', a relation between 'mountain' and 'mountain-ness' as well as a relation between 'fire' and 'fire-ness', and a relation between 'mountain' and 'fire'.[20] These constituents including the relations are conceived of as real things and they all contribute to the meaning of the sentence '[This] mountain possesses fire'.

Now, if this exposition is acceptable as a general depiction, as I think it is, and recalling the basic traits of situation semantics, it is not very difficult to interpret the notion of a jñāna in a rather straightforward fashion. All that is needed is to assign to a jñāna a 'structural content' as an constituent of its total content. I call this part of its whole content 'structural', because it must be such that it is usable with

[20] Even this is a simplified representation.

respect to different objects. Since a jñāna derives its content from the particular object it is related to by *viṣayatā*, and since that object, though being momentary, also exemplifies 'invariants across different objects', it seems obvious to view the structural content of a jñāna as consisting of those entities which different actual objects have in common; i.e. the generic entities. So, I take the structural content of a jñāna to be a construction out of real generic entities which are obtained by abstracting from their various instantiations. In fact, I take the structural content of a jñāna to be just what in situation semantics is called a *type of situation*.

The 'type of situation' is what different jñāna's might have in common. From a different angle, however, there are no two jñāna's that are the same, because they are conceived of as momentary, non-repeatable entities, tied to the cognitive agent in which as well as the time at which they occur, and because they are restrained to the particular object from which they borrow their content. The latter is clearly brought forward by the analysis of the object a jñāna is related to, for this object comprises not only generic entities, but also the particular entities that exemplify these generics with respect to a certain time and place. In other words, the analysis of the relevant object not only presents the structural content of a jñāna, but it also reveals the real particulars filling in the argument roles that come with the 'type of situation' classifying the jñāna. In this way, then, a jñāna can be conceived as tied up to reality, to the particular object that is cognized by someone, somewhere and somewhen.

To summarize, under the interpretation I propose the notion of a jñāna comprises four aspects: a spatio-temporal location, a cognizing agent, a structure or type of situation, and a set of particulars filling in the argument roles of the type of situation. Leaving aside the variables for the spatio-temporal location and the cognizing agent, a cognitive state is simply a pair $< S, f >$, where S is a type of situation, i.e. a construction out of a certain number of uniformities across situations, and f is a function from the argument roles of S to the real individuals obtaining in the actual world.

To conclude this section, I want to touch upon the question as to the 'validity' of a cognition frequently raised in Navyanyāya. It is said that to have a cognition is one thing, but to have a cognition that is *pramā*, i.e. 'correct', is something different, for not all cognitions are correct. An epistemic event is held to be correct if it is 'in accordance with its object', that is to say, if it is *yathārtha*. A remarkable aspect of this conception is that the question as to the correctness of some jñāna is answered by taking into account only the object that is taken to be related to the cognition by *viṣayatā*. The cognition '[This] is silver', for example, is said to be *pramā* with respect to silver, but *apramā* with respect to nacre. In other words, the correctness of a cognitive state $< S, f >$ can only be determined with respect to a particular setting f. This setting is not the whole world, but only a part thereof consisting of exactly those particulars singled out by f, particulars which exemplify certain generic entities. The question as to the correctness of a cognitive state, then, boils down to the question whether the particulars fixed by f instantiate the very same generics as are comprised by S. In a way this does remind of the lines in accordance with which the question whether a state of affairs holds or not is settled in situation semantics. For this theory takes the question whether a state of affairs holds or not primarily to be the question whether a state of affairs holds or not *in a given situation*. The question is decided on not by taking the whole world into account, but with respect to a part of the world. Consequently, different answers will be given, when different situations are taken into account.

The latter rather quick comparison is to be taken just at face value. That is to say, I do not contend that it does provide sufficient arguments on the basis of which it can be decided that Navyanyāya logic itself took recourse to partial objects. This, however, is mainly a philological question which, as far as my opinion is concerned, ought to be answered by spelling out in detail what such a partial object would look like, thereby resorting to the concepts maintained by the Navyanaiyāyikas only. In recent Western logic partial objects come in many colours and, consequently, an unequivocal conception of the notion of a partial object cannot be derived from this logic. In situation semantics, too, the notion of a situation is, to a certain extent, unclear and liable to modifications. This, however, is not to say that partiality itself is a clouded notion, for it is certainly possible to get a clear sight at the general features of a logic based on a partial semantics, no matter what exactly it is that is taken to be a partial object.

Since, for the time being, I am unable to provide a precise notion of the partial object as it, perhaps, has been resorted to by Navyanyāya logicians, I will consider the proposition that Navyanyāya logic is partial in the ontological sense of the word to be a working hypothesis only.

A fruitful one, I think, especially with regard to a reconstruction of Navyanyāya logic in terms of Western logical devices, but nevertheless a hypothesis. The question whether Navyanyāya logic is partial in the semantic sense of the word, however, is altogether a different one liable to decisive arguments as will come up in the next section.

3 Vyāpti

In Navyanyāya logic an inference is conceived of as a mental activity, a process falling apart in several phases, that can be analyzed in terms of a succession of certain jñāna's or mental states. An inference sets out on the cognition of the presence of, as I call it, an indicating property (hetu), 'indicator' for short, at a particular place at some particular moment (pakṣa), given that the conditions which surround the taking of an entity as a *pakṣa* are met. Finally, an inference results in the coming into being of an epistemic event representing the fact that the pakṣa also possesses the probandum or, literally, the 'thing to be inferred' (sādhya). Before the conclusion can be reached however, there are two other phases to be passed through, phases I will indicate as 'recognition of the relevant vyāpti-relation' and 'consideration', respectively. The several phases which constitute an inference are described by Annaṃbhaṭṭa as follows:

T1 A conclusion is an epistemic event to which cause is given by a consideration. A consideration is an epistemic event [having as its content that] the property of the pakṣa is qualified by vyāpti. An epistemic event the content of which is 'that mountain possesses smoke-pervaded-by-fire' is an example of a consideration. The conclusion proceeding from this [consideration] is an epistemic event [with the content]: '[That] mountain possesses fire'. Vyāpti is an invariable going-together [in this case] referred to by: 'Wherever there is smoke, there is fire'. The presence of a pervaded [object] on a mountain, etc., is a property of the pakṣa.[21]

[21]parāmarśa-janyaṃ jñānam anumitiḥ | vyāpti-viśiṣṭa-pakṣa-dharmatā-jñānaṃ parāmarśaḥ | yathā vahni-vyāpya-dhūmavān ayaṃ parvata iti jñānam parāmarśaḥ | taj-janyaṃ parvato vahnimān iti jñānam anumitiḥ | yatra yatra dhūmas tatra-agnir-iti sāhacarya-niyamo vyāptiḥ | vyāpyasya parvatādi-vṛttitvaṃ pakṣa-dharmatā || *Tarkasaṃgraha* 44.

The difference between the expressions ‘Wherever there is smoke, there is fire’ describing the second step, and ‘This mountain possesses smoke-pervaded-by-fire’ describing the third step, does remind of the distinction, made in situation semantics, between situation-meaning and type-meaning. Type-meaning is what allows a situation of a particular type to have meaning, i.e. situation-meaning. According to situation semantics, attunement to type-meaning is what permits an agent to infer soundly what a particular situation means, given that that situation is of the first situation-type involved in the type-meaning. Following this line of thoughts, then, an attunement to the type-meaning expressible by ‘Wherever there is smoke, there is fire’ is what allows an agent to decide that the cognition of a particular smoky place contains the information that there is fire. I will take up this distinction again in the next section and relate it to an alternative interpretation of the Navyanyāya scheme of inference. For now, I want to focus my attention to *vyāpti* in order to get some grip on the role this relation is assigned to in the Navyanyāya analysis of inference.

A first glance at vyāpti reveals this relation as a relation of ‘going invariably together’ or ‘invariable concomitance’ as it is usually rendered, leaving us with the notions ‘invariable’ (niyata) and ‘concomitance’ (sahacāra). An object goes together with another object if there is a locus that both have in common (sāmānādhikaraṇya). The notion ‘related to the same locus’, taken in isolation, is ambiguous.[22] It can be taken to mean that two entities have exactly the same loci, that all the loci of one entity are also loci of the other, or that two entities have at least one locus in common. The latter is what is usually meant by ‘sāmānādhikaraṇya’. Conceived of in this sense, the notion ‘related to the same locus’ expresses a symmetrical relation, one that holds between ‘smoke’ and ‘fire’, but also between ‘fire’ and ‘smoke’. Given two entities that have at least one locus in common, it is said that the first of these stands to the second either in a vyāpti-relation, or in the relation of deviation (vyabhicāra), the converse of a vyāpti-relation. The latter being the case whenever there is a locus possessing the first, but

[22] Compare Goekoop 1967, p. 6.
Another notion frequently used is ‘aikādhikaraṇya’. I am not sure whether this notion is ambiguous in the same way ‘sāmānādhikaraṇya’ is, or that it should taken to mean ‘related to (at least one) common locus’ only. This question though has little importance for the problems at hand.

not the second entity. Putting things together, then vyāpti might be determined as 'a going-together that is not deviation'. A definition of vyāpti having this purport is given by Viśvanātha:

T2 Vyāpti [obtains if] the indicator has a locus in common with the probandum, [and provided that this probandum] is not the counterpositive of an absence resident [in a locus] possessing the indicator.[23]

This definition consists of two conditions of which the first one states that there must be a 'going-together' of the indicator and the probandum. Indirectly, the second one poses a restriction on the locus of the indicator; it must not be a locus that possesses an absence having the probandum as its counterpositive (pratiyogi). In other words, it must not be a locus that possesses the indicator, but lacks the probandum. The second condition seeks to account for the invariable character of the concomitance. A discussion of the function and import of the first condition will be taken up later.

3.1 Grasping Vyāpti

An inference is conceived of by the Navyanaiyāyikas as one of the four available instruments to gain knowledge about what is, in fact, the case. Using some Western terminology this amounts to the demand that an inference not only must be valid, but, in addition, that it must be sound, too.[24] Consequently, the acceptability of the premises to be used in order to infer a conclusion, i.e. to give cause to an epistemic event reflecting the way (a part of) the world really is, has to be ascertained previously. From this it can be understood that in Navyanyāya literature there is not only attention for the definition of vyāpti, but

[23]hetuman-niṣṭha-viraha-apratiyoginā sādhyena hetor aikādhikaraṇyaṃ vyāptir || *Bhāṣā-Pariccheda* 69. Annaṃbhaṭṭa gives a similar definition in *Tarkadīpikā* 44: hetu-samānādhikaraṇa-atyantābhāva- apratiyogi-sādhya-sāmānādhikaraṇyaṃ vyāptiḥ ||

[24]In FPL a distinction is made between valid inferences and sound inferences. A sound inference is a valid inference the premises of which are all true. So, the set of sound inferences is a subset of the set of valid inferences. This distinction, as it seems to me, is not acceptable to a Navyanyāya logic, for, since it excludes unsound premises, it is quite unthinkable that it would allow for valid inferences having unsound premises.

also for the grounds on the basis of which the acceptance of a vyāpti-relation is justified.

At first the knowledge of a vyāpti-relation is led back by Annaṃbhaṭṭa to a repeated observation of a going-together as can be read from *Tarkasaṃgraha* 45. In the *Dīpikā*, however, he points out that this view is inadequate; even if the going-together of smoke and fire is observed a hundred times, then still it cannot be excluded that there is a locus having smoke but lacking fire. How is it possible, he addresses himself, to grasp the vyāpti-relation in which smoke stands to fire, while it is impossible to observe all cases of smoke and fire?[25]

This question takes a particular course if it is looked upon in connection with the Navyanyāya ontology of the whole according to which a whole is not the same entity as, and is therefore to be distinguished from, the sum of its parts. So, even if it were possible to observe all instances of smoke and fire, the conclusion that smoke goes together with fire *in general* would still be impossible, since smoke in general, the generic entity smoke-ness, is not the same as the sum of all the instances of smoke.

Realizing this characteristic of Navyanyāya ontology Annaṃbhaṭṭa concludes that correct knowledge of a vyāpti-relation is not obtained, nor can be obtained, by generalizing from particular instances. Instead it is stated that knowledge of vyāpti arises from a so-called 'extraordinary perception' (sāmānyalakṣaṇā pratyāsatti), a form of perception that gives rise to the perception of a certain generic entity, and that follows immediately after the normal perception (laukikapratyakṣa) of the entity exemplifying the generic entity at hand.[26] So, according to Annaṃbhaṭṭa vyāpti is a relation between *generic* entities. In addition to this he clearly states that knowledge of vyāpti, i.e. knowledge of the invariable going-together of two generic entities, is not the same as the knowledge of the going-together of all instances of these generic entities. The latter being restricted to those and only those that are omniscient.

Knowledge of a vyāpti-relation between two generic entities X and Y is what allows an agent, to make use of some terms borrowed from situation semantics, to infer soundly that a situation being of type X contains the information that there is a situation of type Y. One could

---

[25] *Tarkadīpikā* 45

[26] See *Tarkasaṃgraha* 42.

also say that knowledge of this kind in some way gives direction to what is to be expected. In other words, confronted with a situation of type X, one may expect that a situation exemplifying Y can be found there, too.

The notion underlying vyāpti is 'being related to the same locus'. Strictly speaking, I think, this notion offers the possibility to consider two entities appearing in the same locus one after another as two entities being related to the same locus. This possibility, however, is blocked by the Navyanyāya epistemology. To know two entities as being related to the same locus presupposes the presence at a certain moment of a complex object comprising both entities. For it is only then that the corresponding epistemic event and, consequently, the knowledge of these two entities as being related to the same locus, can be said to have its proper cause. The perception of two entities appearing one after another in the same locus, I think, would be analyzed by the Navyanaiyāyikas in terms of two successive epistemic events each having its own object. Following this line of thoughts the first perception would be analyzed as a jñāna that derives its content from a complex object containing an exemplification of X but lacking an instantiation of Y. But with respect to an assumed vyāpti-relation between X and Y, this is what is called a cognition of deviation of X and Y, and as such it rules out the very acceptability of the vyāpti-relation:

T3 Apprehension of vyāpti [obtains if] there is a cognition of a going-together of the indicator and the probandum and [provided that] a cognition of [their] deviation is lacking.[27]

So, it is to be concluded that knowledge of a vyāpti-relation, i.e. the generic knowledge of a systematic relation between two generic entities X and Y, implies that both entities are to be instantiated simultaneously.

Fragment T3 states the conditions the satisfaction of which justifies the knowledge of a vyāpti-relation. It displays some remarkable similarities with the definition of vyāpti as represented in T2. A difference to be noted though, is that T3, in contrast with T2, leaves ample room for an interpretation according to which the correctness of the knowl-

[27]vyāptes grāhakaṃ hetu-sādhya-sahacāra-darśanaṃ
vyabhicāra-darśana-abhāvas ca | *Maṇikaṇa*, Vyāptigrahopāya-prakaraṇa. Sarma 1960, pp. 34–35.

edge of a vyāpti-relation is *relative with respect to the total amount of knowledge available to a particular agent at a given moment.* Take, for example, someone who never saw fire anywhere but in a kitchen. On the basis of T3 this person would be justified in accepting that 'Wherever there is fire, there is smoke', since he possesses the knowledge of a locus that fire and smoke have in common and he does not know of a locus that possesses fire but lacks smoke. However, since the relation expressed by 'Wherever there is fire, there is smoke' is not accepted as a genuine vyāpti-relation it has to be settled that the absence of a cognition of deviation is a necessary indeed, yet not a sufficient condition for the correctness of the apprehension of vyāpti.

Not being a sufficient condition, a fulfillment of the requirement that there must not be a cognition of deviation allows one to conclude, strictly speaking, that the vyāpti- relation under consideration is 'not false'. Or, what amounts to the same, given that the requirement is met, there is no evidence on the basis of which it can be concluded that the vyāpti-relation is false. This, however, is not to say that the vyāpti-relation is 'true', for the amount of knowledge justifying the conclusion that a vyāpti-relation is 'not false' might grow into a knowledge-state on the basis of which it turns out to be 'false'. It is, of course, also possible that the knowledge grows into evidence on the basis of which the relation can be said to be 'true'. In the remainder of this chapter I will capture this aspect by saying that a vyāpti-relation that is 'not false' is *assertable.* The notion of 'being assertable' is opposed to 'being unassertable' and ought to be distinguished from both 'being true' and 'being false'.

What we have then, is, in fact, an intuitionistic conception of negation, according to which 'not not p' is taken to mean that there is no evidence for concluding that p is false, and, furthermore, that the lack of evidence for 'not p' is not the same as evidence for 'p'. I have resorted to this conception, for, as it turns out, the Navyanaiyāyikas were quite aware of the insufficiency of the criteria of non-deviation. That is to say, they took recourse to some auxiliary devices, notably *tarka* and *upādhi*, in order to obtain more certainty with regard to a vyāpti-relation that is 'not false'. Without going into details, tarka can be conceived of as a means by which it is pointed out that the rejection of a vyāpti-relation is in conflict with other assumptions already adopted. *Tarka*, then, can be looked upon as a device for extending the evidence on the basis of which a vyāpti-relation is 'not false' by taking some previously ascertained assumption into consideration, in such a way that the evidence develops into evidence on the basis of which the vyāpti-relation must be 'true'. The discovery of an *upādhi*, on the other hand, that is to say, the detection of an entity that is present whenever the probandum is, but that is not always present whenever the indicator is, does turn the vyāpti-relation into a false one, or, as I want to present it, the evidence grows into evidence on the grounds of which a stated vyāpti-relation can definitely said to be 'false'.

3.2 Contraposition

In Navyanyāya a distinction is made between 'positive-vyāpti' (anvaya-vyāpti) and 'negative-vyāpti'(vyatireka-vyāpti). As can be read from the Tarkasaṃgraha, this distinction relates to the kind of entities involved: 'A vyāpti-relation between an indicator and a probandum is positive-vyāpti. A vyāpti-relation between the absences of these is negative-vyāpti.'[28] Given the vyāpti-relation 'Wherever there is smoke, there is fire', Annaṃbhaṭṭa's formulation does not give us any clues as to the question whether the negative vyāpti-relation is to be taken as 'Wherever there is absence of smoke, there is absence of fire', or as 'Wherever there is absence of fire, there is absence of smoke', since negative entities are involved in both expressions. But, since negative vyāpti-relations have the same characteristic (anugama) as positive vyāpti- relations, we know that they also have to fulfill the criteria of non-deviation. Hence, we can conclude that only 'Wherever there is absence of fire, there is absence of smoke' expresses a genuine negative vyāpti-relation.

The distinction between positive and negative vyāpti-relations has given several authors cause to assume that the law of contraposition, at least partly, has been recognized in Navyanyāya logic.[29] Commonly, this claim is supported by resorting to the expressions 'Wherever there is smoke, there is fire' and 'Wherever there is absence of fire, there is absence of smoke'. These expressions are translated in the language of FPL thereby yielding something like '$\forall x(\mathrm{S}x \rightarrow \mathrm{F}x)$', and '$\forall x(\neg \mathrm{F}x \rightarrow \neg \mathrm{S}x)$', respectively. Next to this it is noticed that according to the

[28]hetu-sādhyayor vyāptir anvaya-vyāptiḥ tad-abhāvayor vyāptir vyatireka-vyāptiḥ | *Tarkadīpikā* 48. See also *Bhāṣa-Pariccheda* 142-143.

[29]Mullatti 1977, p. 77, Staal 1962, p. 640.

law of contraposition holding in FPL the expressions '$\forall x(\mathrm{S}x \rightarrow \mathrm{F}x)$' and '$\forall x(\neg \mathrm{F}x \rightarrow \neg \mathrm{S}x)$' are logically equivalent; i.e., they have the same truth-value under every interpretation of their letters. Hence, it is stated that, since the logical calculus of FPL applied to the expression '$\forall x(\mathrm{S}x \rightarrow \mathrm{F}x)$' yields the logically equivalent expression '$\forall x(\neg \mathrm{F}x \rightarrow \neg \mathrm{S}x)$', and since Navyanyāya logic accepts 'Wherever there is smoke, there is fire', as well as 'Wherever there is absence of fire, there is absence of smoke' as expressions of genuine vyāpti-relations, the same results are obtained both in FPL as in Navyanyāya logic. And this is taken to be evidential of the claim that, to a certain extent, the principles underlying FPL are respected in Navyanyāya logic, too.

It has not gone unnoticed, though, that already the claim that contraposition holds in Navyanyāya logic runs into obvious difficulties, for there are certain inferences in the context of which it is not allowed to make use of the contrapositive of a vyāpti-relation in order to derive the desired conclusion. These inferences are named *kevala-anvayī* and *kevala-vyatirekī*, two notions that I will render as 'only-positive' and 'only-negative', respectively. An only-negative inference is one the conclusion of which follows on the cognition that the absence of the indicator is the term of a vyāpti-relation having the absence of the probandum as its subject, although there is no cognition of a going-together of the indicator and the probandum [in a locus different from the pakṣa].[30] A stock example of an only-negative inference is, stated elliptically: 'Water has no smell because it differs from what is not different from earth.' In order to derive the conclusion that water has no smell, use is made of the vyāpti-relation 'Whatever differs from what is different [i.e., from earth], has no smell'. To this the Navyanaiyāyikas would not object, though they would disapprove of an inference that employs the contrapositive vyāpti-relation 'Whatever has smell, that does not differ from what is not different from earth'. The reason for this is that, since 'earth' is ontologically conceived of as the one and only substance that possesses 'smell' (gandhavatī), the going-together of 'smell' and the property of 'being different from what differs [from earth]' can only be illustrated by reference to the locus 'earth'. Hence, we run into the difficulty that in order to infer that earth possesses

[30]kevala-vyatireki tad-ucyate, yatra sādhya-sādhana-sāmānādhikaraṇyāgrahe 'pi sādhya-abhāvaṃ prati hetvābhāvasya vyāpakatāgrahādanumitiḥ | *Maṇikaṇa*, Sarma 1960, p. 42.

smell, it is impossible to give an example of a locus that differs from the pakṣa 'earth' and that, at the same time, possesses 'smell' as well as the property 'being different from what differs [from earth]'. The inference breaks down on the requirement that the example cited in the third step of the inference-for-another should involve a locus that is different from the pakṣa and turns into a circular reasoning.

The reason for excluding the contrapositive of a vyāpti-relation figuring in an only-positive inference is altogether a different one. An only-positive inference is described as one involving a probandum that is not the counterpositive of a (constant) absence.[31] That is to say, the absence of the probandum does not occur in the pakṣa, nor in any locus that is different from the pakṣa. An example of this kind of inference is given by: '[The] pot is nameable, because it is knowable, like a cloth.' To reach the conclusion that the pakṣa 'pot' is nameable recourse is taken to the vyāpti-relation 'Whatever is knowable, is nameable'. This relation is unproblematic, but its counterpositive is. The expression 'Whatever is not nameable, is not knowable' states a systematic relation between 'absence of nameability' and 'absence of knowability', both of which are conceived of as *aprasiddha*; i.e., 'hypothetical' or 'unlocatable', entities. It has been noticed that the Navyanaiyāyikas exclude *aprasiddha* terms from philosophical and logical discourse and that this feature of Navyanyāya logic gives cause to the exclusion of a vyāpti-relation like 'Whatever is not nameable, is not knowable'. This, of course, I do not dispute, though I do think that there is a little more to say about this matter, in particular when it is related to the conditions that are imposed on the grasping of a vyāpti-relation (vyāpti-graha). So, let me dwell on this subject a little longer.

The Presupposition of a Vyāpti-relation

To accept a vyāpti-relation, i.e. to know that it is not false, one has to know of a locus that the indicator and the probandum have in common, and one must not know of a locus that possesses the indicator but lacks the probandum (see T3). But with regard to an unlocatable entity the cognition of a locus possessing that entity together with any other entity whatsoever is impossible, since there is no locus that possesses that entity. Hence, knowledge of a vyāpti-relation between, say, 'absence of

[31] atyantābhāvāpratiyogisādhyakaṃ kevalānvayi | *Maṇikaṇa*, Sarma 1960, p. 40.

nameability' and 'absence of knowability' is totally unsupported. That is to say, there is nothing in reality that can be looked upon as the proper cause of the jñāna under consideration. The point I want to make is that, as I see it, the *primary* reason for excluding a relation like 'Whatever is not nameable, is not knowable' is not that 'absence of nameability' and 'absence of knowability' are *aprasiddha* entities, but that a locus possessing both entities is lacking. This notwithstanding, of course, that from the fact that 'absence of nameability' is unlocatable it obviously follows that there can be found no locus possessing that entity, let alone that there can be found a locus that it shares with another entity.

The reason I have for seeking the motivation for excluding certain vyāpti-relations primarily in the lack of a common locus, rather than in the occurrence of terms denoting unlocatable entities, is that the former has a slightly broader scope than the latter. Suppose, for example, that the second condition of T3 is fulfilled, but the first is not. In that case, there is no cognition of deviation (fulfillment second condition), i.e. there is no cognition of a locus possessing the indicator, but lacking the probandum, and, since the first condition is not satisfied, there is no cognition of a locus possessing the indicator as well as the probandum. In short, then, there is no cognition of the indicator *with* the probandum, nor is there a cognition of the indicator *without* the probandum. It seems to me that, consequently, there can be no cognition of the indicator at all, for the cognition of an entity presents that entity either with, or without some other entity. From this it follows that the indicator is *possibly* an unlocatable entity. Now, since unlocatable entities are excluded without exception, it seems reasonable to assume that Navyanyāya logicians would not allow for vyāpti-relations involving a *possible unlocatable* indicator either.

Suppose, to continue, that neither the first, nor the second condition is satisfied. Then there is no cognition of a going-together of the indicator and the probandum. But, although there is no cognition of deviation either, likewise this does not seem to suffice for accepting the alleged vyāpti-relation. For vyāpti is going-together without deviation, and the question whether a going-together is deviation or not seems rather pointless if there is no going-together at all. On the basis of these considerations, then, I propose to take the expression 'There is a locus that the indicator and the probandum have in common' to be the *presupposition* of a vyāpti-relation, such that a vyāpti- relation the presupposition of which is not satisfied is neither true, nor false, that is, *unassertable*, thereby ascribing to Navyanyāya logic a partial semantics. It turns out that the presupposition of a vyāpti-relation as I have stated it explicitly amounts to the first condition of the definition of vyāpti represented in T2. This also offers a possibility to relate the fragments T2 and T3 in a more intrinsic way. While the first criteria of definition T2 states the presupposition of a vyāpti-relation explicitly, the first criteria of T3 states the condition the satisfaction of which assures that the presupposition is fulfilled: given a *cognition* of a locus where the indicator goes together with the probandum, the presupposition 'There is a locus that the indicator and probandum have in common' is true. The latter, since it is a tenet of the Navyanyāya 'logic' of perception that to have a cognition normally 'means' that the object cognized obtains. That is to say, it is assumed that there are systematic relations between mental representations or, as I have called them, structural contents of jñāna's, and actual objects exemplifying these structural contents. Without going into details, it can be readily seen that precisely this feature of Navyanyāya epistemology provides well-motivated reasons for a semantic account of presuppositions and hence for a partial semantics.

Returning to the central issue of this section, the question whether contraposition is a law of Navyanyāya logic seems to have a negative answer, since contraposition is not allowed in the context of an only-negative or an only-positive inference. That is to say, the expressions '$\forall x(\mathrm{S}x \rightarrow \mathrm{F}x)$' and '$\forall x(\neg \mathrm{F}x \rightarrow \neg \mathrm{S}x)$' do not have the same truth-value under every interpretation of their predicate letters. Contrary to what one would expect, this aspect of Navyanyāya logic has not given cause to the rejection of the claim that the law of contraposition is a part of Navyanyāya logic, but, instead, to another misrepresentation concerning the conception of, what occasionally is called, the *restrictions* on the law of contraposition.[32] It is said that contraposition is a principle accepted in Navyanyāya logic, except under certain interpretations of *hetu* and *sādhya*; i.e., those interpretations that give rise to an only-positive, or an only-negative inference. However, most interpretations give cause to another type of inference, called *anvaya-vyatirekī* (positive-negative). With respect to this type of inference it is claimed that the law of contraposition is, indeed, accepted.

[32]See, for example, Mullatti 1977, p. 78-79.

A stock example of a positive-negative inference is: 'There is fire on the mountain, because there is smoke'. The relevant vyāpti-relation is: 'Wherever there is smoke, there is fire'. This relation as well as its contrapositive 'Wherever there is absence of fire, there is absence of smoke' are accepted by the Navyanyāya logicians. However, the presumption that within the context of positive-negative inferences the law of contraposition is valid brings with it the claim that the reasons I have for accepting a vyāpti-relation, say 'Wherever there is smoke, there is fire', are sufficient for accepting its contrapositive 'Wherever there is absence of fire, there is absence of smoke', too. But, given the presupposition of a vyāpti-relation as explicated above, this is not true, as can be readily seen from the following.

In case I know that 'Wherever there is x, there is y', I know that there is a locus where x goes together with y. The knowledge that 'Wherever there is absence of y, there is absence of x', on the other hand, presupposes that the knowledge of a locus possessing both 'absence of x' and absence of y' has been ascertained previously. Obviously, the fact that I know of a locus possessing both x and y, does not imply that I also know of a locus possessing both the absence of y and the absence of x. So, since I can conclude to 'Wherever there is x, there is y' in case I know of a locus possessing both x and y, and given the absence of a cognition of deviation, the very knowledge on the basis of which I can do so, does not allow me to conclude to 'Wherever there is absence of y, there is absence of x', for knowing that there is a locus that x and y have in common, and knowing that there is a locus that the absences of x and y have in common, is just not the same thing. It has to be concluded, then, that, since they have different presuppositions, the vyāpti-relations 'Wherever there is smoke, there is fire', and 'Wherever there is absence of fire, there is absence of smoke' have different truth-conditions. That is to say, I am not allowed to conclude that 'Wherever there is absence of fire, there is absence of smoke' is true as soon as I know that 'Wherever there is smoke, there is fire' is true. Consequently, the law of contraposition is not applicable, not even in the context of a positive-negative inference. But then, this is just what one would expect of a logic based on a partial semantics.

4 A Logic of Inference

4.1 Generic Instantiation

The distinction between a non-verbal type of inference, the 'inference-for-oneself' (svārtha), and a verbal type of inference, the 'inference-for-another' (parārtha) is typical of the Navyanyāya theory of inference having no analogue in the history of Western logic. A feature that can be taken as a support for the thesis that the Navyanyāya theory of inference is better approached starting from a general theory of meaning that takes us outside the realm of sentences and relations between sentences of any language, natural or formal. A general sketch of such a theory has been given in § 2.2 of this chapter. An interpretation of the Navyanyāya scheme based on this theory will be developed in the course of this section. For now, I want to focus my attention to the inference scheme as it is described by Aṇṇambhaṭṭa among others, and to the problems arising whenever this scheme is interpreted taking the Western classical framework as a starting point. These problems I subsume under the heading 'generic instantiation'.

Aṇṇambhaṭṭa tells us that the inference-for-another is to be used by anyone who wants to convince someone else of a conclusion already established by means of an inference-for-oneself. An inference-for-another consists of five steps, or verbal expressions, of which the following is a standard example:[33]

(1) [That] mountain possesses fire.
(2) Because it possesses smoke.
(3) Wherever there is smoke, there is fire, like [in] a kitchen.
(4) And that [mountain] is similar [to a kitchen].
(5) Therefore, it is so [as stated under (1)]

When this inference scheme is looked upon from the perspective of Western classical logic it is most nigh to reduce it to the inference: 'Wherever there is smoke, there is fire. There is smoke on the mountain. Therefore, there is fire on the mountain.' This inference exemplifies the scheme: $\forall x(\mathrm{S}x \rightarrow \mathrm{F}x)$, $\mathrm{Sa} \rightarrow \mathrm{Fa}$, Sa / Fa, a scheme that is valid on the basis of the rules of inference *universal instantiation* and *modus ponens*. For reasons of notational convenience I will not speak of universal instantiation and modus ponens, but I will make use of the

[33] *Tarkasaṃgraha* 45

notion 'generic instantiation', which I will take to mean the successive application of universal instantiation and modus ponens.

Within a classical framework, then, the inference as described by Aṇṇambhaṭṭa, is considered to be valid, because it can be reduced to an inference that exemplifies a scheme which is valid on the basis of generic instantiation. Although in this way the legitimacy of the Navyanyāya inference can be accounted for, it immediately becomes clear that this approach leaves no room whatsoever for a logical account for the alleged necessity of both the example, 'kitchen' in the case at hand, and the fourth step as a whole. So, what we have here is an inference that, from a classical point of view, is clearly valid, but that, due to the example and step 4, does not have the form of a non-redundant valid argument on traditional logical grounds. As far as classical logic concerns, then, both are to be looked upon as being outside the scope of logic.

My objection to this way of approaching Navyanyāya inferences is mainly of a methodological nature. It aims at the tacit assumption that Navyanyāya inferences are valid solely on the basis of the inference-rule of generic instantiation. Under this assumption one is overlooking the possibility that under certain circumstances the application of the rule of generic instantiation could be liable to limitations, or even is not allowed at all. With respect to this, it is significant to note that Western classical logic makes no use of a notion of subject matter, of what in particular an argument is about. As a consequence, in a classical analysis of inferences all that is specific of an argument is lost leaving nothing but underlying logical forms on the one hand, and general rules to operate on these forms on the other. This, as it seems to me, really is the heart of the matter, especially since it is unmistakable that subject matter, embodied in the notion of pakṣa, plays a very important role in Navyanyāya logic.

Since the evidence seems to be otherwise, I have no reason whatsoever to believe that from a Navyanyāya point of view the example as well as step 4 are logically superfluous. The task I appoint to a formal interpretation of Navyanyāya logic, then, is that it does justice to this outlook, that is to say, a formal interpretation of Navyanyāya logic should be able to justify logically both the example and step 4. In what follows I will indicate the starting points on the basis of which such an interpretation, perhaps, could be achieved.

4.2 Inference and Subject Matter

Validity depends on the semantic content of the entities which are taken to be the constituents of an inference. Generally speaking, this is characteristic not only of FPL, but also of situation semantics. In fact, it is a basic starting point of every logical theory. This being so, the question arises as to what features account for the differences between logical theories taking the same precept as their basic principle. Recalling what has been under discussion earlier (see §§ 1, 2 of this chapter), it will not come as a surprise that this question boils down to the one asking for a precise notion of what is held to be the semantic content of the entities making up an inference.

In FPL the semantic content of the (linguistic) entities constituting an inference is derived by resorting solely to the forms of the expressions involved. The language used is designed for stating these forms explicitly and provided with symbolic representations from which the semantic content of the expressions representated can be derived straightforwardly. In other words, the semantic content of an expression is contributed to by its (syntactical) form only, or, to bring the matter to a head, its semantic content, being totally independent of circumstances, is not indexical. Subsequently, the language of FPL consisting of devices for generating symbolic representations and rules for manipulating these can be called a *direct language*. A direct language, then, is one in which the validity of an inference of S from $S_1, \ldots, S_n$ can be determined by taking into consideration the representations of $S_1, \ldots, S_n$ and S only.

A radically different conception is put forward in situation semantics. The disparity proceeds from an assumption taken as very basic in situation semantics, an assumption that is quite incompatible with the nominalistic basis of FPL resorted to in Ideal Language Philosophy. It says that, in general, the semantic content of a representation depends not only on the features of the representation itself, but also on the circumstances in which it arises. This holds, it is said, for linguistic expressions as well as mental events, because both are conceived of as representations in their own right. Since the validity of an inference depends on the semantic content of the representations involved, and since the semantic content of a representation in turn depends on the circumstances in which it occurs, it follows that a valid inference also depends on the embedding circumstances. What we have, then, is a so-called *situated inference*; i.e., a conception of inference that comprises the agent's physical embedding in the world. As a consequence, the language used in situation semantics is not a directly, but a contextually interpreted language. That is to say, the expressions of this *semantically situated language* have meanings in contexts only. A semantically situated language carries with it a rather different notion of validity, because it is a language in which whether or not an inference of S from $S_1 \ldots S_n$ is valid depends on the relation between the contents $I(S_1, c_1), \ldots, I(S_n, c_n)$, and hence can depend on the embedding circumstances $c_1 \ldots c_n$ and c in which the representations occur.[34]

Taking recourse to a semantically situated language, it has to be acknowledged that the same representation, i.e., a construction out of generic entities in case of mental representations, can be a part of, or, what amounts to the same, can contribute to, different contents. Since the relation of entailment is a relation between semantic contents, it follows that a representation S' can be part of a content I' on the basis of which it is allowed to infer the conclusion S, while the very same representation occurring in different circumstances can contribute to another content I'' that does not allow at all for the inference of S. Hence, what is needed is some kind of mechanism for capturing the parameters which, just like the representation itself, contribute to the semantic content and, thus, impose restrictions on the information conveyed by the particular situation s' represented by S', information about something besides s'.

Now, let us suppose that the language used by the Navyanaiyāyikas is not a directly interpreted, but a semantically situated language and try to reconstruct the Navyanyāya scheme of inference in accordance with the requirements set forward by this language. The inference-for-another, as Annaṃbhaṭṭa describes it, sets in with a formulation of what exactly is to be inferred. The second step provides the reason on the basis of which this can be done. It is important to note that the statement by which the reason is presented has a very specific subject matter. It is not about there being smoke somewhere, but it is about a particular place possessing smoke at a particular time. That is to say, the statement of the reason is about the *pakṣa*. From the perspective of a semantically situated language the first two steps taken together can be reconstructed as stating that there is a particular situ-

[34] Compare Barwise 1989, p. 146.

ation comprising a particular property ('smoke'), which situation contains the information that there obtains a different situation which has another property, notably 'fire'.[35] When I, furthermore, assume that the Navyanyāya logicians did not presuppose that the one who had to be convinced did possess any of knowledge relevant with respect to the inference of the conclusion, then the remainder of the inference can be looked upon, in the first place, as providing the means to convince the other that this particular situation, indeed, does contain the information stated in the first step, and secondly, as furnishing a reason on the basis of which it can be decided that there are no decisive arguments to suppose that this particular situation does not allow for extracting that very information after all. In a semantically situated language the latter requirement immediately proceeds from the role assigned to the circumstances with respect to semantic contents and hence from the importance attached to subject matter. It is just one of the typical differences between a directly interpreted and a situated language.

The Example

A basic tenet of situation semantics is that systematic relations between types of situations are what allow one situation to contain information about another situation. In order to be able to recognize a situation s as meaningful, that is, as conveying information about another situation, an agent has to be attuned to a constraint, or type-meaning, comprising the type of the particular situation s and another type of situation. For it is claimed that only an attunement to a type-meaning allows an agent to conceive of a particular situation as having a situation-meaning. Applying this tenet to the Navyanyāya scheme of inference, it can be said that in order to convince someone that a particular situation (pakṣa) containing 'smoke' has the situation-meaning that there also obtains a situation having a particular fire as its constituent, one has to draw his attention to the type-meaning 'Wherever there is smoke, there is fire'. Note that under the interpretation I propose an expression of a vyāpti-relation is interpreted as an expression of a type-meaning. In keeping with this view it is, furthermore, possible to

[35]Note that in situation semantics the notion of a situation has little to do with the notion of (occupying) a spatio-temporal location. That is to say, the theory allows for different situations occupying the same spatio-temporal location. Compare assumption 4 mentioned in § 2.2 of this chapter.

translate 'vyāpti' as 'involvement', 'vyāpya' as '[what] involves', and 'vyāpaka' as '[what] is involved'.[36]

So, the statement of the type-meaning 'Wherever there is smoke, there is fire' provides the one to be convinced with the means to assign a situation-meaning to the pakṣa, the situation under consideration. But, as I have pointed out above, it is required that the vyāpti-relation resorted to is assertable. That is, its presupposition must be satisfied. The example can well be thought of to serve this purpose. By giving an example, i.e., a locus that the indicator and the probandum have in common and that is known as such by philosophers as well as laymen, evidence is provided with on the basis of which it can be said that there is a locus having both the indicator and the probandum, and, consequently, that the assertability of the vyāpti-relation at hand is assured, and hence its use allowed.

The Example and the Similarity

By expressing a vyāpti-relation or type-meaning the assertability of which is ascertained the first desideratum of a situated inference is met. Given an awareness of the type-meaning 'Wherever there is smoke, there is fire' the meaning of a particular situation containing smoke can be determined. But then we run into the second requirement according to which there must not be any reasons on the basis of which the attribution of this very information to the situation under consideration is blocked. Whether there are such reasons or not, is a question that can only be answered by taking recourse to the circumstances embedding this particular situation, the pakṣa, and, occasionally, to those embedding the process of inference taken as a (complex) situation in its own right. According to the interpretation under discussion, it is at this point that the example comes into play for a second time. This time, though, in connection with the similarity put forward in the fourth step.

The similarity pointed out in step four takes two terms the first of these, as I see it, being the pakṣa and the other provided with by the example. The question, then, is: What does the similarity amount to?

[36]So, 'smoke is pervaded by fire', can be written as 'smoke involves fire' and, the other way round, 'fire pervades smoke' can be rendered as 'fire is involved by smoke'. These translations are obviously not philologically motivated, but they follow from the perspective of interpretation I have chosen.

That is to say, in which respect is the pakṣa similar to the example? And, furthermore, it must be asked in which way the knowledge of this similarity contributes to the validity of the inference of the conclusion.

In Navyanyāya it is said that there are two kinds of examples called *sapakṣa* and *vipakṣa*, respectively. A *sapakṣa* example is an example of a locus with respect to which the presence of the probandum has been ascertained.[37] A *vipakṣa* example on the other hand is an example of a locus with respect to which the absence of the probandum has been ascertained.[38] As to the probandum 'fire' a lake is said to be a *vipakṣa* example, i.e., a locus where the probandum definitely does not occur, and a kitchen is viewed as a *sapakṣa* example, because it is a locus where fire does occur. So, the similarity between a mountain and a kitchen could be taken as amounting to the statement that a mountain, just like a kitchen, is a locus that does not possess a constant absence having fire as its counterpositive. That is to say, the mountain is a locus with respect to which the presence of fire is not excluded. That is, as I take it to be, the presence of fire on the mountain is *possible*.

The notion of being a locus where the presence of the probandum is possible, however, can be understood in two different ways. It can be taken to mean that it is not true that the presence of the probandum is impossible because of the physical nature of the pakṣa itself. This, at least, is a necessary condition a pakṣa has to meet in order to be a possible locus of the probandum. But is it also a sufficient condition? In other words, is a pakṣa being a possible locus of the probandum as far as its own physical nature concerns, always a locus where the presence of the probandum is possible?

Given the framework from which a semantically situated language proceeds, this question has to be answered negatively, for it has to be acknowledged that the circumstances can be such that the presence of the probandum is precluded, not by the physical nature of the pakṣa, but by the presence of other aspects constituting the circumstances in which the pakṣa is embedded. With respect to this it is significant to note that Udayana, for example, while discussing causality, reaches the conclusion that the coming into being of an effect is not only dependent on the presence of the relevant causes, but also on the absence of

[37]niścita-sādhyavān sapakṣaḥ || *Tarkasaṃgraha* 50.
[38]niścita-sādhya-abhāvavān vipakṣaḥ || *Tarkasaṃgraha* 51.

'blockades' (pratibandhaka) which prevent the effect from originating.[39] Taking this contention into account, what we have then, is that in order to be a locus where the probandum is possible a pakṣa must be such that (1) the presence of the probandum is not precluded by its own physical nature, and (2) it must not be a locus (which is a part of a situation) containing 'blockades' preventing the presence of the probandum.

The knowledge that with respect to a certain time the pakṣa meets the two requirements mentioned above surely is relevant to the validity of the inference of the conclusion, for the conclusion to be reached is not that the probandum obtains somewhere, but that the probandum is present at the pakṣa. Especially, since in Navyanyāya it is claimed that the ultimate criteria for establishing the validity of an inference is given by the action to which the conclusion leads. That is to say, a conclusion, taken as a (mental) representation of what is in fact the case, is valid if the action resulting from it is successful. This, again, is a reason for resorting to a semantically situated language in order to reconstruct Navyanyāya logic, for to claim that an inference is valid if the action it leads to is successful, cannot go without taking into consideration the circumstances in which the pakṣa is embedded and which, thus, contribute to the semantic content of the pakṣa, the situation under consideration. Since these circumstances are build up out of contingent aspects, it follows that the information conveyed by the pakṣa is also contingent and will vary under different circumstances. Hence, the same representation will contribute to different contents and, consequently, it will give rise to different inferences.

All this amounts to saying that the systematic relations between types of situations S and S' an attunement to which allows an agent, given a situation s of type S, to infer that there also obtains another situation of type S', are conditional on certain background conditions B, conditions that are, or are not, met by the environment. If the circumstances do not meet these conditions, then the systematic relation itself looses its significance with respect to a situation embedded in those circumstances. That is to say, the situation-meaning one would normally assign to that situation on the basis of an acquaintance with the relevant type-meaning cannot be obtained under those particular circumstances.

[39] See *Nyāyakusumāñjali* I, 10, and the suplemental commentary.

By this the starting points, as I believe, from which a symbolic reconstruction of Navyanyāya logic should proceed are given. They come down, to summarize, to the assumption that the Navyanyāya scheme of inference provides the means for recognizing the meaning of the situation under consideration, the pakṣa, by pointing out the relevant vyāpti-relation (a type-meaning: $S \Rightarrow S' \mid B$) as well as an example by which the assertability of this relation is assured. And, furthermore, to the view that the scheme also provides with a device, i.e. the similarity between the pakṣa and the example, by which it is indicated that there are no reasons for believing that the situation under consideration does not have the situation-meaning one would expect solely on the basis of an acquaintance with the type-meaning. Under my interpretation I take the latter to be the statement that the same background conditions B in which the example is thought of to be embedded also obtain in (the environment of) the pakṣa.[40]

The question as to how exactly a symbolic reconstruction of Navyanyāya logic proceeding from these starting points will look like is one I will not answer here. To this topic there are several other questions which will have to be taken up first and which will take us outside the scope of Navyanyāya logic and thereby outside the scope of this chapter.

[40]There are, of course, many parameters figuring in the background which can influence the validity of the application of a general rule in order to reach a conclusion. I have mentioned only two, notably the physical nature of the pakṣa, and the presence of blockades. The elaborate treatises on the definition of vyāpti, however, bring several other parameters in light as, for example, the relations that are in play.

References

Barwise, J.: 1989, *The Situation in Logic*, Stanford, Center for the study of language and information

Barwise, J. and Perry, J.: 1983, *Situations and Attitudes*, MIT Press, Cambridge Mass

Belnap, N. D.: 1967, Tonk, Plonk and Plink, in P. Strawson (ed.), *Philosophical Logic*, pp 132–137, Oxford University Press

Bencivenga, E.: 1987, *Kant's Copernican Revolution*, Oxford University Press, Oxford

Brouwer, L.: 1908, Over de Onbetrouwbaarheid der Logische Principes, *Tijdschrift voor Wijsbegeerte* **2**, 152–158

Carl, W.: 1994, *Frege's Theory of Sense and Reference*, Cambridge University Press, Cambridge

Cozzo, C.: 1994, Are Dummett's Requirements on a Theory of Meaning Sufficient for Rejecting Classical Logic?, *Erkenntnis* **40(2)**, 243–264

Currie, G.: 1994, On the road to Antirealism, *Inquiry* **36**, 465–83

Dummett, M.: 1975, What is a Theory of Meaning?, in S. Guttenplan (ed.), *Mind and Language*, pp 97–138, Clarendon Press, Oxford

Dummett, M.: 1976, What is a Theory of Meaning? (ii), in J. McDowell and G. Evans (eds.), *Essays in semantics*, pp 67–137, Clarendon Press, Oxford

Dummett, M.: 1977, *Elements of Intuitionism*, Oxford Logic Guides, Clarendon Press, Oxford

Dummett, M.: 1978a, The Justification of Deduction, in *Truth and other Enigmas*, pp 290–318, Duckworth

Dummett, M.: 1981, *The Interpretation of Frege's Philosophy*, Harvard University Press, Cambridge, Massachusetts

Dummett, M.: 1987a, Reply to Tennant, in B. M. Taylor (ed.), *Michael Dummett. Contributions to Philosophy*, pp 235–252, Martinus Nijhoff

Dummett, M.: 1987b, Reply to Wright, in B. M. Taylor (ed.), *Michael Dummett. Contributions to Philosophy*, pp 221–234, Martinus Nijjhoff Publishers, Dordrecht

Dummett, M.: 1991a, *Frege and Other Philosophers*, Clarendon Press, Oxford

Dummett, M.: 1991b, *The Logical Basis of Metaphysics*, Duckworth, London

Dummett, M.: 1991c, Objectivity and Reality: Lotze and Frege, in *Frege and other Philosophers*, pp 97–125, Oxford University Press, Oxford

Dummett, M.: 1992a, *Frege. Philosophy of Language*, Duckworth, London, Second Edition

Dummett, M.: 1992b, Frege's Philosophy, in *Truth and Other Enigmas*, pp 87–115, Duckworth

Dummett, M.: 1992c, *Truth and other Enigmas*, Duckworth, London, First impression 1978

Dummett, M.: 1992d, Wittgenstein's Philosophy of Mathematics, in *Truth and other Enigmas*, pp 166–185, Duckworth

Dummett, M.: 1992 [1978]b, Realism, in *Truth and other Enigmas*, pp 145–165, Duckworth

Dummett, M.: 1993, *The Origins of Analytical Philosophy*, Duckworth, London

Dummett, M.: 1994, Wittgenstein on Necessity: Some Reflections, in P. Clark and B. Hale (eds.), *Reading Putnam*, pp 49–65, Blackwell, Oxford

Fine, K.: 1975, Vagueness, Truth and Logic, *Synthese* **30**, 265–300

Frege, G.: 1884, *Die Grundlagen der Arithmetik. Eine logisch-mathematische Untersuchung über den Begriff der Zahl*, Verlag von W. Koebner, Breslau

Frege, G.: 1892, Über Sinn und Bedeutung, *Zeitschrift für Philosophie und philosophische Kritik* **100**, 25–50

Frege, G.: 1894, Rezension von: E. Husserl, Philosophie der Arithmetik, Erster Band, *Zeitschrift für Philosophie und philosophische Kritik* **103**, 313–332

Frege, G.: 1895, Le Nombre Entier, *Revue de Métaphysique et de Morale* **3**, 73–78

Frege, G.: 1918, Der Gedanke. Eine logische Untersuchung, *Beiträge zur Philosophie des deutschen Idealismus* **3**, 58–77

Frege, G.: 1969, *Nachgelassene Schriften*, Felix Meiner Verlag, Hamburg, Edited by Hans Herms, Friedrich Kambartel, Friedrich Kaulbach

Gentzen, G.: 1934, Untersuchungen über das logische Schliessen, *Mathematische Zeitschrift* **39**, 176–210

Goekoop, C.: 1967, *The Logic of Invariable Concomitance in the Tattvacintāmaṇi. Gaṅgeśa's* Anumitinirūpaṇa *and* Vyāptivāda *with Introduction, Translation and Commentary*, Reidel Publishing Company, Dordrecht

Haack, S.: 1978, *Philosophy of Logics*, Cambridge University Press, Cambridge

Kreisel, G.: 1965, Mathematical Logic, in Saaty (ed.), *Lectures on Modern Mathematics*, John Wiley

Landman, F.: 1986, *Towards a Theory of Information*, Ph.D. thesis, University of Amsterdam

Matilal, B. K.: 1968, *The Navya-nyāya Doctrine of Negation. The Semantics and Ontology of Negative Statements in Navya-nyāya Philosophy*, Vol. 46 of *Harvard Oriental Series*, Harvard University Press, Cambridge Massachusetts

Matilal, B. K.: 1986, *Perception. An Essay on Classical Indian Theories of Knowledge*, Clarendon Press, Oxford

McDowell, J.: 1981, Anti-realism and the Epistemology of Understanding, in J. Bouveresse and H. Parret (eds.), *Meaning and Understanding*, pp 225–248, De Gruyter, Berlin

Mohanty, J. N.: 1985, Psychologism in Indian Logical Theory, in B. Matilal and J. Shaw (eds.), *Analytical Philosophy in Comparative Perspective; Exploratory Essays in Current Theories and Classical Indian Theories of Meaning and Reference*, pp 203–211, Reidel Publishing Company, Dordrecht

Mullatti, L. C.: 1977, *The Navya-Nyāya Theory of Inference*, Vol. 30 of *Research Publications Series*, Extension Service & Publications Karnatak University, Dharwad

Notturno, M. A.: 1985, *Objectivity, Rationality and the Third Realm: Justification and The Grounds of Psychologism*, Vol. 16 of *Nijhoff International Philosophy Series*, Martinus Nijhoff, Dordrecht

Perry, J.: 1979, The Problem of the Essential Indexical, *Noûs* **13**, 3–21

Potter, K. H.: 1954, Are the Vaiśeṣika Guṇa's Qualities?, *Philosophy East and West* **4**, 259–264

Potter, K. H.: 1957, More on the Unrepeatability of Guṇa's, *Philosophy*

East and West **7**, 57–60.
Prawitz, D.: 1965, *Natural Deduction*, Almqvist & Wiksell, Stockholm
Prawitz, D.: 1977, Meaning and Proofs: on the Conflict between Classical and Intuitionistic Logic, *Theoria* **43**, 2–40
Putnam, H.: 1980a, The Meaning of 'Meaning', in *Mind, Language, and Reality*, pp 215–271, Cambridge University Press, Cambridge
Putnam, H.: 1980b, Models and Reality, *The Journal of Symbolic Logic* **45**, 464–482
Putnam, H.: 1984, Meaning and Reference, in S. P. Schwartz (ed.), *Naming, Necessity, and Natural Kinds*, pp 119–132, Cornell University Press, Ithaca, New York
Quine, W.: 1986, *Philosophy of Logic*, Harvard University Press, Cambridge, Massachusetts, second edition
Rasmussen, S. A. and Ravnkilde, J.: 1982, Realism and Logic, *Synthese* **52**, 379–437
Sarma, E. R. S.: 1960, *Maṇikaṇa: A Navya-Nyāya Manual. Edited and translated*, The Adyar Library and Research Centre, Adyar, Madras
Staal, J. F.: 1962, Contraposition in Indian Logic, in Nagel and Suppes (eds.), *Logic, Methodolody and Philosophy of Science*, pp 634–649. Also in: J. F. Staal *Universals*, The University of Chicago Press, Chicago 1988. pp. 93–08
Sundholm, G.: 1983, Constructions, Proofs and the Meaning of Logical Constants, *Journal of Philosophical Logic* **12**, 151–172
Sundholm, G.: 1993, Questions of Proof, *Manuscripta* 16
Sundholm, G.: 1994, Vestiges of Realism, in B. McGuiness and G. Olivieri (eds.), *The Philosophy of Michael Dummett*, pp 137–165, Kluwer Academic Publishers
Taylor, R.: 1952, Negative Things, *Journal of Philosophy* **49**, 433–449
Taylor, T. J.: 1992, *Mutual Misunderstanding. Scepticism and the Theorizing of Language and Interpretation*, Routledge, London
Tennant, N.: 1987, *Anti-realism and Logic*, The Clarendon Library of Logic and Philosophy, Clarendon press, Oxford
Thomason, R. H.: 1969, A Semantical Study of Constructible Falsity, *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* **15**, 247–257
Troelstra, A. and van Dalen, D.: 1988, *Constructivism in Mathematics 1*, Vol. 123 of *Studies in Logic and the Foundations of Mathematics*, North-Holland, Amsterdam
van Benthem, J.: 1986, Partiality and Nonmonotonicity in Classical

Logic, *Logique et Analyse* pp 225–247

Veltman, F.: 1981, Data Semantics, in Groenendijk, Janssen, and Stokhof (eds.), *Formal Methods in the Study of Language*, pp 541–565, Mathematical Centre Tract 136, Amsterdam

Veltman, F.: 1985, *Logics for Conditionals*, Ph.D. thesis, University of Amsterdam

Wittgenstein, L.: 1963, *Tractatus Logico-philosophicus*, Suhrkamp Verlag

Wright, C.: 1980, *Wittgenstein on the Foundations of Mathematics*, Harvard University Press, Cambridge, Massachusetts

Wright, C.: 1987, Dummett and Revisionism, in B. Taylor (ed.), *Michael Dummett. Contributions to Philosophy*, Martinus Nijhoff, Dordrecht

Wright, C.: 1993a, *Realism, Meaning & Truth*, Basil Blackwell, Oxford, second edition

Wright, C.: 1993b, Strict finitism, in *Realism, Meaning & Truth*, Blackwell, Oxford, second edition

Samenvatting

Dit proefschrift gaat *over* logica. Het onderzoekt drie verschillende manieren om logica te beschouwen. Drie verschillende manieren om dezelfde vraag te beantwoorden: Hoe kunnen logische principes gerechtvaardigd worden? Het beschouwt een *ontologische* invalshoek in het vierde hoofdstuk, een *transcendentale* invalshoek in het derde hoofdstuk, en een *semantische* invalshoek in de eerste twee hoofdstuken.

In de eerste twee hoofdstukken staat het principe van tweewaardigheid centraal. Volgens dit principe is elke bewering óf waar, óf onwaar. Dit is een semantisch principe, omdat het, zij het in zeer algemene zin, iets zegt over de betekenis van beweringen. De juistheid van dit principe wordt ter discussie gesteld door een filosofische stroming die bekend staat als het *anti-realisme*. Michael Dummett is de meest prominente pleitbezorger van deze stroming. Kenmerkend voor het anti-realisme is niet alleen de verwerping van het principe van tweewaardigheid, maar vooral de consequenties ten aanzien van de logica die daaraan verbonden worden. Het anti-realisme stelt dat, gegeven dat het principe van tweewaardigheid onhoudbaar is, ook het logische principe van het uitgesloten derde verworpen dient te worden. Volgens dit principe is, voor elke bewering p, de zin 'p of *niet*-p' universeel geldig. Dit principe vormt een essentieel element van de klassieke propositie-, en predikaatlogica. Verwerping ervan impliceert dan ook een verwerping van deze logica's.

In hoofdstuk I wordt gedetailleerd ingegaan op de overwegingen op basis waarvan een anti-realist in de voetsporen van Dummett het principe van tweewaardigheid als onjuist verwerpt. De leidraad wordt gevormd door de meta-vraag: 'Hoe kan de vraag "Hoe is het mogelijk algemeen geaccepteerde logische principes te critiseren?" beantwoord worden?' De traditionele opvatting is dat een rechtvaardiging van logische principes mogelijk is op basis van een voorafgaand metafysisch onderzoek. Dummett stelt dat het juist deze opvatting is die een adequaat antwoord op de vraag naar de rechtvaardiging van logische principes in de weg staat. Tegenover de traditionele opvatting stelt Dummett dat metafysica niet vooraf gaat aan logica, maar dat, omgekeerd, vragen met betrekking tot metafysisca beantwoord dienen te worden op basis van een voorafgaand betekenis-theoretisch onderzoek. Bezien vanuit dit perspectief is de fundamentele vraag : Wat is het correcte theoretische model van betekenis? Deze vraag wordt door Dummett niet op directe wijze beantwoord. Hij beperkt zich tot een explicatie van de voorwaarden waaraan een model van betekenis moet voldoen, wil het acceptabel zijn. In hoofdstuk I wordt nader ingegaan op de overwegingen die aan deze voorwaarden ten grondslag liggen, alsmede op de stelling dat, gegeven deze voorwaarden, een semantisch model dat het principe van tweewaardigheid onderschrijft niet geaccepteerd kan worden.

Hoofdstuk II knoopt aan bij een bekend tegenvoorbeeld tegen de anti-realistische stelling dat de verwerping van het principe van tweewaardigheid met zich meebrengt dat ook het principe van het uitgesloten derde onhoudbaar is. Dit hoofdstuk laat zien dat een anti-realist dit tegenvoorbeeld niet zonder meer hoeft te accepteren. Echter, het doel van dit hoofdstuk is niet zozeer de anti-realistische positie te verdedigen, alswel te laten zien dat het anti-realistische verweer zich uiteindelijk betrekt op de aanname dat bepaalde conceptuele onderscheidingen dermate belangrijk zijn dat ze ook in de taal uitgedrukt moeten kunnen worden. In hoofdstuk II gaat het om het onderscheid tussen directe en indirecte evidentie. Beargumenteerd wordt dat, gegeven de verwerping van het principe van tweewaardigheid, het principe van het uitgesloten derde slechts dan kan worden afgeleid als het onderscheid tussen directe en indirecte evidentie niet in de logische taal kan worden uitgedrukt.

Het anti-realistische argument tegen klassieke logica zoals dat stapsgewijs besproken wordt in hoofdstuk I en II is naar zijn aard een *negatief* argument. Dat wil zeggen, het probeert ons ervan te overtuigen dat klassieke logica *niet* te rechtvaardigen is. In hoofdstuk III wordt onderzocht of er ook een positief antwoord gegeven kan worden op de vraag: Wat is *de* correcte logica? Volgens Dummett is dit mogelijk middels een analyse van de notie van een bewijs. Deze analyse moet inzichtelijk maken wat het minste is dat we moeten kunnen *doen* om zoiets als een logisch bewijs te verkrijgen. Een belangrijk uitgangspunt is dat een bewijs beschouwd wordt als het resultaat van een bepaalde wijze van taalgebruik. De basisgedachte is dat onder de vele wijzen waarop we de logische voegwoorden kunnen gebruiken er enkele zijn zonder welke we niet tot een logisch bewijs zouden kunnen komen. Deze 'minimale' vormen van gebruik worden beschouwd als constitutief voor de mogelijkheid van een logisch bewijs. Het volgende criterium ligt dan voor de hand: de interpretatie van een logisch voegwoord, alsmede de logica die door deze interpretatie wordt geïnduceerd, is zonder meer gerechtvaardigd als deze interpretatie overeenstemt met een 'minimale' vorm van gebruik van het desbetreffende voegwoord. Deze benadering leunt zwaar op de veronderstelling dat het mogelijk is om uitsluitend in termen van gebruik op eenduidige wijze vast te leggen wat geldt als een logisch bewijs. In dit hoofdstuk, echter, wordt beargumenteerd dat deze veronderstelling onjuist is: wat wij als een logisch bewijs beschouwen is niet uitsluitend afhankelijk van de wijze waarop wij gebruik maken van de taal. En de conclusie die daaraan verbonden moet worden is dat het op de wijze die Dummett voorstaat niet mogelijk is te bepalen wat *de* correcte logica is, onafhankelijk van hoe de wereld is.

Ook al is het, op z'n zachtst gezegd, twijfelachtig of het mogelijk is een logica te rechtvaardigen onafhankelijk van hoe de wereld is, het is wel degelijk mogelijk de keuze voor een logica te rechtvaardigen door deze te begrijpen tegen de achtergrond van filosofische opvattingen betreffende de werkelijkheid. Dit wordt geïllustreerd in hoofdstuk IV aan de hand van de vraag naar de interpretatie van de logica ontwikkeld door Navyanyāya, een Indiase filosofische stroming met een rijke historie. Betoogd wordt dat wanneer deze logica beschouwd wordt in het licht van de ontologische uitgangspunten zoals die werden onderschreven door Navyanyāya, een interpretatie en beoordeling van Navyanyāya logica in termen van klassieke predikaat logica volstrekt inadequaat is.

Titles in the ILLC Dissertation Series:

ILLC DS-94-01: **Harold Schellinx**
The Noble Art of Linear Decorating

ILLC DS-94-02: **Jan Willem Cornelis Koorn**
Generating Uniform User-Interfaces for Interactive Programming Environments

ILLC DS-94-03: **Nicoline Johanna Drost**
Process Theory and Equation Solving

ILLC DS-94-04: **Jan Jaspars**
Calculi for Constructive Communication, a Study of the Dynamics of Partial States

ILLC DS-94-05: **Arie van Deursen**
Executable Language Definitions, Case Studies and Origin Tracking Techniques

ILLC DS-94-06: **Domenico Zambella**
Chapters on Bounded Arithmetic & on Provability Logic

ILLC DS-94-07: **V. Yu. Shavrukov**
Adventures in Diagonalizable Algebras

ILLC DS-94-08: **Makoto Kanazawa**
Learnable Classes of Categorial Grammars

ILLC DS-94-09: **Wan Fokkink**
Clocks, Trees and Stars in Process Theory

ILLC DS-94-10: **Zhisheng Huang**
Logics for Agents with Bounded Rationality

ILLC DS-95-01: **Jacob Brunekreef**
On Modular Algebraic Protocol Specification

ILLC DS-95-02: **Andreja Prijatelj**
Investigating Bounded Contraction

ILLC DS-95-03: **Maarten Marx**
Algebraic Relativization and Arrow Logic

ILLC DS-95-04: **Dejuan Wang**
Study on the Formal Semantics of Pictures

ILLC DS-95-05: **Frank Tip**
Generation of Program Analysis Tools

ILLC DS-95-06: **Jos van Wamel**
Verification Techniques for Elementary Data Types and Retransmission Protocols

ILLC DS-95-07: **Sandro Etalle**
Transformation and Analysis of (Constraint) Logic Programs

ILLC DS-95-08: **Natasha Kurtonina**
Frames and Labels. A Modal Analysis of Categorial Inference
ILLC DS-95-09: **G.J. Veltink**
Tools for PSF
ILLC DS-95-10: **Giovanna Cepparello**
Studies in Dynamic Logic
ILLC DS-95-11: **W.P.M. Meyer Viol**
Instantial Logic. An Investigation into Reasoning with Instances
ILLC DS-95-12: **Szabolcs Mikulás**
Taming Logics
ILLC DS-95-13: **Marianne Kalsbeek**
Meta-Logics for Logic Programming
ILLC DS-95-14: **Rens Bod**
Enriching Linguistics with Statistics: Performance Models of Natural Language
ILLC DS-95-15: **Marten Trautwein**
Computational Pitfalls in Tractable Grammatical Formalisms
ILLC DS-95-16: **Sophie Fischer**
The Solution Sets of Local Search Problems
ILLC DS-95-17: **Michiel Leezenberg**
Contexts of Metaphor
ILLC DS-95-18: **Willem Groeneveld**
Logical Investigations into Dynamic Semantics
ILLC DS-95-19: **Erik Aarts**
Investigations in Logic, Language and Computation
ILLC DS-95-20: **Natasha Alechina**
Modal Quantifiers
ILLC DS-96-01: **Lex Hendriks**
Computations in Propositional Logic
ILLC DS-96-02: **Angelo Montanari**
Metric and Layered Temporal Logic for Time Granularity
ILLC DS-96-03: **Martin H. van den Berg**
Some Aspects of the Internal Structure of Discourse: the Dynamics of Nominal Anaphora
ILLC DS-96-04: **Jeroen Bruggeman**
Formalizing Organizational Ecology
ILLC DS-97-01: **Ronald Cramer**
Modular Design of Secure yet Practical Cryptographic Protocols
ILLC DS-97-02: **Nataša Rakić**

Common Sense Time and Special Relativity
ILLC DS-97-03: **Arthur Nieuwendijk**
On Logic. Inquiries into the Justification of Deduction