

PER MARTIN-LÖF

TRUTH OF A PROPOSITION, EVIDENCE
OF A JUDGEMENT, VALIDITY OF A PROOF*

One way of giving meaning to the various syntactical entities of a formal language is by modelling it in the way with which we are all familiar: the typical case would of course be the standard modelling of first order predicate logic. How does it proceed? You have a symbol, usually not made explicit, for the type of individuals, and you assign to it a set which is referred to as the individual domain. Similarly, to each individual term you assign an individual, that is, an element of the individual domain, and to each formula you assign a proposition. Finally, you prove that, if a formula is formally derivable, then the proposition which is assigned to it comes out true: intuitionistically interpreted, this means that you assign to each formal derivation a proof of the proposition which is assigned to its end formula. This is a pattern which is followed in all kinds of modelling, most recently in the denotational semantics of programming languages: we assign to the syntactical entities that we are dealing with certain mathematical objects and speak of those objects as the interpretations of the syntactical entities. In model theory, you look upon the entities that are to be interpreted syntactically, that is, taking the linguistic attitude, whereas you look upon their interpretations in the object oriented way in which you ordinarily deal with mathematical objects, that is, you forget about language and handle the objects directly in the way you are used to as a mathematician.

But a moment's reflection is enough to show that you are not at all dealing with these objects in a language free way. How could you? You are after all assigning a mathematical object to the syntactical entity by giving an expression for that object: you always use an expression, a linguistic expression, in order to express the object which is to serve as the interpretation of the syntactical entity. So you may look at the modelling in a different way, the way in which a proof theorist rather looks at a model: namely, think not only of what is to be interpreted as linguistic expressions but also of the interpretations which are assigned to them as linguistic expressions, expressing objects, of course, but look at them linguistically, and then what

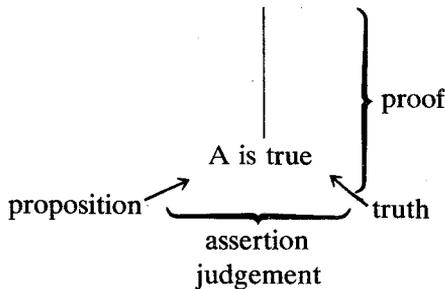
appears to the model theorist as a modelling appears, taking this other attitude, simply as a translation, a translation into another language then, because a translation of course always has to be a translation into another language. So modelling and translation into another language are one and the same thing, depending on whether you take the object oriented attitude towards the interpretation or whether you look at it linguistically. Whatever way you choose to look at it, as modelling or as translation, this is certainly one way of giving meaning to the linguistic expressions of a formal language. But, once you look at it as a translation into another language, it is equally clear that you cannot always give meaning in this way, because you can give meaning in this way only if you have another language to translate into.

Of course, this is the appropriate way to proceed in many situations. If you are dealing with a geometrical axiom system which refutes the parallel axiom, for instance, you make sense of it by modelling it by means of standard geometrical concepts, which may be thought of as a translation into the language of ordinary geometry. Or, if you want to establish the relative consistency of Church's thesis with some intuitionistic axiom system in which it can be expressed, you use the realizability interpretation, which is to say that you model it by means of standard arithmetical notions or translate it into the language of ordinary arithmetic: these are just two different ways of viewing one and the same procedure. But, eventually, you will of course have to have a language which is not given meaning by translating it into another language but has to be given meaning in some other way, and this is the language of the most primitive notions that you are dealing with, because that they are primitive means precisely that they cannot be defined in terms of any other notions.

How are you to proceed then to give meaning to the most primitive notions that you are dealing with? I think the answer is that you must enter on something completely different from modelling or translation, depending on whether you look at it model theoretically or proof theoretically: you must enter on a genuinely semantical or meaning theoretical investigation, which means that you must enter on something that you are not at all prepared for as a mathematical logician, whether model theorist or proof theorist: you must enter on an enterprise which is essentially philosophical or phenomenological, if you prefer, in nature. And, since it is this which is our concern at this workshop on theories of meaning, I think it would be appropriate

to give, or at least outline, *one* example of a theory of meaning, namely, a theory of meaning for the standard language of predicate logic, because if we cannot deal even with that exceedingly simple language it is very unlikely that we should be able to give any substantial theories of meaning for more complicated languages, like fragments of natural language. So I shall outline one particular theory of meaning, intuitionistic or verificationistic theory of meaning, for the language of predicate logic. Of course, pure predicate logic is not sufficient for all of mathematics: in addition to the logical operations you need ordinary inductive definitions, possibly also generalized or transfinite inductive definitions, but the pattern of this kind of meaning theory can certainly be seen already from the explanations for the pure predicate logic.

The fundamental concepts that have to be explained semantically can be read off either from the title of my talk or else from the schematic figure



First of all, we have the notion of proposition. Second, we have the notion of truth of a proposition. Third, combining these two, we arrive at the notion of assertion or judgement. There are various forms that a judgement may exhibit, in general, but I am only going to consider judgements of the particular form which is used for holding a proposition to be true. Fourth, in addition to the notion of judgement, we have the notion of evidence or proof of a judgement, which I have indicated schematically in the figure by means of the vertical line. Fifth, in the very end, I shall also have to consider the notion of correctness or validity of a proof: that is the last notion that enters into the title of my talk. So the semantical or meaning theoretical explanations that have to be supplied in the case of the language of predicate logic are explanations of the notions of proposition, of truth, of

judgement, of proof, and eventually something has to be said about validity of proofs also.

Let me begin with the notion of proposition, and I have to begin exactly with the notion of proposition, because that is the notion which needs to be explained first. I would like to point out that the explanation of a proposition as the expression of its truth conditions, which is the explanation of the notion of proposition that was given by Wittgenstein in the *Tractatus*, who in turn took it as being implicit in Frege's writings, that that explanation, which we ordinarily associate with classical logic, is in fact just as good for intuitionistic logic and in complete agreement with the explanations of the notion of proposition that have been given by the intuitionists, notably by Heyting and Kolmogorov. How is this possible? The intuitionists explain the notion of proposition, not by saying that a proposition is the expression of its truth conditions, but rather by saying, in Heyting's words, that a proposition expresses an expectation or an intention, and you may ask, An expectation or an intention of what? The answer is that it is an expectation or an intention of a proof of that proposition. And Kolmogorov phrased essentially the same explanation by saying that a proposition expresses a problem or task (Ger. *Aufgabe*). Soon afterwards, there appeared yet another explanation, namely, the one given by Gentzen, who suggested that the introduction rules for the logical constants ought to be considered as so to say the definitions of the constants in question, that is, as what gives the constants in question their meaning. What I would like to make clear is that these four seemingly different explanations actually all amount to the same, that is, they are not only compatible with each other but they are just different ways of phrasing one and the same explanation. Let us first look at the truth conditions for the usual logical operations. As we all know, they read:

the proposition	provided that
\perp is true	—
$A \vee B$ is true	A is true or B is true
$A \& B$ is true	A is true and B is true
$A \supset B$ is true	B is true provided that A is true
$(\exists x)A(x)$ is true	$A(a)$ is true for a specific individual a
$(\forall x)A(x)$ is true	$A(x)$ is true for an arbitrary individual x

The first line of this table, which gives the truth conditions for absurdity, should be interpreted as stipulating that absurdity is true on no condition. Expressed in this way, the truth conditions make up nothing but a formal scheme: there is a formal pattern here that we are faced with, and, as with all formal patterns, the question arises, How is it to be interpreted? I mean, How are these truth conditions to be interpreted? One way of interpreting them is certainly the Tarskian way which we all know from the thirties. What is defined then, by Tarski's truth conditions, is what it means for a formula in the language of first order predicate logic, say, to be true in the Tarskian sense. Thus it is the Tarskian notion of truth which is defined by the truth conditions when they are interpreted in his way. However, there is another way of interpreting these truth conditions, and it is only when you interpret them in this other way that they serve to determine the meanings of the logical constants and are in complete agreement with the explanations given by Heyting, Kolmogorov and Gentzen. This is most simply seen by just turning the table of truth conditions displayed above counterclockwise by one right angle. What happens? The vertical line is turned into a horizontal line, what stands in the left column is placed below the horizontal line, and what stands in the right column, the conditions under which the proposition displayed on the same line in the left column is true, above it:

$\frac{A \text{ is true}}{A \vee B \text{ is true}}$	$\frac{B \text{ is true}}{A \vee B \text{ is true}}$	$(A \text{ is true})$
$\frac{A \text{ is true} \quad B \text{ is true}}{A \ \& \ B \text{ is true}}$	$\frac{B \text{ is true}}{A \supset B \text{ is true}}$	
$\frac{A(a) \text{ is true}}{(\exists x)A(x) \text{ is true}}$	$\frac{A(x) \text{ is true}}{(\forall x)A(x) \text{ is true}}$	

If you interpret the truth conditions in this way, you see that they are identical with the introduction rules for the logical constants as formulated by Gentzen. So I have now explained why, suitably interpreted, the explanation of a proposition as the expression of its truth conditions is no different from Gentzen's explanation to the effect that the meaning of a proposition is determined by its introduction rules.

How about the two other explanations, the Heyting explanation and the Kolmogorov explanation, which they both agreed, already in the

thirties, by the way, were the same? Thus it is just a difference of wording whether you prefer to speak with Heyting about an expectation or an intention or with Kolmogorov about a problem or task. What was Heyting's way of explaining the meanings of the logical constants? He would say, or did say, that a disjunction $A \vee B$ expresses the expectation or intention of a proof of A or a proof of B , that a conjunction $A \& B$ expresses the intention of a proof of A and a proof of B , that an implication $A \supset B$ expresses the intention of a hypothetical proof of B from the hypothesis or assumption A , that an existentially quantified proposition $(\exists x)A(x)$ expresses the intention of an individual a and a proof of $A(a)$, and that a universally quantified proposition $(\forall x)A(x)$ expresses the intention of a free variable proof of $A(x)$. In the Kolmogorov interpretation, you just change the word 'expectation' or 'intention' into 'problem' or 'task', and, instead of speaking of fulfilling an expectation or intention, you speak of solving a problem or task. So, in Kolmogorov's wording of this one and single interpretation, you would say that a disjunction $A \vee B$ expresses the problem of solving either A or B , that a conjunction $A \& B$ expresses the problem of solving both A and B , that an implication $A \supset B$ expresses the problem of solving B provided that A can be solved, that an existential proposition $(\exists x)A(x)$ expresses the problem of finding an individual a and solving the problem $A(a)$, and that a universal proposition $(\forall x)A(x)$ expresses the problem of solving $A(x)$ for an arbitrary individual x . Thus you see that, interpreted in this way, Wittgenstein's dictum that a proposition is the expression of its truth conditions is in complete accordance with the Heyting, Kolmogorov and Gentzen explanations, and this is, I would say, the official intuitionistic explanation of the notion of proposition.

Everything that I have said so far was said, or could have been said, already in the thirties. What made these explanations as given in the thirties lack precision somewhat, was the absence of a clear distinction between what Dummett has called a canonical proof and a demonstration of a proposition, and which I shall call a direct as opposed to an indirect proof of a proposition. In explaining the meanings of the logical constants, I did not care whether a proof of a proposition was direct or indirect: nevertheless, we have to make that distinction, because what is explained in the meaning explanations of the logical constants, as I just gave them, is what constitutes a direct proof of a proposition formed by means of one of those constants. Thus a

proposition is defined by what counts as a direct proof of it. In the case of disjunction, for instance, you say that a direct proof of $A \vee B$ consists of a proof of A or a proof of B . But there are of course also indirect proofs of a proposition: for instance, every elimination rule in the sense of Gentzen gives you an indirect proof of its conclusion. If, to begin with, you ask, What is a direct proof of a proposition? the answer is that that depends on the proposition, because it is exactly that feature of a proposition which determines it as such, that is, which gives it its meaning as a proposition. So there is no uniform answer to this question: rather, a separate answer has to be given for each of the logical operations, and that answer is what determines the meaning of the logical operation in question. As I have already said, you explain the meaning of a logical operation by laying down what counts as a direct proof of a proposition formed by means of that logical operation. On the other hand, if you ask, What is a possibly indirect proof of a proposition? the answer is that an indirect proof of a proposition is a method of proving it directly, that is, a method which yields a direct proof of the proposition as result. Thus to know an indirect proof of a proposition is to know how to give a direct proof of it. Here I am implicitly using the identification of knowledge of a way or method of doing something and knowledge how to do it. As you have noticed, I have used the word method in explaining the notion of indirect proof, but I can certainly do without it, if you prefer, and always express myself in terms of knowing how and knowledge how.

The notion of truth of a proposition is explained intuitionistically by explaining what it is to know that a proposition is true: to know that a proposition is true is to know a proof of it, a proof which may in general be indirect. Since to know a possibly indirect proof of a proposition is to know how to give a direct proof of it, to know that a proposition is true is to know how to prove it directly. On the other hand, according to the principle that to know how to do something is the same as to know that it is doable, to know how to prove a proposition directly is the same as to know that it is directly provable. Hence, by transitivity, to know that a proposition is true is the same as to know that it is directly provable. Cancelling know in both members of this identity, we arrive at the conclusion that to be true is the same as to be directly provable, and that truth is identified with direct provability. What is characteristic of this whole analysis, intuitionistic or verificationistic analysis, of the notions of proposition and truth is that

the notion of proof of a proposition is conceptually prior to the notion of truth. And it is this which makes intuitionism into an idealistic philosophy in the knowledge theoretical sense: by saying that a proposition is true, you express that you know that it is true, and there can be no question of a proposition's being true except as the result of someone's knowing it to be true. In this precise sense, the notion of truth is knowledge dependent. Of course, when you look at the combination of words, know that a proposition *A* is true, when you look at it as it stands, it seems that, first of all, you have the proposition *A*, then you have the notion of truth which you apply to it, arriving at the assertion or judgement, *A* is true, and, finally, you apply the notion of knowledge on top of that. That is how it seems, of course, if you just look at the syntactical building of it, but the outcome of the intuitionistic analysis of the notion of truth is that the order of priority is the reverse here: it is the concept of knowing that a proposition is true, that is, of knowing a proof of the proposition, which is the conceptually prior notion, and then the notion of truth is extracted from it by saying that a proposition is true if it is directly provable, that is, if it can be proved by the most direct means. Moreover, the truth conditions for the logical constants, which have the same wording as you are used to, are interpreted in such a way that they appear as direct proof conditions.

This philosophical position, which is no doubt the intuitionistic position, that the notion of proof is conceptually prior to the notion of truth, is not new with intuitionism, certainly. You find it explicitly stated in Husserl's writings, not only after the turn into transcendental idealism, first expounded in *Ideas* and then in all his later publications, but already in the *Logical Investigations* and as late as in the *Formal and Transcendental Logic*, that the notion of truth is transposed into the idea of possibility of evident judgement, that is, something is true if it is possible to judge it with evidence. This is not surprising, of course, because what were the problems that Husserl was dealing with and which led him to develop phenomenology? They were precisely the foundational problems of logic: that is why the *Logical Investigations* were termed logical, although later on they were widened into a more encompassing, not to say all encompassing, philosophy. If you look at this position, namely, that the notion of proof or evidence is prior to the notion of truth, not in the context of logic, but instead in the context of natural experience, that is,

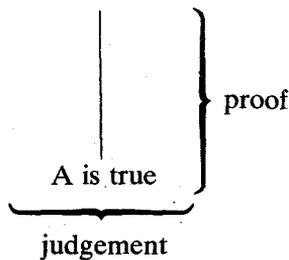
the experience of nature or the natural world, then it is of course the Kantian position, because Kant simply identified nature or world with possible experience: or, in a more careful formulation, nature or the natural world is the totality of everything that can be the object of our experience, that is, it is the totality of all possible experience. And, of course, it is very appropriate to mention Kant in this connection, because we know that Kant was the principal philosophical source of Brouwer, although there is much more to be found in Husserl in the case of logic, but, as far as we know, there was no direct influence from Husserl on Brouwer, although we do know that Heyting, when giving the intuitionistic interpretations of the basic logical notions, was influenced by Husserl via Oskar Becker, apparently.

Why did Husserl speak of transposition of the order of conceptual priority between the notion of truth and the notion of evident judgement? That was because of the line of thought to which he belonged and from which he had started, namely, the tradition from Bolzano and Brentano. And Bolzano is the most clear exponent of the opposite view, namely, that it is the notion of truth which is the prior notion which has to be explained first and therefore has to come first in a systematic exposition of logic, whereas the notion of knowledge of truth, that is, of evident judgement, is posterior to it. That is how Bolzano's *Wissenschaftslehre* is built up: it starts with a treatment of propositions and truths in themselves and only much later on does it treat of knowledge of propositions and knowledge of truths. Similarly, Brentano with his strong influence from Bolzano started from the same order of conceptual priority. And not only did Bolzano influence Brentano: rather, both of them were influenced by the Aristotelian and high scholastic tradition, and it was through that influence that they came to take this particular order of conceptual priority between the notion of truth and the notion of proof or evidence.

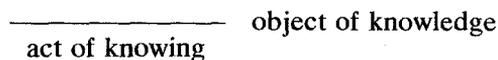
Where do you find the change? Of course, before Bolzano you have Kant, who had already performed this revolution, but in this line of thought the change is midway in Brentano. His point of departure was the Aristotelian and high scholastic tradition that I just mentioned. This had to do with the fact that Bolzano and Brentano were both Catholics. So that was their common point of departure, and they tried to resist Kantian idealism. In ever so many places in their writings there is a point directed towards Kant, but in spite of all this Brentano

was eventually forced by necessity in his analysis of the notions of truth and evidence into the position that it is the notion of evidence which is the prior notion, that is, he was eventually forced into an evidence theory of truth. Although his final position was not the same as the one that I am in the process of expounding, it was certainly the same in the sense that the notion of evidence was taken to be prior to the notion of truth. And then, through Brentano's influence, this was wholeheartedly accepted by Husserl, who, precisely by taking this step, was able to overcome the split between the Kantian branch of German philosophy and the objectivistically oriented branch of German philosophy of which Bolzano and Brentano were the chief exponents.

So much about the notions of proposition and truth: now to the notions of judgement and proof. Again, look at the schematic picture



and ask yourself this time, What is a judgement? and, What is a proof of a judgement? When you are dealing with such basic notions, it is clear that there can be no question of reducing them to any other, more basic notions: rather, you have to satisfy yourself that these notions are the same as certain other notions that you use other words to express, and there can only be the question of revealing the structure into which these notions fit and finding the words for them which make most clear their nature. So look at the picture, apply to it the trick of turning it counterclockwise by one right angle, and compare it with the picture



which is so familiar to everyone who has thought in terms of acts and objects. The word act, of course, is the most general word that we use for all of our acts, and, similarly, the word object has that complete generality. Thus we have acts of conjecturing and doubting and wishing and fearing and so on, and we have the objects of those acts, that is, objects of conjecture and doubt and wish and fear and so on.

One particular kind of acts, and, correlatively, kind of objects that we direct ourselves towards in those acts, are acts of knowing and objects of knowledge, respectively. My answer to the questions, What is a judgement? and, What is a proof of a judgement? is simply that a proof of a judgement is an act of knowing and that the judgement which it proves is the object of that act of knowing, that is, an object of knowledge. This fits completely with the standard explanation of a proof that you get out of an ordinary dictionary: it says that a proof is what shows the truth of a statement. Now, I have been careful to use the word truth in connection with propositions, whereas I try to use evidence in connection with assertions or judgements. With that terminological precaution, the proper explanation of the notion of proof of a judgement is that a proof is that which makes an assertion or judgement evident, or, if you prefer, simply that a proof of a judgement is the evidence for it. Thus proof and evidence are the same. And what is it that makes a judgement evident to you? Before you have understood or grasped the judgement, it is not evident to you, and, when you have grasped it, it is obvious or evident to you. Thus it is simply your act of understanding or grasping it which confers evidence on the judgement, that is, which makes it evident to you. This is one way of seeing that the proof of a judgement is nothing but the act of knowing, or, perhaps better, the act of understanding or grasping, and that what you grasp, namely, the object of knowledge, is the same as what you prove, namely, the assertion or judgement.

How can you argue for such an interpretation of the primitive notions of judgement and evidence? I do not think that you can really argue very much about it: you can have different paths to this insight, and I have taken two, I think, so far. There is a third one, starting from the dictum, which again Husserl stuck to in all of his work, that evidence is nothing but experience of the truth (Ger. *Evidenz ist Erlebnis der Wahrheit*). And what is the experience of the truth of a proposition? That is precisely the act of understanding or grasping its truth, that is, the act of getting to know its truth, and hence we arrive

at the same conclusion via this third path, namely, that the evidence for a judgement is the very act of knowing it.

This would be all there is to it, concerning these fundamental notions of proposition, truth, judgement and proof or evidence, if it were not for the regrettable fact that we make mistakes: we constantly make mistakes, not only in ordinary life, but we make mistakes in proofs as mathematicians although mathematics is allegedly the most rigorous of all sciences. And it is because of the fact that we make mistakes that the notion of validity of a proof is necessary: if proofs were always right, then of course the very notion of rightness or rectitude would not be needed. The mathematician's activity would simply proceed by his propounding his proposition or theorem, giving its proof, and then another theorem, its proof, and so on, and you would only say proof all the time: there would be no need to add some qualifying adjective and speak of valid proof or correct proof or conclusive proof. The notion of correctness or validity, so to say, does not belong to the mathematical activity when things go well, that is, when you do not make mistakes, the notion does not arise: it arises only when you make mistakes, because then there begins a discussion whether the proof, and we then call it alleged proof or seeming proof, is really a proof or if it is wrong, that is, if it is a deceptive proof or an illusory proof. Now, there can be no proof, no conclusive proof, that a proof really proves its conclusion, because, if such a miraculous means were available to us, we would of course always use it and be guaranteed against mistakes. It would be a wonderful situation, but we are human beings and we are not in that position. As human mathematicians we constantly make mistakes, although we try to be careful so that it does not endanger the stability of the whole of our enterprise. So, clearly, there can be no proof in the proper sense of the word that a proof really proves its conclusion: the most there can be is a discussion as to whether a proof is correct or not, and such discussion arises precisely when we make mistakes, or when we have insufficient grounds or evidence for what we are doing.

As should be clear from what I have just said, this notion of validity or conclusiveness or correctness of a proof is a very fundamental notion. Indeed, it is the most fundamental one of all, the one of all the notions that I have discussed which has no other notion before it, because to say that a proof is valid or conclusive or correct, as should be clear from the formulations that I have used, is nothing but saying,

either that it is a proof with an emphatic is, or, better, that it is a real proof and not a deceptive proof: it is a real proof or it is a true proof, if I use the word true or real the way you use it when you say that this is true bread or that this is real bread and not some miserable bread that you cannot still your hunger with. So the notion of validity or conclusiveness or correctness as applied to proof is nothing but the notion of truth or reality with its opposite falsehood or appearance. This notion of truth is certainly not the same as the notion of truth of a proposition which I started by explaining and identified with direct provability or direct verifiability. In order to distinguish this notion of truth or reality from the notion of truth of a proposition, let me call it the metaphysical notion of truth as opposed to the notion of truth of a proposition. So the outcome of what I have said about the notion of validity of a proof is that validity is nothing but the notion of truth or reality applied to the particular acts and objects with which we are concerned in logic, namely, acts of knowing and objects of knowledge.

The last point I want to make is that the knowledge theoretical idealism which is so characteristic of the explanation of the notion of truth of a proposition, that is, the intuitionistic explanation of the notion of truth of a proposition, is entirely compatible with realism, if by realism you mean the philosophical position which takes the notion of truth or reality for granted, realism, of course, signifying reality here. And what is the opposite of that position? That is a position for which the notion of truth or reality in this sense does not exist, which means that the most that I can say about a judgement, for instance, is that it is evident to me: it may not be evident to you, and to you something may be evident which is perhaps even in conflict with what is evident to me, and there is no way of resolving that conflict because there is no notion of correctness to appeal to. So the ordinary discussion as to who is right that we mathematicians embark upon in that situation simply cannot arise. That is a position which is opposed to realism in the sense that I just described: it is a position which we normally call a relativistic or subjectivistic position, which is to say that there is nothing but proof for me or evidence for me and evidence for you without there being any absolute standard to invoke. Maybe this kind of realism could be called metaphysical realism to distinguish it on the one hand from knowledge theoretical realism, that is, the view that the world exists independently of us and our cognitive activity,

which is the opposite of the knowledge theoretical idealism characteristic of the intuitionistic analysis of the notion of truth of a proposition, and on the other hand from the realism with respect to the existence of universals which figured in the medieval debate about the nature of universals. If you agree to use the word realism also for this third position, namely, the position which simply takes the notion of truth or reality for granted, then we mathematicians, whether intuitionists or not, all seem to be realists: we do engage in discussions as to whether a proof is correct or not, and, once we do that, we have already taken the notion of correctness or truth or reality for granted. Instead of metaphysical realism, you could of course use the opposite of the word relativism or subjectivism, which would be absolutism or objectivism, but the words do not matter here so much. The point I wanted to make is that the knowledge theoretical idealism which I began by expounding is perfectly compatible with metaphysical realism in this sense, whereby I mean the position which takes the notion of truth or reality for granted.

NOTE

* This talk was given at the workshop Theories of Meaning organized by the Centro Fiorentino di Storia e Filosofia della Scienza at the Villa di Mondeggi near Florence, 3-7 June 1985. I am greatly indebted to the organizers of the workshop for recording the talk on tape and subsequently making the tapes available to me for transcription.

Department of Mathematics
University of Stockholm
Box 6701
113 85 Stockholm
Sweden