Zermelo, E. (1930). Über Grenzzahlen und Mengenbereiche: Neue Untersuchungen über die Grundlagen der Mengenlehre. *Fundamenta Mathematicae*, **16**, 29–47.

Department of Mathematics
University of California
Los Angeles
CA 90095-1555
U.S.A.

Department of Mathematics
University of Athens
Athens, Greece

email : ynm@math.ucla.edu

5

Truth and knowability: on the principles C and K of Michael Dummett

Per Martin-Löf

Truth is a highly ambiguous term. At least four clearly recognizable senses, all of relevance for this meeting, can be listed, namely,

Tarski's notion of truth of a closed formula, or sentence,
truth of a proposition,
truth of an assertion, or judgement,
truth in the sense of reality, as opposed to appearance.

There is an ambiguity in the term 'assertion': you may use it either generally for a claim, a knowledge claim, or specifically for an affirmation, that is, for a claim of the form 'A is true', where A is a proposition. In this talk, I shall use it consistently in the first of these two senses. Also, I shall use the terms 'assertion' and 'judgement' synonymously. At the bottom of my list is the notion of truth as one pole of the distinction between appearance and reality: it is so to speak the high notion of truth, often capitalized, although I have put it at the bottom. At the top is the notion of truth of closed formulas, or sentences, which is the lowest notion in the sense that it is a purely mathematical notion, determined by Tarski's well-known recursive definition, and I shall not be concerned with that either, but remain in the middle region, dealing exclusively with the notion of truth of a proposition and the notion of truth of an assertion, or judgement.

Now, in his long paper 'What is a theory of meaning? (II)' from 1976, Michael Dummett posed the problem, quoting verbatim, of how the notion of truth, within a theory of meaning in terms of verification, should be explained. The idea is of course that, in a truth-conditional theory of meaning, the notion of truth has to be there from the very beginning, since the meaning explanations of the various logical constants are given in terms of truth conditions. But suppose now that we replace the notion of truth as the basic notion by the notion of proof, or verification. Then, at the most basic level, we shall not speak about truth any longer, but instead about proof, or verification, and there then arises the problem that Dummett formulated: even if the notion of truth of a proposition is no longer the basic notion, we are still interested in how it is to be understood. And, in that same paper, he formulated two principles that ought to be satisfied by the notion of truth of a proposition, or statement, as he himself says, namely,

> C: If a statement is true, there must be something in virtue of which it is true,

and

> K: If a statement is true, it must be in principle possible to know that it is true.

Actually, these two principles form a recurring theme in Dummett's writings. The first principle occurs already in his very early paper 'Truth' from 1959, where the formulation is even more explicit, saying as it does that a statement is true only if there is something in the world in virtue of which it is true. Both principles occur together for the first time in the Postscript that was added to it in 1972, and they then recur in Chapter 13: Can Truth be Defined? of *Frege: Philosophy of Language* as well as in 'What is a theory of meaning? (II)', which is where they are labelled C and K.

It is clear from the label of the first principle, C for Correspondence, that it is meant to be a formulation of the well-known correspondence principle, which as we know goes back to Aristotle and is so basic that it so to speak has got to be right if it is sensibly interpreted. For instance, if I say, 'My fountain pen is blue', there is something in the world in virtue of which that is so, if it really is so, namely, the blueness of my fountain pen. So this is it, in this case, which is there and verifies, or makes true, the proposition that my fountain pen is blue. It is clear that the correspondence principle, understood in this very general and unsophisticated way, is somehow right, and has to be right on any conception, whether it be in terms of a truth-conditional or a verificationist, a classical or an intuitionist theory of meaning.

Then there is the second principle K, where I suppose K stands for Knowability, or at least something having to do with Know, which says that, if a proposition is true, it must be in principle possible to know that it is true. As you see, this is a principle which is quite different from C, and, whereas C is so to speak readily accepted, when you look at K, I think you immediately get some feeling of uneasiness: could the 'true' that you have in the conditional clause possibly be the same 'true' as you have in the main clause? It sounds somehow strange to say that, if a proposition is true, then, from that alone, it follows that it is in principle possible to know that it is true, in the same sense of 'true' as you have in the conditional clause: it seems somehow unlikely. At least, this has left me with uneasiness, and the purpose of this talk is to try to resolve the difficulties which are inherent in the principle K, and actually also to emend it in such a way that it becomes acceptable.

The key to resolving the perplexities surrounding the principle K turns out to be the very distinction between the notion of truth of a proposition and the notion of truth of an assertion, or judgement, with which I started. First of all, I should say that you cannot hope to explain these two notions, truth of a proposition and truth of a judgement, in isolation: they are two concepts that fit into a certain conceptual structure, where also other notions are involved, and, if we want to clarify them, we shall have to display, as it were, this little conceptual structure, or conceptual system, and see how the various pieces fit together and what functions they fulfil in it. The key elements of this conceptual structure are the ones that are displayed in the following table:

Non-epistemic concepts	Epistemic concepts
proposition	judgement
proof (verification) of a proposition	proof (demonstration) of a judgement
truth of a proposition	truth (correctness) of a judgement

So we shall have to clarify the notion of proposition and the notion of judgement, and we shall have to clarify the notion of proof of a proposition as opposed to the notion of proof of a judgement. Here we have a good terminological possibility, because in English we have both the term 'proof' and the term 'demonstration', and 'demonstration' is quite unambiguously associated with making something evident, which is to say that it is an ideal word to use on the epistemic side, demonstration of a judgement, and then we get the term 'proof' free for propositions, or, if you prefer, you could also use verification in connection with propositions. Finally, we shall have to clarify the two notions with which I started, namely, the notion of truth of a proposition and the notion of truth of a judgement, and, if one finds it inconvenient to use truth in both cases here, although it is sometimes unavoidable, one can decide to use correctness, or objective correctness, in connection with assertions, or judgements: that is what Dummett usually does. But, of course, this means already deciding to make a technical distinction between truth and correctness, because no doubt ἀληθής and ὀρθός were used essentially synonymously in connection with the Greek δόξα and, similarly, in scholastic philosophy, you had the Latin *judicium verum seu rectum*. But now that we have the two words, this is a convenient technical terminology.

To begin with, I would like to say something preliminary about the distinction between propositions and judgements, before properly answering the questions, 'What is a proposition?' and, 'What is a judgement?' Now propositions are the things that are held true, or sometimes held false, and the things on which the logical operations operate: the connectives operate on propositions and the quantifiers on propositional functions. Judgements, on the other hand, are what we demonstrate: in each step of a chain of reasoning, or demonstration, proceed from some previously demonstrated judgements to a new judg which is evident on the grounds of the previous ones, such a step being an inference with the previously made judgements as premises and the newly made judgement as conclusion. The forms of judgement are totally different from the logical operations. First of all, we have the affirmative form of judgement

$$A \text{ is true,}$$

where A is a proposition. This is the only form of judgement that I shall need to consider in the course of this talk, but there are also hypothetical judgements, or consequences, general judgements, and hypothetico-general judgements, which all have their own characteristic forms. A Gentzen sequent is an example of a hypothetico-general judgement, that is, a judgement which is both hypothetical and general. These are some forms of judgement that are used in predicate logic, but there are many more forms of judgement, and, just as we cannot limit in advance our logical operations, we cannot limit in advance our forms of judgement: indeed, in type theory, there are several other forms of judgement, in particular, the form of judgement which is used to say that something is an object of a certain type, and, just as crucially, that two objects of a certain type are the same, where 'same' means definitionally or intensionally the same.

In the preceding, I made the distinction between propositions and judgements in a preliminary fashion by simply giving examples of some well-known forms of proposition and some well-known forms of judgement, but, by doing so, I have of course not really defined what a proposition is and what a judgement is. So what is a judgement? Well, the notion of judgement, and everything actually that stands in the right-hand column of my table, is an epistemic notion, which means that the notion of knowledge is crucially involved. The simplest answer to the question of what a judgement is seems to be to say that a judgement is defined by laying down what it is that you must know in order to have the right to make it. Or, using the term 'assertion' rather than 'judgement', an assertion is a knowledge claim, and hence, in order to clarify the assertion, you have to clarify what knowledge it is that you claim to have when you make the assertion. So, however you phrase the explanation, the crucial question is, 'What knowledge?'

Now, once we have fixed the notion of judgement in this way, the notion of demonstration of a judgement, which is located on the second line of the right-hand column of my table, is defined simply by saying that a demonstration is what makes a judgement known, or evident: a demonstration is a chain of reasoning, and what it purports to do is to make the final judgement of that chain known, or evident. There are many words that you can choose among here, and from a logical point of view it is immaterial which of these terms you choose, because they are but different labels of one and the same piece in the conceptual structure, and you may label that piece in any way you want: the only important thing is how it functions in the structure. The natural labels here are to say known, evident, demonstrated, justified, or warranted: this is the term usually adopted by Dummett, contrasting as he does an assertion's being

warranted with its being correct, which is the next notion to be analysed, or you may say reasoned, or grounded. So the notion of evidence here comes before the notion of truth, or correctness, of a judgement in the conceptual order.

But now, having the notion of evidence, or knownness, how do we define the notion of truth, or correctness, of a judgement? Well, the proper conceptual connection seems to be this: a judgement is by definition true, or correct, if it *can* be known, or made evident. You see, evident means known, which is to say, actually known, but a judgement is true, or correct, if it is knowable, evidenceable, demonstrable, justifiable, warrantable, or groundable, whichever you prefer. The crucial notion that comes in here is the notion of possibility, and it is of course a question of possibility in principle. So the difference between, on the one hand, known, evident, demonstrated, and so on, and, on the other hand, knowable, evidenceable, demonstrable, and so on, is nothing but the difference between actuality and potentiality. Now this definition of the notion of truth, or correctness, of a judgement validates Leibniz's principle of sufficient reason. The most widespread formulation of it has several ingredients, but, if we restrict ourselves to what has to do with the truth of judgements, then what Leibniz's principle of sufficient reason says is that, if a judgement is true, then it can be known. A judgement is not true unless there exists a reason for it, that is, unless a reason for it can be given: that is the content of the principle of sufficient reason. And why does it hold? Well, it holds because of the definition of the notion of truth of a judgement: truth of a judgement is simply defined as knowability, and therefore the principle holds. This was also Leibniz's own view, that the principle of sufficient reason is contained in the definition of the notion of truth.

Now, as an indication that the conceptual connections have been properly made here, I would like to say a few words about Descartes' criterion of truth. Stated as briefly as possible, it says that, if a judgement is evident, then it is true: *si quid intellectui meo sit evidens, illud omnino est verum.* There is no surer sign of the truth of a judgement than our having made it evident to ourselves: that is the gist of Descartes' truth criterion. So evidence implies truth, or correctness, of a judgement. Now, as Brian McGuinness said in his introduction to this meeting, Descartes had to invoke the veracity of God in order to justify his truth criterion, because why does it hold, according to Descartes? Well, it holds because he took it as an axiom that God does not deceive us. But, at least to my mind, it would be very strange if one should have to invoke the notions of God and deception in order to see that the evidence of a judgement entails its truth. Things of this sort normally hold on purely conceptual grounds, and you see now how it comes out: truth is simply defined as evidenceability, and hence Descartes' truth criterion, saying as it then does that, if a judgement has been made evident, then it can be made evident, follows from the principle that, if something has been done, then it can be done. This, on the other hand, is the truly fundamental metaphysical principle which was given the succinct scholastic formulation *ab esse ad posse valet consequentia*, a formulation which in its turn probably derives from the short passage ἐξ ἐνεργείας ἡ δύναμις in Aristotle's

Metaphysics, Book Θ, Chapter IX. So Descartes' truth criterion, fundamental as it may seem to be, is actually a consequence of this even more basic principle, the *ab esse ad posse* principle.

Another effect of this definition of the notion of truth of a judgement is that the traditional Platonic characterization of knowledge as justified true opinion, *δόξα ἀληθὴς μετὰ λόγου,* opinion true with justification, or by aid of justification, does not look natural any longer when the notion of truth receives the conceptual determination that I have just given to it. Indeed, since true is the same as justifiable, justified true opinion becomes justified justifiable opinion. But, if an opinion is justified, it is superfluous to say that it is justifiable by the *ab esse ad posse* principle. Hence 'justifiable' can be omitted from the formulation, and we get the simpler characterization of knowledge as justified opinion. Also, although *δόξα* is traditionally rendered by opinion, it is equally well translated by judgement, so a piece of knowledge is the same as a justified, or demonstrated, judgement. Presumably, the reason for the more complicated formulation is that, from Plato onwards, the notions of knowledge and truth have been associated with infallibility, and, if you include infallibility in the notion of truth of a judgement, then you cannot argue from evidence to truth in this simple way by the *ab esse ad posse* principle, and that is precisely why Descartes had to invoke the veracity of God at this point. Now, as a matter of fact, our demonstrations are not infallible: a demonstration purports to make something evident to us, and it is the best guarantee that we have, but it is not infallible. We do sometimes make mistakes in our demonstrations, and hence, if you include infallibility in the notion of truth of a judgement, then the step from evidence to truth cannot be taken any longer. That means that the problems that have to do with infallibility have to be moved to another level, so to speak, and that is the level that I put at the bottom of my list, that is, the highest level that has to do with the notion of truth in the sense of reality as opposed to falsehood in the sense of appearance, illusion, or deception, and that will be completely left out of my talk.

This finishes the semantical explanations of the concepts occurring in the right-hand column of my table, that is, the epistemic concepts that are associated with the notion of judgement. There remain the non-epistemic concepts in the left-hand column of the table, which is to say, the notion of proposition, the notion of proof, or verification, of a proposition, and the notion of truth of a proposition. So what is a proposition? Well, in a truth-conditional theory of meaning, a proposition is defined by its truth conditions, whereas, in a verificationist theory of meaning, this explanation is replaced by saying that a proposition is defined by its proof conditions, or verification conditions, which state what a proof, or verification, of the proposition looks like. Now it has sometimes been said, for example, by Dummett in his paper 'Truth' from 1959, that the difference between a classical and an intuitionist, or constructivist, account of the meanings of the logical constants is that truth conditions are replaced by assertion conditions. But observe that that is not what I am saying here: I am saying that truth conditions are replaced by proof conditions, or verification conditions. Now the notion of assertion condition is also important, but the role of an assertion condition is to determine the meaning of an assertion, or judgement, the concept that we had at the top of the right-hand column of the table of concepts to be explained: an assertion, or judgement, is defined by its assertion condition, that is, by laying down what it is that you must know in order to have the right to assert it.

As concerns the notion of proof of a proposition, we must distinguish between proofs, or verifications, of the forms that enter into the meaning explanations of the various logical constants on the one hand, and arbitrary proofs, or verifications, on the other. We all know the Brouwer–Heyting–Kolmogorov explanations of the meanings of the logical constants, which run according to the pattern: a proof of a conjunction $A \& B$ is a pair consisting of a proof of A and a proof of B, and similarly for the other logical operations. But we also have to allow proofs which are not directly of one of the forms that enter into the meaning explanations of the logical constants, just as, when we let the natural numbers be defined by the first two Peano axioms, 'Zero is a natural number', and, 'If n is a natural number, the successor of n is a natural number', some innocent person may come and ask, 'But what about $2+2$, is it not a natural number?' The answer is of course that, when you give an inductive definition, like that of the natural numbers, it is tacitly understood that something should count as a natural number even if you may need to calculate it a few steps to get it into zero or successor form, and similarly here, in the Brouwer–Heyting–Kolmogorov explanations, a proof in general may have to be calculated before you get it into the form, or one of the forms, that define the proposition in question. We then have two terminological possibilities, either to call proofs of the forms that enter into the meaning explanations of the logical constants simply 'proofs', in which case we would in general only have a method of proof, or to call proofs of the forms prescribed by the meaning explanations 'canonical proofs', or 'direct verifications', in which case we also have to allow non-canonical proofs, or indirect verifications. Choosing the latter alternative, a non-canonical proof, or indirect verification, becomes clearly the same as a method of canonical proof, or direct verification. So we have these two terminological possibilities.

Now, if a proposition is defined in this way by its proof conditions, then, when you come to the next question in the left-hand column of my table, which is to say, 'What is a proof of a proposition?', the answer is exceedingly simple, because a proposition was defined precisely by laying down how its proofs are formed, which means that there is nothing more that needs to be said: Indeed, once we have understood the proposition, we already know what a proof of it is, a canonical proof in the first place, and then a proof in general is a method such that, when you execute it, you obtain a canonical proof as result.

There now remains in the left-hand column only the notion of truth of a proposition, which appears on the third and last line. So we must ask ourselves, 'How is the notion of truth of a proposition to be defined?' This is precisely the problem of Dummett's that I started by quoting, namely, of how the notion of truth, within a theory of meaning in terms of verification, should be explained.

The answer is most simply given in the form of the chain of equations

$$\begin{aligned} A \text{ is true} &= \text{there exists a proof of } A \\ &= \text{a proof of } A \text{ can be given} \\ &= A \text{ can be proved} \\ &= A \text{ is provable,} \end{aligned}$$

in which the equality sign signifies sameness of meaning. So here again the notion of possibility in principle comes in, but now it is in connection with the notion of truth of a proposition, whereas previously it was in connection with the notion of truth, or correctness, of a judgement. And 'proof' is here to be understood in the sense of 'canonical proof', which means that the truth of a proposition is equated with the possibility of coming up with a canonical proof, or direct verification, of it. So here I have chosen the first of the two terminological alternatives that I mentioned. Now what are given in the chain of equalities are but different permissible readings of one and the same form of judgement 'A is true'. After all, I have to follow my own official explanations, and, since 'A is true' is a form of assertion, or judgement, its meaning is determined by laying down its assertion condition, that is, by laying down what it is that you must know in order to have the right to make a judgement of this form, and, in this case, the explanation is that, to have the right to make a judgement of the form 'A is true', you must know a proof of A, a proof which is in general non-canonical, that is, which is in general merely a method such that, when you execute it, you get a canonical proof as result. Now that is the official meaning explanation, but it is clear from that meaning explanation that you may allow yourself to read 'A is true' in these different ways, which are of course quite similar, actually, to the reading that Kleene used in his realizability interpretation: forgetting about all other differences, Kleene read the proposition that there exists a realizer of A, where A is an arithmetical formula, as 'A is realizable'.

Now let me return to my original promise of clarifying Dummett's principles C and K. If you first look at C, 'If a statement is true, there must be something in virtue of which it is true', you will see that it is in complete agreement with what I have said about the notion of truth of a proposition. Indeed, the verificationist definition of truth is that a proposition is true if there exists a proof of it, so, if we just call that something in virtue of which a statement is true its proof, or verification, then C is nothing but the definition of truth that I just gave. That means of course that the intuitionist, or verificationist, notion of truth is really a version of the correspondence notion of truth, truth as agreement with reality: the only novelty is that we call that thing in reality, or in the world, which has to be there in order for the proposition to be true, its proof, or verification.

Let us now finally turn to the principle K, 'If a statement is true, it must be in principle possible to know that it is true'. So remember the principle of sufficient reason, which says that, if a judgement is true, then it can be known. Now apply the principle of sufficient reason to a judgement of the particular form 'A is true', where A is a proposition. Then what we get is that, if a judgement of the form 'A is true' is correct, then this judgement can be known, but that is the same as saying that the proposition A can be known to be true. So we have now achieved in the main clause exactly what you find in the principle K, but there is a fundamental difference in the conditional clause, which no longer takes the simple form 'if the proposition A is true', but the more complicated form 'if the assertion, or judgement, "A is true" is correct'. This means that two truth operators have turned out to be involved here: one is the truth of the given proposition, and the other is the truth, or correctness, of the judgement which is obtained by applying the first truth operator to the given proposition. So this is the corrected form of the principle K that we have arrived at:

> If a judgement of the form 'A is true' is correct, then the proposition A can be known to be true.

Now, unfortunately, this reads a bit awkwardly, but it may be rephrased in the following way, if only we accept the principle that a judgement of the form 'A is true' is correct if and only if the proposition A really is true. This is a principle that I think everybody accepts: the only difference that you find between the realist and the idealist is in the sense that they give to the qualifier 'really' that appears here. The realist takes that notion as a primitive notion that cannot be reduced to anything else, whereas, on the analysis that I have given, the notion of reality that comes in here is nothing but the notion of knowability. In any case, the principle is acceptable as it stands, and hence we can replace saying that the judgement that A is true, where A is a proposition, is correct by saying that the proposition A really, or in reality, is true. If we make that replacement, we arrive at the following

> Emendation of K: If a proposition really is true, then it can be known to be true.

This is the amended version of the principle K that I propose. It agrees entirely with the principle K in the main clause, but has a crucial modification in the conditional clause, and it is an almost immediate consequence of the principle of sufficient reason.

Acknowledgements

This paper is an only slightly altered transcript of the talk that I gave at the conference, to whose organizers I am indebted for a tape recording which lasted until the power failed. After that, I have relied on a battery-powered tape recording, kindly given to me by Margaret MacDougall.

Bibliography

Dummett, M.A.E. (1959). Truth. *Proceedings of the Aristotelian Society*, new series, **59**, 141–62. Postscript (1972). In *Logic and philosophy for linguists: a book of readings* (ed. J. M. E. Moravcsik), pp. 220–5. Mouton, The Hague.

Dummett, M.A.E. (1973). *Frege: philosophy of language*. Duckworth, London.

Dummett, M.A.E. (1976). What is a theory of meaning? (II). In *Truth and meaning* (ed. G. Evans and J. McDowell), pp. 67–137. Clarendon Press, Oxford.

Department of Mathematics
University of Stockholm
106 91 Stockholm
Sweden

PART II

Formalism and naturalism