MICHAEL DUMMETT

TRUTH AND OTHER ENIGMAS

1. *Truth* (1959)

FREGE HELD THAT truth and falsity are the references of sentences. Sentences cannot stand for propositions (what Frege calls 'thoughts'), since the reference of a complex expression depends only on the reference of its parts; whereas if we substitute for a singular term occurring in a sentence another singular term with the same reference but a different sense, the sense of the whole sentence, i.e. the thought which it expresses, changes. The only thing which it appears *must* in these circumstances remain unchanged is the truth-value of the sentence. The expressions 'is true' and 'is false' look like predicates applying to propositions, and one might suppose that truth and falsity were properties of propositions; but it now appears that the relation between a proposition and its truth-value is not like that between a table and its shape, but rather like that between the sense of a definite description and the actual object for which it stands.

Frege's starting point: truth-values as what sentences refer to.

To the objection that there are non-truth-functional occurrences of sentences as parts of complex sentences, e.g., clauses in indirect speech, Frege replies that in such contexts we must take ordinary singular terms as standing, not for their customary reference, but for their sense, and hence we may say that in such a context, and only then, a sentence stands for the proposition it usually expresses.

If someone asks, 'But what kind of entities are these truth-values supposed to be?', we may reply that there is no more difficulty in seeing what the truth-value of a sentence may be than there is in seeing what the direction of a line may be; we have been told when two sentences have the same truth-value—when they are materially equivalent—just as we know when two lines have the same direction—when they are parallel. Nor need we waste time on the objection raised by Max Black that on Frege's theory certain sentences become meaningful which we should not normally regard as such, e.g., 'If oysters are inedible, then the False'. If sentences stand for truth-values, but

Truth-values like directions of lines, and Max Black's oysters.

12. †Dummett: Truth

Why Dummett grants that sentences have a reference.

there are also expressions standing for truth-values which are not sentences, then the objection to allowing expressions of the latter kind to stand wherever sentences can stand and vice versa is grammatical, not logical. We often use the word 'thing' to provide a noun where grammar demands one and we have only an adjective, e.g., in 'That was a disgraceful thing to do'; and we could introduce a verb, say 'trues', to fulfil the purely grammatical function of converting a noun standing for a truth-value into a sentence standing for the same truth-value. It may be said that Frege has proved that a sentence does not ordinarily stand for a proposition, and has given a plausible argument that *if* sentences have references, they stand for truth-values, but that he has done nothing to show that sentences do have references at all. This is incorrect; Frege's demonstration that the notions of a concept (property) and a relation can be explained as special cases of the notion of a function provides a plausible argument for saying that sentences have a reference.

The complaint is about the words 'truth' and 'falsity'.

What *is* questionable is Frege's use of the words 'truth' and 'falsity' as names of the references of sentences; for by using these words rather than invented words of his own he gives the impression that by taking sentences to have a reference, with material equivalence as the criterion of identity, he has given an account of the notions of truth and falsity which we are accustomed to employ. Let us compare truth and falsity with the winning and losing of a board game. For a particular game we may imagine first formulating the rules by specifying the initial position and the permissible moves; the game comes to an end when there is no permissible move. We may then distinguish between two (or three) kinds of final position, which we call 'Win' (meaning that the player to make the first move wins), 'Lose' (similarly) and possibly 'Draw'. Unless we tacitly appeal to the usual meanings of the words 'win', 'lose' and 'draw', this description leaves out one vital point—that it is the object of a player to win. It is part of the concept of winning a game that a player plays to win, and this part of the concept is not conveyed by a classification of the end positions into winning ones and losing ones. We can imagine a variant of chess in which it is the object of each player to be checkmated, and this would be an entirely different game; but the formal description we imagined would coincide with the formal description of chess. The whole theory of chess could be formulated with reference only to the formal description; but which theorems of this theory interested us would depend upon whether we wished to play chess or the variant game. Likewise, it is part of the concept of truth that we aim at making true statements; and Frege's theory of truth and falsity as the references of sentences leaves this feature of the concept of truth quite out of account. Frege indeed tried to bring it in afterwards in his theory of

The board game, and chess played to be checkmated.

assertion—but too late; for the sense of the sentence is not given in advance of our going in for the activity of asserting, since otherwise there could be people who expressed the same thoughts but went in instead for denying them.

An account must give the point, as with valid argument.

A similar criticism applies to many accounts of truth and falsity or of the meanings of certain sentences in terms of truth and falsity. We cannot in general suppose that we give a proper account of a concept by describing those circumstances in which we do, and those in which we do not, make use of the relevant word, by describing the *usage* of that word; we must also give an account of the *point* of the concept, explain what we use the word *for*. Classifications do not exist in the void, but are connected always with some interest which we have, so that to assign something to one class or another will have consequences connected with this interest. A clear example is the problem of justifying a form of argument, deductive or inductive. Classification of arguments into (deductively or inductively) valid and invalid ones is not a game played merely for its own sake, although it *could* be taught without reference to any purpose or interest, say as a school exercise. Hence there is really a problem of showing that the criteria we employ for recognising valid arguments do in fact serve the purpose we intend them to serve: the problem is not to be dismissed—as it has long been fashionable to do—by saying that we use the criteria we use.

One classification, several points: ethics and empty names.

We cannot assume that a classification effected by means of a predicate in use in a language will always have just *one* point. It may be that the classification of statements into true ones, false ones, and, perhaps, those that are neither true nor false, has one principal point, but that other subsidiary ends are served by it which make the use of the words 'true' and 'false' more complex than it would otherwise be. At one time it was usual to say that we do not call ethical statements 'true' or 'false', and from this many consequences for ethics were held to flow. But the question is not whether these words are in practice applied to ethical statements, but whether, if they were so applied, the point of doing so would be the same as the point of applying them to statements of other kinds, and, if not, in what ways it would be different. Again, to be told that we say of a statement containing a singular term which lacks reference that it is neither true nor false is so far only to be informed of a point of usage; no philosophical consequences can yet be drawn. Rather, we need to ask whether describing such a statement as neither true nor false accords better with the general point of classifying statements as true or false than to describe it as false. Suppose that we learn that in a particular language such statements are described as 'false': how are we to tell whether this shows that they use such statements differently from

The people who call empty-name statements 'false'.

ourselves or merely that 'false' is not an exact translation of their word? To say that we use singular statements in such a way that they are neither true nor false when the subject has no reference is meant to characterise our use of singular statements; hence it ought to be possible to describe when in a language not containing words for 'true' and 'false' singular statements would be used in the same way as we use them, and when they would be used so as to be false when the subject had no reference. Until we have an account of the general point of the classification into true and false we do not know what interest attaches to saying of certain statements that they are neither true nor false; and until we have an account of how the truth-conditions of a statement determine its meaning the description of the meaning by stating the truth-conditions is valueless.

Frege's other account: 'It is true that P' says just P.

A popular account of the meaning of the word 'true', also deriving from Frege, is that ⌜It is true that P⌝ has the same sense as the sentence P. If we then ask why it is any use to have the word 'true' in the language, the answer is that we often refer to propositions indirectly, i.e., without expressing them, as when we say 'Goldbach's conjecture' or 'what the witness said'. We also generalise about propositions without referring to any particular one, e.g., in 'Everything he says is true'. This explanation cannot rank as a definition in the strict sense, since it permits elimination of 'is true' only when it occurs attached to a 'that'-clause, and not when attached to any other expression standing for a proposition or to a variable; but, since every proposition can be expressed by a sentence, this does not refute its claim to be considered as determining uniquely the sense of 'is true'. It might be compared with the recursive definition of '+', which enables us to eliminate the sign '+' only when it occurs in front of a numeral, and not when it occurs in front of any other expression for a number or in front of a variable; yet there is a clear mathematical sense in which it specifies uniquely what operation '+' is to signify. Similarly, our explanation of 'is true' determines uniquely the sense, or at least the application, of this predicate: for any given proposition there is a sentence expressing that proposition, and that sentence states the conditions under which the proposition is true.

Not a strict definition, yet it fixes the sense of 'is true'.

First consequence: no room for truth-value gaps.

If, as Frege thought, there exist sentences which express propositions but are neither true nor false, then this explanation appears incorrect. Suppose that P contains a singular term which has a sense but no reference: then, according to Frege, P expresses a proposition which has no truth-value. This proposition is therefore not true, and hence the statement ⌜It is true that P⌝ will be *false*. P will therefore not have the same sense as ⌜It is true that P⌝, since the latter is false while the former is not. It is not possible to plead that ⌜It is true that P⌝ is itself neither true nor false when the singular term occurring in P lacks a reference, since the *oratio obliqua* clause ⌜that P⌝ stands for the proposition expressed by P, and it is admitted that P does have a sense and express a proposition; the singular term occurring in P has in ⌜It is true that P⌝ its indirect reference, namely its sense, and we assumed that it did have a sense. In general, it will always be inconsistent to maintain the truth of every instance of 'It is true that p if and only if p' while allowing that there is a type of sentence which under certain conditions is neither true nor false. It would be possible to evade this objection by claiming that the 'that'-clause in a sentence beginning 'It is true that' is not an instance of *oratio obliqua*; that the word 'that' here serves the purely grammatical function of transforming a sentence into a noun-clause without altering either its sense or its reference. We should then have to take phrases like 'Goldbach's conjecture' and 'what the witness said' as standing not for propositions but for truth-values. The expression 'is true' would then be exactly like the verb 'trues' which we imagined earlier; it would simply convert a noun-phrase standing for a truth-value into a sentence without altering its sense or its reference. It might be objected that this variant of Frege's account tallies badly with his saying that it is the *thought* (proposition) which is what is true or false; but we can express this point of Frege's by saying that it is the *thought*, rather than the *sentence*, which primarily stands for a truth-value. A stronger objection to the variant account is that it leans heavily on the theory of truth-values as references of sentences, while the original version depends only on the more plausible view that clauses in indirect speech stand for propositions. In any case, if there are meaningful sentences which say nothing which is true or false, then there must be *a* use of the word 'true' which applies to propositions; for if we say ⌜It is neither true nor false that P⌝, the clause ⌜that P⌝ must here be in *oratio obliqua*, otherwise the whole sentence would lack a truth-value.

The gap can't be pushed into 'It is true that P' as well.

The 'trues' escape, and what it costs.

'False' needs negation, and no rule forms the negation.

Even if we do not wish to say of certain statements that they are neither true nor false, this account cannot give the *whole* meaning of the word 'true'. If we are to give an explanation of the word 'false' parallel to our explanation of 'true' we shall have to say that ⌜It is false that P⌝ has the same sense as the negation of P. In logical symbolism there exists a sign which, put in front of a sentence, forms the negation of that sentence; but in natural languages we do not have such a sign. We have to think to realise that the negation of 'No one is here' is not 'No one is not here' but 'Someone is here'; there is no one rule for forming the negation of a given sentence. Now according to what principle do we recognise one sentence as the negation of another? It is natural to answer: The negation of a sentence P is that sentence which is true if and only if P is false and false if and only if P is

true. But this explanation is ruled out if we want to use the notion of the negation of a sentence in order to explain the sense of the word 'false'. It would not solve the difficulty if we did have a general sign of negation analogous to the logical symbol, for the question would then be: How in general do we determine the sense of the negation, given the sense of the original sentence?

'And' by when we may assert it: the trained dog.

We encounter the same difficulty over the connective 'or'. We can give an account of the meaning of 'and' by saying that we are in a position to assert ⌜P and Q⌝ when and only when we are in a position to assert P and in a position to assert Q. (This is not circular: one could train a dog to bark only when a bell rang *and* a light shone without presupposing that it possessed the concept of conjunction.) But, if we accept a two-valued logic, we cannot give a similar explanation of the meaning of 'or'. We often assert ⌜P or Q⌝ when we are not either in a position to assert P or in a position to assert Q. I use the word 'we' here, meaning mankind, advisedly. If the history master gives the schoolboy a hint, saying, 'It was either James I or Charles I who was beheaded', then the schoolboy is in a position to assert, 'Either James I or Charles I was beheaded' without (perhaps) being in a position to assert either limb of the disjunction; but it is not this sort of case which causes the difficulty. The *ultimate* source of the schoolboy's knowledge derives from something which justifies the assertion that Charles I was beheaded; and this is all that would be required for the proposed explanation of the word 'or' to be adequate. Likewise, the explanation is not impugned by cases like that in which I remember that I was talking either to Jean or to Alice, but cannot remember which. My knowledge that I was talking either to Jean or to Alice derives ultimately from the knowledge that I had at the time that I was talking to (say) Jean; the fact that the incomplete knowledge is all that survives is beside the point. Rather, the difficulty arises because we often make statements of the form ⌜P or Q⌝ when the ultimate evidence for making them, in the sense indicated, is neither evidence for the truth of P nor evidence for the truth of Q. The most striking instance of this is the fact that we are prepared to assert *any* statement of the form ⌜P or not P⌝, even though we may have no evidence either for the truth of P or for the truth of ⌜Not P⌝.

The schoolboy and Jean or Alice aren't where 'or' resists.

The real trouble: 'P or not P' with no evidence either way.

The truth-table appeal assumes the bivalence it should justify.

In order to justify asserting ⌜P or not P⌝, we appeal to the truth-table explanation of the meaning of 'or'. But if the whole explanation of the meanings of 'true' and 'false' is given by 'It is true that p if and only if p' and 'It is false that p if and only if not p', this appeal fails. The truth-table tells us, e.g., that from P we may infer ⌜P or Q⌝ (in particular, ⌜P or not P⌝); but *that* much we already knew from the explanation of 'or' which we have

rejected as insufficient. The truth-table does not show us that we are entitled to assert ⌜P or not P⌝ in every possible case, since this is to assume that every statement is either true or false; but, if our explanation of 'true' and 'false' is all the explanation that can be given, to say that every statement is either true or false is just to say that we are always justified in saying ⌜P or not P⌝.

So the truth-table account of the connectives fails.

We naturally think of truth-tables as giving the explanation of the sense which we attach to the sign of negation and to the connectives, an explanation which will show that we are justified in regarding certain forms of statement as logically true. It now appears that if we accept the redundancy theory of 'true' and 'false'—the theory that our explanation gives the whole meaning of these words—the truth-table explanation is quite unsatisfactory. More generally, we must abandon the idea which we naturally have that the notions of truth and falsity play an essential rôle in any account either of the meaning of statements in general or of the meaning of a particular statement. The conception pervades the thought of Frege that the general form of explanation of the sense of a statement consists in laying down the conditions under which it is true and those under which it is false (or better: saying that it is false under all other conditions); this same conception is expressed in the *Tractatus* in the words, 'In order to be able to say that "p" is true (or false), I must have determined under what conditions I call "p" true, and this is how I determine the sense of the sentence' (4.063). But in order that someone should gain from the explanation that P is true in such-and-such circumstances an understanding of the sense of P, he must already know what it means to say of P that it is true. If when he enquires into this he is told that the only explanation is that to say that P is true is the same as to assert P, it will follow that in order to understand what is meant by saying that P is true, he must already know the sense of asserting P, which was precisely what was supposed to be being explained to him.

Frege and Tractatus 4.063 on sense, and the circle.

Supplement or give up: no criterion of truth, but a theory.

We thus have either to supplement the redundancy theory or to give up many of our preconceptions about truth and falsity. It has become a commonplace to say that there cannot be a criterion of truth. The argument is that we determine the sense of a sentence by laying down the conditions under which it is true, so that we could not first know the sense of a sentence and then apply some criterion to decide in what circumstances it was true. In the same sense there could not be a criterion for what constitutes the winning of a game, since learning what constitutes winning it is an essential part of learning what the game is. This does not mean that there may not be in any sense a theory of truth. For a particular bounded language, if it is free of ambiguity and inconsistency, it must be possible to characterise the true sentences of the language; somewhat as, for a given game, we can say

which moves are winning moves. (A language is bounded if we may not introduce into it new words or new senses for old words.) Such a characterisation would be recursive, defining truth first for the simplest possible sentences, and then for sentences built out of others by the logical operations employed in the language; this is what is done for formalised languages by a truth-definition. The redundancy theory gives the general form of such a truth-definition, though in particular cases more informative definitions might be given.

Truth's role, and the twofold division of states of affairs.

Now we have seen that to say for each particular game what winning it consists in is not to give a satisfactory account of the concept of winning a game. What makes us use the same term 'winning' for each of these various activities is that the point of every game is that each player tries to do what for that game constitutes winning; i.e., what constitutes winning always plays the same part in determining what playing the game consists in. Similarly, what the truth of a statement consists in always plays the same rôle in determining the sense of that statement, and a theory of truth must be possible in the sense of an account of what that rôle is. I shall not now attempt such an account; I claim, however, that such an account would justify the following. A statement, so long as it is not ambiguous or vague, divides all possible states of affairs into just *two* classes. For a given state of affairs, either the statement is used in such a way that a man who asserted it but envisaged that state of affairs as a possibility would be held to have spoken misleadingly, or the assertion of the statement would not be taken as expressing the speaker's exclusion of that possibility. If a state of affairs of the first kind obtains, the statement is false; if all actual states of affairs are of the second kind, it is true. It is thus *prima facie* senseless to say of any statement that in such-and-such a state of affairs it would be neither true nor false.

Bets and commands, and the gap in the conditional bet.

The sense of a statement is determined by knowing in what circumstances it is true and in what false. Likewise the sense of a command is determined by knowing what constitutes obedience to it and what disobedience; and the sense of a bet by knowing when the bet is won and when it is lost. Now there may be a gap between the winning of a bet and the losing of it, as with a conditional bet; can there be a similar gap between obedience and disobedience to a command, or between the truth and falsity of a statement? There is a distinction between a conditional bet and a bet on the truth of a material conditional; if the antecedent is unfulfilled, in the first case the bet is off—it is just as if no bet had been made—but in the second case the bet is won. A conditional command where the antecedent is in the power of the person given the order (e.g., a mother says to a child, 'If you go out, wear your coat') is always like a bet on the material conditional; it is equivalent to the command to ensure the truth of the material conditional, viz., 'Do not go out without your coat'. We cannot say that if the child does not go out, it is just as if no command had been given, since it may be that, unable to find his coat, he stayed in in order to comply with the command.

Commands have one kind of consequence only.

Can a distinction parallel to that for bets be drawn for conditional commands where the antecedent is not in the person's power? I contend that the distinction which looks as if it could be drawn is in fact void of significance. There are two distinct kinds of consequence of making a bet, winning it and losing; to determine what is to involve one of these is not yet to determine completely what is to involve the other. But there is only one kind of consequence of giving a command, namely that, provided one had the right to give it in the first place, one acquires a right to punish or at least reprobate disobedience. It might be thought that punishment and reward were distinct consequences of a command in the same sense that paying money and receiving it are distinct consequences of a bet; but this does not tally with the rôle of commands in our society. The right to a reward is not taken to be an automatic consequence of obedience to a command, as the *right* to reproach is an automatic consequence of disobedience; if a reward is given, this is an act of grace, just as it is an act of grace if the punishment or reproach is withheld. Moreover, any action deliberately taken in order to comply with the command (to avoid disobedience to it) has the same claim to be rewarded as any other; hence to determine what constitutes disobedience to the command is thereby to determine what sort of behaviour might be rewarded, without the need for any further decision. If the child stays in because he cannot find his coat, this behaviour is as meritorious as if he goes out remembering to wear it; and if he forgets all about the order, but wears his coat for some other reason, this behaviour no more deserves commendation than if he chooses, for selfish reasons, to remain indoors. Where the antecedent is not in the person's power, it is indeed possible regard the conditional command as analogous to the conditional bet; but since obedience to a command has no consequence of its own other than that of avoiding the punishment due for disobedience, there is not for such commands any significant distinction parallel to that between conditional bets and bets about a material conditional. If we regarded obedience to a command as giving a right to a reward, we could then introduce such a distinction for commands whose antecedent was in the person's power. Thus the mother might use the form. 'If you go out, wear your coat', as involving that if the child went out with his coat he would be rewarded, if he

Reward is an act of grace, so staying in counts too.

Even outside the child's power, no distinction worth drawing.

went out without it he would be punished, and if he stayed indoors—even in order to comply with the command—he would be neither punished nor rewarded; while the form, 'Do not go out without your coat', would involve his being rewarded if he stayed indoors.

Statements are like commands, not like bets.

Statements are like commands (as we use them) and not like bets; the making of a statement has, as it were, only one kind of consequence. To see this, let us imagine a language which contains conditional statements but has no counterfactual form (counterfactuals would introduce irrelevant complications). Two alternative accounts are suggested of the way in which conditionals are used in this language: one, that they are used to make statements conditionally; the other, that they represent the material conditional. On the first interpretation, a conditional statement is like a conditional bet: if the antecedent is fulfilled, then the statement is treated as if it had been an unconditional assertion of the consequent, and is said to be true or false accordingly; if the antecedent is not fulfilled, then it is just as if no statement, true or false, had been made at all. On the second interpretation, if the antecedent is not fulfilled, then the statement is said to be true. How are we to settle which of these two accounts is the correct one? If statements are really like bets and not like commands; if there are two distinct kinds of consequence which may follow the making of a statement, those that go with calling the statement 'true' and those that go with calling it 'false', so that there may be a gap between these two kinds of consequence; then we ought to be able to find something which decides between the two accounts as definite as the financial transaction which distinguishes a bet on the truth of the material conditional from a conditional bet. It is no use asking whether these people *say* that the man who has made a conditional statement whose antecedent turns out false said something true or that he said nothing true or false: they may have no words corresponding to 'true' and 'false'; and if they do, how could we be sure that the correspondence was exact? If their using the words 'true' and 'false' is to have the slightest significance, there must be some difference in their behaviour which goes with their saying 'true' or 'neither true nor false' in this case.

What would settle the two accounts: a difference in behaviour.

The distinction is empty: one decision fixes the sense.

It is evident on reflection that there is nothing in what they do which could distinguish between the two alternative accounts; the distinction between them is as empty as the analogous distinction for conditional commands whose antecedent is not in the person's power. In order to fix the sense of an utterance, we do not need to make two separate decisions—when to say that a true statement has been made and when to say that a false statement has been made; rather, any situation in which nothing obtains which is taken as a case of its being false may be regarded as a case of its being

true, just as someone who behaves so as not to disobey a command may be regarded as having obeyed it. The point becomes clearer when we look at it in the following way. If it makes sense in general to suppose that a certain form of statement is so used that in certain circumstances it is true, in others false, and in yet others nothing has been said true or false, then we can imagine that a form of conditional was used in this way (von Wright actually holds that *we* use conditionals in this way). If P turns out true, then ⌜If P, then Q⌝ is said to be true or false according as Q is true or false, while if P turns out false we say that nothing was said true or false. Let us contrast this with what Frege and Strawson say about the use in our language of statements containing a singular term. If there is an object for which the singular term stands, then the statement is true or false according as the predicate does or does not apply to that object, but if there is no such object, then we have not said anything true or false. Now do these accounts tell us the sense of sentences of these two kinds?—that is, do they tell us how these statements are used, what is *done* by making statements of these forms? Not at all, for an essential feature of their use has not yet been laid down.

Two gappy forms: the conditional and the singular statement.

Someone uttering a conditional statement of the kind described may very well have no opinion as to whether the antecedent was going to turn out true or false; that is, he is not taken as having misused the statement or misled his hearers if he envisages it as a possibility that that case will arise in which he is said not to have made a statement true or false. All that he conveys by uttering the conditional statement is that he excludes the possibility that the case will arise in which he is said to have said something false, namely that the antecedent is true and the consequent false. With the case of a singular statement it is quite different. Here someone is definitely either misusing the form of statement or misleading his hearers if he envisages it as a possibility that that case will arise in which what he said will be said to be neither true nor false, namely that the singular term has no reference. He conveys more by making the statement than just that he excludes the possibility of its being false; he commits himself to its being true.

The conditional conveys only what it excludes.

The singular statement commits you to its truth.

Are we then to say that laying down the truth-conditions for a sentence is not sufficient to determine its sense, that something further will have to be stipulated as well? Rather than say this we should abandon the notions of truth and falsity altogether. In order to characterise the sense of expressions of our two forms, only a twofold classification of possible relevant circumstances is necessary. We need to distinguish those states of affairs such that if the speaker envisaged them as possibilities he would be held to be either misusing the statement or misleading his hearers, and those of which this is not the case: and *one* way of using the words 'true' and 'false' would be to

Only a twofold classification is needed for sense.

True in two ways, false in two ways.

call states of affairs of the former kind those in which the statement was false and the others those in which the statement was true. For our conditional statements, the distinction would be between those states of affairs in which the statement was said to be false and those in which we said that it would either be true or else neither true nor false. For singular statements, the distinction would be between those states of affairs in which we said that the statement would either be false or else neither true nor false, and those in which it was true. To grasp the sense or use of these forms of statement, the twofold classification is quite sufficient; the threefold classification with which we started is entirely beside the point. Thus, on *one* way of using the words 'true' and 'false', we should, instead of distinguishing between the conditional statement's being true and its being neither true nor false, have distinguished between two different ways in which it could be true; and instead of distinguishing between the singular statement's being false and its being neither true nor false, we should have distinguished between two different ways in which it could be false.

Finer distinctions matter only inside compounds.

This gives us a hint at a way of explaining the rôle played by truth and falsity in determining the sense of a statement. We have not yet seen what point there may be in distinguishing between different ways in which a statement may be true or between different ways in which it may be false, or, as we might say, between degrees of truth and falsity. The point of such distinctions does not lie in anything to do with the sense of the statement itself, but has to do with the way in which it enters into complex statements.

The 'not' that flips the consequent, and the failed table.

Let us imagine that in the language of which the conditional statements we considered form part there exists a sign of negation, i.e., a word which, placed in front of a statement, forms another statement; I call it a sign of negation because in most cases it forms a statement which we should regard as being used as the contradictory of the original statement. Let us suppose, however, that when placed in front of a conditional statement ⌜If P, then Q⌝, it forms a statement which is used in the same way as the statement ⌜If P, then not Q⌝. Then if we describe the use of the conditions by reference to a twofold classification only, i.e., in the same way as we describe a material conditional, we shall be unable to give a truth-functional account of the behaviour of their sign 'not'. That is, we should have the tables:

P	Q	⌜If P, then Q⌝	⌜Not: if P, then Q⌝
T	T	T	F
T	F	F	T
F	T	T	T
F	F	T	T

A third value X gives their 'not' a table that works.

in which the truth-value of ⌜Not: if P, then Q⌝ is not determined by the truth-value of ⌜If P, then Q⌝. If, on the other hand, we revert to our original threefold classification, marking the case in which we said that no statement true or false had been made by 'X', then we have the tables:

P	Q	⌜If P, then Q⌝	⌜Not: if P, then Q⌝
T	T	T	F
T	F	F	T
F	T	X	X
F	F	X	X

which can be quite satisfactorily accounted for by giving the table for 'not':

R	⌜Not R⌝
T	F
X	X
F	T

(I have assumed that the statements P and Q take only the values T and F.) It now becomes quite natural to think of 'T' as representing 'true', 'F' 'false' and 'X' 'neither true nor false'. Then we can say that their symbol 'not' really is a sign of negation, since ⌜Not P⌝ is true when and only when P is false and false when and only when P is true. We must not forget, however,

That is the only justification the third value has.

that the justification for distinguishing between the cases in which a conditional was said to have the value T and the cases in which it was said to have the value X was simply the possibility, created by this distinction, of treating 'not' truth-functionally. In the same way if we have in a language an expression which normally functions as a sign of negation, but the effect of prefacing a singular statement with this expression is to produce a statement whose utterance still commits the speaker to there being an object for which the singular term stands, it is very natural to distinguish between two kinds of falsity a singular statement may have: that when the singular term has a reference, but the predicate does not apply to it, and that when the singular term lacks a reference. Let us represent the case in which the singular term has no reference by the symbol 'Y', and let us suppose S to be a singular statement. Then we have the table:

Two kinds of falsity for the singular statement: Y.

S	⌜Not S⌝
T	F
Y	Y
F	T

Here again it is natural to think of '*T*' as representing 'true', '*F*' 'false' and '*Y*' 'neither true nor false'.

Designated and undesignated values, the logicians' term.

There is no necessity to use the words 'true' and 'false' as suggested above, so that we have to interpret *X* as a kind of truth and *Y* as a kind of falsity. Logicians who study many-valued logics have a term which can be employed here: they would say that *T* and *X* are 'designated' truth-values and *F* and *Y* 'undesignated' ones. (In a many-valued logic those formulas are considered valid which have a designated value for every assignment of values to their sentence-letters.) The points to observe are just these.

Four morals: finer values are for compounds only.

(i) The sense of a sentence is determined wholly by knowing the case in which it has a designated value and the cases in which it has an undesignated one. (ii) Finer distinctions between different designated values or different undesignated ones, however naturally they come to us, are justified only if they are needed in order to give a truth-functional account of the formation of complex statements by means of operators. (iii) In *most* philosophical discussions of truth and falsity, what we really have in mind is the distinction between a designated and an undesignated value, and hence choosing the names 'truth' and 'falsity' for particular designated and undesignated values respectively will only obscure the issue. (iv) Saying that in certain circumstances a statement is neither true nor false does not determine whether the statement is in that case to count as having an undesignated or a designated value, i.e., whether someone who asserts the statement is or is not taken as excluding the possibility that that case obtains.

What correspondence got right, and realism defined.

Baffled by the attempt to describe in general the relation between language and reality, we have nowadays abandoned the correspondence theory of truth, and justify our doing so on the score that it was an attempt to state a *criterion* of truth in the sense in which this cannot be done. Nevertheless, the correspondence theory expresses one important feature of the concept of truth which is not expressed by the law 'It is true that *p* if and only if *p*' and which we have so far left quite out of account: that a statement is true only if there is something in the world *in virtue of which* it is true. Although we no longer accept the correspondence theory, we remain realists *au fond*; we retain in our thinking a fundamentally realist conception of truth. Realism consists in the belief that for any statement there must be something in virtue of which either it or its negation is true: it is only on the basis of this belief that we can justify the idea that truth and falsity play an essential rôle in the notion of the meaning of a statement, that the general form of an explanation of meaning is a statement of the truth-conditions.

To see the importance of this feature of the concept of truth, let us envisage a dispute over the logical validity of the statement 'Either Jones was

brave or he was not'. *A* imagines Jones to be a man, now dead, who never encountered danger in his life. *B* retorts that it could still be true that Jones was brave, namely if it is true that if Jones *had* encountered danger, he would have acted bravely. *A* agrees with this, but still maintains that it does not need to be the case that either 'Jones was brave' = 'If Jones had encountered danger, he would have acted bravely' nor 'Jones was not brave' = 'If Jones had encountered danger, he would not have acted bravely' is true. For, he argues, it might be the case that however many facts we knew of the kind which we should normally regard as grounds for asserting such counterfactual conditionals, we should still know nothing which would be a ground for asserting either. It is clear that *B* cannot agree that this is a possibility and yet continue to insist that all the same either 'Jones was brave' or 'Jones was not brave' is true; for he would then be committed to holding that a statement may be true even though there is nothing whatever such that, if we knew of it, we should count it as evidence or as a ground for the truth of the statement, and this is absurd. (It may be objected that there are assertions for which it would be out of place to ask one who made them for his evidence or grounds; but for *such* assertions the speaker must always either be in a position to make or in a position to deny them.) If *B* still wishes to maintain the necessity of 'Either Jones was brave or he was not', he will have to hold either that there must be some fact of the sort to which we usually appeal in discussing counterfactuals which, if we knew it, would decide us in favour either of the one counterfactual or of the other; or else that there is some fact of an extraordinary kind, perhaps known only to God. In the latter case he imagines a kind of spiritual mechanism—Jones's character—which determines how he acts in each situation that arises; his acting in such-and-such a way reveals to us the state of this spiritual mechanism, which was however already in place before its observable effects were displayed in his behaviour. *B* would then argue thus: If Jones *had* encountered danger, he would either have acted bravely or have acted like a coward. Suppose he had acted bravely. This would then have shown us that he was brave; but he would *already* have been brave before his courage was revealed by his behaviour. That is, either his character included the quality of courage or it did not, and his character determines his behaviour. We know his character only indirectly, through its effects on his behaviour; but each character-trait must be *there* within him independently of whether it reveals itself to us or not.

A and B on Jones: the counterfactual needn't be true either way.

Why B can't grant it: truth beyond all evidence is absurd.

The extraordinary fact: character as spiritual mechanism.

Anyone of a sufficient degree of sophistication will reject *B*'s belief in a spiritual mechanism; either he will be a materialist and substitute for it an equally blind belief in a physiological mechanism, or he will accept *A*'s

A's ground, and the small retreat from realism.

conclusion that 'Either Jones was brave or he was not' is not logically necessary. His ground for rejecting *B*'s argument is that if such a statement as 'Jones was brave' is true, it must be true in virtue of the the sort of fact we have been taught to regard as justifying us in asserting it. It cannot be true in virtue of a fact of some quite different sort of which we can have no direct knowledge, for otherwise the statement 'Jones was brave' would not have the meaning that *we* have given it. In accepting *A*'s position he makes a small retreat from realism; he abandons a realist view of character.

The test for a realist account of a kind of statement.

In order, then, to decide whether a realist account of truth can be given for statements of some particular kind, we have to ask whether for such a statement *P* it must be the case that if we knew sufficiently many facts of the kind we normally treat as justifying us in asserting *P*, we should be in a position either to assert *P* or to assert ⌜Not *P*⌝: if so, then it can truly be said that there must either be something in virtue of which *P* is true or something in virtue of which it is false. It is easy to overlook the force of the phrase 'sufficiently many'. Consider the statement 'A city will never be built on this spot'. Even if we have an oracle which can answer every question of the kind, 'Will there be a city here in 1990?', 'In 2100?', etc., we might never be in a position either to declare the statement true or to declare it false. Someone may say: That is only because you are assuming the knowledge of only finitely many answers of the oracle; but if you knew the oracle's answers to *all* these questions, you would be able to decide the truth-value of the statement. But what would it mean to know infinitely many facts? It could mean that the oracle gave a direct answer 'No' to the question, 'Will a city ever be built here?': but to assume this is just like *B*'s assumption of the existence of a hidden spiritual mechanism. It might mean that we had an argument to show the falsity of ⌜A city will be built here in the year *N*⌝ irrespective of the value of *N*, e.g., if 'here' is the North Pole: but no one would suggest that it must be the case that either the oracle will give an affirmative answer to some question of the form 'Will there be a city here in the year . . . ?' or we can find a general argument for a negative answer. Finally, it could mean that we were *able* to answer every question of the form, 'Will there be a city here in the year . . . ?': but having infinite knowledge in *this* sense will place us in no better position than when we had the oracle.

'Sufficiently many': the city never built, and the oracle.

Bivalence only for the decidable, and the limit isn't trivial.

We thus arrive at the following position. We are entitled to say that a statement *P* must be either true or false, that there must be something in virtue of which either it is true or it is false, only when *P* is a statement of such a kind that we could in a finite time bring ourselves into a position in which we were justified either in asserting or in denying *P*; that is, when *P*

is an effectively decidable statement. This limitation is not trivial: there is an immense range of statements which, like 'Jones was brave', are concealed conditionals, or which, like 'A city will never be built here', contain—explicitly or implicitly—an unlimited generality, and which therefore fail the test.

The intuitionist transfer: the odd perfect number.

What I have done here is to transfer to ordinary statements what the intuitionists say about mathematical statements. The sense of, e.g., the existential quantifier is determined by considering what sort of fact makes an existential statement true, and this means: the sort of fact which we have been taught to regard as justifying us in asserting an existential statement. What would make the statement that there exists an odd perfect number true would be some particular number's being both odd and perfect; hence the assertion of the existential statement must be taken as a claim to be able to assert some one of the singular statements. We are thus justified in asserting that there is a number with a certain property only if we have a method for finding a particular number with that property. Likewise, the sense of a universal statement is given by the sort of consideration we regard as justifying us in asserting it: namely we can assert that every number has a certain property if we have a general method for showing, for any arbitrary number, that it has that property. Now what if someone insists that either the statement 'There is an odd perfect number' is true, or else every perfect number is even? He is justified if he knows of a procedure which will lead him in a finite time either to the determination of a particular odd perfect number or to a general proof that a number assumed to be perfect is even. But if he knows of no such procedure, then he is trying to attach to the statement 'Every perfect number is even' a meaning which lies *beyond* that provided by the training we are given in the use of universal statements; he wants to say, as *B* said of 'Jones was brave', that its truth may lie in a region directly accessible only to God, which human beings can never survey.

Insisting anyway wants a truth only God can reach, like B's.

The operators explained by when we may assert them.

We learn the sense of the logical operators by being trained to *use* statements containing them, i.e., to assert such statements under certain conditions. Thus we learn to assert ⌜*P* and *Q*⌝ when we can assert *P* and can assert *Q*; to assert ⌜*P* or *Q*⌝ when we can assert *P* or can assert *Q*; to assert ⌜For some *n*, *F*(*n*)⌝ when we can assert ⌜*F*(0)⌝ or can assert ⌜*F*(1)⌝ or . . . We learn to assert ⌜For every *n*, *F*(*n*)⌝ when we can assert ⌜*F*(0)⌝ and ⌜*F*(1)⌝ and . . . ; and to say that we can assert all of these means that we have a general method for establishing ⌜*F*(*x*)⌝ irrespective of the value of *x*. Here we have abandoned altogether the attempt to explain the meaning of a statement by laying down its truth-conditions. *We no longer explain the*

Excluded middle falls, but not for a third value.

sense of a statement by stipulating its truth-value in terms of the truth-values of its constituents, but by stipulating when it may be asserted in terms of the conditions under which its constituents may be asserted. The justification for this change is that this is how we in fact learn to use these statements: furthermore, the notions of truth and falsity cannot be satisfactorily explained so as to form a basis for an account of meaning once we leave the realm of effectively decidable statements. One result of this shift in our account of meaning is that, unless we are dealing only with effectively decidable statements, certain formulas which appeared in the two-valued logic to be logical laws no longer rank as such, in particular the law of excluded middle: this is rejected, not on the ground that there is a middle truth-value, but because meaning, and hence validity, is no longer to be explained in terms of truth-values.

Intuitionist anti-realism, and whether it generalises.

Intuitionists speak of mathematics in a highly anti-realist (anti-platonist) way: for them it is *we* who construct mathematics; it is not already *there* waiting for us to discover. An extreme form of such constructivism is found in Wittgenstein's *Remarks on the Foundations of Mathematics*. This makes it appear as though the intuitionist rejection of an account of the meaning of mathematical statements in terms of truth and falsity could not be generalised for other regions of discourse, since even if there is no independent mathematical reality answering to our mathematical statements, there is an independent reality answering to statements of other kinds. On the other hand the exposition of intuitionism I have just given was not based on a rejection of the Fregean notion of a mathematical reality waiting to be discovered, but only on considerations about meaning. Now certainly someone who accepts the intuitionist standpoint in mathematics will not be inclined to adopt the platonist picture. Must he then go to the other extreme, and have the picture of our creating mathematics as we go along? To adopt this picture involves thinking with Wittgenstein that we are *free* in mathematics at every point; no step we take has been forced on us by a necessity external to us, but has been freely chosen. This picture is not the only alternative. If we think that mathematical results are in some sense imposed on us from without, we could have instead the picture of a mathematical reality not already in existence but as it were coming into being as we probe. Our investigations bring into existence what was not there before, but what they bring into existence is not of our own making.

A third picture: reality coming into being as we probe.

Whether this picture is right or wrong for mathematics, it is available for other regions of reality as an alternative to the realist conception of the world. This shows how it is possible to hold that the intuitionist substitution of an account of the *use* of a statement for an account of its truth-

No subjective idealism, but truth dethroned from meaning.

conditions as the general form of explanation of meaning should be applied to all realms of discourse without thinking that we create the world; we can abandon realism without falling into subjective idealism. This substitution does not of course involve dropping the words 'true' and 'false', since for most ordinary contexts the account of these words embodied in the laws 'It is true that p if and only if p' and 'It is false that p if and only if not p' is quite sufficient: but it means facing the consequences of admitting that this is the *whole* explanation of the sense of these words, and this involves dethroning truth and falsity from their central place in philosophy and in particular in the theory of meaning. Of course the doctrine that meaning is to be explained in terms of use is the cardinal doctrine of the later Wittgenstein; but I do not think the point of this doctrine has so far been generally understood.

POSTSCRIPT (1972)

The article still seems to me substantially correct, but I should like to add the following emendations and glosses:

Second thoughts on Frege: the fault is truth-values as objects.

(1) The remark about Frege at the end of the third paragraph is quite misleading. The really questionable part of Frege's doctrine is not that sentences have references, nor that these references are truth-values, but that truth-values are objects. If truth-values are not objects, then the relation between a sentence and its truth-value is only analogous to, not identical with, the relation between a name and its bearer, just as is the relation between a predicate and the concept for which it stands. It is true that the notion of the incompleteness of a function is more readily intelligible than, and can be used to illuminate, that of the incompleteness of a concept or relation; but that does not require that concepts and relations be taken as actually special cases of functions, rather than merely analogues of them.

The game comparison kept, and a misreading headed off.

(2) The comparison between the notion of truth and that of winning a game still seems to me a good one. The text as it stands might, however, give rise to a certain misunderstanding, though this would be a misinterpretation of what I intended at the time of writing. Suppose that we have, for each of a large range of games, a characterisation of the conditions under which one player or side is said to have won the game. Now we ask whether it is sufficient, in order to convey to someone previously quite unacquainted with it what the notion of winning is, to add to this characterisation the

Two senses of 'game': when the bare intention suffices.

mere observation that the participants in a game play with the intention of winning. The correct answer is, I think, that there is a wider and a narrower sense to the expression 'to play a game', and that, in the context of the wider sense, this simple supplementation is sufficient, while, in that of the narrower sense, it is not. Someone who came from a culture from which the practice of playing games was entirely lacking could hardly understand the bare remark that the intention of the players is to win, since he would want to know what further purpose this intention subserved: we should therefore have to explain to him the character of games as a social institution, which is a complicated matter. (A child who thinks that there is no point in playing a game unless he wins may be said not yet to have grasped fully the concept of a game; yet to say that the intention of the players is to win does not allow for there being any such point.) All this relates, however, only to the narrower, i.e. strict, sense of 'game'. Suppose that two countries, engaged in a political dispute, were to agree to resolve it, on precise terms, not by war, but by a contest between their respective chess champions. The two champions could then naturally be said to play a game of chess; but, since the usual social surroundings of a game would be absent, in the narrower sense it would also be correct to say that they weren't *playing* and that it was not a mere game that they were engaged in. When we are thinking of games in that wider sense of 'game' which would allow such a contest to be called a 'game', *all* that can be said about winning, beyond the characterisation of what counts as winning, is that it is what each player has the intention to do.

The chess champions settling a dispute.

Truth is far more complicated: the practice of assertion.

In this respect, truth is an enormously more complicated notion than that of winning. The misunderstanding to which I referred would be the idea that all that had to be added to a characterisation of the conditions under which a sentence of each of a number of languages was true, in order to explain the notion of truth, would be the flat observation that, in uttering a sentence assertorically, a speaker did so with the intention of uttering a true sentence. We have seen that, looked at in one way, the corresponding thesis for the notion of winning may be held to be correct; but there is no way in which the thesis would be held correct for the notion of truth, and it was not my aim to suggest this. What has to be added to a truth-definition for the sentences of a language, if the notion of truth is to be explained, is a description of the linguistic activity of making assertions; and this is a task of enormous complexity. What we can say is that any such account of what assertion is must introduce a distinction between correct and incorrect assertions, and that it is in terms of that distinction that the notion of truth has first to be explained.

(3) It should be noted that stipulating that every instance of 'It is true that p if and only if p' is to hold will not succeed in determining, even for one who already understands the language relative to which the stipulation is made, the application of the predicate 'true' to sentences of that language unless the language is such as to confer a definite meaning on every conditional formed by taking any arbitrary sentence of the language as antecedent, since, obviously, in order to apply the stipulation to some sentence P, we must be able to understand the conditional ⌜If P, then it is true that P⌝. Hence, when P is a sentence such that we attach no determinate sense to conditionals having P as antecedent, the stipulation will not succeed in telling us when P is true. For instance, we have virtually no use in English for conditionals whose antecedents are themselves conditionals: so we can obtain no help towards resolving the disputed question when an indicative conditional ⌜If Q, then R⌝ of English should be considered true by appeal to the principle that we should accept ⌜If, if Q, then R, then it is true that if Q, then R⌝.

(4) In the text it is argued (quite correctly, I still believe) that uses of 'true' and 'false' which involve that, in recognisable circumstances, a sentence (assumed to be neither ambiguous nor vague) will be neither true nor false relate only to the behaviour of that sentence as a constituent of compound sentences, and, in particular, to the sense of its negation. There is, however, a distinct but prior consideration of the same kind. We are accustomed, in general, to distinguish between an assertion's being *correct* and the speaker's having a *warrant* for making it: the text assumes implicitly that this distinction has already been drawn. But we need to ask on what basis we make any such distinction. For instance, a sentence in the future tense used to express the intention of the speaker contrasts with the corresponding declaration of intention (in the form 'I intend to . . .'). The same contrast exists between the genuine future tense used to make a prediction, and the future tense used to express present tendencies. (This latter occurs, e.g., in an announcement of the form 'The wedding announced between A and B will not now take place'. Such an announcement cancels, but does not falsify, the earlier announcement, and is not itself falsified if the couple later make it up and get married after all; if this were not so, the 'now' would be superfluous.) With both pairs, the conditions under which an assertion of either form would be warranted coincide: but the conditions for their truth differ. In both pairs, the first member may actually be false, although the assertion was warranted; or, conversely, be true, although it was not warranted. I should maintain that this distinction also derives its

significance from the behaviour of sentences as constituents of more complex ones. Thus, for each pair, a conditional which has one member of the pair as antecedent has a quite different sense from one having the other member as antecedent; or, again, the sense of the past future ('I was going to . . .') varies according as we take the future ingredient as constituting the plain future tense or the future tense which expresses the existence of a present intention or tendency. If we are concerned with an assertoric sentence which cannot appear as a constituent of a more complex sentence, we have no need of a distinction between cases in which an assertion made by means of it would be warranted and those in which it would be correct. As already noted, the conditionals of natural language are, or approximate to being, just such sentences: it is for just that reason that philosophers have found it so hard to say what should be reckoned as part of the truth-condition for such a conditional and what as merely part of the grounds for asserting it.

(5) Some people have been surprised by my characterising the content of an assertion in terms of what it excludes, i.e. what would show it to be wrong, rather than in terms of what establishes it as correct. I find their surprise surprising in its turn, since the example shows so clearly why the former approach is better. It is obvious that one who makes a conditional assertion does not wish to rule out the antecedent's being false, and that one who asserts a singular statement does not wish to allow for the term's lacking a reference: but, if we tried to contrast the two cases in terms of what established the assertions as correct, we should quickly find ourselves involved in disputes about when a conditional statement is to be said to be true. Of course, we can talk instead about what is required to be the case by an assertion; but this notion relates, once again, to how we recognise the assertion as incorrect. The reason is similar to what is said in the text about obedience and disobedience: our notions of right and wrong, for assertions as for actions, are asymmetrical, and it is the apparently negative notion which is primary. There is a well-defined consequence of an assertion's proving incorrect, namely that the speaker must withdraw it, just as there is a well-defined consequence of disobedience; there is not in the same way a well-defined consequence of an assertion's proving correct, or of obedience. When we talk of a previous assertion having been right, we are usually primarily concerned with the speaker's having been justified in making it, just as, when obedience is in question, it is always a matter of deliberate compliance.

(6) 'The making of a statement has, as it were, only one kind of consequence.' This remark is essentially correct, but the 'as it were' is important: assertions do not have the same kind of relation to determinate consequences that commands and bets do.

The final section revised: superficial and deep grounds.

(7) It is the final section of the article which stands most in need of revision. In this section I am concerned with deep grounds for rejecting the law of excluded middle as applied to certain statements, whereas, in the earlier part, I had been concerned with superficial grounds for rejecting it. The superficial case is that in which it is held that, in certain *recognisable* circumstances, a statement will be neither true nor false. The upshot of the preceding discussion was that, whenever some recognisable circumstance is regarded as settling the truth-value of the statement, that circumstance must determinately confer on the statement either a designated or an undesignated value, and that that is all that is relevant to the content of an assertion of the statement on its own; the point of describing the sentence as neither true nor false whenever some recognisable circumstance obtains can, therefore, relate only to the content of more complex sentences in which the given sentence occurs as a constituent.

The deep case: both sides accept the in-virtue-of principle.

The deep case is that in which it is agreed, for all recognisable circumstances, whether they determine the statement as true or as false (and there are none in which it is held to be neither); but no effective method exists which will in all cases bring about circumstances of one or the other kind. The text describes the principle that a statement can be true only if there is something in virtue of which it is true as a realist one: but the point of the dispute over 'Jones was brave' is that both parties accept the principle. The anti-realist uses it to infer that the statement is not necessarily either true or false; the realist uses it to infer that that which makes it true or false cannot be identified with that by which we recognise it as true or as false, when we are able to do so. For the realist, our understanding of the statement consists in our grasp of its truth-conditions, which determinately either obtain or fail to obtain, but which cannot be recognised by us in all cases as obtaining whenever they do; for the anti-realist, our understanding consists in knowing what recognisable circumstances determine it as true or as false.

The 'absurd' of 1959 withdrawn for a weaker principle.

The text stigmatises as absurd the proposition that a statement might be true even though there was nothing such that, if we knew of it, we should count as evidence of its truth. This was intended to mean 'nothing of the sort which we ordinarily use as evidence for the truth or falsity of such a statement'. The claim in the text is a very bold one, and should be rejected by a realist, who might, and I think ought to, agree to the following weaker principle: that a statement cannot be true unless it is in principle capable of

'Capable of being known': by us, or by a greater being.

being known to be true. This principle is closely connected with the first one: for that in virtue of which a statement is true is that by which the statement might be known to be true. The fundamental difference between the anti-realist and the realist lies in this: that, in the second principle, the anti-realist interprets 'capable of being known' to mean 'capable of being known *by us*', whereas the realist interprets it to mean 'capable of being known by some hypothetical being whose intellectual capacities and powers of observation may exceed our own'. The realist holds that we give sense to those sentences of our language which are not effectively decidable by appealing tacitly to means of determining their truth-values which we do not ourselves possess, but which we can conceive of by analogy with those which we do. The anti-realist holds that such a conception is quite spurious, an illusion of meaning, and that the only meaning we can confer on our sentences must relate to those means of determining their truth-values which we actually possess. Hence, unless we have a means which would in principle decide the truth-value of a given statement, we do not have for it a notion of truth and falsity which would entitle us to say that it must be either true or false.

The realist's analogy, the anti-realist's illusion of meaning.

The 1959 text was frankly anti-realist, Dummett now says.

The text of the article espouses a frankly anti-realist position; it says, in effect, that a realist interpretation is possible only for those statements which are in principle effectively decidable (i.e. those for which there is no serious issue between the realist and anti-realist). I am no longer so unsympathetic to realism: the realist has a lot more to say for himself than is acknowledged in the article. The dispute is still a long way from resolution. On the one hand, it is unclear whether the realist's defence of his position can be made convincing; on the other, it is unclear whether the anti-realist's position can be made coherent. I remain convinced, however, that the issue between realism and anti-realism, construed roughly along the present lines, is one of the most fundamental of all the problems of philosophy; convinced also that very few people are thinking about it in the terms which seem to me to be the right ones.

Part II

Rule-following and interpretation