DAVID LEWIS
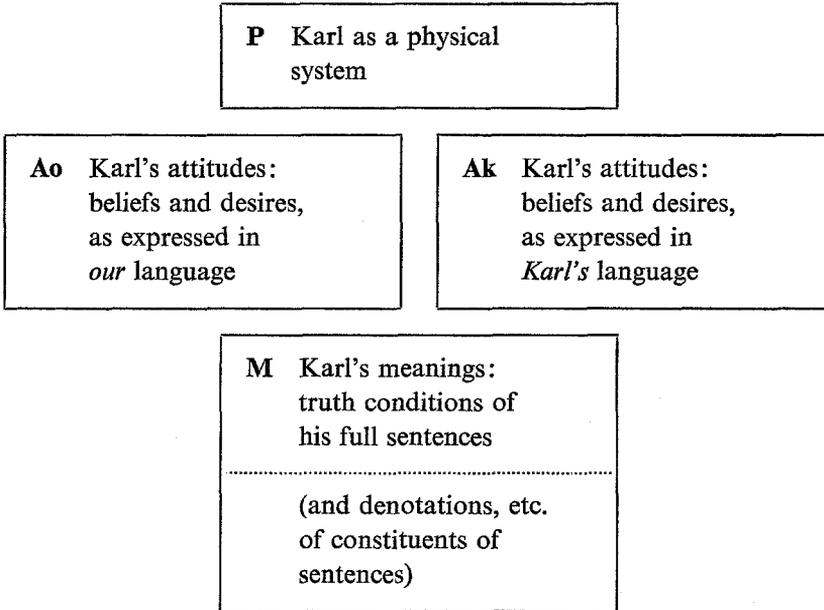
RADICAL INTERPRETATION

Imagine that we have undertaken the task of coming to know Karl as a person. We would like to know what he believes, what he desires, what he means, and anything else about him that can be explained in terms of these things. We seek a two-fold interpretation: of Karl's language, and of Karl himself. And we want to know his beliefs and desires in two different ways. We want to know their content as Karl could express it in his own language, and also as we could express it in our language. (For instance, we want to know whether 'ownsnay isyay itewhay', is a sentence that Karl could use to express something or other that he believes; and also whether Karl believes that, as we would put it, snow is white. Of course, Karl's language just might turn out to be the same as ours – that is part of what we want to find out – but the two questions are independent even so.)

Imagine also that we must start from scratch. At the outset we know nothing about Karl's beliefs, desires, and meanings. Whatever we may know about persons in general, our knowledge of Karl in particular is limited to our knowledge of him as a physical system. But at least we have plenty of that knowledge – in fact, we have all that we could possibly use. Now, how can we get from that knowledge to the knowledge we want?

I can diagram the problem of radical interpretation as follows. **Given P, the facts about Karl as a physical system, solve for the rest.**

P, our ultimate data base, gives us the whole truth about Karl as a physical system. It tells us how Karl moves, what forces he exerts on his surroundings, what light or sound or chemical substances he absorbs or emits. It tells us the same things about all of Karl's material parts, great or small, permanent or temporary. It tells us all the masses and charges of the particles that compose him, and all the magnitudes and directions of the fields and potentials and radiation that pervade him. It tells us not only his present physical state but also his physical history; and not only the actual particular physical facts but also the nomic or counterfactual or causal dependences among them. It tells us higher order facts, if need be:

```
┌─────────────────────────────┐
│  P   Karl as a physical     │
│      system                 │
└─────────────────────────────┘
```

```
┌──────────────────────────┐   ┌──────────────────────────┐
│  Ao  Karl's attitudes:   │   │  Ak  Karl's attitudes:   │
│      beliefs and desires,│   │      beliefs and desires,│
│      as expressed in     │   │      as expressed in     │
│      our language        │   │      Karl's language     │
└──────────────────────────┘   └──────────────────────────┘
```

```
┌─────────────────────────────┐
│  M   Karl's meanings:       │
│      truth conditions of    │
│      his full sentences     │
│  ·························   │
│      (and denotations, etc. │
│      of constituents of     │
│      sentences)             │
└─────────────────────────────┘
```

as that there exist some or other states of Karl, of unspecified character, that realize such-and-such patterns of causal relations to one another and to such-and-such specified physical states. And in case the material parts of Karl interact with any psionic fields, astral bodies, entelechies, or what-not, then **P** must tell us the physics and the physical states of those things as well.

Both **Ao** and **Ak** are to be specifications of Karl's propositional attitudes – in particular, of Karl's system of beliefs and desires. (I limit my attention to these attitudes in the hope that all others will prove to be analyzable as patterns of belief and desire, actual or potential; but if not, then whatever attitudes resist such analysis also should be included in **Ao** and **Ak**.) To specify the propositional content of Karl's beliefs and desires, we may specify how to express the believed or desired proposition, in some language, by a sentence in a context. **Ao** specifies Karl's beliefs and desires as expressed in *our* language; **Ak** specifies them as expressed in *Karl's* language; until we find out what the sentences of Karl's language mean, the two sorts of information are different. We take Karl's beliefs and desires to admit of degree, with the zero and unit of desire fixed arbitrari-

ly. Also we allow them to vary with time. Thus **Ao** and **Ak** will consist of ascriptions of the form: *Karl* $\begin{Bmatrix} believes \\ desires \end{Bmatrix}$ *,to degree d, at time t, the proposition expressed, in context c, by the sentence* '- - - -' *of* $\begin{Bmatrix} our \\ Karl's \end{Bmatrix}$ *language.* Take this, if you wish, as a single relation between Karl, a degree (*i.e.* a number), a time, a context, and a sentence of our language or Karl's. I would prefer to take the seeming reference to a proposition at face value: Karl is related (to a degree, at a time) to a proposition, and that proposition in turn is related (in a context) to a sentence. For if our interest is in the philosophy of mind and of language, then the pursuit of ontological parsimony seems to me an unnecessary distraction. Propositions may be dispensable or they may not, but at least they seem harmless. True, Quine has rightly warned us againt the question-begging myth of the museum ([16]: 27–29); but the myth begs questions not because it countenances propositions or other abstract entities, but rather because it uncritically takes for granted that our mental and semantic relations to these entities are determinate and need no explanation.

M, the third component of our desired interpretation of Karl, is to be a specification, in our language, of the meanings of expressions of Karl's language. Primarily, M specifies the truth conditions of full sentences of Karl's language (perhaps relative to contexts of utterance). I leave it open, here, just what a 'truth condition' is; in particular, whether it should specify only the actual truth value of the sentence, or whether it should also specify what the truth value would be in various counterfactual situations not too remote from actuality, or whether it should specify the truth value at all possible worlds whatever. Secondarily, M specifies a way of parsing the sentences of Karl's language, the denotation or sense or comprehension or what-not of the constituents from which sentences may be compounded, and the way that the denotation (or whatever) of a compound depends on that of its constituents. In short, it specifies the syntactic and semantic rules of a grammar capable of generating Karl's sentences plus the truth conditions thereof.

It should be obvious by now that my problem of radical interpretation is not any real-life task of finding out about Karl's beliefs, desires, and meanings. I am not really asking how *we* could determine these facts. Rather: how do *the facts* determine these facts? By what constraints, and

to what extent, does the totality of physical facts about Karl determine what he believes, desires, and means? To speak of a mighty knower, who uses his knowledge of these constraints to advance from omniscience about the physical facts **P** to omniscience about the other facts determined thereby, is a way of dramatizing our problem – safe enough, so long as we can take it or leave it alone. The real-life knower has all the problems of our fictitious knower, and more besides: he does not have all of **P** to draw on, and he may be limited in endurance, intelligence, or memory. (On the other hand, he does not aspire to omniscience about **Ao, Ak,** and **M.**) But these further obstacles to his investigation are irrelevant to our real topic.

If I ask *how* **P** determines all the rest, my question requires the presupposition that **P** *does* determine all the rest. Or, at least, that **P** determines all the rest to the extent that anything does – that where determination by **P** leaves off, there indeterminacy begins. In other words, I am presupposing that there cannot possibly be two Karls exactly alike with respect to **P** but differing somehow with respect to **Ao, Ak,** or **M.** (It does not matter that there might be two equally correct ways to resolve some mental or semantic indeterminacy, so long as *both* ways are available for *both* Karls. The two Karls still do not differ.) This basic presupposition of our enterprise is a sort of minimal materialism. 'Minimal', first, because we have allowed **P** to include the physics of astral bodies or what-not if need be; and second, because even if all the mental and semantic facts about Karl are determined by the physical facts, it does not follow that they can be stated in the language of physics. (I think they can be, if the language of physics is taken fairly broadly, but that is another story.)

What are the constraints by which the problem of radical interpretation is to be solved? Roughly speaking, they are the fundamental principles of our general theory of persons. They tell as how beliefs and desires and meanings are normally related to one another, to behavioral output, and to sensory input.

The general theory of persons serves as a schema for particular theories of particular persons. A particular theory of Karl, for instance, may be constructed by ascribing particular beliefs, desires, and meanings to him. That is: by filling in **Ao, Ak,** and **M.** But not just any filling-in will do. The relations of **Ao, Ak,** and **M** to one another and to **P** must conform – for the most part, more or less – to the principles of the general theory. Else the particular theory of Karl is inconsistent. In this way, the general

theory provides the constraining power to make radical interpretation possible.

Our general theory of persons, like a term-introducing scientific theory, has a mixed character. On the one hand, it implicitly defines its key theoretical concepts: in particular, the concepts of belief, desire, and meaning. On the other other, it uses these concepts to make an empirical claim about human beings – a claim so well confirmed that we take it quite for granted. If we disentangle the definitional content and the empirical content, we have something roughly like this. Definitional content: something may count as a person's system of beliefs, desires, and meanings if and only if it is a system that more or less conforms to the principles of the theory. Empirical content: for any human being (with certain exceptions) there will exist a system of beliefs, desires, and meanings correctly so-called – that is, one that conforms to the principles. That is: almost whatever P may be, within the limits of human possibility, the problem of radical interpretation should have a solution.
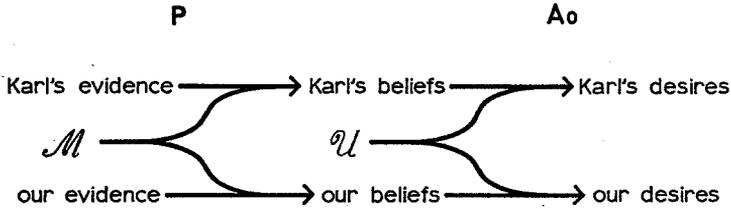
The concepts of belief, desire, and meaning are common property. The theory that implicitly defines them had better be common property too. It must amount to nothing more than a mass of platitudes of common sense, though these may be reorganized in perspicuous and unfamiliar ways. Esoteric scientific findings that go beyond common sense must be kept out, on pain of changing the subject.

I have said, rather loosely, that the fundamental principles of our common-sense theory of persons implicitly define such concepts as belief, desire, and meaning. Actually, I would like to claim something stronger: that the implicit definitions can be made explicit (in the way I suggested in [8], [10], and [13]), and that the explicit definitions so obtained would be analytic. If so, then our constraining principles would themselves have a status akin to analyticity: Karl might have no beliefs, desires, or meanings at all, but it is analytic that if he does have them then they more or less conform the constraining principles by which the concepts of belief, desire, and meaning are defined. But it would not be appropriate to press this claim here. For the question of whether analyticity is a legitimate notion is part of the broader question of the extent to which semantic facts are determinate, and that question is part of the very problem of radical interpretation that we are right now considering. I do not think I need to claim analyticity; it is enough that the constraining

principles should be very firmly built into our common system of belief.

The principles that I would like to put forward as constraints on radical interpretation are the following six. They are not independent, and some may be entirely redundant given the others. On the other hand, I may well have overlooked important ones that should have been listed.

The *Principle of Charity* constrains **Ao**, or the relation between **Ao** and **P**: Karl should be represented as believing what he ought to believe, and desiring what he ought to desire. And what is that? In our opinion, he ought to believe what *we* believe, or perhaps what we would have believed in his place; and he ought to desire what we desire, or perhaps what we would have desired in his place. (But that's only our opinion! Yes. Better we should go by an opinion we *don't* hold?) A crude version of the Principle of Charity might just require that, so far as other constraints allow it, the beliefs and desires ascribed to Karl by **Ao** should be the same as our own beliefs and desires. "We will try for a theory that finds him consistent, a believer of truths, and a lover of the good (all by our own lights, it goes without saying)" as Davidson puts it ([3]: 97). But it would be more charitable to make allowances for the likelihood that Karl's circumstances – his life history of evidence and training, recounted in physical terms in our data base **P** – may have led him understandably into error. We should at least forbear from ascribing to Karl those of our beliefs and desires which, according to **P** and our notions of reason, he has been given no reason to share. We should even ascribe to him those errors which we think we would have made, or should have made, if our evidence and training had been like his. Perhaps an improved Principle of Charity would require Karl's beliefs and ours to be related as follows: there must exist some common inductive method $\mathscr{M}$ which would lead to approximately our present systems of belief if given our life histories of evidence, and which would likewise lead to approximately the present system of beliefs ascribed to Karl by **Ao** if given Karl's life history of evidence according to **P**. As for desires: there must exist some common underlying system of basic intrinsic values $\mathscr{U}$ which would yield approximately our systems of desires if given our systems of beliefs, and which would likewise yield approximately the system of desires ascribed to Karl by **Ao** if given the system of beliefs ascribed to Karl by **Ao**, Diagrammatically: there must exist $\mathscr{M}$ and $\mathscr{U}$ such that:

P   Ao

Karl's evidence ⟶ Karl's beliefs ⟶ Karl's desires

$\mathcal{M}$ $\mathcal{U}$

our evidence ⟶ our beliefs ⟶ our desires

approximately. That 'approximately' is important: our common-sense theory of persons tells us that Karl's beliefs and desires may differ from ours not just because of Karl's different evidence, but also because of the effects of Karl's different training on his underlying inductive method and basic intrinsic values. He may or may not have learned to beware of hasty generalization, or to like raw eel. If our common-sense theory (without benefit of esoteric scientific knowledge) told us just what these effects of training were, we could build them into a still better version of the Principle of Charity. But if not, we must idealize them away, and then not apply the idealized Principle too stringently.

The *Rationalization Principle* constrains the relation between **Ao** and **P**: Karl should be represented as a rational agent; the beliefs and desires ascribed to him by **Ao** should be such as to provide good reasons for his behavior, as given in physical terms by **P**. Thus if it is in **P** that Karl's arm goes up at a certain time, **Ao** should ascribe beliefs and desires according to which it is a good thing for his arm to go up then. I would hope to spell this out in decision-theoretic terms, as follows. Take a suitable set of mutually exclusive and jointly exhaustive propositions about Karl's behavior at any given time; of these alternatives, the one that comes true according to **P** should be the one (or: one of the ones) with maximum expected utility according to the total system of beliefs and desires ascribed to Karl at that time by **Ao**. A precondition: those ascribed beliefs and desires should be coherent enough to permit the comparison of expected utilities of alternative ways of behaving. (See, for instance, the exposition of decision theory in Jeffrey [7]; or the less formal account of rationalization of behavior in Davidson [1].)

Would such use of the machinery of decision theory go against my requirement that the constraining principles should come from our common-sense theory of persons, not from esoteric science? I think not. Decision theory (at least, if we omit the frills) is not esoteric science, however un-

familiar it may seem to an outsider. Rather, it is a systematic exposition of the consequences of certain well-chosen platitudes about belief, desire, preference, and choice. It is the very core of our common-sense theory of persons, dissected out and elegantly systematized. The same goes for my previous mention of underlying inductive methods and systems of basic intrinsic values: to the extent that unfamiliar theories are involved, they are not scientific theories but philosophical systematizations of parts of the common-sense theory of persons. The machinery derived from such systematizations is common property, whether or not the accompanying jargon is.

There is an ambiguity in the term 'behavior'. Note that I am using it to refer to raw behavior – body movements and the like – given by **P** alone; not to refer to behavior specified partly in terms of the agent's intentions. The latter is given jointly by **P** and **Ao**. That Karl's fingers move on certain trajectories and exert certain forces is what I call 'behavior'; that he signs a check is not. If anyone prefers, however, we could restate the Rationalization Principle in terms of the second sense of 'behavior': the beliefs and desires ascribed to Karl by **Ao** should be such as to provide good reasons – preferably, decision-theoretic explanations – for his non-raw behavior as given jointly by **P** and **Ao**. (This second version would fit better the account of rationalization in Davidson [1].)

The Principle of Truthfulness constrains the relation between **Ao** and **M**: Some of the beliefs and desires ascribed to Karl by **Ao** – in particular, some of those that pertain to speech behavior or responses to it on the part of Karl and his usual partners in conversation – should constitute Karl's part of the pattern of attitudes that I have elsewhere ([9], [12]) called a convention of truthfulness in Karl's language, under the truth conditions assigned by **M** to sentences thereof. For instance, if **M** assigns to 'Ionlay!' the truth condition that a lion is present, then some of the beliefs and desires ascribed to Karl by **Ao** should be as follows: (1) a desire not to utter 'Ionlay!' unless a lion is present; (2) a belief that his partners have a like desire; (3) a belief that a lion is present, at times when he hears 'Ionlay!' uttered; (4) a belief that his partners respond in the same way to utterances of 'Ionlay!'; (5) a belief that his partners expect him to have the beliefs and desires (1)–(4), or at least that they have no contrary expectations about him; (6) a belief that they expect him also to have the belief (5), or at least that they have no contrary expectations; and

so on. (Rather similar constraints on the relations between Karl's atti-
tudes and his meanings could be taken from the discussions of meaning
in Grice [6], Schiffer [14], or Stenius [15].)

The Principle of Generativity constrains **M**: **M** should assign truth con-
ditions to the sentences of Karl's language in a way that is at least finitely
specifiable, and preferably also reasonably uniform and simple. Perhaps
we may add that it should do so by means of syntactic and semantic rules
that fit some standardized format: the format of a Tarski-style truth
theory, as proposed by Davidson in [2] and elsewhere; or a categorically
based transformational grammar with intensional semantic functions, as
I propose in [11]; or the sort of semantically based grammar proposed
by some transformational grammarians, with some sort of specification
of truth conditions for the base structures.

The Manifestation Principle constrains the relation between **P** and **Ak**,
and to a slight extent also **Ao**: Karl's beliefs, as expressed in his own lan-
guage, should normally be manifest in his dispositions to speech behavior.
The sentences (in context) that he could be made to utter should normally
be among that express propositions that he believes to a high degree. So
unless **Ao** ascribes to Karl some special reason for deception or secrecy,
it should be possible to read off the beliefs in **Ak** from the dispositions-to-
utter given in **P**. (I do not see quite how to state a companion Manifesta-
tion Principle for the desires in **Ak**, but I rather think that there should be
one.)

The Triangle Principle constrains the three-way relation between **Ao**, **M**,
and **Ak**: Karl's beliefs and desires should be the same whether expressed
in his language or in ours. Suppose that **M** assigns a certain truth condi-
tion to a sentence s (in context c) in Karl's language, and suppose that a
sentence s' (in context c') of our language has the same truth condition.
Then if **M** is correct (and assigns truth conditions of a sufficiently rich
sort), we ought to be entitled to regard s (in c) and s' (in c') as expressing
the same propositions in their respective languages. If so, the status of s
(in c) in **Ak** should be the same as that of s' (in c'), in **Ao**: if **Ak** ascribes to
Karl a certain degree of belief in the proposition expressed by s (in c), then
Ao should ascribe to him the same degree of belief in the proposition ex-
pressed by s' (in c'), and likewise for Karl's degrees of desire.

How might we use these six constraining principles to solve the problem
of radical interpretation? (I still mean the unreal problem of advancing

from omniscience about **P** alone to omniscience also about **Ao**, **Ak**, and **M**.) I should like to contrast three rather different methods.

Method 1 is meant to borrow as much as possible from discussions of radical interpretation in Davidson [3], [4], and [5]. Davidson's problem of radical interpretation is not the same as mine, but can be treated as a sub-problem of mine. Given Karl's beliefs as expressed in his own language (part of **Ak**), Davidson asks how to solve both for Karl's beliefs as expressed in our language (part of **Ao**) and for the truth conditions of full sentences of Karl's language (part of **M**). As a by-product we may also get parsings of Karl's sentences and some sort of meanings for their constituents (the rest of **M**), since these are needed to play an auxiliary role in generating the truth conditions of the full sentences. Davidson's suggested method of attacking this sub-problem (which he regards as only a sketch of the main features of a fully adequate method) is to fill in the beliefs in **Ao** and the truth conditions (and auxiliary machinery) in **M** simultaneously, subject to three constraints: the Triangle Principle, and the Principles of Charity and Generativity. We must strike a balance as best we can between the demands of Charity on **Ao** and the demands of Generativity on **M**; it is not to be expected that we can satisfy both constraints perfectly.

To incorporate Davidson's method of solving the sub-problem into a method for solving my problem of radical interpretation, we must first use the Manifestation Principle to fill in the beliefs in **Ak**. But for that we need some input from **Ao**: so we must proceed by successive approximations. Method 1 goes as follows.

First step: fill in **Ao**, tentatively, by means of the Principle of Charity. We may or may not draw on **P** at this point; it depends on which version of the Principle of Charity we use.

Second step: using the tentative filling-in of **Ao** to exclude the cases in which Karl seems to have reason for deception or secrecy, and using Karl's dispositions to speech behavior as given in **P**, fill in the beliefs in **Ak** by means of the Manifestation Principle. This brings us up to the beginning of Davidson's sub-problem.

Third step: simultaneously revise the beliefs and **Ao** and fill in **M**, in such a way that the Triangle Principle is satisfied and a satisfactory balance is struck between the demands of the Principle of Charity on the beliefs in **Ao** and the demands of the Principle of Generativity on **M**.

Fourth step: using Karl's behavior of all sorts, as given in **P**, and using

the revised beliefs in **Ao**, revise the desires in **Ao** by means of the Rational-
ization Principle. Then fill in the desires in **Ak** by means of the Triangle
Principle. This gives a complete, but tentative, solution.

Fifth step: see whether the subsequent revision of **Ao** would have made
any difference to the outcome of the second step. If so, start over at the
second step using the revised **Ao**; if not, a stable solution has been reached.
If the outcome at the second step is not too sensitive to the details of **Ao**,
and if the revision of **Ao** at the third and fourth steps is not too great, then
we may hope for a stable solution after not too many rounds of iteration.

I doubt the adequacy of Method 1; mostly because it wastes the con-
straining power of the Principle of Truthfulness, but also because it fails
to constrain the ascribed beliefs in **Ao**, as well as the desires, by means of
the Rationalization Principle. Too much emphasis goes to language as a
vehicle for manifestation of belief and belief as manifest in language; not
enough either to language as a social practice or to belief as manifest in
non-linguistic behavior. It may well be that my full problem of radical
interpretation (for Karl, or in general) does have a fairly determinate
solution by virtue of the constraints that are left out of Method 1, and yet
that Davidson's sub-problem – the problem that has to be solved as the
third step of Method 1 – is severely indeterminate if taken by itself.

Method 2, the method I am inclined to favor, is almost opposite to
Method 1 in its plan of attack. The idea is to start by securing **Ao**, go on
to **M**, and fill in **Ak** only at the end. If we proceed in that order, David-
son's sub-problem never arises. Two other sub-problems arise instead:
the problem of determining Karl's attitudes (as expressed in our language)
on the basis of certain of the physical facts about him, and then the prob-
lem of determining his meanings on the basis of certain of these attitudes.
I hope that these are sub-problems that can safely be taken by themselves,
without loss of determinacy. Method 2 goes as follows.

First step: using **P** both as a source of information on Karl's behavior
and as a source of information on his life history of evidence, fill in **Ao**
completely by means of the Rationalization Principle and the Principle
of Charity. No special attention is given to Karl's language at this step;
his speech behavior is merely included along with all the rest of his be-
havior to be rationalized.

Second step: using **Ao** to give information about those of Karl's atti-
tudes that pertain to speech behavior, fill in **M** in such a way that, first,

the relation of these attitudes to the truth conditions of the full sentences conforms to the Principle of Truthfulness; and second, the demands of the Principle of Generativity on the truth conditions and auxiliary apparatus of M are satisfied as well as they can be.

Third step: given **Ao** and **M**, fill in **Ak** by means of the Triangle Principle. (The Manifestation Principle should then be satisfied automatically; it is redundant, given the Principle of Truthfulness, the Rationalization Principle as applied to speech behavior, and the Triangle Principle.) This completes the solution.

Method 3, finally, is the obvious holistic non-method: try to fill in **Ao**, **Ak**, and **M** all at once, satisfying all six of our constraining principles or balancing them off as best we can. That is the method we would have to fall back on if we decided that no sub-problem could be separated out without loss of some determinacy, and hence gave up hope of solving the full problem step by step. That would be no great defeat. Our fictitious mighty knower will not mind trying all possible complete solutions to find out which are the ones that fit the constraints! After all, we are not really interested in a practical method of finding out anything; we are interested in the determination of the mental and semantic facts by the physical facts via the constraining principles. Plans of attack and the determinacy of various sub-problems are side issues.

The things that matter are: (1) what the problem of radical interpretation is; (2) the set of constraints by which it is solved, and the source of their constraining power; (3) the presupposition that the physical facts determine the mental and semantic facts, somehow, to the extent that anything does; and (4) the extent of the determinacy.

As regards the extent of determinacy, all that I have said so far is meant to be neutral between optimism and pessimism. I have only tried to say what a solution to the problem of radical interpretation is – namely, a filling-in of **Ao**, **Ak**, and **M** – and what constraints it must satisfy to be correct. That leaves it open whether there is one correct solution or many; and, if many, how different two correct solutions can be.

Indeterminacy might come in more or less virulent forms. It is worth distinguishing, for the worst sort is also the least credible – by my lights, always far less credible than the hypothesis that we have misconceived the whole problem!

It seems hopeless to deny, in the face of such examples as those in [16]:

30–39, that the truth conditions of full sentences in **M** do not suffice to determine the rest of **M**: the parsings and the meanings of the constituents of sentences. At least, that is so unless there is something more than our Principle of Generativity to constrain this auxiliary syntactic and semantic apparatus.

It also seems hard to deny that a more general indeterminacy can arise because no solution fits all the constraints perfectly, and many different ways to strike a balance give many different compromise solutions. The 'unsharp analyticity' of some so-called definitions in physics, or the confused desires of a compulsive thief, might exemplify this indeterminacy of compromise.

Grant these two sorts of indeterminacy, and set them aside. Could indeterminacy of beliefs, desires, and truth conditions also arise because two different solutions both fit all the constraints perfectly? Here is the place to hold the line. This sort of indeterminacy has not been shown by convincing examples, and neither could it be shown – to me – by proof. *Credo*: if ever you prove to me that all the constraints we have yet found could permit two perfect solutions, differing otherwise than in the auxiliary apparatus of **M**, then you will have proved that we have not yet found all the constraints.

Princeton University

BIBLIOGRAPHY

[1] Donald Davidson, 'Actions, Reasons, and Causes', *Journal of Philosophy* **60** (1963), 685–700.
[2] Donald Davidson, 'Truth and Meaning', *Synthese* **17** (1967), 304–233.
[3] Donald Davidson, 'Mental Events', in Lawrence Foster and J. M. Swanson (eds.), *Experience and Theory*, University of Massachusetts Press, 1970.
[4] Donald Davidson, The Locke Lectures, given at Oxford University, 1970.
[5] Donald Davidson, 'Conceptual Relativism', two lectures given at the University of London, 1971.
[6] H. P. Grice, 'Meaning', *Philosophical Review* **66** (1957), 377–388.
[7] Richard Jeffrey, *The Logic of Decision*, McGraw-Hill, 1965.
[8] David Lewis, 'An Argument for the Identity Theory', *Journal of Philosophy* **63** (1966), 17–25.
[9] David Lewis, *Convention: A Philosophical Study*, Harvard University Press, 1969.
[10] David Lewis, 'How to Define Theoretical Terms', *Journal of Philosophy* **67** (1970), 427–446.
[11] David Lewis, 'General Semantics', *Synthese* **22** (1970), 18–67.

[12] David Lewis, 'Languages and Language', in a forthcoming volume of *Minnesota Studies in the Philosophy of Science.*

[13] David Lewis, 'Psychophysical and Theoretical Identifications', *Australiasian Journal of Philosophy* **50** (1972), 249–258.

[14] Stephen Schiffer, *Meaning*, Oxford University Press, 1972.

[15] Erik Stenius, 'Mood and Language-Game', *Synthese* **17** (1967), 254–274.

[16] W. V. Quine, *Ontological Relativity and Other Essays*, Columbia University Press, 1969.