

## RADICAL INTERPRETATION LECTURE 4

1. Davidson's attitude towards the indeterminacy of reference is relaxed (see e.g. 'A coherence theory of truth and knowledge', in his *Subjective, Intersubjective, Objective*). Like Quine (see e.g. 'Speaking of objects' in his *Ontological Relativity*), he doesn't think that there are facts about reference that we can never know but rather that there is nothing *to* know. The choice between schemes of reference is simply a choice between different ways of measuring something: it is like the choice between measuring temperatures using Fahrenheit or Celsius.

2. On reflection that sort of attitude is natural given that the data of radical interpretation are the subject's *uses of sentences*. When you say what the words of a language 'refer to' you are not pointing to new semantic facts: you're just summing up a variety of facts about the uses of sentences where those words appear (cf. Wittgenstein *Philosophical Investigations* §37.) It is not surprising that there is more than one way to give such a summing-up: there is more than one way to give a summing-up of *any* very extensive range of facts. What implications does this have for *causal* theories of reference, as in Kripke (*Naming and Necessity*) and Putnam (ch. 2 of *Reason, Truth and History*)?

3. Davidson takes radical interpretation to mean interpretation of a *language*: from data about what *sentences in the target language the subject holds true* we are supposed to extract a theory telling us the meanings of those sentences and the others in that language. Lewis's approach ('Radical interpretation' in his *Philosophical Papers* Vol 1) is more general. For Lewis radical interpretation tries to extract a much broader theory from a different and larger database. Lewis takes radical interpretation to be an exercise in the understanding of *another person* in a broader sense – of that person's beliefs and desires – given the totality of *the physical data*.

4. More specifically, Lewis describes the investigation into his imaginary subject, 'Karl', via the inter-relations of four classes of facts:

- **P**: The class of physical facts
- **Ao**: Karl's attitudes (degrees of beliefs, desire etc.) expressed in our language
- **Ak**: Karl's attitudes expressed in his language
- **M**: The truth-conditions of the sentences in Karl's language

**Ao** includes facts like this: that Karl believes to degree *d* the proposition expressed in English by 'Snow is white'. **Ak** includes facts like this: that Karl desires true to degree *d* the proposition expressed in his language by 'Oogle doogle'. Lewis (as those examples show) doesn't deal with monolithic mental states like belief or desire, which you either have or don't, but rather with degrees of confidence (credence) and degrees of desire (utility). Radical interpretation is then the exercise of trying to find out as much as you can about the facts in **Ao**, **Ak** and **M** given *all* the facts in **P** plus certain constraints. Note that Davidson's radical interpretation is a matter of trying to determine **M** given the facts in **Ak**.

5. Before saying more about those constraints I should emphasize that Lewisian radical interpretation is not part of an epistemological argument. It is not about whether

## RADICAL INTERPRETATION LECTURE 4

we can know what Karl's mental states are. By idealizing the investigation in such a way that the enquirer has perfect access to *all* the physical facts (and if you like perfect memory, perfect powers of reasoning etc.) Lewis is effectively making it into a metaphysical question: any 'facts' about **Ao**, **Ak** and **M** that could not be fixed under these circumstances are simply not facts at all (cf. Kripke's method in *Wittgenstein on Rules and Private Language* ch. 1).

6. Lewis mentions six constraints: Charity, Rationalization, Truthfulness, Generativity, Manifestation and Triangle. Charity, Rationalization, Truthfulness and Manifestation capture general (and partly definitional) claims about how people's beliefs (a) respond to inputs and (b) together with desires generate behavioural (including verbal) outputs. For instance, Charity says that given the same physical inputs as you – training and perceptual evidence – Karl will form the same beliefs as you; and Rationalization says that Karl's beliefs and desires cause him to act as you would expect: if, e.g., what he wants most is ice cream and if he is confident that there is ice cream over there (but nowhere else), he will go over there. Generalized, Charity is naturalistic epistemology. Generalized, Rationalization is decision theory.

7. In making these assumptions, Lewis's point is not to endorse substantive conditions that for some reason are underwritten by some a priori guarantee. Of course, it is all too possible that Karl will fail to satisfy Charity, Rationalization, Truthfulness and Manifestation. Lewis's point is rather that the further Karl is from satisfying them, the further he is from having mental states like beliefs or desires at all. Charity, Rationalization, Truthfulness and Manifestation can therefore be regarded as partially definitive of the mental. (For more on this approach see Lewis, 'How to define theoretical terms', *Philosophical Papers Vol. I.*)

8. The Generativity and Triangle constraints are 'internal' i.e. they do not involve **P**, at least not directly. The Generativity constraint essentially says that it should be possible to give a finite specification of the facts in **M**, though L doesn't say why – perhaps the idea is that the language must be learnable for Karl (cf. lecture 3, 7-8). The Triangle constraint connects **M**, **Ao** and **Ak**: it says that if **M** identifies the truth condition of a sentence  $S_K$  of Karl's language with that of a sentence  $S_O$  of our language then **Ak** assigns an attitude to the proposition expressed in Karl's language by  $S_K$  if and only if **Ao** assigns that same attitude to the proposition expressed in our language by  $S_O$ . We can think of the Triangle constraint as partly definitive of propositions i.e. as the objects of Karl's attitudes.

9. Lewis doesn't say much on the extent of the indeterminacy that radical interpretation faces, other than to affirm that it extends at least as far as indeterminacy of reference and possibly also mental indeterminacy in some difficult cases (p. 118). The main contribution in this paper is not so much in specifying the semantic indeterminacy that radical interpretation reveals but in showing how a naturalistic philosopher can use it as an instrument to enquire not only into language but also into the mental.