

There's a bit less in Lecture II than in the flanking lectures: in addition to being shorter, it starts with recap of Lecture I, and ends with a mix of recap of Lec I and anticipation of Lec III. I want to start with some unfinished biz from last time.

## 1. BACK TO THE CARNAP ARGUMENT

Where does Carnap's argument for eliding analyticity and necessity misfire? (I'll break with K, interpreting analyticity not as necessary apriority, but as apriority in virtue of meaning.)

Recall his argument:

The following consideration shows that the concept of L-truth thus defined is in accord with the convention 2-1 and hence is an adequate explicatum for logical truth. If S holds in every state-description, then the semantical rules of ranges suffice for establishing this result. [. . .] Therefore, the semantical rules establish also the truth of S because, if S holds in every state-description, then it holds also in the true state-description and hence is itself true. If, on the other hand, S does not hold in every state-description, then there is at least one state-description in which S does not hold. If this state-description were the true one, S would be false. Whether this state-description is true or not depends upon the facts of the universe. Therefore, in this case, even if S is true, it is not possible to establish its truth without reference to facts. (11)

We can treat "being establishable to be true obo semantical rules alone" as amounting to analyticity, L-truth as amounting to metaphysical necessity.

So he defends first the claim that if S is necessary, it's analytic; and second, the claim that if S isn't necessary, it isn't analytic.

He argues as follows:

- If S is necessarily true, then the semantical rules suffice for establishing this; and trivially, that it is actually true—hence, it's analytic. Hence no necessary a posteriori.
- Conversely, if S isn't necessarily true, there's some world at which not-S. If that world is described only with true propositions, S is false. But which world is described only with true propositions can only be determined a posteriori—hence, not analytically. Hence can't be known analytically/a priori whether S. Hence no contingent a priori.

Where do these arguments go wrong?

- The first argument plausibly goes wrong by conflating epistemic and metaphysical necessity. Granted that semantical rules suffice for establishing that ‘Nixon = the winner’, still that’s intuitively contingent.

Alternatively, perhaps it’s not accurate to interpret C’s state-descriptions as metaphysically possible worlds, but rather better to take them as epistemically possible worlds. In that case, the argument is unimpeachable, but it’s misleading of him to present himself as allied with Leibniz and Wittgenstein.

- The second argument is a bit more subtle. Note that it’s not strictly correct to say that “which world is described only with true propositions can only be determined a posteriori”. This all depends on how the world is given. If it’s given in terms of some qualitative description, that seems right. I can know that the world described only with truths is a world with an even number of trees only by counting. But I can know a priori that the world described only with true propositions is *this* world, or the *actual* world.

More generally, I can know a priori that the actual F = the F. One might take ‘the F’ and ‘the actual F’ as absolutely synonymous if one ignored the epistemic/metaphysical intension distinction: these clearly have the same epistemic intension, so one might think the claim is necessarily true. But it ain’t. Hence the ep/met conflation is at the root of this argument as well.

## 2. SEMANTICS AND METASEMANTICS

Lecture II of N&N treats primarily of questions in *metasemantics*, questions pertaining to the nonsemantic facts which ground the semantic facts. (“Ground”? Think of certain facts as holding *in virtue of* other facts holding, as the former facts being *constituted* by the latter facts. For instance, facts about what a certain organism is doing are grounded in facts about what the molecules that are part of that organism (and its environment?) are doing.)

What would a metasemantical theory look like? Considered extremely vaguely, every metasemantic theory makes room for the idea that the semantic properties of an expression are determined by how “we” “use” it. Here there are a pair of questions, who “we” are, and what “use” amounts to.

- A. Are “we” a group of people? Better be, if the aim is to ground the meaning of an expression considered as a *social* object (otherwise, one person would set the standard—who?). Alternatively, perhaps the aim is to ground *idiolectic* meaning, in which case it might be that “we” is just “one”, the person whose idiolect is at issue. However, others might get into the act, as we shall see below.
- B. Is “use” entirely “internal”, in the sense that it is exhausted by mental facts? If so, can these facts all be brought to consciousness, or are some (or all) of the mental facts that constitute use rather matters of unconscious processing of the sort that cognitive scientists attempt to describe? Or does use involve some external world component as

well, such as how one interacts with the world? If there is such an external-world component, should we include *sense*-like aspects of meaning under those that are externally grounded, or only *reference*-like aspects? And are these external-world components entirely social, or entirely concerned with how one interacts with the non-social world, or a mix?

Finally, we might wonder whether *more* goes into the metasemantical base than just use. For instance, David Lewis argued that there may be cases in which use cannot break ties among a number of candidate semantic theories. In cases in which this allowed for an undesirably high lack of determinacy of semantic properties, Lewis appealed to “reference-magnets”, intrinsically grounded differences in eligibility for reference.

(Note that Wittgenstein is sometimes claimed to have undergone a conversion from the view that meaning is about truth-conditions to the view that “meaning is use”. Clearly these are not competing views: one can endorse a use-theoretic metasemantics and a truth-conditional semantics.)

K’s aim is not to build or refute any *general* metasemantic theory; he is concerned solely with metasemantic theories which address the following issue: what nonsemantic facts ground the fact that one’s uses of N refer to o?

He has both a negative and a positive aim. The negative aim is to argue against a certain “theory”; the positive aim is to advance in its place a certain “picture”. (The picture/theory distinction is perhaps not especially clear.)

With that said, it’s important to note the following qualifications:

- K does not think that the theory is false for all cases. Rather his aim is to argue that it is false for some cases. Indeed, he acknowledges that the theory is correct in certain localized or restricted instances;
- K notices that his objections to the theory rely on certain schematic features of the theory being interpreted in a certain way. If they are interpreted in other ways -- in particular, ways that are compatible with the picture -- the initial theory can be resuscitated in a somewhat different form.

### 3. METASEMANTICAL DESCRIPTIVISM

What then is the theory he seeks to refute? He states it using six theses and a “condition on the satisfaction of the other theses”. We can ignore the sixth thesis. The five theses, then, cleaned up quite a bit, are as follows:

- 1) Suppose that, for a name ‘n’, one believes ‘Fn’. Then say that the predicate ‘F’ *corresponds* for one to ‘n’.

- 2) Among the predicates which correspond for one to 'n', there are some ('F', 'G', 'H', let's suppose), such that one believes ' $F_n \wedge G_n \wedge H_n \wedge [\forall x:F_x \wedge G_x \wedge H_x](x = n)$ ': say that these predicates *define* 'n' for one.
  - 3) If the predicates which define 'n' for one are uniquely satisfied by o, one's uses of 'n' refer to o. (Two comments: first, K gives a slightly more intricate formulation of this thesis, not so significant for present purposes; second, K is unclear here: he talks about "the properties", but does he mean the corresponding or the defining predicates?)
  - 4) If the predicates which define 'n' for one are not uniquely satisfied, one's uses of 'n' do not refer.
  - 5) 'If n exists,  $F_n \wedge G_n \wedge H_n$ ' is a priori for one.
- C) The predicates which define 'n' must not ineliminably exploit the notion of reference.

We can call this theory *metasemantical descriptivism*.

- In answer to question (A) above, metasemantical descriptivism is understood as a theory to ground idiolectic meaning. And certainly the view that one's own referential practices might be partly grounded in those of others is far from explicit in the statement of the theory.
- In answer to question (B), MD seems to make use entirely internal. Reference is grounded entirely in (i) one's attitudes toward sentences concerning a name, and (ii) one's choice to regard certain of these as "definitional".

Distinguish MD from what we might call *semantical descriptivism*, the view that a proper name is synonymous with a description. Last time we discussed K's big concern about that latter theory, that it rules out that proper names are rigid designators. (We expressed some reasons for skepticism about that objection, which we need not consider once again.) It is entirely compatible with metasemantical descriptivism that proper names are rigid designators. Metasemantical descriptivism is a theory about how reference gets fixed, not a theory about what the referent or the (metaphysical) intension of a name is.

Importantly, as we will see, K assumes of the predicates which define 'n' for one, that these concern what David Lewis called "famous deeds and distinctive peculiarities". As we will see, if other predicates are allowed into the mix, K's complaints may have less force.

#### 4. WORRIES ABOUT METASEMANTICAL DESCRIPTIVISM

What are the concerns?

##### Concerning (2)

Most of the time, we do not have sufficiently rich opinions about a person to be able to uniquely identify them.

*Example:* Feynman/Gell-mann. One's only opinion about these two might be that each was an eminent physicist. The predicate 'is an eminent physicist' does not uniquely identify either, as each satisfies it. Still, one's uses of 'Feynman' refer to Feynman.

*Example:* Cicero/Catiline. One's only opinion about the former is that he denounced the latter, one's only opinion about the latter is that he was denounced by the former. We can define each "holistically" in terms of being a pair one of whom denounced the other. Still, plenty of people have denounced one another, so this holistic definition does not single the pair out. Still, one's uses of 'Cicero' refer to Cicero.

*Example:* Einstein/ToR. One's only opinion about Einstein is that he invented the ToR, one's only opinion about the ToR is that it was Einstein's theory. But plenty of people-theory pairs are such that the former invented the latter.

To these we could add: Looking someone up in the phone book. My only opinion about N.N. is that he lives in Toronto, so do plenty of other people. Still I can use 'N.N.' to refer to N.N.

### Concerning (3)

One can refer, using 'n', to something other than the unique satisfier of the predicates that define 'n' for one.

*Example:* Gödel/Schmidt. One uses 'proved the incompleteness of arithmetic' to define 'Gödel'. Suppose, however, Gödel did not prove the incompleteness of arithmetic, some man named Schmidt did, then Gödel killed him and stole the credit for his work; so the unique satisfier of the predicate that defines 'Gödel' for one is not Gödel. Even then, one's uses of 'Gödel' would refer to Gödel.

*Example:* Peano/Dedekind. One uses 'formulated the Peano Axioms' to define 'Peano'. However, Dedekind actually did this. Still, one's uses of 'Peano' refer to Peano.

*Example:* Einstein/atom bomb. One uses 'invented the atom bomb' to define 'Einstein'. Einstein did not do that. Still, one's uses of 'Einstein' refer to Einstein.

### Concerning (4)

One can use 'n' to refer to o even though none of the predicates I use to define 'n' are uniquely satisfied.

*Example:* Feynman/Gell-mann. The predicate one uses to define 'Feynman', 'is an eminent physicist', is not satisfied by only one person.

*Example:* Gödel's good luck. One uses 'proved the incompleteness of arithmetic' to define 'Gödel'. Suppose, however, no one proved the completeness of arithmetic: Gödel luckily found a stack of paper on which the proof had materialized by raw chance.

## Concerning (5)

We can conceive that Gödel did not prove the incompleteness of arithmetic, therefore it is not a priori that if Gödel existed, he proved the incompleteness of arithmetic.

### “Direct reference”

Sometimes it is stated that N&N establishes or at least strongly motivates a “direct reference theory” of names. This label is more trouble than it’s worth. What does it mean?

“Direct”—as opposed to “indirect”, or “mediated”, by some other act of course (surely not an act of reference! but what act then?). More specifically, as opposed to “mediated” by a description, or by a sense. But mediated how? In order of decreasing plausibility, what might be intended by this label is:

- The rejection of *semantical* descriptivism, or at least the acceptance of the claim that the metaphysical intension of a name differs from the metaphysical intension of a description in being constant or “rigid”;
- The rejection of *metasemantical* descriptivism;
- The adoption of a “Millian” theory of the meanings of names, in effect a *reference-only* theory combined with a rigid metaphysical intension.

The first of these is as I have argued very plausible; the third is as I have argued very implausible on ground of its absurd prediction that ‘Hesperus = Phosphorus’ is a priori (Kripke agrees); soon we will turn to the assessment of the second.

Don’t confuse these issues. Best to eschew this heinous terminology!

## 5. THE COMPETING PICTURE

K borrows an idea from Strawson, to the effect that if you use ‘n’ to refer refer to o, I can exploit this and intend to use ‘n’ to refer to the same thing you use ‘n’ to refer to.

Someone, let’s say, a baby, is born; his parents call him by a certain name. They talk about him to their friends. Other people meet him. Through various sorts of talk the name is spread from link to link as if by a chain. A speaker who is on the far end of this chain, who has heard about, say Richard Feynman, in the market place or elsewhere, may be referring to Richard Feynman even though he can’t remember from whom he first heard of Feynman or from whom he ever heard of Feynman. He knows that Feynman is a famous physicist. A certain passage of communication reaching ultimately to the man himself does reach the speaker. He then is referring to Feynman even though he can’t identify him uniquely. He doesn’t know [this that and the other]. So he doesn’t have to know these things, but, instead, a chain of communication going back to Feynman himself has been established, by virtue of his membership in a community

which passed the name on from link to link, not by a ceremony that he makes in private in his study: 'By "Feynman" I shall mean the man who did such and such and such and such'. (91–2)

This isn't supposed to be a "theory". Still, the idea is clear: "A certain passage of communication reaching ultimately to the man himself does reach the speaker. He then is referring to Feynman even though he can't identify him uniquely".

Accordingly, the competing picture seems to be the following:

- 1) One uses 'n' to refer to o if one has engaged in a reference-fixing initial baptism of o under the name 'n';
- 2) One uses 'n' to refer to o if one has acquired the name 'n' from someone who used 'n' to refer to o.

Are these clauses exhaustive according to K? Dunno!

### The "causal theory of reference"

Some folks say that K endorses in Lecture II a "causal theory of reference". This label is more trouble than it's worth:

- Nothing he says is about the "theory of reference", if that's understood as a sort of semantical theory, rather it's at best a "causal theory of metareference".
- The initial baptism in the first clause need not involve any causation. For instance, Kripke thinks you can commit an "alone in one's study" sort of baptism of Gödel (91), in which no causation is relevant; consider also Neptune [\*\*\*]. Indeed, one can initially baptize mathematical objects ('Nancy'), which of course have no causal properties whatever.
- Even in the paradigm "reference borrowing" cases, it's a bit misleading and vague to appeal to "causation": first, you aren't *causing* me to refer to anyone; second, it's more specific than mere causation.

Eschew this terminology. (Even more embarrassing, the "new theory of reference", not new then considered as a theory of reference, not new anymore considered as a theory of metareference.)

### Worries

- Clearly type 1 cases involve some specific mental act in order to establish that they are cases of this type. How could it be that no mental act is needed to establish that a type 2 case is not a type 1 case? Clearly there must be some kind of mental act here. Sadly, K's story leaves us with nothing.

- Not *merely* having picked up 'n' from one who uses it to refer to o suffices for one's uses of 'n' to refer to o. After all, I might use 'n' for all sorts of willful purposes. At least I need some kind of *intention* to use it to refer to the guy who someone I picked it up from used it to refer to. How is the intention to be phrased? As K points out as a concern for Strawson, I needn't recall who I picked it up from (against 'the guy Bill used 'n' to refer to'). Perhaps 'the guy who I picked it up from' is enough?
- What if the name permutes? One loses one's accent German accent and gains an Upstate NY accent over time so the name is no longer 'FRahnk' but 'Freahnk'. I didn't pick up the name 'Freahnk' from anyone, I picked up the name 'FRahnk'. Perhaps we need to appeal to one's own past uses. We could appeal to the apparatus of a "mental file" or "dossier" of information. I have a certain persisting mental file, I use 'Freahnk' to refer to the guy at the start of that file.
- What happens when I have *two* uses of 'Alex', from different chains that enable me to refer to different Alexes, but I don't recall which guy I got either from? Here sometimes people appeal to files also. When I use 'Alex'-1, I'm "activating" a certain "dossier" of information that updates in a certain way; when I use 'Alex'-2, I activate a different dossier full of different information. Perhaps my intention then is to refer to the guy who *this* mental file was activated in response to.
- Worries about the social chain: Madagascar (Evans) used to be used by the Madagascans to refer to mainland Africa, later it was used to refer to Madagascar the island. The shift occurred because of a big collective shift in which place was the source of information people took as relevant to updating their mental files.

Nice question what exactly the truth is here but certainly K's view is too simple.

## 6. ASSESSMENT OF THE WORRIES ABOUT MD

### FDDP

Recall that all the preds one regarded as definitional in the objections were "FDDP" preds. Those *can* be used to determine referent in initial baptisms, but most of the time reference is not determined in this way.

Still, nothing in the formulation of MD required that the preds be of any particular kind! If my principal aim in using a name is *social* and *cooperative*, I will want the information I communicate it to concern the same thing everyone else uses it to communicate info about.. And I can certainly help myself to our common practice to attain reference to entities about which my knowledge is thin. What does this have to do with whether I use predicates to fix reference?

Not obviously anything. Why can't I use a predicate like 'the guy I heard of under the name 'Feynman'' to fix reference on Feynman? This predicate will achieve reference to Feynman under exactly the conditions one will refer to Feynman on K's picture. More complicated scenarios may require more complicated predicates.

Still, as Nozick pointed out to K (fn38 p88), if we can ascertain a priori which conditions  $\phi$  must satisfy in order for one's use of 'n' to refer to it in these or those circumstances, perhaps these conditions can be exploited as the preds that fix reference for 'n' for one.

Certainly K's method suggests that we'd \*better\* be able to do this. We didn't go about doing any experiments to ascertain whether 'Gödel' refers to Gödel rather than to Schmidt—thought experiments did all the work. The science of reference looks for all the world like an a priori discipline. (David Lewis impressed this point upon me in my proseminar.)

### X-phi

Certain philosophers (Machery, Mallon, Nichols, and Stich) have been suggesting that the science of reference *ought not* to be an a priori discipline. They have determined that intuitions about Gödel/Schmidt vary on the basis of cultural origin. Apparently Anglo-Americans tend to share K's judgement that 'Gödel' refers to 'Gödel', while Asian-Americans tend to disagree, thinking it refers to Schmidt.

Just what this shows I have no idea!

## 7. DESCRIPTIONS V INTENSIONS

Where are we?

*If* the true metasemantics of reference can be determined a priori, and *if* the true metasemantics of reference is not too complex, there may be hope for the description theory as piggy-backing on the true metasemantics of reference. However, the points above suggest difficulty for either of these claims. Maybe speakers don't know the MS a priori, but only have a tacit grasp of it which informs their judgement about assorted cases. And maybe the facts that these judgements track are too complex to be captured in a description, or at least a description in the language that the average speaker grasps.

None of this would cast doubt on the apparatus of the epistemic intension, however. After all, the epistemic intension:

- Is determined on the basis of dispositions to judgement about certain indicative conditionals. There is no question that there are such judgements: if sometimes the judgements are indeterminate or it comes out that there is no extension, this is a fact about the epistemic intension. Complexity can perhaps be idealized away from; indeed, since the apparatus of the metaphysical intension depends on similar idealization, if this apparatus is to get off the ground, the idealization had better be available. And, ad hominem, since this apparatus is key to K, it had better get off the ground.
- What exactly is determining these judgements—something directly available to consciousness, something more deep in the mind—is not relevant. The epistemic intension is, accordingly, significantly less demanding psychologically than is the description.

Accordingly it does not seem that any of K's arguments against MD cut against the epistemic intension; indeed, the very fundamentals of K's case against MD—the use of judgements about reference in worlds considered as actual—depend on the coherence of the epistemic intension.

So perhaps the following is the true metasemantics of reference. What one's use of 'n' refers to is the o such that one judges the following, where w is the actual world:

- If w is actual, n is the o;

Where what determines the fact that one makes this judgement is some aspect of one's psychology, perhaps entirely conscious, perhaps not.

## 8. WHAT HANGS ON ALL THIS?

First off, it would be nice to get the true metasemantics of names. The topic seems to have some intrinsic interest, inasmuch as there is a widely accepted semantic theory which is nontrivial and makes interesting predictions (by contrast, the semantics of 'and' as requiring mutual truth of the constituents for truth of the compound is widely accepted but a bit less interesting).

Second, and much more importantly, the issue is whether there is such a thing as the synthetic a priori.

Suppose:

- i. K is right and MD needs to be junked most of the time. Ordinary uses of 'Feynman' have *no* connection to *any* definition.
- ii. Any a priori knowledge about Feynman must come from grasp of a definition of 'Feynman'.
- iii. Nonetheless we still have a bunch of a priori knowledge about Feynman, as K thinks we do.

It would seem to follow that there is such a thing as "rational intuition", or a Kantian synthetic a priori.

Here the availability of the epistemic intension undermines premiss (ii).

## 9. PREVIEW

K canvasses a certain progression of thought:

- A. We have noticed that whenever C-fibers fire, there is pain, and vice versa.

- B. One explanation of this is that pain = the firing of C-fibers; another is that pain  $\neq$  the firing of C-fibers, but there is some special explanation of why they are always found together.
- C. The latter is an unacceptable Cartesian dualism; hence pain = the firing of C-fibers.
- D. It can't be that  $\Box$ (pain = the firing of C-fibers), since it's not a priori that pain = the firing of C-fibers, and necessity is the same as a priority.
- E. Accordingly, the identity between pain and the firing of C-fibers is *contingent*.
- F. Contingent identities are no problem, though: the inventor of bifocals = the first postmaster general of the US, but  $\neg\Box$ (the inventor of bifocals = the first postmaster general of the US).

Supposing that 'pain' and 'the firing of C-fibers' are rigid, K will of course disagree that (E) is coherent. So K will think that either  $\Box$ (pain = the firing of C-fibers), or pain  $\neq$  the firing of C-fibers. He is going to argue against the former. His argument will appeal to a version of (D). Of course K rejects (D) in full generality in the strongest terms, but there are certain very special features of the present case which make it hold in this case.