11/7/2006

Naturalism Week 9 Notes--Sellars

1. First, a reminder about why we care about this essay—how it fits into the larger issues of the course. Sellars is a *scientific* naturalist about (what is expressed by) alethic *modal* vocabulary—since when one uses it one is "about the business of *explaining*", [See quote #6 on handout, from §80] and according to the *scientia mensura*, "in the dimension of describing *and explaining*", science is the measure of all things. But he is *not* a *descriptive* naturalist about modality—nor (one wants to say, *a fortiori*) about *deontic* modal expressions, that is, (what is expressed by) *normative* vocabulary. For although describing and explaining come as an indissoluble package, they are quite different uses of language. Descriptivism about alethic modality, Sellars says, involves making a mistake structurally like that made by descriptivism about the normative, which is just the naturalistic fallacy.

2.

a) We saw last week how Sellars distinguishes *descriptive concepts* from mere *labels*: by their *inferential articulation*.
b) It seems, though, that any expression that can be used to form *declarative sentences* will be in this way distinguished from mere labels. So the question arises of what rules out *declarativism* about *description*. Sellars *must* rule it out, if he is to claim that modal vocabulary is *not descriptive*, since it meets the weak requirements of declarativism.
c) We also looked at Sellars's answer to this question. It takes the form, not of a criterion that could demarcate descriptive uses of vocabulary from *all* others, but only from one class—the one that matters for his claim about modal vocabulary. He distinguishes being used as a *premise from which* one reasons from being used as a *principle in accordance with which* one reasons. And his argument makes it clear that he is committed to the claim:

   Principles in accordance with which one reasons do not describe the world.

   Though he doesn't argue for this, we can think of Lewis Carroll's "Achilles and the Tortoise" argument that we should not think of *modus ponens* as a further *premise from which* we reason, on pain of an infinite regress, no stage of which constitutes a justifiable inference. And it seems right to think of *modus ponens* as doing something different from describing how things are. In fact, it seems right to think of it as articulating the framework within which alone conditionals can mean what they mean. (It is a further step, perhaps requiring further collateral premises, to infer from this characterization that *modus ponens* is "part of the meaning" of the conditional.)
d) So the claim is that alethic *modal* vocabulary does *not describe*. If it did, it would offer premises from which we reason. Instead, it expresses principles in accordance with which we reason. Those principles are invoked in *explaining* why one description applies, given that another one does. [See quote #6 on handout, from §80.] And those principles are an

essential part of the framework that provides the necessary context in which descriptive expressions can *describe*, and not just *label*, bits of the world, that is, can *state facts about* objects, rather than just *labeling* those objects. It is this whole framework, comprising both the *explanatory* principles according to which we reason and the *descriptive* premises from which we reason, over which the *scientific* naturalist asserts the hegemonic authority of 'science': that scientific, empirical, inductive method articulates, develops, and applies the *normative standards* for criticizing and improving that descriptive-explanatory framework.

e) I argued last time that:
    i. The distinction between premises we reason from and principles we reason in accordance with is real; and
    ii. Any inferentially articulated constellation of doxastic commitments must include some of both.
    iii. But within those limits, there are recipes for trading off premises for principles and principles for premises.
f) That is, the distinction between premises from which we reason and principles in accordance with which we reason is not a *hard* one, but a *soft* one, in the sense that it is a genuine distinction between two ways of *treating* something, but we are not *obligated* by how things are to treat any particular item as belonging in the one category rather than the other. Provided only that something remains in each category, we are *permitted* to treat things one way or the other, depending on other desiderata of our theory.
g) If that is right, then we should understand Sellars as *proposing* and *recommending* that we treat alethic modal vocabulary as expressing principles in accordance with which we reason, rather than premises from which we reason.
h) I sketched hurriedly, at the end of the session last time, how this could work for (what is expressed by) *deontic* modal vocabulary, that is, *normative* vocabulary.
i) So we should ask: What sorts of methodological, theoretical, or philosophical advantages does such a strategy for dealing with (what is expressed by) alethic modal vocabulary have?

3. One big one is that it provides a way to reconcile scientific *naturalism* with *empiricism* on the topic of modality, by showing as both *incorrect* and *not obligatory* (hence *optional* and *discardable*) empiricism's traditional *descriptivism* about the modal.

a) I will pick up one major strand out of all the many things that are going on in CDCM, to focus on this week. (Some others will come up as well.) That is the line of thought in MNI,
b) leading from the great sea change that the three phases (so far) of the modal revolution have accomplished,
c) through the question of its rationale, noting how astonished people in the heyday of Carnap and Quine—the audience Sellars is addressing—would have been to have what they found most questionable now taken as

unproblematic conceptual raw material for addressing potentially puzzling phenena—say, in naturalism about semantics or psychology, or in ontology and metaphysics. In this connection can observe that a central innovation of the classical American pragmatists, vis-à-vis their Enlightenment empiricist-naturalist ancestors, is precisely that they took notions of tendency, disposition, habit as *basic*, looking to the *refinement* of them through *selection*, whether by *evolution* or by *learning*, (and whether in the organism or its environment, or in their interaction). But the *logistical* tradition from Russell through Quine was always *empiricist* in a much more traditional sense (Frege and Wittgenstein, on the German side, and Moore on the English, are outliers in this regard).

d) the rejection of some initially plausible accounts,
e) to the Kant-Sellars thesis about modality.

4. [Tell here the story from "Modality, Normativity, and Intentionality" and the fourth of my Locke lectures, rehearsed in the body of "Pragmatism, Inferentialism, and Modality in Sellars's Arguments Against Empiricism", starting from the astonishing change of attitude (*volte face*) toward modality in the last third of the twentieth century, considering and rejecting two stories about how that might be thought to be justified (advances in logic and the realization that science can't do without modal claims), and ending with the Kant-Sellars thesis about modality: in using ordinary empirical descriptive vocabulary (the only sort the empiricist admits as intelligible) one always already knows how to *do* everything one needs to know how to do in order to deploy modal vocabulary. So one can never be in the predicament empiricists like Hume and Quine envisaged at the outset of their arguments concerning the intelligibility of modal concepts: understanding empirical *descriptive facts*, but not the inferential *rules* in virtue of which they stand in *explanatory* relations to one another.]

5. I think the Kant-Sellars thesis about modality (and, indeed, the corresponding thesis, endorsed by both thinkers, for the *normative* as well) is correct. And Sellars's denial of *descriptivism* about the modal is *one* way of thinking about modality that underwrites that thesis, and so secures its advantages—in particular, showing how one can both be a *scientific naturalist* about modality *and* an *empiricist*, so long as one is not a *descriptivist empiricist*. But is in order to endorse the Kant-Sellars thesis about modality, is it *necessary* to insist that in stating laws of nature one is not *saying* how things are, in the sense of *describing* how things are? I don't see that it is.

6.

a) Sellars sets up his discussion dialectically, in the form of "a sympathetic reconstruction of the controversy in the form of a debate between a Mr. C (for Constant Conjunction) and a Mr. E (for Entailment) who develop and qualify their views in such a way as to bring them to the growing edge of the problem." [Introduction]
b) Following a time-honored philosophical methodology (one favored not only by Plato and Hegel, but also by McDowell), Sellars discerns a presupposition common to the two sides, diagnoses this as what makes

their genuine insights show up as irreconcilable, and offers as a therapeutic suggestion the rejection of that shared assumption.

c) That assumption, of course, is *descriptivism* about the modal. The empiricist Mr. C and the rationalist Mr. E (whose views are to be synthesized in Kantian fashion) agree that modal claims *describe* some feature of the world. They just disagree over what feature that is. The empiricist things it is a *descriptive regularity*, and the rationalist thinks it is an *inferential commitment*.

d) Mr. C has trouble distinguishing lawlike from accidental regularities, and explaining why the latter, but not the former, can both be established *inductively*, on the side of *circumstances* of appropriate application, and support *counterfactual* reasoning of the kind codified by subjunctive conditionals (rather than subjunctive identicals), on the side of appropriate *consequences* of application.

e) Mr. E faces the problem that when we state laws of nature, we don't seem to be saying anything at all about our inferential practices or commitments—seem and intend, indeed, to be saying things that could be true even if no-one had ever had such practices or commitments. Modal claims do not, in fact, *describe* patterns of inference. [Cf. quotes #9,10 on handout.]

"Idealism is notorious for the fallacy of concluding that because there must be minds in the world in order for us to have reason to make statements about the world, therefore there is no sense to the idea of a world which does not include minds; the idea, that is, that things might have been such that there were no minds." [§101]

f) Once Sellars gives up his pretense of even-handedness, and devotes himself to fixing up the rationalist position, (Cf.: "It is now high time that I dropped the persona of Mr. E, and set about replying to the challenge with which Mr. C ended his first critique of the entailment theory." [§85]), he invokes the distinction between what one *says* by making a modal assertion and what one thereby *does*, in the sense of what one "conveys" or "conventionally implies."

"statements involving modal terms have the force of *prescriptive* statements about the use of certain expressions in the object language. Yet there is more than one way to *'have the force of* a statement, and a failure to distinguish between them may snowball into serious confusion as wider implications are drawn." [§81]

In asserting a modal claim, one *does* endorse a pattern of inference—the one licensed by that "principle in accordance with which we reason". But one does not *say that* one endorses that pattern of inference, one does not *describe oneself as* endorsing that pattern of inference. Endorsing it is part of the *force* of what one *does*, not part of the (descriptive) *content* of what one *says*.

g) Thus, in saying that every integer is the sum of at most 27 primes I have *conveyed* or *conventionally implied* that *I, Bob, believe that* every integer is the sum of at most 27 primes. But I have not *said that*. I have not said *anything* about myself and what I believe. And we can see that the

contents of these two claims are not the same, because they function very differently as antecedents of conditionals. To use an argument-form we have seen before:

i. *If* every integer is the sum of at most 27 primes, *then* every integer is the sum of at most 27 primes,

expressing the "stuttering inference", is true.

ii. If *I, Bob, believe that* every integer is the sum of at most 27 primes, then every integer is the sum of at most 27 primes,

is, alas, far from as certain. (One "route to God" is to postulate a being for whom the contents of these two claims *is* the same.)

7.

a) What is *said* vs. what is *conveyed* or conventionally implied by what is said. Q: Is the latter happily thought of as a kind of *pragmatic force*? (So, something stripped off by embedding.) Or should it go into a different box? Here too, we want to know what kind of a difference this is—a factual one? What settles where something should be put? Think of Steven Neale's claim that all definite descriptions are really Russellian, and that 'referential' descriptions are to be understood in terms of *speaker's reference*, and accordingly are not a *semantic* affair at all. Or the corresponding claim that 'if…then__' always *says* what the horseshoe says it does, and that any different *implications* we find are really a matter of *connotation*, or pragmatic or conventional implications of implicatures. Q: How is *saying* something related to *describing* the world? How is it related to the property at the base of the *declarativism* toward which the Geach-Frege embedding test drives us?

b) What if the embedding test does *not* strip off what is 'conveyed' or conventionally implied? Not *all* perlocutionary acts are insulated by embedding, for instance, rudeness.

c) On the other hand, the pragmatic implications of the "Bob believes that…" sort *are* stripped off by embedding.

8. Sellars "*Does* describe, but does something *more*." [ref.]

a) The conditional p□→q (cf. "c-box"), a *counterfactually robust*, true subjunctive conditional, gives us a new kind of *sayable*. But it is *not* supposed to be a *description* of the world. So Sellars *must* be *rejecting declarativism* about description.

b) It *does* embed as the antecedent of a conditional.

c) And we *should* be able to understand its *embedded* use in terms of its *free-standing* use as codifying inferences, in the sense of being a principle in accordance with which we reason—its *ingredient sense* in terms of its *free-standing* sense.

d) For we *can* use "principles in accordance with which we reason" as premises from which we reason, and look at their *consequences*.

e) But, think of my T&A stuff on *assertibility* conditions and *truth* conditions. What if we can*not* specify the ingredient sense from the free-standing one, just as compounds like conditionals are not assertibility-functional (as

shown by the 'will'-'foresee' example). If there is *not* a way to *derive* the *ingredient* sense of modal vocabulary from the *free-standing* sense needed to specify its role in codifying principle *in accordance with which* we reason:

i. How *is* the ingredient sense determined?
ii. How, if at all, is that question or issue related to the *descriptiveness* of the vocabulary?

f) So, for instance, "I *will* Φ," thought of as a *prediction*, *is* descriptive. But "I foresee that I will Φ," is *not*—or at least, it describes at most *me*, not the rest of the world. (Cf. in my T&A, ESPs: expression-statement pairs. The 'foresee' claims are on the expression side.) So the 'will'-'foresee' case can seem like a *displaced* description, where *two different sorts* of description, one of the world, one of me, *systematically* have the *same* circumstances of application in free-standing cases, but *different* ingredient senses. If that is all right, something need *not* be *non*-descriptive in order to fail in the derivability of ingredient sense from free-standing sense.

g) Q: So, what to do? A: One might think to look to the *consequences* of application. For these distinguish *what* one is describing (what properties one is attributing): I can draw inferences about *me* from the 'foresee' statement, and about the world (as well as ones about me) from the 'will' statement.

h) Q: Can one use *this* strategy—looking to the consequences of application—to get ingredient senses for subjunctive conditionals, even when one understands their free-standing use as non-descriptive? For one *can* talk about the *consequences* of endorsing an inference or pattern of inference. And it seems we can talk about the consequences of the inference endors*ed*, not just the consequences of endors*ing* it. We would usually distinguish these two by the embedding test. But is that *circular*, and so methodologically forbidden, in the present context? After all, we are asking about what *conditionals* mean here. How can we appeal to them as having unproblematic contents so as to use the embedding test? We still can distinguish the *force* of endorsing an inference from saying *that* I endorse it (though not, perhaps, from saying *that* it is a good one). (Think of the ESP consisting of this pair.)

i) Q: So is there any *other* way of distinguishing between what follows from the (pattern of) inference endors*ed* by asserting a subjunctive conditional and what follows from the endors*ing* of it (as an *act*), *besides* the embedding test? A: Yes.

i. Try 1: By abstracting from the effects of *other* commitments. But what about other attributions to me—A1 abstract from these *too*, i.e. for any collateral commitments attributed to … . Just collateral premise that are non-ascriptional, used as premises.
ii. Try 2: Just look at inferences I am committed to *before* I *say* something that makes that explicit—i.e. what *is* made

explicit rather than the making explicit of it. So …some inference whether *by this attribution*, apart from *his* collateral commitments, *S* is attributed to it just by asserting p□→q.

2 cases:

iii. Can use—*how* we use—the *right* consequences, *without* appealing to embedding test, and

iv. *If* we can, does that let us go from FSS to IS? A: for *conditionals,* arguing I may not *need* to get *ascription* (hence negation, the same way).

j) So this is a promising line to keep p□→q a) a description of the world *or* are (the 2 alternatives) *but* b) …an IS. If it *can* get both (i) and (ii), then we are *not* being *declarativist* about *descriptivism*.

9. Sellars correctly wants a view that will entitle him to say both:
   a) Modal claims do not just describe regularities; and
   b) Modal claims do not describe proprieties of inference.

While also being entitled to say:

   c) It is the correctness of inferences connecting descriptive terms, inferences without which those terms could not mean what they mean, that makes modal claims *true*. [But cf. quote #19 from §91.]

His way of doing that is to say that modal claims do not have the job of describing how the world is. This gets him (a) and (b). To get (c) without denying (b), he must talk about what is "conventionally implied by" modal claims, rather than what is "said by" such claims.

   d) But even if he is entirely right in this, it leaves wholly open the question of what *is* said by modal claims.
   e) Should we think that this is settled by the claim that they do not describe? This would require that the only sort of *saying that* is a *describing as*. But that seems to be precisely the sort of *descriptivism* about the *discursive* that he is principally concerned to deny.
   f) Further—though Sellars does not discuss this issue explicitly—modal claimables embed perfectly well as antecedents of conditionals, complements of propositional attitude ascriptions, and so on. We don't have to be *declarativists* about *description* (the converse of the claim in (e)) to be declarativists about what is *said*—in the sense of taking it that declarative sentences *do **say** something*.
   g) In any case, why isn't it both
      i. open to Sellars; and
      ii. The best line to take here,

to say that the modal claim ∀x[Fx□→Gx] says *that* something's being G *follows from* its being F? That is, why can't we treat possession of one property's *following from* possession of another property, one kind of *fact* following from another kind of *fact*, as being part of how the world is and can be *described* as being? Compare: facts about what is *incompatible* with what. Surely these are things we can *discover*, as we can the ground-level facts that *are* incompatible with or consequences of others.

h) To do that is not in the first instance to describe any possible *acts of inferring* or *uses of linguistic expressions* as appropriate or inappropriate, correct or incorrect. It is true that, *when combined* (as auxiliary hypotheses or collateral premises) *with further claims* about what some linguistic expressions mean, the claim in (g) has *consequences* of this sort. But those consequences are *not* part of what is *said* by a claim of the form "that something melts at 1083.4° C. is a consequence of, follows from, its being pure copper."

i) Such a line, in effect, is what the idiom of possible worlds does when applied in the way of intensional semantics (the second phase of the modal revolution). For it says that in saying $\forall x[Fx \Box\!\!\rightarrow Gx]$ we are saying that every causally or physically (or 'scientifically') possible world in which something is F is a world in which it is G (that for every x, the set of Fx-worlds is a subset of the set of Gx-worlds). This is thought of as a straightforward *description* of the modal space of possibilities.

j) In order to take this line about what is *said by* modal claims, we need not deny Sellars's claim that in uttering such a claim we are doing *something more* than describing things, that we are *also* endorsing an inference. As we have seen, I do *that* when endorse *any* claim. So doing that "something more" is compatible with *also* saying *that* things are thus-and-so, in the *narrow* sense of *describing* things as being thus-and-so ('narrow' because there is a wider sense of 'saying' that is not restricted to any sort of description). And we can grant that the description in question is *more* than a mere description of *this* world—which is the claim of Mr. C that Sellars is most concerned to deny.

k) Viewed this way, the second phase develops the first phase of the modal revolution in such a way as to *broaden* the notion of <u>description</u> beyond the use that *empiricists* had permitted. Its *license* to do so is *precisely* the Kant-Sellars thesis about modality: description in the *narrow*, actualist, empiricist sense depends for its intelligibility on the possibility of doing something more, of inferential connections among such *narrow, actualist descriptions* that are made explicit by modal vocabulary. But *once that point is taken on board*, there is no reason remaining not to take that modal vocabulary as *descriptive*, albeit in a *broader* sense than that term was employed in before.

10. Mark made a very good point last time, which I let go by too quickly: Sellars's distinction between *subjunctive identicals* and *subjunctive conditionals* is not nearly as straightforward as he suggests.

a) When it is drawn, as I did, in terms of the example of "All of the coins in my pocket are copper," and "Copper melts at 1084° C.," it seems sharp enough. Though both do support counterfactuals, the former supports counterfactuals about objects that are *in fact* in my pocket, and the latter about *any* copper.

b) But if we look at generalizations such as "The Baltic is less salty than the Atlantic," and "Homeotherms have higher metabolic rates than

poikylotherms," we seem to have intermediate cases. Past, probable-future, and at least some non-actual water samples drawn from the two oceans will conform to the former. And the latter, while almost certainly not an exceptionless generalization across *possible* species (and perhaps not even for some actual dinosaurs) is true of more than just the currently extant species.

c) But it seems to me I did—briefly—say the right thing about this distinction, while not emphasizing that saying that involves conceiving of it somewhat differently from the way Sellars presents it. The issue is whether there is some specification, typically a description, such that one proceeds by determining what things *actually* satisfy that description, and then looking at counterfactuals that are true of *them* in other worlds. That is a *de re* procedure. Or whether one looks at what things satisfy that specification in *other* worlds, and then at the counterfactuals true of *them*. That is a *de dicto* procedure. Almost any given generalization can be read either way—that is, is subject to both kinds of readings. We can think of the distinction, accordingly, not so much as concerning the generalization as concerning the connection between it and counterfactuals that it supports. Sellars is just pointing out that his claims—paradigmatically about what *induction* supports—are intended to address only the counterfactuals *de dicto* that a generalization supports.

d) This is a distinction of the order in which two operations are performed (so, using the precise word for such an issue, a "*scope* distinction"). If the generalization is $\forall x[Fx(\Box)\rightarrow Gx]$:

   i. One can *first* figure out which things are F in *this* world, and *then* follow *them* to other possible worlds in order to assess the truth-values of counterfactuals involving *them*—the *de re* method of assessing counterfactuals on the basis of that generalization; or

   ii. One can *first* move to the world with respect to which one wants to assess the truth-value of some counterfactual, and *then* see what things are F in *that* world.

e) Carnap (and Russell, I think)—Hempel discusses this view in *Aspects of Scientific Explanation*—argues that genuine laws, of the kind (the best sort of) science aims at and employs in its explanations, do not make any reference to *particular* objects, events, and so on. Thus, they do not refer to the Earth, to North America, to Julius Caesar, to me, or to now. The point is sometimes put by saying that they are "purely universal". But the distinction aimed at *could* be the one above—or rather, the distinction between extracting counterfactuals from generalizations in the *de re* way and in the *de dicto* way could be what lies behind the thought Carnap (and Russell, and others—Reichenbach, Nagel?) were expressing. For what they forbid, in the statement of laws, is the use of *proper names*, *demonstratives*, and *indexicals*, all of which are (and are paradigms of) *rigid designators*. And that is to say that they *must* be read *de re* (and so

have problematic A-intensions, in Jackson's sense). If this is right, then there is a progression of interpretation:

i. Carnap (and the others) offer as a *necessary* condition of being a "law-like" rather than an "accidental" generalization, that proper names, demonstratives, and indexicals, as expressions picking out particular things, do not occur in the statement of law-like generalizations.

ii. Sellars renders this distinction in terms of that between "subjunctive identicals" and "subjunctive conditionals", namely, in terms of whether counterfactual consequences can only be derived via *identities* linking the objects of the counterfactual with objects the antecedent predicate of the generalization actually applies to.

iii. I then read Sellars's point in terms of the order of application (hence *scope*) of two operations: moving to a counterfactual world, and applying the predicate of the antecedent.

f) The next question, then, is: when the distinction is read as I suggest above, in what sense is it *true* that induction supports only *de dicto* counterfactuals?

g) The idea is that if we look not at the probability of *all* of the coins in my pocket being copper being increased by checking them, one-by-one (since finding that the 9 checked so far of the 10 are *does* increase the probability that *all* of them are, assuming independence, and regardless of the antecedent probability of each), but at the probability that the *next* coin in my pocket that we check will be copper, if the generalization is really accidental, then we can look to the case where each is independent, and there is some antecedent probability for each, and finding that the first 9 checked of the 10 are copper says *nothing* about the probability of the 10^{th}. To think that it does is the gambler's fallacy. (Compare coin-flipping.)

h) But does this point really generalize as it is claimed to (on the reading I offered) to *de re* and *de dicto* counterfactuals supported by generalizations generally?

11. Sellars's account of "primary statistical induction" as moving from the observation that n/m of observed As are Bs to the prediction that n/m of any class of unobserved As will be Bs is one that, I hope and trust, hardly anyone today would accept. Bayes' Theorem tells us instead how to use the initial observation to adjust our prior probability distributions, which need not be flat. Suppose the question is how many adult African elephants weigh more than 2000 pounds. We may have observed that n/m of them, chosen at random out of a large population, do, and hence have good reason to suppose that n/m of them do in general. But we may also know that when elephants travel, they tend to arrange themselves in order of size, with the largest leading the way. Then if our next observation is to be of the first 5 of a troop of 10 to pass through a gate, we can be pretty sure Jumbo will be among them, along with his nearly-Jumbo friends.

12. t