

10/3/2006

Jackson Notes—Week 4

- 1) The discussion for this first week on Jackson is in 4 parts:
 - I) Introduction
 - II) Chapter One
 - III) Classical Intensional Semantics
 - IV) Two-Dimensional Intensional Semantics
- 2) For Part I—Introduction:
 - a) On Locke Lectures. How the lecturers are picked. History: 37 in 56 years (3 from Pitt). Famous catastrophes (Lorenzen, Sellars, Putnam). Challenge: Late-career dissertation (2-4 years to write six lectures/chapters).
 - b) Jackson's are a good set. We'll spend 2 weeks on Chs 1-3, and then one week on 5-6. I'll argue he'd have done better expanding the material in 1-3 to fill all 6. But the treatment of ethics in 5-6 shows how one sophisticated physicalist approaches a key kind of *normative* discourse.
 - c) FJ's LLs have not yet been really digested by the profession. Stalnaker has a pedestrian discussion, and Stich a truly lazy one. Beaney is pretty good, but doesn't go very deep.
 - d) There are a lot of arguments here, some of them summing up the thought of others (e.g. Papineau). I'll try to help in sifting them into major and minor ones (there is a lot of response to possible objections, which matter for getting the exact shape of the claims and justifications right, but are not the principal arguments driving the work).
 - e) The key moves are applications of what is called (I think misleadingly) "two-dimensional modal logic", the origins of which can be found in Pavel Tichy, David Kaplan, and Stalnaker, but which are applied philosophically to greatest effect by the "Canberra Planners": Jackson (following Armstrong), his star student David Chalmers (who gets to "pan-psychist dualism"), and Martin Davies.
 - f) The overall line of thought is, I think, quite original, even though it is assembled largely from off-the-shelf components. It leads from global supervenience through an entailment thesis to a very strong claim of *a priori* reducibility. Cf. philosophical argumentation beginning with premises that are undeniable, and proceeding by steps that are individually compelling to conclusions that are unbelievable. [Mention Belnap's policy about incredible claims: he doesn't believe them.] A rabbit is finally pulled from the hat, and one wants to know whether it was there all along, without our realizing it, or if not, at which stage exactly it was smuggled in unnoticed somewhere else along the line.
 - g) The final position is distinctive in itself. We have seen that classical Carnap-Nagel *reductionism* (of the sort characteristic of high-church unity-of-science reductive physicalism) is the conjunction of two kinds of claims: *definability* of terms of the target vocabulary by those of the base vocabulary, and *derivability* of the true laws statable in the target vocabulary from those statable in the base vocabulary. There are good reasons (the many-levels and multiple-realization arguments) to think claims of this sort are too strong to be plausibly *true*. *Global supervenience* is a substantive claim that, in the interest of plausible truth, gives

up *both* of these—and may for that very reason be too weak to yield an *interesting* sort of physicalism or other scientific naturalism. Token-token identity theories such as Davidson's anomalous monism try to split the difference by giving up on *derivability*, and weakening *definability* to co-reference. Jackson is going to give up even co-reference on the side of definability, in the way global supervenience does (cf. Haugeland), but will reinstate *derivability* in a very strong form: *a priori* knowable *entailments* of all *facts* (not just laws) statable in the target vocabulary by all the *facts* statable in the base vocabulary. We can diagram the positions in the form of a table:

	<i>Definability</i>	<i>Derivability</i>
Carnap-Nagel Reducibility:	Strong, type-type	Of Laws
Anomalous Monism:	Weak, token-token coreference	No
Jackson:	No	<i>A Priori</i> entailment of Facts
Global Supervenience:	No	No

- h) So Jackson has a distinctive and unique suggestion for a Goldilocks position intermediate between reducibility (too hot) and global supervenience (too cold).
- i) Notice that for classical Carnap-Nagel reducibility, the type-type biconditional definitions provided *bridge principles* connecting the two vocabularies, which, together with the laws couched in the base vocabulary, are to allow the *logical derivation* of the laws statable in the target vocabulary. Since Jackson does not assume definability, he cannot appeal to such bridge principles in funding the *a priori* entailment of the facts expressible in the target vocabulary by the facts expressible in the base vocabulary. Q: What takes their place (does the work done by the bridge principles)? And how does it become available *a priori* (hence in a world-independent manner)?
- j) With two exceptions, my discussion today will be expository rather than critical. The two exceptions, both involving arguments from Chapter One, are:
 - i) Jackson's version of Papineau's response to the Hempel-Crane-Mellor worry about what 'physics' is in the formulation of physicalism. The criticisms here are not a big deal; this discussion is really just a warm-up.
 - ii) Jackson's argument that global supervenience of the psychological (Ψ) on the physical (Φ) has as a consequence that the physical facts *entail* the psychological facts. This is a key move in his journey from global supervenience to a fairly strong (original) form of reductionism.
- k) t

For Part II—Jackson's Chapter One:

- 3) Basic setting of the project:
 - a) "Serious metaphysics" is attempting to solve the *location problem*.
 - b) The *location problem* is for any target range of facts (vocabulary), either *locate* it (find a place for it) in terms of the facts statable in a base vocabulary, or *eliminate* it.
 - c) "Entry by entailment thesis": *Locating* the target facts is showing that they are *entailed* by the base facts.

d) The examples Jackson offers early on are:

- i) Statements of Jones's and Smith's heights *entail* that (say) Jones is taller than Smith. Here we have well-ordered magnitudes, real numbers, and a statement about their ordering. Presumably the ordering facts are part of the meanings of the numerical expressions. So no further premises are needed.
- ii) Solidity: "The story in favoured terms...tells us that these lattice-like arrays of molecules *exclude* each other, the intermolecular forces being such as to prevent the lattices encroaching on each other' spaces. And that is what it takes, according to our concept, to be solid." [3] Here we seem to appeal to a definition or meaning, which "identifies solidity with being disposed to resist encroachment" [p3 note 3].
- iii) Density: "Though density is a different property from either mass or volume (since density cannot be identified with either mass or volume), there is a clear sense in which density is **not an additional feature of reality over and above mass and volume** [BB: I think this is a particularly important phrase for FJ's motivations], and we can capture this by noting that the account of how things are in terms of mass and volume implicitly contains, in the sense of entailing, the account of how things are in terms of density." [4] Given the definition of density as mass/volume.

These all are more or less *analytic entailments*, underwritten by *reductive definitions*, in the form of biconditionals with the target vocabulary on one side, and only base vocabulary on the other. But he will need something weaker for, e.g., semantics in terms of behavioral dispositions. In general, these classical reductive examples are OK to introduce the *genus* he is interested in, but since he gives up everything that would come under the heading of *definability* (as indicated in the table above), these examples cannot be representative of the view he will eventually endorse. What goes in the place of these definitions for him—that is, what does the corresponding work of bridge principles?

4) On *picking out the physical* (contra Crane&Mellor, and Hempel), FJ thinks we can get a good enough idea in 3 ways [7]:

- a) Kinds: "They will be broadly of a kind with those that appear in current physical science..." This seems more promising, but notice that this is a meta-kind: the kinds appealed to in future physics will be kinds *like* the kinds appealed to in contemporary physics. First, do we know what we mean by "same kind of kind" here? Of course, almost anything will be like the kinds of physics in *some* ways. (The kinds of economics are like physical kinds in being associated with mathematically tractable magnitudes, for instance.) Second, does the history of science suggest that this claim is *true*? Most nineteenth-century physicists would have insisted that explanations that invoked *efficacy of place* of the sort common in general-relativistic geometrodynamics were *nothing at all* like their own explanations, and were just the sort of thing physics had properly learned to reject.
- b) "Ostensive definition": by "pointing to some exemplars of non-sentient objects—tables, chairs, mountains, and the like—and then saying that by physical

properties and relations, they mean the kinds of properties and relations needed to give a **complete account** of things like them. Their clearly non-trivial claim is then that the kinds of properties and relations needed to **account for** the non-sentient are enough to account for everything, or at least everything contingent.” [FJ acknowledges that this is Papineau’s idea, from his 1993 *Philosophical Naturalism*.] Everything turns on what is meant here by a “complete account”. We won’t pick out the vocabulary FJ wants if that means anything like “an account of *all the facts* about those objects.” Tables and chairs are artifacts, produced by *people* for *reasons*, and are made the particular ways they are because of quite complicated historical facts. Mountains are climbed by people for complicated psychological and sociological reasons. (The Nepalese and Tibetans *never* “climbed mountains”—i.e. just to reach the summit—before Europeans brought that practice to them.) Is there any way of saying what sense of “complete account” is intended here, without begging the question about what should count as *physical* vocabulary (by making implicit appeal to a distinction between the facts included in this “complete account” as being the *physical* facts)? (We’ll discuss the idea of a “complete account” further below.)

c) Size: “They can characterize the physical properties and relations as those that are needed to **handle** everything below a certain size.” [7] “Physicalism is the clearly non-trivial claim that the **kinds** [BB: see above] of properties and relations that are enough to **account for everything below a certain size**, and in particular, below the size needed to have semantic or psychological properties, are, in suitable combinations, enough to account for everything, or anyway everything semantic and psychological.” [8] Here, too, care is needed. It is a fact about things smaller than the index size that they are parts of larger things. The claim is trivial if *those* facts are part of the “everything” appealed to. Perhaps then we must restrict ourselves to facts that can be stated without referring to anything larger than the index size. But that would require doing quantum mechanics without saying anything about even the possibility of *measurement* or *observation* at the macro-level. Without that, we not only get an unrecognizable remnant of QM, but have burned all the bridges to the macro level—all the bridges that would need to be invoked to make the conclusion plausible.

d) Q: So, how are we to evaluate FJ’s (and Papineau’s) suggestions here? [Do three weak arguments add up to one good one? Cf. Rawls.] Can three characteristics that individually seem either circular or to pick out the wrong class be put together to get a useful grip on a vocabulary? This is not just a rhetorical question: sometimes the best one can do is roughly triangulate to locate a view. (Cf. Wittgenstein: “stand roughly there” can be wholly in order.)

5) *Completeness of science 1:*

a) “[W]e know [BB: I think this claim to *know* is probably a little too strong, on *any* nontrivial reading of “whole story.”] that science can in principle tell us **the whole story** about physical objects.” [2] It is not so easy to say in what sense this is true. At most, it seems, it can tell us the whole story *that can be specified in physical terms*. But how does what is true of physics contrast with any other vocabulary? Doesn’t the vocabulary of, say, *etiquette*, permit us to tell the whole

story about anything at all—so long as we restrict ourselves to facts statable in the vocabulary of etiquette? Doesn't even astrology do this?

b) Re a reductive account of *semantics* in physicalistic terms: "We know enough as of now to be able to say...that it will look something like the story I gave a glimpse of—a story about masses, shapes, causal chains, **behavioral dispositions of language users**, evolutionary history, and the like—and...that it will not contain terms for truth, reference, and meaning." [2] Notice that if it is really restricted to the language of physics, those "behavioral dispositions" of what are in fact (though they cannot be specified as such in the physicalistic base vocabulary) language users must be specified entirely in the language of physics. These will not look *anything* like dispositions in our ordinary (philosophical) sense of the term. For *none* of the generalizations that would group responses together as exercises of what we would think of as *one* disposition (e.g. to respond to visible red things sometimes by using the term 'red') can be formulated at this level. Consider Dennett's story (from "True Believers") about the secretary of state's press conference. Not even the disposition to say 'ouch' upon stubbing one's toe would be visible at this level. Is it at all plausible that there is an *entailment* concerning what a sign-design token *refers to*, or when it would be *true*, taking this sort of description as its premises? What sort of auxiliary hypotheses would one need to add to produce one?

c) "Physicalism...claims that a **complete account** of what our world is like, its nature (or, on some versions, a complete account of everything contingent about our world), can in principle be told in terms of a relatively small set of favoured particulars, properties, and relations, the 'physical' ones. [6]

d) Next we will see what he means by 'completeness':

6) Completeness of Science and Supervenience:

a) "It is the physicalists' claim to have a *complete* story about the nature of our world which commits them to our world having a psychological nature if and only if that nature is entailed by the world's physical nature...The physicalists' distinctive doctrine is...that the world is entirely physical in nature, that it is nothing but, or nothing over and above, the physical world, and that a full inventory of the instantiated physical properties and relations would be a full inventory *simpliciter*." [9]

b) "One particularly clear way of showing *incompleteness* is by appeal to independent variation. What shows that three co-ordinates do not provide a complete account of location in space-time is that we can vary position in space-time while keeping any three co-ordinates constant. Hence, an obvious way to approach completeness is in terms of the *lack* of independent variation...A body's mass and volume completely specifies its [average] density because you cannot have a difference in density without a difference in at least one of mass and volume. But **lack of independent variation is supervenience**: density supervenes on mass and volume. This suggest that we should look for a suitable supervenience thesis to capture the sense in which physicalism claims completeness." [9] Notice that here the completeness of account being claimed is *not* physics' offering accounts that are complete *within its own domain*. Nor can a supervenience thesis support the idea that physics gives accounts that are

complete in extending to *other* domains. Rather, the idea is that the accounts physics gives of physical phenomena are complete in that everything else we might want an account of (“everything there is”) supervenes on what physics gives accounts of. This is what I last time called the “claim of extramural authority”, rather than the “claim of intramural excellence.”

7) Supervenience:

- a) “What we need to capture physicalism’s distinctive claim...is a *contingent* global supervenience claim.” [12]
- b) “(B) Any world which is a *minimal* physical duplicate of our world is a duplicate *simpliciter* of our world,” where a minimal physical duplicate is what you get if you ‘stop right there’...Thus a minimal physical duplicate of our world is a world that (a) is exactly like our world in every physical respect (instantiated property for instantiated property, law for law, relation for relation) and (b) contains nothing else in the sense of nothing more by way of kinds or particulars than it *must* to satisfy (a). Clause (b) is a ‘no gratuitous additions’ or ‘stop’ clause. Thesis (B) is a claim about the nature of our world expressed in terms of a claim about a very limited range of worlds, namely the minimal physical duplicates of our world.”[13] This is really very weak, since it allows that a world that was only *slightly* different from ours—one in which, say, I parked in a different parking space yesterday, might differ wildly in psychological or semantic properties, with, say, no-one having any psychological properties at all. FJ allows that “some physicalists want to make a bolder claim...among the worlds which contain the same basic laws and ingredients as our world, any two physical duplicates are duplicates *simpliciter*.” But he sticks with the weak one as strong enough for his purposes.
- c) The ‘must’ in clause (b) of thesis (B) is itself modal, and we should ask what worlds it quantifies over. The question is what it is and is not *possible* to have in a world with the physical constitution of ours. Is it *possible* that such a world should lack psychological properties? If so, then these will be thrown out with the ectoplasm. It seems that would be *logically* possible. Is it *physically* possible? *Psychologically* possible? This issue is underlined by the notion of minimal psychological duplicate considered below.

8) “Physicalism is associated with various asymmetry doctrines.” [14] He goes on to argue that minimal physical duplicates of our world must be psychological duplicates, but minimal psychological duplicates need not be physical duplicates, claiming that this is “obvious”. But

- a) this depends on how rich the psychological vocabulary is. If we are semantic externalists, and include claims about what *can* be thought, as well as what *is* thought—that is, have modal vocabulary available to (and we surely need at least *disposition* talk in our psychology)—then this is not at all clear. And
- b) I’ve argued that no such asymmetry holds for *semantic normative* vocabulary.
- c) In any case, what do we mean by minimal psychological duplicate: it must have no properties etc. beyond what is *necessary* for the psychological duplication. But what is that? If we are functionalists, we might think that *nothing* about our physiology is *necessary* for thought, in the sense that creatures could be having just those thoughts,

but realized in quite a different physiology. So is what is necessary for minimality of psychological duplication that we have *some one or other* of the possible physiologies? In that case functionalism shows that the psychological does not determine the physical.

- 9) Here is FJ's big claim—the punchline—of the first chapter: “It is easy to show, given that (B) follows from physicalism, that if physicalism is true, then the psychological account of our world is entailed by the physical account of our world.” [24]
 - a) “Let Φ be the story as told in purely physical terms, which is true at the actual world and all the minimal physical duplicates of the actual world, and false elsewhere.” [25] There are several questions here:
 - i) Is there such a ‘sentence’, even in an infinitary language? Let us grant that there is a story Φ' , that is true at all and only the physical duplicates of our world. Can we put in a sentence what we need to in order to get a sentence true at all and only the *minimal* physical duplicates?
 - ii) If there is such a sentence, is it possible to formulate it *in the language of physics*? FJ insists it is: “ Φ is a hugely complex, **purely physical** account of our world.”[25]
 - iii) If it can be formulated in the language of physics, does it follow from physics alone that it is true?
 - b) Here is the argument: “Let Ψ be any true sentence which is about the psychological nature of our world...Now if (B) is true, every world at which Φ is true [BB: viz., the *minimal* physical duplicates of our world] is a duplicate *simpliciter* of our world, and so *a fortiori* a psychological duplicate of our world. But then every world at which Φ is true is a world at which Ψ is true—that is, Φ **entails** Ψ .” [25]
 - c) On the face of it, this is a *very* different sense of ‘entails’ from the sense in which settling the mass and volume entails a value for density. This is *semantic* entailment in the sense of inclusions of sets of validating possible worlds. (Note that Tarskian model-theoretic entailment uses sets of *models*, algebraic structures, rather than *worlds*, to define entailment.) FJ will have to argue (in Lecture 3), that entailments in this sense are knowable *a priori* (!), and express *conceptual analyses*.
 - d) But look first at how he has gotten to the claim that *all* Φ -worlds are Ψ -worlds. What we knew before was that all Φ' worlds that were *minimal* are Ψ worlds. But that all Φ -worlds meeting some additional condition are Ψ -worlds is *not* sufficient for *entailment* of Ψ by Φ .
 - e) FJ requires: “ Φ must contain some such clause as ‘and that is all’—the ‘stop’ clause—in order to be true only at minimal physical duplicates of the actual world. So when I say physicalists are committed to the story about our world as told in purely physical terms entailing *inter alia* its psychological nature, I am ruling that a clause like ‘and that is all’ when attached to a purely physical story preserves its purely physical character.” [26] Is this legitimate? FJ says this is a ‘ruling’, as though it were a legitimate matter for *stipulation*. And that, at least, seems wrong about the character of the claim.
 - i. We might grant, for the sake of argument, that the physicist can say in her own language something like: “And this is the *whole* physical story. There are no more physical facts beyond the ones included in the story I have told you about the world—no more particles, no more physical properties possessed by particles, no more forces....”

- ii. But that is *not* the force of FJ's 'stop' clause. That says "And nothing else is true except what *must* (necessarily) be true in order for these physical facts to hold." So it says **w is a Φ -world if and only if it is a Φ' -world and it is also a σ -world, for some σ expressed in any vocabulary whatsoever** (not just the language of physics), which is not contained in Φ' , **if and only if every Φ' -world is a σ -world.**
- iii. That is, formulating the 'and that is all' clause FJ wants requires quantifying over *all possible vocabularies*, over *everything sayable* (or at least all possible *descriptions* in any vocabulary at all).
- iv. Now it is not clear whether this actually makes sense.
- v. But even if it does, you certainly cannot say *that* in the language of physics. And in *whatever vocabulary* one *could* say it, its *truth* will certainly not *follow* from the truth of *physics+ Φ'* . This is not an issue on which FJ has the authority to make a '**ruling**'. The right word is 'hope', or perhaps, 'commitment'. And there are pretty good reasons to think that the claim that the "stop-clause" can be formulated in the language of physics is just false.
- vi. Further, the psychological sentence Ψ is such a σ . So the 'stop' clause ($\Phi-\Phi'$) has as a consequence that Ψ is true in a Φ -world only if *every* Φ' -world is a Ψ -world. If that is right, the *minimal* physical duplicates claim seems to collapse into just global supervenience on the physical, restricted to the *actual* physical facts.
- f) A not wholly irrelevant side issue: FJ seems to think that the claim p : "The average house today is less than 1000 square meters," is *equivalent* to a conjunction of claims about the sizes of n particular houses and a "that's all the houses" claim (and a division by n). But the truth *conditions* of the claims are *not at all* alike. For there are lots of other ways p *could* be true. It is at most *entailed* by the claim he makes. Compare: $\forall x[Fx \rightarrow Gx]$ is *not* equivalent to $Fa_1 \& Ga_1 \& \dots Fa_n \& Ga_n \& \forall z[z=a_1 \vee z=a_2 \vee \dots z=a_n]$. Again, the latter claim only entails the former. And note that the latter still needs a *full, wide-open* quantifier. So why not just use that to begin with? (Anil Gupta took me to task for making this mistake after an injudicious (and unnecessary) move, taken from Stephen Leeds, in my early "Truth Talk" essay.)
- g) A key point is that FJ needs *two* "that's all" claims in Φ :
 - i. "that's all the *physical* facts," and
 - ii. "that all": just the physical facts *and* whatever other facts *must necessarily* hold *if* those are the physical facts *and all* the physical facts.
- (i) is more or less OK. But (ii) quantifies both over *all* (accessible) *worlds* ("must necessarily") *and* over all *kinds* of facts (individuated by the *vocabularies* in which they are expressible, which provide the individuating-and-identifying fact *sortals*). If Φ is to be *physical* facts, then "all the worlds" must be all the *physically possible* worlds. And then we seem to have a *circularity* worry. (ii) *cannot* be restricted to the physical, on pain of trivializing the claim.
- h) FJ says: "Those unhappy with the **ruling** will have to say that physicalists are committed to the story about our world as told in purely physical terms (except for the stop clause) entailing *inter alia* its psychological nature." [26] That is, on this rendering, it is Φ' that entails Ψ , not Φ . But (B), at least, gives us *no* reason to think *this* entailment holds. (B) simply does not address *this* issue. That is, we just have to

move to a completely different view from FJ's, if his "ruling" is rejected as false or unwarranted.

For Part III—Intensional Semantics:

10) Propositional *modal* logic is what you get if you add to the propositional calculus (formulated, say, with the conditional C and negation N), two modal operators, one expressing *necessity* (L), and the other *possibility* (M), related by the principle that $Lp \approx NMNp$ and (so) $Mp \approx NLNp$, subject to the new rule of *necessitation*: if $\vdash \neg p$ then $\vdash \neg Lp$. This says that if p is logically valid, so is Lp.

Some Axioms:

$$A0: \quad LCpq \vdash \neg CLpLq.$$

[Note that it is a notorious modal *fallacy* to think that $LCpq \vdash \neg CpLq$ —that if, necessarily if p then q, and p, then necessarily q.]

$$A1: \quad Lp \vdash \neg p.$$

$$A2: \quad p \vdash \neg \neg Lp.$$

$$A3: \quad Lp \vdash \neg \neg Lp.$$

Some Systems:

$$K: \quad A0$$

$$T: \quad A0 + A1$$

$$B: \quad A0 + A1 + A2$$

$$S4: \quad A0 + A1 + A3$$

$$S5 \quad A0 + A1 + A2 + A3$$

Algebraic Conditions:

Reflexivity: $\forall x[xRx]$.

Symmetry: $\forall x,y[xRy \rightarrow yRx]$.

Transitivity: $\forall x,y,z[(xRy \ \& \ yRz) \rightarrow xRz]$.

A Kripke frame is an algebraic relational structure $\langle W, R \rangle$ consisting of a *domain* W of possible worlds, and a dyadic *relation* R of *accessibility* or *relative possibility* between worlds.

Kripke Frames:

K is validated by *all* Kripke frames.

T is validated by all *reflexive* Kripke frames.

B is validated by all *reflexive* and *symmetric* Kripke frames.

S4 is validated by all *reflexive* and *transitive* Kripke frames.

S5 is validated by all *reflexive*, *symmetric*, and *transitive* Kripke frames.

11) From modal logic to intensional semantics:

- Extensions and intensions. Intensional concepts.
- Lewis's "General Semantics". Adverbs (attributive and non-attributive) as an example of the mathematical characterization of *meanings*:

General Semantics:**Syntax:**

Basic categories: sentences S and singular terms T.

Derived categories: $X \rightarrow Y$ (more generally, $X_1 \dots X_n \rightarrow Y$) takes item of basic or derived syntactic category X and produces from it an item of basic or derived category Y.

Examples: $T \rightarrow S$ takes a term as input, yields a sentence as output; unary predicates are of this derived syntactic category. $(T \rightarrow S) \rightarrow S$ takes a unary predicate as input, yields a sentence as output; quantifiers are of this derived syntactic category.

Semantics:

Basic semantic interpretants: sets of possible worlds for sentences, objects for terms. (Or, sets of assertibility conditions for sentences and recognition conditions for terms; or inferential roles for sentences and substitutional roles for terms; or...). Sentence p entails q just in case the set of possible worlds associated with p (intuitively: the worlds in which it is true) is a subset of the set of possible worlds associated with q : all p -worlds are q -worlds.

Semantic interpretants of derived categories: Semantic interpretant of a unary predicate, $f_p(T \rightarrow S)$, is a function from objects to sets of possible worlds (intuitively, all the worlds in which the object has the property).

Example: Adverbs take unary predicates (e.g. ‘walks’) into unary predicates (e.g. ‘walks slowly’), so are of derived syntactic category $(T \rightarrow S) \rightarrow (T \rightarrow S)$. So their semantic interpretants should be functions *from*: functions from objects to sets of possible worlds *to*: functions from objects to sets of possible worlds.

Adverbs can be divided semantically into *attributive* and *non-attributive*, depending on whether φ -ing A-ly entails φ -ing: anyone who walks slowly walks, but not everyone who walks in their imagination walks. And now we can represent this semantic distinction precisely and algebraically, in terms of the set-theoretic relations between the domains and ranges of the functions-from-functions-to-functions semantically associated with the different classes of adverbs.

- c) Content and character (Kaplan) for indexicals, hence the need for *double* indexing. *Character* (e.g. the repeatable lexical type ‘I’) + circumstances of *utterance* (who is talking) determine *content* (the object to be tracked rigidly through worlds), which in general is a function from context of *evaluation* to a semantic value (in this case, a person in the world, usually different from the one in which the indexical was uttered, at which some ‘I’ claim is being evaluated).
- d) Kripke on reference-fixing, rigid designators, *a posteriori* necessities, and contingent *a priori*.
- e) Putnam’s Twin Earth, H_2O/XYZ case (recall my qualms about the notion of intrinsic property that is needed to get the notion of duplicate required for these cases), and semantic externalism.

For Part IV—Two Dimensional Intensional Semantics:

12) The basic ideas:

- a) Abraham Lincoln (that old modal logician): “Q: If we all agree to call a tail a ‘leg’, how many legs would horses have? A: Four, because you can’t change how many legs horses have by deciding to talk differently.”

- b) Two dimensional modal logic—better called “two-dimensional intensional semantics”—is what you get if you deny the Lincoln truism (or, less tendentiously, decide to explore the dimension of counterfactual that one is being invited into by the question, and which Lincoln’s answer rejects).
- c) Generalize Kaplanian *character* on the model of Kripkean reference-fixing.
- d) *De dicto* and *de re* for temporal case (“The President of the U.S. will be a woman by 2010.”) and for believers (“McCarthy believed the first sentence of the Communist Manifesto was true.”).
- e) Application of *de dicto / de re* to two modal dimensions: Either first fix reference and then follow it across possible worlds (*de re*, C-intension) or first move across possible worlds and then fix reference (*de dicto*, A-intension).
- f) A-intensions of ‘water’ and ‘red’.
- g) Claim is that every sentence expresses *two* propositions, a C-proposition and an A-proposition.
- h) Relations among A-propositions are *a priori*, since they are independent of what world is the actual one.

13) Jackson passages on C-intensions (propositions) and A-intensions (propositions) (from late in Chapter Two):

- a) [48]: “We can think of the various possible particulars, situations, events, or whatever to which a [descriptive] term applies in two different ways, depending on whether
 - i) we are considering what the term applies to under various hypotheses about which world is the actual world, or
 - ii) whether we are considering what the term applies to under various counterfactual hypotheses.
- b) In the first case, we are considering, for each world w , what the term applies to in w , given or under the supposition that w is the actual world, our world. We can call this the A-extension of the T in world w —‘A’ for actual—and call the function assigning to each world the A-extension of T in that world, the A-intension of T.
- c) In the second case, we are considering, for each world w , what T applies to in w given whatever world is in fact the actual world, and so we are, for all worlds except the actual world, considering the extension of T in a counterfactual world. We can call this the C-extension of T in w —‘C’ for counterfactual.
- d) There is no ambiguity about the extension of the term at the actual world, as the A and C-extensions at the actual world must, of course, be the same.”
- e) [49]: “‘Water’ is a rigid designator for the kind common to the watery exemplars we are, or the appropriate baptizers in our language community were, acquainted with. This is what we grasp when we come to understand the word.”
- f) [49]: “In sum, the A-extension of the term ‘water’ in a world is the watery stuff of our acquaintance in that world [which may be XYZ], and the C-extension is H_2O .”
- g) Here is the final move in the argument of Chapter Two: [50]: “When a term’s A-extension and C-extension differ at some worlds—when it is a two-dimensional term, as we might say in honour of the role of two-dimensional modal logic in

making all this explicit—there is a crucial difference between the epistemic status of a term's A-extension and C-extension. To know a term's C-extension we need to know something about the actual world...By contrast, we did know the A-extension of 'water' at every world, for it's A-extension does not depend on the nature of the actual world. Ignorance about the actual world does not matter for knowledge about the A-extensions of words. For the A-extension of T at a world w is the extension of T at w given that w is the actual world, and so does not depend on whether or not w is in fact the actual world....[K]nowledge of the A-intension of T does not require knowledge of the nature of the actual world."

- h) [51]: "What we can know independently of knowing what the actual world is like can properly be called '*a priori*'. The sense in which conceptual analysis involves the *a priori* is that it concerns A-extensions at worlds, and so A-intensions, and accordingly concerns something that does, or does not, obtain independently of how things actually are."