 Last edited Saturday, February 3, 2001

Empirical Critique — Diagnosis

I. Introduction

A. Where are we?

1. We're investigating the (empirical) truth of the ontological sub-reading of the negative reading of 'formality,' in our analysis of the formal symbol manipulation (FSM) construal.
2. Last Thursday, we went through a variety of examples: internal and mathematical. There are also a variety of external examples, covered in the notes, which you might want to look at. They demonstrate more examples of the same basic conclusion: that, in case of input and output, the state of affairs that a structure designates can play a causal role in affecting, or can be causally affected by, how the structure is "manipulated."
3. The provisional moral (informally) drawn from all these cases—internal, mathematical, and external—is that they are *counter-examples* to the ontological reading of formality. Since they are echt computational situations, that suggests that the FSM thesis must be wrong.

B. Reactions

1. As was evident from the strong reactions in the class, these examples didn't sit comfortably with everyone. There are at least two (related) reasons why that could be so:
   a. They genuinely violated what people think it is to be computational, or (more likely) what it is to be formal.
   b. They reflect the depth of our presumption that computation *must* be formal—because that is what we have always been taught. I.e., the reaction may have stemmed from a sense that these examples (in spite of what I claimed) *surely must* be formal.
2. Regarding the second reaction: note that *from the fact that you can give an account of how they work, effectively*—i.e., in a roughly, structural or causal or "syntactic" way—it doesn't follow that they are "formal," in the sense of working "independently of semantics"
3. After all, we are *assuming* that computers and computation can be physically constructed. As a result, it follows that there will *always* be a non-semantical account of how they work! But that in no way secures the truth of the "formality" thesis, for two reasons:
   a. It could be that, although the account of how the system works doesn't refer to anything semantical, it is nevertheless true that the way in which the computation works (to which the account refers) is nevertheless *not independent* of the semantics.
   b. Or, it could be that although the account of how they works isn't framed in semantical terms, it nevertheless *refers to semantical relations*—just in a non-semantical (i.e., causal or reductive) way.
4. If the latter were true, someone might think that that would be evidence that we could give a physicalist (effective, causal) reduction of semantics (at least in this computational case).
5. However that is not true, either. All we would have is *one case* of a causal account of some-

thing genuinely semantic.[1] That isn't a full account. What would count as reduction would be a causal account of semantic phenomena *in general*.

C. Diagnosis
   1. So we need to be very careful, to lay out just exactly what has been demonstrated, what is being claimed, etc.
   2. Last time, everything was very anecdotal—very bottom up.
   3. Today, we need to back off from examples for a bit, and lay out the conceptual structure of what has been going on

II. Conceptual situation

A. There are four issues, outstanding:
   1. First, there are a suite of different issues under consideration
      a. For things, in particular, must clearly be distinguished among
         i. State(s) of affairs **S** that play(s) a causal or effective role in engendering behaviour.
         ii. Properties $\mathbf{P}_{\mathbf{eff}}$ of **S**, in virtue of which it plays the causal/effective role that it does.
         iii. State(s) of affairs **D**, that **S** signifies or refers to or designates or models or …
         iv. Properties $\mathbf{P}_{\mathbf{sem}}$ of **S** in virtue of which it signifies or refers to or designates **D**, or (which may or may not be equivalent) the relation $\mathbf{R}_{\mathbf{sem}}$ holding between **S** and **D**.
      b. Note that 'semantic' and 'syntactic' are *higher-order* properties, holding of $\mathbf{P}_{\mathbf{eff}}$ and $\mathbf{P}_{\mathbf{sem}}$, (or $\mathbf{R}_{\mathbf{sem}}$), respectively.
      c. 'Semantic' can also be used as a regular (first-order) property, holding of **D**.
   2. Second, there is a question of what the "independent of semantics" claim is saying.
      a. What it is claimed *of* is **S**'s playing a causal or effective role, in virtue of properties $\mathbf{P}_{\mathbf{eff}}$, (potency properties—paradigmatically 'syntax")—or perhaps more generally, these properties $\mathbf{P}_{\mathbf{eff}}$ themselves (or, more particularly, **S**'s exemplification of $\mathbf{P}_{\mathbf{eff}}$).
      b. What it *claims* is that this playing of a causal role is independent of either:
         i. States of affairs **D**; or
         ii. The properties $\mathbf{P}_{\mathbf{sem}}$ of **S** in virtue of which it signifies (refers to, designates, etc.) **D**
      c. I will call the former reading (i) the "state of affairs" reading, and the latter (ii) the "property" reading.
      d. Last Thursday, we focused on the "state of affairs" reading. Later, we will come back to the property reading—which may be closer to what most people think formality means.
      e. For now, though, it will be important for us to continue to focus on the state of affairs reading first. In a moment we will be able to explain why this is so.
   3. Third, there is a question of what "independent" means. As we will see next week, there are both "extensional" and "intensional" readings of "independent" (as well as a host of different *strengths* of independence, of the same sort that characterise necessity: ontological, nomological, metaphysical, conceptual or logical, etc.
   4. Fourth, there is a question of what *establishes* the semantic relation $\mathbf{R}_{\mathbf{sem}}$, or established semantic properties $\mathbf{P}_{\mathbf{sem}}$, of **D**

---

[1] i.e., in philosophical terms, we would have a *single token reduction*.

a. Some people *assume* that **S** has property $\mathbf{P}_{sem}$, or stands in relation $\mathbf{R}_{sem}$ to **D**, *only in virtue of being so interpreted by a (human) user*—and (equivalently) that computers or computational systems cannot "mean" or "refer" or otherwise be "semantic" on their own.
b. But to *assume* such things won't do, for our project.
c. The reason is simple: the *reason* that most people hold such a view (viz., that computers aren't semantically "original," to use Haugeland's phrase) is *because computers are formal*.
d. What we are trying to figure out, though, is *whether* computers are formal.
e. To *assume* that computers aren't semantically original would be to assume that they are formal, which would be viciously circular.

B. Discussion
1. Given this last discussion, about semantic originality, it start to become clearer why we are focusing on the "states of affairs" reading (of the formality claim).
2. The basic reason is that by focusing on the designated states of affairs **D**, and on the causal role that **D** can sometimes play in engendering causal behaviour, something of great importance becomes manifest that I said last week we were going to demonstrate: that

> ◆ C**omputers are involved in their subject matters**

3. Now if you ask people what it would take, in general, for a system to be able to have original semantics—i.e., to be able to "denote" or "mean" or "refer" or "signify" on its own, and not merely have "derivative" semantics, attributed from the outside, the most common response is that the system would have to be *involved in the realm that it signifies*.
4. So if we can show (as indeed we are showing) that computers are involved in their subject matters, that at least undermines the pre-theoretic assumption that it is obvious that computational semantics are entirely derivative.[2]

C. Plan
1. Today, therefore, we will stick with the "state of affairs" reading: i.e., that the ontological reading of the independence claim says that how computers work is independent of the semantic situations or states of affairs **D** that those computational states designate or refer to (or otherwise bear '⇒' relations to).

---

[2]It should be noted that there are other views as to what is required, in order for semantics to be original. Perhaps the strongest are: (i) that the system would have to be conscious (e.g., Searle), and (ii) that it would have to be subject to *norms* (e.g., Haugeland). If we have time, we will get back to these two criteria at the end of the semester. However just because they are stronger (require more than mere involvement), it doesn't follow that, if one were to adopt them (as conditions on semantic originality), it would thereby be easy to show that computers were (anti-semantically) formal, after all. To do that, one would have to show that computers, in virtue of being computers, were *intrinsically* non-conscious or non-normatively-governed. And that is not so easy to do. Normally, such a conclusion would be argued *based on their formality*. I.e., it would be argued that, because computers are restricted by a stronger constraint (formality), they aren't candidates for these more powerful capacities (consciousness or semantic originality). But if we have successfully argued against those narrower restrictions, it is not obvious that anything would be readily available with which to bar them, in principle, from the greater achievement of consciousness or normative existence.

2. It is this reading, I claim, that the examples we rehearsed last Thursday show to be false.
3. Today we will analyse those counter-examples in more depth, in order to reveal their basic structure—and thereby understand better *why* the "states of affairs" reading is false.
4. What we uncover, by doing that, will be of great utility as we move forward:
   a. Partly it will give us a much better understanding of what computation is genuinely like, which we can use, at the end, when we start to construct a better alternative;
   b. It will also give us a leg up on understanding why people originally thought that computers were ontologically formal (even if, in fact, they are not).

III. Diagnosis

A. OK, given that preamble, turn back to the ontological counter-examples
B. We can get to the heart of things with two distinctions and two theses:
   1. Distinctions: distinguish two "boundaries"
      a. Physical: between the system and its surrounding environment—i.e., a distinction between "inside" and "outside"; and
      b. Semantic: between symbols and their referents
   2. Two theses or propositions, stated in terms of these boundaries:
      a. Alignment: that physical and semantic boundaries line up, with all the symbols inside, and all the referents outside.
      b. Isolation: that this putatively aligned boundary is a "moat"—a barrier or gulf across which "dependence" of various forms (causal, semantic, explanatory) does not reach.
   3. Can see paradigmatic example of FSM in these terms:
      a. A pristine inner world of symbols—a "private kingdom", as it were, thought *both* to work (ontologically) and to be analysable (theoretically) in isolation, without distracting influence from the messy, unpredictable exterior.
      b. It is this image that supports not only the conjunction of the ontological and conceptual readings of antisemantical formality.
C. Critique, part I
   1. First (easier) challenge: retain **alignment**, of coincident boundaries, but relax **isolation** no interaction.
      a. This is classical realm of I/O, of transduction, etc.
      b. Note: this critique of formality is by no means new:
         i. In philosophy: cf. Levitt et al's critique of Fodor: I-T and T-O processing, instead of internal (T-T).[3]
         ii. In cognitive science: cf. Harnad's notion of a "Total Turing Test" (TTT):
         iii. In Artificial Intelligence: cf. the "robot" reply to Searle's Chinese room.
   2. Is violation of **isolation** enough to defeat the ontological (horizontal) reading of FSM?
      a. That is not wholly clear.

---

[3] In *Fodor and His Critics*.

Fodor, when asked, once, whether transducers were computational (and hence formal), is reputed to have said: "Please don't ask me about transducers. I am particularly busy right now."

b. This is why transducers and computation are such uneasy bedfellows.
c. Suppose transducers violate *formality*; maybe they should thereby be counted as not *computational*.
d. I.e., we could say that computational operations, except for input and output, are formal (in the negative, ontological sense)
e. If that is all there were to it, it might be possible to rescue the formality thesis

D. Critique, part II
1. My aim, throughout the critique, has been to challenge **alignment**.
2. Specifically: it is a claim that physical and semantic boundaries *cross-cut*
3. Four types of case needed to show this:
a. Internal symbol, internal referent — e.g., quotation, meta-levels, e-mail, etc.
b. Internal symbol, external referent — e.g., internal representation of Pluto
c. External symbol, internal referent — e.g., specifications, proofs of correctness, etc.
d. External symbol, external referent — e.g., skull-and-cross-bones signifying radioactivity
4. Note, however, that breaking down **alignment** not enough to defeat FSM
5. Rather, breaking down **alignment** undermines **isolation**

E. Transducers
1. Given that **alignment** is false, one must then break down the notion of transduction (that about which Fodor is "too busy to think") into two different kinds:
a. Physical transducers: operations or modules that mediate (cross the boundary) between the inside and the outside—i.e., between a system and its surrounding environment; and
b. Semantic transducers: operations or modules that mediate (cross the boundary) between symbols and their referents
2. The only remotely plausible notion of transducer, in practice, are **physical**.[4]
a. Cf. vision, touch, smell, articulation, wheels, muscles, and the like
b. Systems that mediate between the internals of a system and the "outside" world
3. The notion of transducer that challenges FSM, however (and vexed Fodor), are **semantic.**
a. "Antisemantics is challenged as much by disquotation as by driving around"
4. So the only way to preserve FSM (on this reading) would be by excluding (from computation) the operations of *semantic* transducers.
a. This would be a *spectacularly infeasible* way to draw and quarter any real-life, system
b. Also, to *define* transducer in this semantical way, and then to remove them from consideration (as not being genuinely computational), would make computation (minus such transducers) antisemantical *tautologically*.
c. So the FSM claim on computation would be true because it was *stipulated*, emptying it of interest.
d. Furthermore, not only are these cases of "semantic transduction" well-behaved; they also seem, at least intuitively, *to be as "formal" as any other kind of operation.*
e. If that is so, then either

---

[4]One can argue that not even *this* reading correctly gets at the notion of transduction. See lecture notes 5b.

i. Those systems either are not formal, after all,
ii. *Or else the word 'formal' has never meant independence of syntax and semantics in the way that the horizontal claim construes it*

F. Moral
1. It is time to collect together what we have learned.
2. The moral is not (or anyway not just): that computers are involved with *external* subject matters (as various recent "situated" movements might imply).
3. Rather, moral is that computers are engaged in the world *as a whole*—in a world that indiscriminately includes themselves, their own internal states and processes.
4. It is that *total* involvement—constitutive of the defeat of **alignment**—that defeats FSM
5. It is that *total* involvement that our eventual positive reconstruction will have to honour.

—— *end of file* ——