Phil 100D: Philosophy of Mind Professor Aaron Zimmerman

Handout #10: Dennett

1. Two Kinds of Metaphysics of Mind: Realism v. Interpretationism

Realism: "Realism likens the question of whether a person has a particular beief to the question of whether a person is infected with a particular virus—a perfectly objective internal matter of fact about which an observer can often make educated guesses of great reliability"(556).

Interpretationism: "Interpretationism likens the question of whether a person has a given belief to the question of whether a person is immoral, or has style, or talent, or would make a good wife" (556-7).

Dennett's Hybrid: "While belief is a perfectly objective phenomenon…it can be discerned only from the point of view of one who adopts a certain predictive strategy, and its existence can be confirmed only by an assessment of the success of that strategy" (557). "All there is to being a true believer is being a system whose behavior is reliably predictable via the intentional strategy, and hence all there is to really and truly believing that p (for any proposition p) is being an intentional system for which p occurs as a belief in the best (most predictive) interpretation" (564).

Question: Dennett claims that his theory of belief and desire (or preference and credence) walks a line between realism and anti-realism about these states of mind. Is that true? Dennett use the reality of theoretical posits as an example — e.g. centers of gravity or the equator. There are "real patterns" that objectively constrain interpretations, but there are several divergent interpreations of an individual's behavior, which meet the constraints imposed by these patterns, and no fact of the matter which of these "candidate" interpretations is objectively correct.

> "We can be sure in advance that no intentional interpretation of an individual will work to perfection, and it may be that two rival schemes are about equally good, and better than any others we can devise. That this is the case is itself something about which there can be a fact of the matter. The objective presence of one pattern (with whatever imperfections) does not rule out the objective presence of another pattern (with whatever imperfections)" (563).

Indeed, in cases of psychopathology, the patient may exhibit few of these "real patterns" and there might therefore be no fact of the matter at all about what the patient believes (see p. 563).

Questions: Do your beliefs and desires have this degree of reality? How do you know what you believe? Do you beliefs about what you believe have a kind of authority that allows you to select one of these interpretations of your behavior as better than all the others?

2. How the intentional strategy works

The astrological strategy—predicting a person's behavior based on her birthday. Unreliable and no conceivable mechanism. Is impugned too by the realization that people born on the same day act differently. (There must be more to it than this, but that's how Dennett characterizes it.)

Physical Strategy—Predicting a system's behavior based on chemical or physical make-up. This has great efficacy when applied to predict the dynamics of systems like the solar system. And it has some efficacy when applied to animals. Mood might be predicted on the basis of hormonal state and other bio-chemical factors (e.. neurotransmitter levels). And behavior is affected by mood. But it is not currently possible to explain or predict specific intentional acts on the basis of biochemical information alone.

The Design Strategy—e.g. alarm clocks. Functional and evolutionary explanations. How do these integrate with biochemical explanations in, e.g., in cognitive neuroscience explanations of sound perception?

The Intentional Strategy: (a) An input side—we assume the animal or person believes all available facts that are relevant to them (especially those made relevant by their biological design); we assume the agent believes things that their make actions rational in light of attributed desires; in turn we attribute desires for things crucial for the person or animal to act as it is designed to act and we attribute to it those desires that would make its actions rational in light of the attributed beliefs.

Most generally—attribute to the organism the beliefs and desire it **ought** to have given its place in the environment. (We see here that the verdicts arrived at from the design stance provide inputs for interpretations constructed from the intentional stance.)

Criticism: Dennett says that we adopt the intentional stance on ourselves or another person, and we assign them beliefs and desires on the basis of our observations of what they've said and done (given our background knowledge of how they're designed biologically and the environment in which they're functioning) we start by assuming that they are perfectly rational and then relax the assumption when necessary.

"On starts with the ideal of perfect rationality and revises downward as circumstances dictate" (p. 560)

But I think Dennett is wrong on this (though I agree with most of what he says otherwise). Instead**, the person employing the intentional strategy will typically assume that the person whose behavior he or she is trying to explain or predict is as imperfect as he or she (and so neither more nor less "rational" than he or she is) unless given positive evidence to believe otherwise**. This is an element of "simulationist" accounts of the attribution of beliefs and desires, which argue that we attribute these states to others by imagining what we would believe or want were we in the situation we observe the other to occupy.

3. **Further Problems with Dennett's Metaphysics of Belief**: Dennett says what it is to have beliefs and desires is to be effectively attributed them in an effort to predict thing's behavior, but this is too liberal—thermostats, etc. Lectern example. (This is similar to the "too many minds" problems Block raises for funcationalists.)

First answer—must be some predictive gain. Rules out attributing beliefs and desires to the lectern, but not necessarily the thermostat.
Problem: still too weak.

Second answer—behavioral flexibility which requires internal complexity.
Problem—Are there no internal constraints?

"The cumulative effect of enriching…connections between the device and the world in which it resides is to enrich the semantics…the less amenable our device becomes to serving as the control structure of anything else…the class of indistinguishably satisfactory models of the formal system embodied in its internal states gets smaller and smaller as we add such complexities: the more we add the richer or more demanding or specific the semantics of the system, until eventually we reach systems for which a unique semantic interpretation is practically (but never in principle) dictated…*what makes some internal feature of a thing a representation could only be its role in regulating the behavior of an intentional system*" (564-5).

Dennett—"But there is no magic moment in the transition from a simple [system like a] thermostat to a system that really has an internal representation of the world around it."

Questions: what implications does this account of belief have for our accounts of the evolution of minds? Do beliefs evolve gradually? Are there some organisms of whom it is neither true nor false that they have representations of their environments?

Question—What do we say about vagueness?

4. **The Reality of Patterns**—The superscientist Martian and the Earthling have a competition to see who can better predict Mr. Gardner's behavior (pp. 562-3).

Question: Does the fact that the Earthling can predict the other man's behavior in a manner that would shock the Martian (as the Earthling's prediction are nearly as good if not as good as the Martian's despite the earthling's not making any of the calculations the Martian must make) show that the beliefs and desires the Earthling attributes to Mr. Gardener are "real"?

5. Dennett's claims about irrationality—cognitive pathology—no fact of the matter. Must descend to physical stance. True? What about trying to understand what is going on in their minds at the intentional level?

6. Dennett's mild critique of Fodor. Dennett ends by arguing that LOT is an empirical view for which we don't have a ton of evidence. Fodor agrees with this. (Dennett says others don't, but he doesn't cite anyone.). Dennett is right, though, that Fodor's arguments for the LOT hypothesis are relatively conceptual or a priori and it is my sense that scientists are moving away from the LOT. It is no longer the "only game in town."

"I'd like some baked beans please."
"Yes sir. How many?"

Dennett's prior critique of LOT: "Being language using creatures, it is inevitable that we should often come to believe that some particular, actually formulated, spelled and punctuated sentence is true, and that on other occasions we should come to want such a sentence to *come* true, but these are special cases of belief and desire and as such may not be reliable models for the whole domain" (p. 559).

Question: Might Fodor admit that tokens in the LOT are vague and further specifiable?