15. Millikan: Biosemantics

Martín Abreu Zavaleta

June 18, 2014

1 False representations

Millikan, like Dretske, Chisholm and Brentano, thinks that what distinguishes the "directness" of our mental states from other kinds of directness is the fact that our mental states can misrepresent how things are. Like Dretske, she thinks that we can ultimately can explain what is special about our mental states in completely natural terms, but she disagrees in the details.

In particular, she offers a damaging observation to Dretske's theory: Dretske and others think that natural representation (as opposed to mental representation) can be achieved by causal means, and furthermore, that we can come to understand mental representation in terms of natural representation under normal circumstances.

For instance, we can think that a given brain state represents the state of there being a cow because under normal circumstances it is activated or present whenever its bearer perceives cows. But what are the normal circumstances? As we saw when we introduced the disjunction problem, it's hard to restrict the class of relevant things. And saying that these circumstances are statistically normal conditions won't help, for to gather such statistics, we would have to know already of the relevant comparison classes.

She thinks that Dretske's proposal doesn't work. Dretske assimilates having the function of representing R to being a natural sign of R when the system functions normally, but Millikan objects:

> Now the production of natural signs is undoubtedly an accidental side effect of normal operation of many systems. From my red face you can tell that either I have been exerting myself, or I have been in the heat, or I am burned. But the production of an accidental side effect, no matter how regular, is not one of a system's functions; that goes by definition. More damaging, however, it simply is not true that representations must carry natural information. Consider the signals with which various animals signal danger. Nature knows that it is better to err on the side of caution, and it is likely that many of these signs occur more often in the absence than in the presence of any real danger.

Millikan is making two points: first, that in most cases in which natural signaling occurs, it is a side effect rather than one of the functions of the signaling element. She gives the example of the red face, but we can think of others: take our compass. We said that the pointer of the compass indicated the direction of the north pole because the position of the pointer depends causally on the location of the north pole. But Millikan would say that the signaling here is just a side effect of the causal interaction between the metal in the pointer and the magnetic field of the north pole. If we

had to talk about a natural function for the pointer in the compass, we would say that its function is to interact in such and such ways with magnetic fields, and as a side effect, the pointer is directed towards the north pole in ordinary circumstances. But, by definition, an accidental feature or side effect of a system cannot be part of its function, so functions as used by Dretske won't help much.

Her second point is more damaging. Dretske defines the natural function of a given element in a system in terms of the events that cause that element to appear in ordinary circumstances. Millikan offers a simple counterexample: suppose that the sound made by a certain animal signals danger. It would be in its best interest to signal danger more often in the absence than in the presence of real danger, in which case danger would not cause the signaling in ordinary circumstances. So Dretske's account doesn't explain how the sound can signal danger.

Still, she agrees with Dretske about the appropriateness of the general strategy: she thinks that what makes something a representation is that it has a certain function. However, she understands this differently:

> the way to unpack this insight is to focus on representation consumption, rather than representation production. It is the devices that use representations which determine these to be representations and, at the same time, determine their content. If it really is the function of an inner representation to indicate its represented, clearly it is not just a natural sign, a sign that you or I looking on might interpret. It must be one that functions as a sign or representation for the system itself.

Millikan's aim in this paper is to explain why this is the right way to understand representation.

Preliminaries: what Millikan means by 'normal'

When Millikan talks about biological functions, she doesn't think that biological functions are determined only by origin. It may be, for instance, that a given element in a system is preserved due to their performance of new functions that they were not selected for. For instance, perhaps fingers were selected because of their capacity to hold branches of trees or something like that, but they may acquire new functions later, for instance, the function of handling very small things, in which case this will also be part of their function. Moreover, the notion of 'design' should be understood in a way that makes room for designs capable of improving themselves or to be altered by experience. For instance, we have good reason to think that our brain is like this.

Finally, when Millikan talks about normal conditions, she is not thinking about statistically normal conditions: "the normal condition for performance of a function is a condition, the presence of which must be mentioned in giving a flu normal explanation for performance of that function." For instance, normal conditions for discriminating colors are not the same as normal conditions for discriminating smells. Some functions are only performed rarely or under statistically rare conditions, so the notion of statistical normality is not the best for our purposes.

2 Information producers vs. information consumers

Let's divide a system with intentional states into two aspects: one which produces representations and one which consumes or uses the representations. For instance, suppose we have some element in the system that produces natural information in Dretske's sense. That information will be useless unless it is understood by the system, and in particular, unless the consumer part of the system can

interpret these bearers of information adequately. But then, assuming that the understandings of the bearers are systematically related to the structures of the signs taken to carry some specific piece of information, we can obtain a set of rules determining the meanings of each of those signs or information carriers. The meaning of those signs will be whatever allows the consumer part of the system perform its full function.

Millikan expresses this point as follows:

> [First,] the representation and the represented [must] accord with one another, so [by a certain rule] is a normal condition for proper functioning of the consumer device as it reacts to the representation [...] The content hangs only on there being a certain condition that would be normal for performance of the consumer's function—namely, that a certain correspondence relation hold between sign and world. (pp. 502–503)
>
> Second, represented conditions are conditions that vary, depending on the form of the representation, in accordance with specifiable correspondence rules that give the semantics for the relevant system of representation.

She illustrates with two examples. First, beavers splash water with their tails to signal danger. Splashing means danger for the beaver because the fact that the splash is related to danger is a normal condition for the fulfillment of the interpretive mechanism's function, namely, to interrupt activities and hide. If it wasn't the case that in normal conditions for the fulfillment of the function of self-preservation, splashing was correlated with danger, then splashing wouldn't mean danger.

Notice that the account doesn't require that there is danger whenever the beaver splashes the water, or even that there is danger in most such cases. It suffices that the splashes are correlated with danger often enough.

Millikan examines an example to show how her approach is superior to others. There are some species of bacteria in the northern hemisphere for which oxygen-rich water is toxic. They run away from that water by attending to their magnetosomes, tiny magnets that pull towards the north. The function of the magnetosomes seems to be to make the bacterium move into oxygen-free water, and similarly, we would like to think that the pull of the magnetosome represents the location of oxygen-free water. Clearly, what causes the magnetosome to pull one way rather than other is the magnetic field of the north pole, and it is indeed a function of the magnetosome to signal the north pole, so according to Dretske's account, the pull of the magnetosome signals the north pole, which is the wrong result.

Can Millikan's view do better? She thinks it does. The reason that the pull of the magnetosome represents the location of oxygen-free water rather than the north pole is that only the correlation between oxygen-free water and the pull of the magnetosome will allow the consumer part of the system perform its proper function: to move away from the oxygen-rich water. If there was no such relation, the consumer part would not be able to perform that function.

As Millikan points out, the same sort of representational state may represent different things in different systems, which seems to be an advantage. To a frog, a black point in the retina may represent food, but to a bug that same black point may represent a potential mate. Similarly, many different things may represent the same thing for a single system.

Of course, individual beliefs may not have proper functions like the ones required by Millikan's account, and whatever represents in our brains the claims of economics and computer science can't be thought to have been evolutionarily selected. Millikan's reply is that it is the whole brain that

gets selected, and it has the mechanisms to learn. We can institute the picture of proper functioning even for the things we learn, in pretty much the way advocated by Dretske.

Millikan presents some replies to putative objections, but we won't examine them. Instead, we will examine more general objections to teleological accounts of intentionality.

3   General objections

3.1   Causal indeterminacy

Fodor and others raise the following kind of objection: evolutionary selection processes are not fine-grained enough to track meaning or other intentional things. Consider, for instance, the following case: frogs snap at anything that is small, dark, and moves in order to feed themselves. Presumably, this mechanism is adaptive and was selected because it was good at getting frogs to feed themselves. However, let's suppose that the only thing that frogs eat are flies, all of which happen to be small, dark things that move.

What should we say that the state represents? Should we say that the frog's state represents flies, or that it represents small, dark things that move? In other words, the objection is that causal relations are not fine-grained enough to account for our seemingly determinate contents. Perhaps it doesn't matter to the frog whether flies are different from small, dark things, but these kind of fine-grained distinctions matter to us. We often make distinctions between two different properties that have the same causal profiles, and indeed it seems to be part of the defining characteristics of our mental states that they can do such fine grained distinctions.

As we saw, Dretske answers this objection by appealing to counterfactuals, and Millikan appeals to the notion of normal conditions and proper functions. Do their replies succeed? Recall that Dretske and Millikan's aim is to naturalize intentionality: to give a scientifically respectable account of intentionality, in terms that don't appeal to non-physical or non natural things.

Appealing to counterfactuals in this context might seem suspicious. For one thing, we may always ask: what is the physical fact that makes a counterfactual judgment true or false? Whether counterfactuals have physical verifiers is not very clear, but let's suppose that they do. In this case, Dretske hasn't offered a full account of intentionality yet, for he only offers an account in terms of causal relations, yet he accepts that it is only causal relations plus some counterfactual truths that constitute a complete account. He has yet to offer an account of the natural features that make those counterfactual truths.

On the other hand, many people have seen the notion of a proper function as suspect. For instance, it may seem like we can't define the notion of a proper function without the notion of a normal condition. But the way that Millikan understands normal conditions is by reference to the notion of a proper function. This is blatantly circular, but perhaps we could escape the circularity by defining either notion independently. How this is to be done is not very clear.

3.2   Swampman

Donald Davidson invites us to imagine the following case. Out of nowhere, by some freak quantum accident, a being appears that happens to be an exact physical replica of, say, Davidson at the time. It undergoes exactly the same physical and chemical processes at that time, but there is no historical

relation between that being—call it swamp man—and Davidson. Swampman has no history before it appeared.

Given that Swampman undergoes exactly the same brain processes as Davidson, it would seem that whatever mental description is true of Davidson is also true of the swampman. However, if the teleological theories we have seen so far are correct, this would not be the case. Swampman's brain and brain states were not selected by any process whatsoever—whether natural or unnatural—and neither do they have functions or proper functions in the biological sense. Those functions are defined historically, but Swampman has no history before the moment of its appearance. If teleological theories of intentionality were true, then we could not accurately describe Swampman as having the same mental states as Davidson, but this doesn't seem to be the right result.