24.09x Minds and Machines

Hilary Putnam, 'The meaning of 'meaning''

Excerpts from Hilary Putnam, 'The meaning of 'meaning'', *Minnesota Studies in the Philosophy of Science* 7:131-93 (1975).

This excerpt from Putnam's original (and lengthy) paper has three parts. In the first part, Putnam explains a pair of important terms from semantics, 'extension' and 'intension'. In the second part, Putnam clarifies "the traditional notion of a mental state" and states an assumption Putnam thinks many philosophers have made, "the assumption of methodological solipsism" (this is mentioned at the start of lecture 17). In the third part, Putnam presents his famous "Twin Earth" example, designed to show that "meanings ain't in the head". We will use this example to argue that (in the sense explained in the lectures) *beliefs* ain't in the head.

> *Meaning and extension*. Since the Middle Ages at least, writers on the theory of meaning have purported to discover an ambiguity in the ordinary concept of meaning, and have introduced a pair of terms—*extension* and *intension*, or *Sinn* and *Bedeutung*, or whatever—to disambiguate the notion. The extension of a term, in customary logical parlance, is simply the set of things the term is true of. Thus, "rabbit," in its most common English sense, is true of all and only rabbits, so the extension of "rabbit" is precisely the set of rabbits. Even this notion—and it is the least problematical notion in this cloudy subject—has its problems, however. Apart from problems it inherits from its parent notion of *truth*, the forgoing example of "rabbit" *in its most common English sense* illustrates one such problem: strictly speaking, it is not a term, but an ordered pair consisting of a term and a "sense" (or an occasion of use, or something else that distinguishes a term in one sense from the same term used in a different sense) that has an extension. Another problem is this: a "set," in the mathematical sense, is a "yes-no" object; any given object either definitely belongs to S or definitely does not belong to S, if S is a set. But words in a natural language are not generally "yes-no": there are things of which the description "tree" is clearly true and things of which the description "tree" is clearly false, to be sure, but there are a host of borderline cases. Worse, the line between the clear cases and the borderline cases is itself fuzzy. Thus the idealization involved in the notion of extension—the idealization involved in supposing that there is such a thing as the set of things of which the term "tree" is true—is actually very severe.

Recently some mathematicians have investigated the notion of a fuzzy set—that is, of an object to which other things belong or do not belong with a given probability or to a given degree, rather than belong "yes-no." If one really wanted to formalize the notion of extension as applied to terms in a natural language, it would be necessary to employ "fuzzy sets" or something similar rather than sets in the classical sense.

The problem of a word's having more than one sense is standardly handled by treating each of the senses as a different word (or rather, by treating the word as if it carried invisible subscripts, thus: "rabbit$_1$"—animal of a certain kind; "rabbit$_2$ — coward; and as if "rabbit$_1$" and "rabbit$_2$" or whatever were different words entirely). This again involves two very severe idealizations (at least two, that is): supposing that words have discretely many senses, and supposing that the entire repertoire of senses is fixed once and for all. Paul Ziff has recently investigated the extent to which both of these suppositions distorts the actual situation in natural language; nevertheless, we will continue to make these idealizations here.

Now consider the compound terms "creature with a heart" and "creature with a kidney." Assuming that every creature with a heart possesses a kidney and vice versa, the extension of these two terms is exactly the same. But they obviously differ in meaning. Supposing that there is a sense of "meaning" in which meaning = extension, there must be another sense of "meaning" in which the meaning of a term is not its extension but something else, say the "concept" associated with the term. Let us call this "something else" the *intension* of the term. The concept of a creature with a heart is clearly a different concept from the concept of a creature with a kidney. Thus the two terms have different intension. When we say they have different "meaning," meaning = intension.

…

"Psychological state" and methodological solipsism. …[W]e need …to clarify the traditional notion of a psychological state. In one sense a state is simply a two-place predicate whose arguments are an individual and a time. In this sense, *being five feet tall*, *being in pain, knowing the alphabet*, and even *being a thousand miles from Paris* are all states. (Note that the time is usually left implicit or "contextual"; the full form of an atomic sentence of these predicates would be "*x is five feet tall at time t*," "x is *in pain at time t,*" etc.) In science, however, it is customary to restrict the term state to properties which are defined in terms of the parameters of the individual which are

fundamental from the point of view of the given science. Thus, being five feet tall is a state (from the point of view of physics); being in pain is a state (from the point of view of mentalistic psychology, at least); knowing the alphabet might be a state (from the point of view of cognitive psychology), although it is hard to say; but being a thousand miles from Paris would not naturally be called a *state*. In one sense, a psychological state is simply a state which is studied or described by psychology. In this sense it may be trivially true that, say, knowing *the meaning of the word "water"* is a "psychological state" (viewed from the standpoint of cognitive psychology). …

When traditional philosophers talked about psychological states (or "mental" states), they made an assumption which we may call the assumption of methodological solipsism. This assumption is the assumption that no psychological state, properly so called, presupposes the existence of any individual other than the subject to whom that state is ascribed. (In fact, the assumption was that no psychological state presupposes the existence of the subject's body even: if *P* is a psychological state, properly so called, then it must be logically possible for a "disembodied mind" to be in *P*.) This assumption is pretty explicit in Descartes, but it is implicit in just about the whole of traditional philosophical psychology. Making this assumption is, of course, adopting a *restrictive program*—a program which deliberately limits the scope and nature of psychology to fit certain mentalistic preconceptions or, in some cases, to fit an idealistic reconstruction of knowledge and the world. Just how restrictive the program is, however, often goes unnoticed. Such common or garden variety psychological states as *being jealous* have to be reconstructed, for example, if the assumption of methodological solipsism is retained. For, in its ordinary use, *x is jealous of y* entails that *y* exists, and *x is jealous of y's regard for z* entails that both *y* and *z* exist (as well as *x*, of course). Thus *being jealous* and *being jealous of someone's regard for someone else* are not psychological states permitted by the assumption of methodological solipsism. (We shall call them "psychological states in the wide sense" and refer to the states which are permitted by methodological solipsism as "psychological states in the narrow sense.") The reconstruction required by methodological solipsism would be to reconstrue *jealousy* so that I can be jealous of my own hallucinations, or of figments of my imagination, etc. Only if we assume that psychological states in the narrow sense have a significant degree of causal closure (so that restricting ourselves to psychological states in the narrow sense will facilitate the statement of

psychological *laws)* is there any point to engaging in this reconstruction, or in making the assumption of methodological solipsism. But the three centuries of failure of mentalistic psychology is tremendous evidence against this procedure, in my opinion.

…

Are meanings in the head? That psychological state does not determine extension will now be shown with the aid of a little science fiction. For the purpose of the following science-fiction examples, we shall suppose that somewhere in the galaxy there is a planet we shall call Twin Earth. Twin Earth is very much like Earth; in fact, people on Twin Earth even speak *English*. In fact, apart from the differences we shall specify in our science-fiction examples, the reader may suppose that Twin Earth is exactly like Earth. He may even suppose that he has a *Doppelgänger*—an identical copy—on Twin Earth, if he wishes, although my stories will not depend on this.

Although some of the people on Twin Earth (say, the ones who call themselves "Americans" and the ones who call themselves "Canadians" and the ones who call themselves "Englishmen," etc.) speak English, there are, not surprisingly, a few tiny differences which we will now describe between the dialects of English spoken on Twin Earth and Standard English. These differences themselves depend on some of the peculiarities of Twin Earth.

One of the peculiarities of Twin Earth is that the liquid called "water" is not H_2O but a different liquid whose chemical formula is very long and complicated. I shall abbreviate this chemical formula simply as XYZ. I shall suppose that XYZ is indistinguishable from water at normal temperatures and pressures. In particular, it tastes like water and it quenches thirst like water. Also, I shall suppose that the oceans and lakes and seas of Twin Earth contain XYZ and not water, that it rains XYZ on Twin Earth and not water, etc.

If a spaceship from Earth ever visits Twin Earth, then the supposition at first will be that "water" has the same meaning on Earth and on Twin Earth. This supposition will be corrected when it is discovered that "water" on Twin Earth is XYZ, and the Earthian spaceship will report somewhat as follows:

> "On Twin Earth the word 'water' means XYZ."

(It is this sort of use of the word "means" which accounts for the doctrine that extension is one sense of "meaning," by the way. But note that although

"means" does mean something like *has as extension* in this example, one would *not* say

"On Twin Earth the meaning of the word 'water' is XYZ"

unless, possibly, the fact that "water is XYZ" was known to every adult speaker of English on Twin Earth. We can account for this in terms of the theory of meaning we develop below [not included in this excerpt]; for the moment we just remark that although the verb "means" sometimes means "has as extension," the nominalization "meaning" *never* means "extension.")

Symmetrically, if a spaceship from Twin Earth ever visits Earth, then the supposition at first will be that the word "water" has the same meaning on Twin Earth and on Earth. This supposition will be corrected when it is discovered that "water" on Earth is H_2O, and the Twin Earthian spaceship will report:

"On Earth the word 'water' means H_2O."

Or rather, they will report: "On Twin Earth (the Twin Earthian name for Terra—H.P.) the word 'water' means H_2O."

Note that there is no problem about the extension of the term "water." The word simply has two different meanings (as we say) : in the sense in which it is used on Twin Earth, the sense of $water_{TE}$, what we call "water" simply isn't water; while in the sense in which it is used on Earth, the sense of $water_{E}$, what the Twin Earthians call "water" simply isn't water. The extension of "water" in the sense of $water_{E}$ is the set of all wholes consisting of H_2O molecules, or something like that; the extension of water in the sense of $water_{TE}$ is the set of all wholes consisting of XYZ molecules, or something like that.

Now let us roll the time back to about 1750. At that time chemistry was not developed on either Earth or Twin Earth. The typical Earthian speaker of English did not know water consisted of hydrogen and oxygen, and the typical Twin Earthian speaker of English did not know "water" consisted of XYZ. Let $Oscar_1$ be such a typical Earthian English speaker, and let $Oscar_2$ be his counterpart on Twin Earth. You may suppose that there is no belief that $Oscar_1$ had about water that $Oscar_2$ did not have about "water." If you like, you may even suppose that $Oscar_1$ and $Oscar_2$ were exact duplicates in appearance, feelings, thoughts, interior monologue, etc. Yet the extension of the term "water" was just as much H_2O on Earth in 1750 as in 1950; and the

extension of the term "water" was just as much XYZ on Twin Earth in 1750 as in 1950. $Oscar_1$ and $Oscar_2$ understood the term "water" differently in 1750 although they were in the same psychological state, and although, given the state of science at the time, it would have taken their scientific communities about fifty years to discover that they understood the term "water" differently. Thus the extension of the term "water" (and, in fact, its "meaning" in the intuitive preanalytical usage of that term) is not a function of the psychological state of the speaker by itself.

But, it might be objected, why should we accept it that the term "water" had the same extension in 1750 and in 1950 (on both Earths)? The logic of natural-kind terms like "water" is a complicated matter, but the following is a sketch of an answer. Suppose I point to a glass of water and say "this liquid is called water" (or "this is called water," if the marker "liquid" is clear from the context). My "ostensive definition" of water has the following empirical presupposition: that the body of liquid I am pointing to bears a certain sameness relation (say, *x is the same liquid as y*, or *x is the same*$_L$ *as y*) to most of the stuff I and other speakers in my linguistic community have on other occasions called "water." If this presupposition is false because, say, I am without knowing it pointing to a glass of gin and not a glass of water, then I do not intend my ostensive definition to be accepted. Thus the ostensive definition conveys what might be called a defeasible necessary and sufficient condition: the necessary and sufficient condition for being water is bearing the relation same$_L$ to the stuff in the glass; but this is the necessary and sufficient condition only if the empirical presupposition is satisfied. If it is not satisfied, then one of a series of, so to speak, "fallback" conditions becomes activated.

The key point is that the relation same$_L$ is a *theoretical* relation: whether something is or is not the same liquid as *this* may take an indeterminate amount of scientific investigation to determine. Moreover, even if a "definite" answer has been obtained either through scientific investigation or through the application of some "common sense" test, the answer is *defeasible:* future investigation might reverse even the most "certain" example. Thus, the fact that an English speaker in 1750 might have called XYZ "water," while he or his successors would not have called XYZ water in 1800 or 1850 does not mean that the "meaning" of "water" changed for the average speaker in the interval. In 1750 or in 1850 or in 1950 one might have pointed to, say, the liquid in Lake Michigan as an example of "water." What changed was that in 1750 we would have mistakenly thought that XYZ bore the relation

same$_L$ to the liquid in Lake Michigan, while in 1800 or 1850 we would have known that it did not (I am ignoring the fact that the liquid in Lake Michigan was only dubiously water in 1950, of course) .

Let us now modify our science-fiction story. I do not know whether one can make pots and pans out of molybdenum; and if one can make them out of molybdenum, I don't know whether they could be distinguished easily from aluminum pots and pans. (I don't know any of this even though I have acquired the word "molybdenum.") So I shall suppose that molybdenum pots and pans *can't* be distinguished from aluminum pots and pans save by an expert. (To emphasize the point, I repeat that this could be true for all I know, and a fortiori it could be true for all I know by virtue of "knowing the meaning" of the words *aluminum and molybdenum*.) We will now suppose that molybdenum is as common on Twin Earth as aluminum is on Earth, and that aluminum is as rare on Twin Earth as molybdenum is on Earth. In particular, we shall assume that "aluminum" pots and pans are made of molybdenum on Twin Earth. Finally, we shall assume that the words "aluminum" and "molybdenum" are *switched* on Twin Earth: "aluminum" is the name of *molybdenum* and "molybdenum" is the name of *aluminum*.

This example shares some features with the previous one. If a spaceship from Earth visited Twin Earth, the visitors from Earth probably would not suspect that the "aluminum" pots and pans on Twin Earth were not made of aluminum, especially when the Twin Earthians *said* they were. But there is one important difference between the two cases. An Earthian metallurgist could tell very easily that "aluminum" was molybdenum, and a Twin Earthian metallurgist could tell equally easily that aluminum was "molybdenum." (The shudder quotes in the preceding sentence indicate Twin Earthian usages.) Whereas in 1750 no one on either Earth or Twin Earth could have distinguished water from "water," the confusion of aluminum with "aluminum" involves only a part of the linguistic communities involved.

The example makes the same point as the preceding one. If Oscar$_1$ and Oscar$_2$ are standard speakers of Earthian English and Twin Earthian English respectively, and neither is chemically or metallurgically sophisticated, then there may be no difference at all in their psychological state when they use the word "aluminum"; nevertheless we have to say that "aluminum" has the extension *aluminum* in the idiolect of Oscar$_1$ and the extension *molybdenum* in the idiolect of Oscar$_2$. (Also we have to say that Oscar$_1$ and Oscar$_2$ mean different things by "aluminum," that "aluminum" has a different meaning on

Earth than it does on Twin Earth, etc.) Again we see that the psychological state of the speaker does not determine the extension (or the "meaning," speaking preanalytically) of the word.

Before discussing this example further, let me introduce a *non*-science-fiction example. Suppose you are like me and cannot tell an elm from a beech tree. We still say that the extension of "elm" in my idiolect is the same as the extension of "elm" in anyone else's, viz., the set of all elm trees, and that the set of all beech trees is the extension of "beech" in both of our idiolects. Thus "elm" in my idiolect has a different extension from "beech" in your idiolect (as it should). Is it really credible that this difference in extension is brought about by some difference in our concepts? My concept of an elm tree is exactly the same as my concept of a beech tree (I blush to confess). (This shows that the identification of meaning "in the sense of intension" with concept cannot be correct, by the way.) If someone heroically attempts to maintain that the difference between the extension of "elm" and the extension of "beech" in my idiolect is explained by a difference in my psychological state, then we can always refute him by constructing a "Twin Earth" example—just let the words "elm" and "beech" be switched on Twin Earth (the way "aluminum" and "molybdenum" were in the previous example). Moreover, suppose I have a *Doppelgänger* on Twin Earth who is molecule for molecule "identical" with me (in the sense in which two neckties can be "identical"). If you are a dualist, then also suppose my *Doppelgänger* thinks the same verbalized thoughts I do, has the same sense data, the same dispositions, etc. It is absurd to think *his* psychological state is one bit different from mine: yet he "means" *beech* when he says "elm" and I "mean" *elm* when I say elm. Cut the pie any way you like, "meanings" just ain't in the *head!*