Hilary Putnam, *Representation and Reality*. Cambridge, Mass.: MIT Press, 1988.

In his new book, 'Representation and Reality', Hilary Putnam argues against the view that intentional idioms (with as paradigms the propositional attitudes) can be the subject of a reductionistic scientific approach. Although the book contains some arguments against scienticism in general, and a chapter against eliminativism, the core of the book is directed against functionalism, especially forms of functionalism that try to reduce intentional idioms to computational predicates or relations.

Putnam's case rests upon two main lines of argument. The first of these departs from the Quinian issues of meaning holism and the indeterminacy of translation/interpretation. The second elaborates Putnam's own viewpoint that 'meanings aren't in the head'.

Preliminarily, Putnam argues that for a real reduction of intentional idioms and relations, the computational state or relation these idioms are reduced to must be nomologically related. This means that every physically possible physical organism which instantiates the intentional predicate must also instantiate the computational predicate. Moreover, Putnam demands that the reducing predicate be definable in finitely many words, which means that the computational predicate cannot be an infinite disjunction.

Putnam takes meaning holism to imply a) that different individuals can be in the same intentional state (e.g. share the same belief or give the same meaning to the same word) while being in a different functional/computational state and b) that we, in order to notice that they are in the same intentional state, make use of our -hardly to formalize- intuitive standards of rationality.

Meaning holism says that in order to ascribe meanings to some subject's words or propositions, one has to interpret his discourse as a whole. But this interpretation is indeterminate: one can invent different 'translation manuals' that do the same job differently (translation manual B can ascribe a different content to S's utterance U than translation manual A by revising some of the beliefs that A supposed S to entertain). Translation manuals can, however, differ in reasonableness: we intuitively feel some translation manuals to do the job better. The same aspect of reasonableness occurs in our decisions on synonimity. In order to decide wether people are giving the same meaning to a word

or are sharing the same belief, we interpret their discourses by obeying a 'principle of charity'. What charity amounts to is that we tend to construe our interpretations of someone's discourse such that this someone's beliefs are as similar to ours as possible. We are likely to prefer 'interpretation manuals' which preserve beliefs and meanings over -by Quine's indeterminacy thesis- equally possible 'interpretation manuals' which do not. This accounts for the fact that we (or Putnam) don't want to say that the meaning of the term 'electron' in Bohr's 1900 theory differs from the meaning of the same term in his 1934 theory, despite there being lots of differences in the beliefs associated with the term in the two theories (Putnam sometimes suggests that if we would count the two occurrences of 'electron' as differing in meaning, this would imply that they were incommensurable. If they were, Putnam contends, Bohr's early theory would be completely unintelligible for us. We couldn't even talk about it reasonably. I think Putnam here too easily shifts from 'different in meaning' to 'incommensurable', in the sense of 'making understanding or communication impossible').

Charity however, is constrained by the demand that our interpretation be reasonable: beliefs do not have to be preserved at any cost (e.g. by ascribing the subject we interpret inconsistent or very strange beliefs). We do not translate 'phlogiston' as 'radiant energy', for example. The difference between the electron and the phlogiston cases is, according to Putnam, a difference we intuitively feel by appealing to our little understood 'general intelligence'.

A functionalist who agrees with Putnam that the arguments just given indeed show that two people (or intelligent beings) can be in the same intentional state without being in the same computational state, could still reply, in Putnam's mind, that there nevertheless could exist a computational 'equivalence relation' between their respective states. There could exist an algorithm that, given two beliefs and an arbitrary amount of information about the people who have them, could decide on their carrying the same content.

A first problem with this claim is that it demands that our general intelligence, our present intuitive judgments of reasonableness be formalized. Moreover, Putnam argues, such an algorithm, in order to discriminate all future (or worse, all physically possible) cases of belief sharing beings, would have to foresee all future (or worse, all physically possible) modes of discourse and standards of rationality, because to be able to decide whether two believers have a belief in common, one has to understand and interpret their discourses. In order to conceive

of such a relation, we should be able to conceive all future or possible modes of discourse and standards of rationality. But this is surely too immense a task for a human being. This implies that, even if there objectively would exist such an algorithm, no one could recognize it. Putnam who takes this impossibility of 'rationalizing' rationality in toto to be analogous to Gödel's incompleteness theorem.

If all of this is right, functionalists are not able to reduce (in a proper sense of 'reduce') intentional states to non-disjunctive, finitely defined computational states.

Putnam thinks that if functionalism cannot but allow disjunctions as its predicates, the functional organizations it will propose to be models of the mind, will be realized by every physical system. In an Appendix, he construes a proof of the following theorem: 'Every ordinary open system is a realization of every abstract finite automaton'. The basic idea (roughly sketched) is that given a succession of functional (more precisely automaton-) states, say ABAB, and a physical system that is in physical states a,b,c,d, one can always define a correspondence, by identifying disjunctions of physical states with the functional states (a v c with A and b v d with B). Once this correspondence is established, the arbitrary physical system is an instantiation of the functional model.

I think this argument poses a real threat to many functionalists. For their main theoretical construct is precisely the propositional attitude (or its image in 'the language of thought'). Given this prominent role, one should require functionalists to provide type- identifying criteria for propositional attitudes. Putnam's argument however, points out that they will fail to do so (if they restrict themselves to functional/computational language to state the criteria in).

The second line of argument goes back to Putnam's earlier writings, especially 'The Meaning of Meaning'(Putnam, 1975). It tells that computationalism, by its commitment to methodological solipsism, must be wrong since meanings are not in individuals' heads: they are constituted partly by society and partly by the environment.

Putnam illustrates the fact that the environment contributes in determining the meaning of terms with his famous twin-earth example. One is demanded to imagine a planet which is exactly the same as ours (which implies that it contains replica of the people living on earth at the same time), except for the fact that the water there, though apparently the same as our H2O, is in fact a different substance, XYZ. Putnam claims that, even before people or their twin earth replica discovered the chemical con-

stitution of their 'water', the term had a different reference, and thus a different meaning, on both planets. This he explains by invoking that natural kind terms always contain an element of indexicality, i.e. 'water' is defined by 'focusing' on a sample in one's environment, to which all instantiations must bear a similarity-relation (namely that they have the same structure and behavior as that substance).

Society plays a role in determining meaning because of a 'division of linguistic labour'. This has to do with the fact many people rely on experts to pick out the referents of common natural kind terms, such as 'gold', which implies that it is not 'what is in those people's head' that determines the reference of the terms. To say that the meanings are individuated the head of the expert would also be mistaken, for different experts might use different theories or tests in order to discriminate real from apparent referents. Which expert's understanding would constitute 'the' meaning of the term? The reasonable thing to do, according to Putnam, is to give up trying to individuate meanings.

An obvious, but powerful objection to Putnam's twin-earth argument is to state that Putnam's views commit him to the opinion that we didn't know the meaning of 'water' before we knew modern chemistry, and in general, that we will never know the meanings of our common words before the ultimate scientific truth about their referents is revealed (cf. Fodor(1980)). Putnam replies that this objection exploits an ambiguity in the meaning of 'knowing the meaning of'. 'Knowing the meaning of' can, according to Putnam, mean a)knowing how to translate, b) knowing the reference of and c) being able to use the word in discourse. In sense c), Putnam argues, we have always known the meaning of 'water'. What is dubious about this answer, is that Putnam's main point in the H2O/XYZ example hinges on separating meaning and use (despite our and twin earthlings' identical use of the term water in e.g. the year 1200, it had a different meaning on each planet even then), while his countering of the mentioned objection consists in identifying meaning and use!

Besides these main lines of argument against computationalist functionalism, 'Representation an Reality' contains treatments of more specific versions of functionalism proposed by Jerry Fodor, Ned Block and David Lewis. It also has a chapter arguing against philosophers who want to eliminate intentional idioms. Putnam's position is that these philosophers would have to give up the notion of truth too, which he takes as a reductio of their position. The last chapter contains some sketchy remarks on

Putnam's (not yet fully developed) alternative to functionalism, or reductionism or scientific realism in general, his 'internal realism'.

The body of the book however, is constituted by the arguments against functionalism. I believe that (especially the line of argument first sketched) poses real challenges to the position attacked.

I venture that even if this book had not been written by Putnam, it would have been welcomed as a major contribution.

REFERENCES

Fodor,(1980), "Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology", *The Behavioral and Brain Sciences*, Vol. 3.

Putnam(1975), "The Meaning of 'Meaning'",in Gunderson, ed., *Minnesota Studies in the Philosophy of Science* , Vol. 7 (Minneapolis: University of Minnesota Press).

Erik Myin
Research Associate NFSR
Vrije Universiteit Brussel