

# 7. Introducing Functionalism

Martín Abreu Zavaleta

June 5, 2014

## 1 Multiple realization

Recall *Identity Theory*, which said that to have a certain mental state  $m$  is just to have a certain brain state  $b$ . Putnam (in “The nature of mental states”) and others point out that this view has some important difficulties.

We may start with the observation that defenders of identity theory usually claim that the physical states that mental states are to be identified with are *physical-chemical* states. As Putnam observes, this seems too strong. Here’s why:

He [the identity theorist] has to specify a physical-chemical state such that *any* organism (not just a mammal) is in pain if and only if (a) it possesses a brain of a suitable physical-chemical structure; and (b) its brain is in that physical-chemical state. This means that the physical-chemical state in question must be a possible state of a mammalian brain, a reptilian brain, a mollusk’s brain (octopuses are mollusk, and certainly feel pain), etc. At the same time, it must *not* be a possible (physically possible) state of the brain of any physically possible creature that cannot feel pain. (p. 77)

Putnam is calling our attention to the apparent fact that such condition is very hard to satisfy: if we found aliens that behaved exactly like us, and yelled ‘ouch!’ every time we pinch them with needles, but are made of goo, identity theory tells us that they wouldn’t feel pain.

We don’t even need to go that far. Octopuses presumably feel pain, but their nervous systems seem to be very different from our own!<sup>1</sup> According to identity theory, if octopuses don’t have the physical-chemical kind of states that bring about pain in humans, that means they don’t experience pain.

But remember that identity theory is a general view about the nature of all mental states, not only pain. So, according to identity theory, unless something has the kind of physical-chemical states that bring about beliefs, desires, or even hunger in us (humans), it won’t have beliefs, desires, or hunger. This sounds like a bad consequence for the friends of identity theory.

This kind of objection is usually known as the argument from *multiple realizability*. We can define what it is for a state to be multiply realizable as follows:

**Multiple-realizability:** A property  $P$  is multiply-realizable just in case it can be instantiated by different things in virtue of their having different properties.

---

<sup>1</sup>For instance, two thirds of an octopus’s neurons are in its arms, rather than its brain.

For instance, the property of *being in pain* is multiply realizable because it can be instantiated in humans in virtue of the stimulation of C-fibers, in octopuses in virtue of the stimulation of O-fibers, in martians in virtue of the inflammation of M-nodes, et cetera. Because pain is multiply realizable, there need not be any particular physical property that everything must have in order to be in pain.

Most people believe that *mental states are multiply realizable*. It is because this thesis—the thesis that mental states are multiply realizable—is highly plausible that they came to formulate theories like *Functionalism*, which we will examine now. **Question:** Is behaviorism with the claim that mental states are multiply realizable?

## 2 Functionalism

We can generally describe functionalism as follows:

**Functionalism:** Mental states are constituted by their causal role; that is, by their causal relations to other mental states, as well as their relations to inputs and outputs.

What does this mean? Let's take a look at some things that we can identify by the roles they play. For instance, a water pump: anything that can move a fluid from one place to another in certain ways is a water pump. Or think of a piston: the job of a piston is to transfer the energy in a cylinder to a crankshaft by making a certain movement. Pistons are usually made out of metal, but at least in principle, there could be pistons made out of wood, plastic, or even rock (though perhaps all these pistons wouldn't last long).

What makes a piston a piston is not the material of which it is made, but the role that it plays in the functioning of certain kinds of engines. Notice that in our characterization of a piston above, we didn't only describe the movements that it makes. Instead, we described the relations that it has to other parts of the engine, like the cylinder or the crankshaft.

The idea of functionalism is that mental states can be defined in very much the way we defined a piston: by describing their functional roles in a given system. For instance, here is a very simplistic functional characterization of pain: being in pain is being in a state produced by sitting on a tack, that itself produces the state of being annoyed and the output of yelling 'ouch!' **Question: How is this different from behaviorism?**

Over the course of the years, different philosophers have offered different methods for functionally characterizing a mental state—though note that these methods can be used to give functional characterizations of anything whatsoever! Because of its historical influence and because it was the earliest formal characterization, we will take a look at functional characterizations of mental states in terms of machine states.

### Putnam's machine functionalism

In his seminal paper "The nature of mental states" Putnam offers probably the first version of machine functionalism:

The hypothesis that "being in pain is a functional state of the organism" may now be spelled out more exactly as follows:

1. All organisms capable of feeling pain are Probabilistic Automata.

2. Every organism capable of feeling pain possesses at least one Description of a certain kind (i.e. being capable of feeling pain *is* possessing an appropriate kind of Functional Organization).
3. No organism capable of feeling pain possesses a decomposition into parts which separately possess Descriptions of the kind referred to in (2).
4. For every Description of the kind referred to in (2), there exists a subset of the sensory inputs such that an organism with that Description is in pain when and only when some of its sensory inputs are in that subset.

In order to understand what this means, we must know what probabilistic automata and descriptions are, in the sense relevant to Putnam's view.

### Software, machine tables, and functional states

We can define a machine by means of its machine table: a complete description of all the relations between its internal states with inputs, outputs, and other mental states. In order to illustrate the idea, it will be useful to introduce an abstract characterization of computing machines: *Turing machines*.

A Turing machine is made up four components:

1. A tape divided into square, unbounded in both directions.
2. A head, positioned at one of the squares at any given time. It can scan and print.
3. A finite set of internal states.
4. A finite set of symbols (or alphabet).

The machine operates according to the following general rules:

- i. At each time, the machine is in one of its internal states and scans the square where it's positioned.
- ii. What the machine does at a given time is completely determined by its internal state at that time and the symbol its head is scanning at that time.
- iii. Depending on its internal state, the machine can do one of three things:
  - a. The head erases a symbol and prints a symbol (possibly the same symbol it erased)
  - b. The head moves to the right, to the left, or stays in place.
  - c. The machine enters one of its internal states (possibly the same state it was already in).

A *machine table* is a complete set of instructions that defines a program. Every program can be defined by means of a machine table. Here is a machine table for a very simple addition program using a Turing machine:

Input	Present state	Print	Change to state	Move
1	Start	1	Start	right
	A	1	A	right
	B	0	Halt	stay
0	Start	1	A	right
	A	0	B	left

Note that the machine table specifies the following things:

- How the machine responds to input.
- How the states of the machine are related to each other.
- The conditions under which the machine will give certain output.

With all these notions in place, we can offer a simple characterization of **Putnam's view**: *to be capable of being in a certain mental state is to be describable by a particular kind of machine table, and to be in pain is to be in a certain state*

What Putnam calls a *Description* is simply a very long sentence stating all the information that we can find in a machine table, and no more. For instance, a description of the machine characterized by the table above would say that the machine has four states, such that if the machine is in one of them (start) and it receives input 1, then it will remain in the same state and move to the right, if the machine is in another of them (A), and receives input 1, then it will remain in the same state and move to the right, and so on.

We almost have all we need to understand Putnam's view! The only thing we're missing is the notion of a *Probabilistic automaton*. You may have noticed that our machine table above leaves no room for chance: whenever the machine is in a given state and receives a certain input, it will determinately do a certain thing. Instead of doing this, we could have merely assigned a probability that the machine does a certain thing if it receives a certain input and is in a certain state.

For instance, we could have said that if the machine receives input 1 and is in state Start, then it will print 1 with a probability of .7 or print 0 with a probability of .3, it will remain in state Start with a probability of .6, or change to A with probability .3, or change to B with probability .1, and so on. The sum of the probabilities of the alternatives must always be 1!

So when Putnam says that all organisms capable of feeling pain are probabilistic automata, what he means is that they could be fully described in terms of a description like the one above. Moreover, only certain descriptions will characterize the kind of things that can feel pain (perhaps a very simple description won't make something capable of feeling pain, but a more complex one will).

Condition 3 merely requires that the object that receives such description is not itself composed by objects that also receive that kind of description. For instance, if it turned out that a bee hive behaves in the ways described by the machine table that characterizes pain, and that the bees themselves are describable in those ways, Putnam doesn't want to say that the bee itself is capable of experiencing pain!

### **Advantages of this view**

Notice that the program characterized by the machine table above can be implemented in many different ways! It can be carried out by a machine modeled literally after a Turing machine, but it could also be carried out by a purely mechanical machine (like Babage's Analytical Engine). It can even be carried out by a purely hydraulic machine, or by a machine made out of LEGOs. Thus, at least in principle, it seems that functionalism can account for the multiple realizability of mental states.