

Naturalizing Representation

Where we left off:

Thinking of intentional states in terms of content...

“Intentional states represent objects and states of affairs in the same sense of “represent” that speech acts represent objects and states of affairs”

“...to know what an intention is, or what any other Intentional state with a direction of fit is, we do not need to know its ultimate ontological category but rather we need to know: first, what are its conditions of satisfaction; second, under what aspect(s) are those conditions represented by the Intentional content; and third, what is the psychological mode – belief, desire, intention, etc. of the state in question??” (John Searle)

The Representational Theory of Mind...

- Intentional states are relations to mental representations...
- Their intentionality can be understood in terms of the semantic properties of those representations
- To believe that my sister is cooler than me is to be appropriately related to a mental representation whose propositional content is that *my Sister is cooler than me*. To worry that my sister is cooler than me is to bear an appropriately different relation to that same content.
- Thinking / reasoning / inferring etc. can be understood as a sequence of representational states, where the (causal or rational) sequence between them can be explained in terms of semantic or syntactic properties of the states in question.
- When everything is working well, semantic and syntactic properties constrain the transitions between representations...

What do we want from a theory of representation?

- (i) To make sense of inference

All red boxes contain smarties
This is a red box
This box will contain a smartie.

The cat is on the mat
If I want to sit on the mat I will either have to move the cat or sit on top of it.

- (ii) To make sense of behaviour more broadly

Why did you bring an umbrella with you?
Because I believed it would rain later.

I stepped out of the road because I saw the bike coming, and I didn't want the bike to hit me.

- (iii) Compatibility with Science

“To a large extent, empirical theories of cognition can and do take the notion of mental content as an explanatory primitive. But this is a kind of explanatory loan... If it turns out that the notion of mental representation cannot be given a satisfactory explication – if, in particular, no account of the nature of the

(mental) representation relation can be given that is consistent with the empirical theory that assumes it – then, at least in this response, that empirical theory must be regarded as ill founded, and hence as a less than adequate response to the drive for the kind of thorough intellectual understanding that motivates scientific theory in the first place. (Cummins pp.2-3)”

“I’m not really convinced that it matters very much whether the mental is physical, still less that it matters very much whether we can prove that it is. Whereas, if it isn’t literally true that my wanting is causally responsible for my reaching, and my itching is causally responsible for my scratching, and my believing is causally responsible for my saying... if none of this is literally true, then practically everything I believe about anything is false it’s the end of the world.” (Jerry Fodor 1989 “Making Mind Matter More” *Philosophical Topics*)

The drive to naturalize intentionality

What is it for a cognitive state to have a content?

What is it for a cognitive state to have some specific content (“That’s a bike whizzing towards me”)?

Cummins’ methodology: what do orthodox computational theories *need* from the nature of mental representations? If they are true, what must mental representation be like?

a. The goal

Fred Dretske “A Recipe for Thought”

“...philosophical naturalism is motivated by a constructivist model of understanding. It embodies something like an engineer’s ideal, a designer’s vision, of what it takes to really understand how something works.... If you want to know what intelligence is, you need a recipe for creating it out of parts you already understand.” (p.491)

“...what we are trying to build is a system that exhibits that peculiar array of intentional properties that characterizes thought. We are, in particular, trying to build systems that exhibit what Chisholm describes as the first mark of intentionality, the power to say that so-and-so is the case when so-and-so is not the case, the power to misrepresent how things stand in the world.” (p.493)

The compass Exhibits (original) intentionality: it tells you where the North Pole is.

Criterion: it distinguishes between contingently co-extensional features: polar bears and the north pole. It tracks the latter but not the former. I.e. it exhibits referential opacity.

- Tracking
- Tracking under an aspect (intensionality / intentionality)
- Tracking under an aspect *and getting it wrong* (misrepresentation / the mental/ thought)

b. Causal theories of mental representation

- What makes it the case that the content of a thought is “cow”?
 - o Some item of mental content “X” means X because “X”s are caused by Xs.
 - o In a nutshell: thoughts about cows are about cows because cows cause the mental representation of cows.

- Basic insight: nomic covariation is required for things to carry information.

“The Crude Causal Theory says, in effect, that a symbol expresses a property if it’s nomologically necessary that *all* and *only* instances of the property cause tokenings of the symbol.” (Fodor 1987 p.100)

- Why all and only?
 - Because otherwise the term would refer to the subset that *do* cause the tokening of the term.

These theories are also naturalist (supposing you have a natural account of causal relations...) because causal relations are not themselves intentional.

Compare:

Smoke “means” fire, because fire causes smoke.

These spots mean measles, because measles causes these spots

Thoughts of type C are about cows because cows cause thoughts of type C.

c. Problems for Causal Theories

The **omniscience problem**: the requirement that all instances cause the relevant term to be tokened seems to require that the system be omniscient...

- Any reasonable intentional system will miss instances, and that’s ok...
- A nomic covariation can be informative about a thing, even in the absence of the thing itself on a particular occasion.

(RELATEDLY) The **disjunction problem**: if the content of a thought is determined by whatever causes thoughts with that type of content, then content which is caused by several different kinds of objects just “means” whatever motley set of items happen to cause it.

These spots mean measles or syphilis because measles or syphilis causes these spots

Thoughts of type C are about cows or horses in low lighting because cows or horses in low lighting cause thoughts of type C.

Why is that a problem?

- Because we don’t want to allow that *any* kind of cause gets to determine the content of the thought. Sometimes we want to be able to say that the content of the thought is *mistaken*. If the content is *by definition* whatever causes the content, then we are unable to say that the system ever makes a mistake in this way.

Other variations on this problem / counterintuitive verdicts that seem to come out of a simple causal theory:

- Thoughts of type C are about retinal stimulation of a certain kind because retinal stimulation of a certain kind causes thoughts of type C.
- Thoughts of type C are about taking LSD because taking LSD causes thoughts of type C.
- Thoughts of type C are about the question “what’s that over there?” because the question “what’s that over there?” causes thoughts of type C.

What we want to be able to say:

Thoughts of type C are about cows, even though sometimes they are caused by horses in low lighting conditions, or taking LSD. When something other than a cow causes a thought of type C, the subject *misrepresents* its environment.

FIX #1: Normal conditions: only items which cause thought under normal conditions determine the content of the thought.

Thoughts of type C are about cows because they are caused by cows and cows only *under normal conditions* – rules out LSD, poor lighting

But still need a principled condition for singling out *the* cause of X, distinguishing *normal* from abnormal conditions.

Mere statistical regularity won't do the trick *e.g.* population that learns what mouse means via vole

FIX #2: Asymmetric dependence: The capacity of the non-content determining cause to token the content relies on the capacity of the content-determining cause to token the content

i.e. horses in low light only cause thoughts of type C because cows cause thoughts of type C. If cows no longer caused thoughts of type C, nor would horses.

It can be tricky to get this right: consider Broomhilda and the mice, which she rarely sees but has been trained to recognise on the basis of shrews. Mice wouldn't cause thoughts of type M unless shrews also caused thoughts of type M...

d. Function as a solution to the disjunction problem

What is the difference between a thermometer and a paper clip given that we could (in theory) read the temperature from a paper clip?

Our assignment of a “job” to the former

“If an information-carrying element in a system could somehow acquire the function of carrying information, and acquire this function in a way that did not depend on our intentions, purposes, and attitudes, then it would thereby acquire (just as a thermometer or a compass acquires) the power to misrepresent the conditions it had the function of informing about. Such functions would bring about a detachment of meaning from cause.” (p.495, emphasis added.)

- A fuel gauge represents that a tank is full of fuel, even if that is caused by the tank being full of water, because that is its *function*. Its function is what allows it to misrepresent.
- Turns our attention to looking for naturalistically acceptable sources of function.
- Phylogenetic sources: evolution of an organism \Rightarrow biologically oriented theories of mental representation
- Ontogenetic sources: demands that arise for an individual organism within a particular environment.

Suggested reading:

- *Cummins, Robert, *Meaning and Mental Representation* (Cambridge, MA: MIT Press, 1988). Ch 6-8
- Dretske, Fred. 'A recipe for Thought'. In the CHALMERS anthology on the reading list.
- Dennett, Daniel, 'Intentional Systems', *The Journal of Philosophy*, 68, no. 4 (1971): 87-106.
- Dretske, Fred, *Explaining Behavior: Reasons in a World of Causes* (Cambridge, MA: MIT Press, 1988), chs. 3-5.