

Jerry Fodor – *Psychosemantics*

Chapter 1 : Introduction : The Persistence of the Attitudes

Introduction

- Fodor begins with a quotation from *A Midsummer's Night Dream*, act 3 scene 2, involving Demetrius, Hermia and, in absentio, Lysander. Hermia has deduced, from Lysander's disappearance, that Demetrius, his rival in love, has killed him.
- Fodor shows how the deduction has gone. Hermia believes herself beloved of Lysander because he has told her and, *ceteris paribus*, we can deduce how people feel from what they say. So, Lysander wishes Hermia well; but, if so, he wouldn't voluntarily abandon her at night in a dangerous wood. Since he has in fact deserted her, it must be involuntarily, so must plausibility have come to harm. The prime suspect is his rival Demetrius as it is plausible that love-rivals do not wish one another well and Hermia believes that Demetrius believes that Lysander alive is an impediment to his aims. She also believes the plausible argument that if x wants P, believes that not-P unless Q and that he can bring about Q, then he will, *ceteris paribus*, attempt to do so. She further believes that those who attempt are usually successful. So, Hermia suspects Demetrius of murdering Lysander and we, the audience, following her largely unspoken chain of reasoning and sharing many of her psychological beliefs, sympathise with her suspicions.
- In fact, the *ceteris paribus* clause has been sprung by a peripatetic sprite and the chain of reasoning relying on the interconnected web of beliefs, desires and actions that Hermia, Shakespeare and we share has gone wrong.
- Fodor points out, however, that while commonsense¹ belief/desire psychology may fail occasionally, it is (1) often right (2) deep and (3) indispensable. It has come under philosophical attack and Fodor now addresses whether it is worth saving by looking further at these three points in its favour.

1. How Often It Works

- Our relations with one another are mediated by applications of commonsense psychology (henceforth known as “the theory”, and break down when its predictions fail. Such failures are the stuff of theatre, but its ubiquitous success goes unnoticed.
- The theory works so well, it goes unnoticed. I may agree to meet someone I've never met before some months hence and thousands of miles away, and all go well. Failure to keep the appointment is more likely to do with transportation difficulties than failure of the theory. We cope with one another by its use better than we do with less complex machines.
- The theory tells us how to infer people's intentions from their utterances; not just for our intimate friends but complete strangers; and not in the lab but in the field. If we could perform the simpler function of predicting the weather as well, we'd never get wet.
- What about *ceteris paribus* ? Some philosophers say the theory is trivially incapable of disconfirmation, and so false or vacuous.

¹ Note that this is a less pejorative term for “folk psychology”.

- If we consider the defeasibility of an ordinary utterance, such as a promise, it holds true unless the speaker is lying, saying unintelligible words, talking in his sleep, etc. We might view promises encumbered by these *ceteris paribus* clauses as equivalent to saying “I’ll do this ... unless I don’t”. Nothing can disconfirm it.
- While many philosophers are attracted by this sort of argument, it can’t be right because commonsense psychological generalisations work most of the time and enable us to make predictions, so cannot be *empty*.
- Fodor thinks that all the special sciences, other than basic physics, rely on such *ceteris paribus* clauses. A meandering river may fail to erode its outside bank if it freezes, dries up or the world comes to an end. Do the *ceteris paribus* clauses in this case make geology false or vacuous ?
- No – because the geological generalisation is not of the form “P or not-P”. It is complex matter how the special sciences can be both hedged and informative, supportive of counterfactuals and exceptions; and this relates to why we have special sciences and not just basic physics. The *ceteris paribus* clauses mean “ ... in any nomologically possible world where the operative generalisations are satisfied”, and this is stronger than “P in any world where not not-P”. Hence, commonsense psychology relies on *ceteris paribus* clauses no more than geology.
- However, is this similarity superficial ? Can we, as Davidson suggests, “perfect” the generalisations of the special sciences by enumerating the conditions under which their generalisations are supposed to hold ?
- Fodor thinks that, in this case, the only “real” science left would be basic physics. His point is that to enumerate the exceptions to a geological generalisation, we have to step outside geology. Explaining failure of erosion because of a change in the weather is to invoke a meteorological event rather than a geological one. All that can be said within geology is that, if the generalisation failed to be satisfied, then the operative idealisations must somehow have failed to be satisfied. And so with commonsense psychology; failure to turn up as announced indicates something went wrong.
- Exceptions to generalisations are inexplicable in the vocabulary of a special science, which is why it is *special*, but may be explicable in the vocabulary of another science; most familiarly in terms of a more basic science. This is what the hierarchical arrangement of the sciences provides. So, while we cannot discharge the *ceteris paribus* clauses from within the special science, we have no reason to doubt that we can do so in the vocabulary of some lower-level science (neurology, biochemistry or, at worst, physics).
- The world is describable as a closed causal system, if at all, only in terms of our most basic science, and this is irrelevant to psychology or geology.
- The moral of the above is that the predictive adequacy of commonsense psychology is rationally indisputable and is not obtained by cheating. To find out where I’ll be next Tuesday, it’s no use asking a physicist – you have to ask me !

2. The Depth of the Theory

- Fodor asks whether commonsense psychology is just a collection of Granny’s nostrums, such as “money can’t buy happiness” and agrees that these are not worth saving². However, the bulk of commonsense psychology is non-platitudinous and shows the deductive structure characteristic of real science.

² See the end of the Introduction for Fodor’s aims in this paper.

There are two reasons for this (1) the theory's generalisations are defined over unobservables and (2) these generalisations lead to predictions via iteration and interaction rather than direct instantiation³.

- Justifying (2) first - the example of Hermia shows that she's relying on an implicit theory of the interaction of beliefs, preferences and behaviours – ie. an implicit decision theory. She's no behaviourist, so thinks Demetrius's behaviour is intricately caused by his mental states. She doesn't think that rivals necessarily kill, ie. $(x)(y)(xRy \rightarrow xKy)$, but that the true and counterfactual-supporting generalisation is that rivals prefer one another's discomfiture. Behaviour can only be deduced from this generalisation with the aim of assumptions about beliefs and preferences.
- With respect to (1), it's a deep fact about the world that the most powerful etiological⁴ generalisations hold of unobservable causes, so this is a good test of a deep theory. Fodor gives the example of commonsense *meteorology*, of the "red sky at night" variety, which only involves instantiation and modus ponens, which explains its inefficacy and shallowness. On the contrary, commonsense psychology passes the depth test by assuming that behaviour comes at the end of an arbitrarily long and tangled causal chain of unobservable mental events. We are born and remain both mentalists and realists until persuaded out of it by bad philosophy.

3. The Theory's Indispensability

- We have no alternative to commonsense psychology if we want a theory of counterfactual-supporting generalisations to explain our behaviour and its causes.
- Because commonsense psychology is so close to us, its value is often overlooked. Commonsense psychological generalisations bridge the gap between exchanges of utterances and subsequent behaviour. Without it we can't even describe utterances as *words*, or behaviour as *acts*. *Word* is an irreducibly psychological category⁵.
- Fodor specifies four conditions that require satisfaction if we want to avoid commonsense psychology, and claims they cannot be satisfied. (1) Behaviour in uttering x counts as an event of type T_i . (2) Fulfilling the import of utterance x counts as an event of type T_j . (3) "Events of type T_i are consequent on events of type T_j " is roughly true and counterfactually supporting. (4) Categories T_i and T_j are not irreducibly psychological.
- Fodor's point is that events of types T_i and T_j (utterances, statements, actions) are psychological, so fail to satisfy (4).
- Philosophers and psychologists used to hope that the conceptual apparatus of *behaviour* could be replaced by that of *movements*. As a result, psychological generalisations would be seen to be the consequence of movements being contingent on environmental and / or organic variables. While this contingency is true, the generalisations don't follow from it because behaviour consists of

³ Hopefully the examples will explain what this means !

⁴ What does this mean ?

⁵ With respect to the irreducibility, Fodor says this is "as far as anyone knows". He focuses on the acoustic properties of words, claiming that there are no acoustic properties that all and only tokens, even fully intelligible tokens, of a particular word share, which explains the then difficulty of producing working speech-recognition software. I'm not clear on the truth, importance or relevance of this aside.

actions. Since actions cross-classify movements, we need to know the action type, not the motion type⁶.

- It is generally assumed that the situation will be remedied in due course, as a completed physics would subsume motions of everything, including organisms, and eventually we will be able to have counterfactually-supporting generalisations about the motion of organisms under the description of physics.
- Fodor thinks this is not only optimistic but misleading. A completed physics would have no place for the behavior any more than behaviour. It would not recognise organisms as such, but only matter.
- However, even if psychology were dispensable in principle, this is no argument for dispensing with it, any more than it is sensible to dispense with rivers as a natural kind, and with geology in favour of basic physics. Psychology would have to be dispensable in fact, not just in principle, and it is clear that we cannot explain ourselves other than in a vocabulary saturated with belief/desire psychology. We can't abandon the propositional attitudes because we don't know how to.
- The rest of Fodor's book is an attempt to vindicate commonsense psychology by showing that, or better how, we can have a respectable science whose ontology contains states with the properties attributed to the attitudes by common sense. He now explores what properties these are.

The Essence of the Attitudes

- How do we know whether a psychology is of the belief-desire variety with the propositional attitudes in its ontology ? Inter-theoretic identification – distinguishing reduction from elimination from reconstruction is problematical; even behaviourists were unsure whether they wanted to eliminate the mental or equate it with the behavioural.
- Fodor stipulates that a psychology is common-sensical about the attitudes just in case it postulates states, entities, events and so on that :-
 - i. Are semantically evaluable
 - ii. Have causal powers
 - iii. The generalisations of commonsense belief/desire psychology are true of
 He takes these stipulations as intuitively plausible and squabbling about intuitions as vulgar ! He now explains these stipulations.

(i) Semantic Evaluation

- Beliefs are true or false, desires can be fulfilled or frustrated and so on. According to Fodor, what makes a belief true is its relation to the non-psychological world rather than to other beliefs. Fodor terms the evaluation of the belief against the world “semantic”.
- That beliefs, desires etc. are semantically evaluable is a puzzle as virtually nothing else is. Propositions are, but this is unsurprising as propositions are what beliefs and desires are attitudes *towards*. Semantic evaluability is the most philosophically troublesome element of belief/desire psychology and the part requiring most defence.

⁶ Fodor's example is of the generalisation “the burnt child avoids the fire”, but the movement that constitutes avoidance depends on a multitude of factors, such as the relative positions of the child and the fire. I'm unclear what this is supposed to show.

- Content and semantic evaluability are intimately connected. If you know the content of a belief, you know what it is about the world that determines its semantic evaluation.
- Propositional attitudes have their content *essentially*⁷. To pick out an attitude we need only to specify (1) the *type* (belief, desire, hunch ..) (2) the *content* (“2 is a prime number”, ...) of the attitude. Any propositional-attitude psychology must individuate its states in this way.

(ii) Causal Powers

- Commonsense psychological explanation relies on three sorts of mental causation. (1) Causation of behaviour by mental states, (2) proximal stimulation of mental states by environmental events and (3) causation of mental states by one another. The latter is the most interesting, giving chains of thought – a complex mental event in which one semantically evaluable mental states gives rise to another until, often, belief is fixed.
- All realist psychologies acknowledge the causal powers of mental states. Functionalists hold that the identity of a mental state is determined by its causal powers. Fodor doesn’t think this is correct, though it would be significant if it were, for it would follow that the causal powers of a mental state determine its content (given that propositional attitudes have their contents essentially).
- Any view vindicating belief/desire psychology attributes contents and causal powers to the very same mental things it takes to be semantically evaluable.
- Further, causal relationships respect content relationships. For instance, if someone believes Fa then, *ceteris paribus*, he believes $(\exists x)(Fx)$. Presumably, since $Fa \supset (\exists x)(Fx)$ (by existential introduction), belief that Fa causes belief that $(\exists x)(Fx)$ ⁸.
- Fodor refers back to the introductory section, where Hermia believed the generalisation “that if x wants P , believes that not- P unless Q and that he can bring about Q , then he will, *ceteris paribus*, attempt to do so”. In this case, we can say that we have *explained* x ’s attempt to bring about Q given our knowledge of his beliefs and desires about P and Q . What is typical of the explanation is (a) appeal to causal relations between semantically evaluable mental states and (b) the existence of content relationships amongst the mental states appealed to.
- Fodor points out that such explanations would fail if there were equivocation about the schematic letters – it’s important that what’s believed can’t be had unless (...) is what’s desired. The same contents often occur in causally related mental states, with causal relationships respecting semantic ones.
- The similarity between trains of thought and arguments arises from the parallelism between causal powers and contents. Fodor gives an example of Sherlock Holmes’s deductive train of thought from *The Speckled Band*. The quotation shows not just the psychology – the sequence of mental states that led Holmes to suspect and then believe the method of the murder – but assembles *premises* that are plausibly sufficient for a *conclusion* to be drawn by himself and others. What connects the causal history aspects of the story to the plausible-inference aspects is the fact that the thoughts that fix a belief that P are often *grounds* for believing

⁷ Look up essentialism !

⁸ Need to re-read this section to see whether I’ve missed anything subtle – eg. whether the causation isn’t between mental states but between the world and mental states.

that P. If this – the general harmony between the semantical and causal properties of thoughts - were not the case, there would be little point in thinking.

- Such thoughts raise a number of philosophical issues about the notion of content that cause some philosophers to despair of commonsense psychology but which need to be accounted for if a deep fact about the cognitive mind is not to be left out. These questions include :-
 1. What sorts of content relations are preserved in the generalisations⁹ that subsume typical cases of belief/desire causation ?
 2. How could the mind be constructed that such generalisations are true of it ?
 3. What sort of mechanism could have states that are both semantically and causally connected ?
 4. How can causal connections respect semantic ones in such a mechanism ?

(iii) Generalisations Preserved

- Fodor considers, by way of summary, the properties required of a psychological theory that vindicates commonsense belief/desire explanations. These must :-
 1. Permit the assignment of content to causally efficacious mental states
 2. Recognise behavioural explanations in which covering generalisations quantify over the contents of mental states that they subsume
 3. Not recognise generalisations that are crazy from the viewpoint of common sense
 4. Have attitudes with causal powers more or less as common sense supposes
 5. Be at least approximately true.
- Fodor claims to be unphased by the fact that common sense is often mistaken about the attitudes as about anything else. He hopes that there is more to the mind than is dreamed of by common sense. Since Freud we have learnt that, contrary to common sense, much of what's in the mind is unconscious or unlearned. He doesn't have a list of commonsense generalisations that must be preserved by any theory ontologically committed to bona fide propositional attitudes.
- However, we have no reason to doubt much of commonsense psychology. Any psychology of action that is committed to the attitudes must acknowledge the causal relations amongst beliefs, desires and behavioural intentions explicated by decision theory. Similarly, any psycholinguistics which attributes beliefs, desires and communicative intentions to speakers and hearers must entail an infinity of theorems of the form “‘x’ is the form of words standardly used to communicate the belief that x”.
- And again, any psycholinguistics that appeals to the propositional attitudes of speakers and hearers to explain their verbal behaviour must entail that they know at least one such theorem for each sentence ‘x’ in their language. Consequently, psychology already has an infinite amount of common sense to vindicate¹⁰.
- Self-confident essentialism¹¹ is currently fashionable and some philosophers have modal intuitions about whether there could be *cats* in a world in which all

⁹ I need to check that I have given sufficient weight to *generalisations* in my account of Fodor.

¹⁰ Some of this argument must be tongue in cheek. What is it about ?

¹¹ Essentialism : Fodor was writing in 1987. Essentialists maintain that some of an object's properties are essential to it, the rest are accidental. The essential properties of a thing are those it needs to possess to be the thing it is. There are two forms of essentialism – definitional and modal. The latter has it that x is essentially F iff necessarily whatever is x has the property F; alternatively, x must be F to exist at all. See Kripke, *Naming and Necessity* for a *de re* modal epistemology (*Concise Routledge*

domestic felines were robots or *Homer* in a world where nobody wrote the *Illiad* or *Odyssey*. Fodor doesn't aspire to their enviable epistemic condition. He doesn't know how much commonsense psychology would have to be true for there to be beliefs and desires, but he's not worried as he has no doubts that lots of it *is* true.

RTM (Representational Theory of Mind)

- The main purpose of Fodor's book is to argue that we have every reason to believe that it is possible to have a scientific psychology that vindicates commonsense belief/desire explanation. There already exists (in Fodor's view) a more or less empirical theory that is ontologically committed to the attitudes, is quite probably approximately true, is a vindication of the attitudes and is the only scientific belief/desire psychology in the field. Defending this theory and defending the commonsense attitudes is the same thing.
- Fodor doesn't think that recent arguments against belief/desire explanation are cogent against this theory, which the rest of this chapter is briefly to explain.
- The theory is RTM (Representational Theory of Mind). The fundamental thesis of the theory is that there is a language of thought, an infinite set of mental representations which function both as the immediate objects of propositional attitudes and as the domains¹² of mental processes. Fodor says that RTM is made up of two claims :-
 - *Claim 1* (The nature of propositional attitudes) : For any organism O, and any attitude A towards the proposition P, there is a (computational/functional) relation R and a mental representation MP such that MP means that P, and O has A iff¹³ O bears R to MP.
 - *Claim 2* (The nature of mental processes) : Mental processes are causal sequences of tokenings of mental representations.
- What *Claim 1* means is that to believe that P is to have a mental symbol that means that P tokened in your head; to have such a token in your "belief box". Similarly for hopes tokened in your "hope box", the different boxes, shorthand for *functional* states, reflecting the different causal roles of the attitudes.
- Similarly, *Claim 2* means that a train of thought is a causal sequence of tokenings of mental representations that express the propositional objects of the thoughts. To have a train of thought "because *x* is the case I'll do *y*" is to have a tokening of a mental representation that means *y* caused by one that means *x*.
- Fodor's reasons for thinking RTM more or less true are as follows. (1) Our best science is our best guess at what there is and what it's made of, and versions of RTM underlie all current psychological research on mentation. (2) RTM suggests a plausible mechanism for the parallelism between causal relations amongst mental states and semantic relations between propositional objects, and the fact that trains of thought are largely truth-preserving relies on this symmetry.
- The trick is to use the "computer metaphor", since computers show how to connect semantical with causal properties for *symbols* and therefore we can use this as an analogy for *thoughts* if having a propositional attitude involves tokening a symbol. This is the only major advance of contemporary cognitive science over

Encyclopaedia of Philosophy). Hence, returning to Fodor, if it is essential to being *Homer* that he wrote the *Illiad* and the *Odyssey*, he would not exist in a world in which these works did not exist.

¹² I need to express this in my own words when I understand better what he's on about.

¹³ Fodor says that this biconditional is true strong, but not so as to undermine the spirit of the proposal.

18th and 19th century mentalism. Associationism failed because there was no way for a rational mental life to arise from the causal relations between thoughts recognised by the “laws of association”, as the closing pages of *Ulysses* illustrate / parody.

- This is how the new story goes :-
 1. The connection between the causal and semantic properties of a symbol is via its *syntax*.
 2. The syntax of a symbol is one of its higher order physical properties, metaphorically, an abstract feature of its shape¹⁴.
 3. The shape of a symbol is a potential determinant of its causal role, so we can see how there could be environments where the causal role of a symbol correlates with its syntax, where symbol tokens interact causally *in virtue of* their syntactical structures.
 4. The syntax of a symbol determines the causes and effects of its tokenings much as the geometry of a key determines which locks it can open.
- Modern logic tells us that certain semantic relations amongst symbols can be mimicked by their syntactical relations, which is ultimately what proof-theory is about. Within famous¹⁵ limits, when there is an entailment relation between two propositions, there will be a parallel between the semantic relation between the two symbols and their syntactical relation whereby one symbol can be derived from the other. Hence, we can build machines, again within the famous limits, such that the machine :-
 1. Operates solely by transforming symbols
 2. Is sensitive solely to the syntactic properties of the symbols
 3. Operates solely by altering the shapes of the symbols

Even so, the machine transforms one symbol into another only if the propositions expressed by the symbols stand in relevant *semantic* relations to one another, eg. of a valid argument between premises and conclusion. Such machines – computers – just *are* environments in which the syntax of a symbol determines its causal role by respecting its content.
- This connects with RTM’s ontological commitment to mental representations as follows. Computers one means of mediating between the causal and semantic properties of symbols. Hence, if the mind is a sort of computer, we can see how we can get a theory of mental processes that succeeds, where all previous attempts have failed, in explaining how there can be non-arbitrary content relations among causally-related thoughts. For this to work, there need to be mental representations. The parallelism evident in computers between the syntax and semantics of symbols that brings causal role and content into phase won’t help the theory of *mind* unless there are *mental* symbols possessing both semantical and syntactical properties. We need mental symbols because only symbols have syntax and our only theory of mental processes that isn’t known to be false pictures the mind as a syntax-driven machine.
- Fodor now considers the objection of Churchland et al that commonsense belief/desire psychology is a sterile theory that hasn’t advanced much since Homer or at all since Jane Austin. He agrees that there is some truth in the charge, as, though commonsense psychology may be *implicit* science, it isn’t *research* science. Progress is to be found in the explicit vindication, where there has been

¹⁴ What does this mean ? See Fodor’s end-notes here and elsewhere.

¹⁵ Does this refer to Goedel’s theorems, or Turing’s halting theorems ?

enormous progress; and not just in our understanding of memory and perception concerning the fixation of belief, nor about language concerning the communication of belief. The real progress is that we are on the verge of understanding *how the causal processes of the mind can be semantically coherent*; in other words, *how rationality is mechanically possible*. Such problems can't even be posed unless we suppose, in accord with commonsense belief/desire psychology, that there exist mental states with both semantic contents and causal roles. A good theory invites questions that have answers.

- As was noted above, the biconditional in *Claim 1* won't do, and Fodor finishes the chapter by polishing his claims for RTM.
- *Claim 1* had it that both the following are true, whereas both are on occasion false
 - a) For each tokening of a propositional attitude, there is a tokening of a corresponding relation between an organism and a mental representation.
 - b) For each tokening of that relation, there is a corresponding tokening of a propositional attitude
- How does Fodor escape from this problem ? He makes an analogy with explicit science co-opting the conceptual apparatus of common sense. Is water H₂O ? Not if the operative notion of water is the commonsense one that treats water as the stuff we use in our daily lives. Only in the theory's *core cases* of *chemically pure* water is H₂O, but "chemically pure" isn't a commonsense notion. Still, common sense is right in saying that there is water in the stuff we normally use and that gives it the properties we rely on.
- What's the relevance of this to RTM ? Fodor thinks that RTM vindicates commonsense psychology in the cases that RTM identifies as core. In these cases, what commonsense psychology takes as tokenings of propositional attitudes are indeed tokenings of a relation between an organism and a mental representation. The other cases, where the biconditional fails in either direction, are "derivative" – and this is exactly what we'd expect from scientific precedent. Fodor now considers a couple of examples in either direction both to see what the fuss is about and to tighten up *Claim 1*.

Case 1 : Attitudes without Mental Representations

- This is from Dennett. A designer of a chess-playing program refers to a rival's as "wanting to get its queen out early". Dennett points out that this is a usefully predictive mental attitude, but that there is at no level in the program any explicit representation tokened of this attitude. This attitude is an *emergent property* of the computational processes that have "engineering reality". Dennett thinks there's no reason to think that the relation between belief-talk and psychological-process talk will be any more direct.
- Fodor starts off by pointing out that Dennett's problem *isn't* that some propositional attitudes of commonsense psychology are *dispositional*. I may be said to believe something I've never heard of just because I *would* believe it if I were shown the proof. Dispositional beliefs couldn't correspond to *occurrent* tokenings of relations to mental representations. However, Claim 1 can be easily patched up by saying that for a *dispositional* belief there is a corresponding *disposition* to token a mental representation.
- This leaves open the question when are the attributions of dispositional beliefs *true* ? According to Fodor, one's dispositional beliefs should be identified with the closure of one's occurrent beliefs under principles of inference one currently

accepts. RTM can live with the uncertainty of such a closure as *dispositional* attitudes play no causal role in *actual* mental processes. Therefore, RTM can afford to be operationalist about dispositional beliefs provided it is firm about occurrent ones.

- However, Dennett’s point was not that the chess-playing program had a *potential* belief but an *actual* operational one, even there is no tokening of any symbol that expresses it. Similarly, behavioural commitment to modus ponens might betoken that this is inscribed in brain writing, but it needn’t as this rule might be complied with rather than literally followed¹⁶.
- In Dennett’s example, an attitude is emergent from its own implementation, and this suggests a way of saving *Claim 1*. While the machine doesn’t explicitly represent “get your queen out early” it does explicitly represent some more detailed rules of play – those that have “engineering reality” – for which the strong form of *Claim 1* is satisfied. However, this won’t solve the problem as *none* of the principles by means of which a computational system operates need be explicitly represented by formulae tokened in the device. In particular, the program need not be explicitly represented; the point being that any function computed by an algorithm can be hard-wired. So, what use is the computer metaphor to RTM ?
- Worse, not all the rules of inference that a computational system runs on *can* be represented *just* explicitly in the system; some have to be “realised in the hardware” or the machine won’t run at all.
- Fodor agrees with this, but defers considering whether cases where programs are hardwired defeat *Claim 1* until after considering Case 2.

Case 2 : Mental Representations without Attitudes

- The feature of computers borrowed by RTM so far has been the mechanisation of *rationality*, utilising a syntactically-driven machine to exploit the parallelisms between the syntactic and semantic properties of symbols. Some versions of RTM also borrow a theory of *intelligence*, wherein intelligent behaviour exploits a cognitive architecture consisting of *hierarchies* of symbol processors. The top of each hierarchy is a complex capacity while at the bottom is only the sort of unintelligent operations characteristic of a Turing machine. The middle layers form an explanatory reduction or analysis of intelligent capacities into dumb ones.
- Fodor gives an example of this sort of thing – tying one’s shoe-laces, in which a homunculus follows instructions in a manual and in which various tasks are broken out amongst gangs of wage slaves. At the top of the hierarchy are states corresponding to propositional attitudes (knowing how to tie one’s shoes), but the middle and bottom levels there is lots of symbol-processing that corresponds to nothing that *people*, as opposed to nervous systems, do. While these lower states satisfy *Claim 1*, in that they involve causally efficacious tokenings of mental representations, they do not correspond to the tokening of an attitude as judged by common sense. So, how can we modify *Claim 1* to avoid counter-examples by sub-personal information processing ?

Vindication Vindicated

¹⁶ I need to consider the import of this.

- Fodor says that these objections are not serious. What counts for the vindication by RTM of belief/desire explanation is that in the *core* cases, rather than in *all* cases, there should correspond a tokening of a mental representation for what common sense counts as the tokening of an attitude (and vice versa). RTM has to be able to say what it counts as a core case just as chemistry has to say what counts as chemically pure water if it is to avoid “water is H₂O” being disconfirmed by water in a river.
- Fodor’s core cases are covered by the slogan “No intentional causation without explicit representation”. That is, RTM is committed to the explicit representation of all thoughts featuring as episodes in mental processes, otherwise it is false. This follows from *Claim 2*, where mental processes are seen as causal sequences of transformations of mental representations.
- This gets us out of the counter-examples. The rules that determine the transformations of explicit tokenings of mental representations need not be explicit but can emerge from the explicitly represented procedures of implementation, or from the hardware. The “laws of thought” (modus ponens, “get the queen out early”, ...) *may* be explicitly represented, or not in these cases, but the “data structures” (the contents of thoughts) *have to be* represented.
- So, in Dennett’s case “better get the queen out early” never constitutes an episode in the mental life of the machine – ie. at the program level – it never consults this rule in its operation. However, this is no problem for the core cases covered by RTM. What is *essential* is that the actual or potential positions on the board be explicitly represented, *because* the machine’s computations are defined over these. These computations constitute the machine’s mental processes, so for a representational theory of chess playing to be true of it they have to be causal sequences of explicit representations. The relevance of the computer metaphor to RTM arises from the connection between the constitution of mental processes as computational and there being a language of thought¹⁷.
- Finally, what about the second counterexample – the sub-personal cases where there are tokens of mental representation without tokens of attitudes ? What is needed is that scientific psychology is ontologically committed to beliefs and desires, *not* that the folk-psychology inventory of propositional attitudes be exhaustive and co-extensive with that of RTM. What RTM needs to do is show how intentional states could have causal powers. It is no problem for molecular physics that it classifies glass as a liquid while common sense would not.
- In summary, common sense belief/desire explanation is vindicated if some good theory of the mind is committed to entities like the attitudes which are both semantically evaluable and etiologically involved. RTM is such a theory, so if it is true, common sense is vindicated. RTM needs to provide an independently confirmed empirical case for mental processes as causal sequences of transformations of mental representations. This is the job of modern cognitive psychology, which is almost totally dedicated to this end. A more strictly philosophical problem turns on the semantic assumptions made by RTM, and this is the subject of the rest of Fodor’s book.

¹⁷ Examine Fodor’s endnote on this point !