ψυχη
PSYCHE

Yes, It Does: A Diatribe on Jerry Fodor's *The Mind Doesn't Work that Way*[1]

Susan Schneider
Department of Philosophy
University of Pennsylvania
423 Logan Hall
249 S. 36th Street
Philadelphia, PA 19104-6304
sls@sas.upenn.edu

PSYCHE 13 (1), April 2007
REVIEW OF: J. Fodor, 2000. *The Mind Doesn't Work that Way*. Cambridge, MA: MIT Press. 144 pages, ISBN-10: 0262062127

Abstract: *The Mind Doesn't Work That Way* is an expose of certain theoretical problems in cognitive science, and in particular, problems that concern the Classical Computational Theory of Mind (CTM). The problems that Fodor worries plague CTM divide into two kinds, and both purport to show that the success of cognitive science will likely be limited to the modules. The first sort of problem concerns what Fodor has called "global properties"; features that a mental sentence has which depend on how the sentence interacts with a larger plan (i.e., set of sentences), rather than the type identity of the sentence alone. The second problem concerns what many have called, "The Relevance Problem": the problem of whether and how humans determine what is relevant in a computational manner. However, I argue that the problem that Fodor believes global properties pose for CTM is a non-problem, and that further, while the relevance problem is a serious research issue, it does not justify the grim view that cognitive science, and CTM in particular, will likely fail to explain cognition.

1. Introduction

> The current project, it seems to me, is to understand what various sorts of theories of mind can and can't do, and what it is about them in virtue of which they can or can't do it. If there are constraints on what CTM can explain, then it should be a primary research goal to try and figure out exactly what they are. I don't see any other strategy that research could reasonably pursue (Jerry Fodor, personal communication).

I concur: the scope and limits of computational psychology is one of the most important issues in contemporary cognitive science. And for a proponent of the Computational Theory of the Mind (CTM), like Jerry Fodor, a most pressing issue is whether there are limits on what CTM can explain. Much of Fodor's book is an 'expose' of the problems that confront CTM. Indeed, it is an expose of certain theoretical problems in cognitive science more generally, with the aim of keeping cognitive science honest.

I'd like to say at the outset that this is a truly excellent book. As always, Jerry Fodor has generated a highly original and thought provoking bundle of theses; this book offers over a hundred pages of clever philosophical arguments that will surely be discussed for years to come. However, I believe we should ultimately reject many of the central claims of this book. My overarching view is that the problems that Fodor worries plague CTM divide into two kinds, and neither sort of concern provides good reason to maintain that the success of cognitive science will likely be limited to the modules. The first sort of problem concerns what Fodor has called "global properties"; features that a mental sentence has which depend on how the sentence interacts with a larger plan (i.e., set of sentences), rather than the type identity of the sentence alone. The second problem concerns what many have called, "The Relevance Problem": the problem of whether and how humans determine what is relevant in a computational manner. Herein, I argue that the problem that Fodor believes global properties pose for CTM is a non-problem, and that further, while the relevance problem is a serious research issue, it does not justify the grim view that cognitive science, and CTM in particular, will likely fail to explain cognition. Thus, I shall urge that cognitive science reject Fodor's recommendation:

> For the time, concentrate one's research efforts in those areas of cognitive processing where the effects of globality are minimal; minimal enough, indeed, so that they can be ignored salve (not just reasonable adequacy but) a reasonable degree of scientific insight. (2000, p. 53)

So, what follows is, after a brief overview of the book (section one), a fairly critical reaction to it (sections two through five).

2. The Central Thread

The key dialectical thread of the book is that CTM is in deep, deep, trouble as a general theory about how the cognitive mind works. This will surprise some, for Fodor is the leading proponent of CTM, the view that the cognitive operations of the mind are computational, and more specifically, the view that information processing operations are defined on semantically interpretable strings of symbols. But Fodor thinks that CTM will fail as a general account of how the mind works, because it will not explain non-modular or "central" systems. This is actually not a new position; way back in *The Language of Thought* and *The Modularity of Mind* Fodor voiced worries about the central systems, which, in contrast to the modules, seemed to defy computational explanation. This same sort of worry fuels Fodor's skepticism today.

By "central system" Fodor has in mind a subsystem in the brain in which information from the different sense modalities is integrated, conscious deliberation occurs, and behavior is planned. A central system need not be a spatially contiguous area in the brain, although if it exists, it would be a functionally isolatable subsystem, which the modules feed into. A key characteristic of a central system is that it is "informationally unencapsulated": that is, its operations can draw from information from any cognitive domain. The domain general nature of the central systems is key to human reasoning, in which anything may be relevant to anything else. For instance, given certain contexts, the belief that Italian wine is excellent may be related to the thought that it is snowing today. Indeed, it is precisely our ability to connect the apparently unrelated that is responsible for the creativity and flexibility of human thought – consider, e.g., metaphorical thinking and analogical reasoning.

Fodor contrasts the central systems with the more domain specific sort of processing that the modules engage in. According to Fodor, modules are (*inter alia*):

> (1) Informationally encapsulated—the algorithms that characterize computations in the modules only have access to proprietary information (namely, the information in the module's domain);
>
> (2) Fast—modules are able to perform a particular function very quickly;
>
> (3) Domain Specific—modules are only concerned with a very narrow kind of input;
>
> (4) Mandatory—the algorithms that the modules compute are automatically applied.[2]

The central systems are not modular, on Fodor's view. This is because, by definition, the central systems are informationally unencapsulated. As Fodor explains: "As Kant pointed out, something in your head has to integrate all this stuff, and it's non-modular by definition." [3]

With the distinction between central systems and modules in mind, we can now turn to Fodor's concerns about the scope and limits of cognitive science. As noted, Fodor suspects that only the modules will be explained by cognitive science: in contrast, the central systems seem to defy computational explanation. He writes that this is due to "The Frame Problem" (p. 38). Most philosophers consider the Frame Problem to be the problem of how to locate a computational theory of how humans determine what is relevant. (In light of this, it is sometimes called "The Relevance Problem". Herein, I'll use this expression to avoid confusion with AI's Frame Problem, which differs in important ways).[4] The Relevance Problem is often conveyed in the following way. If one wanted to get a machine to determine what is relevant, it would need to walk through virtually every item in its database, in order to determine whether a given item is relevant or not. This is an enormous computational task, and it could not be accomplished in a quick enough way for a system to act in real time. Of course, humans make quick decisions about relevance all the time. So, it looks like human domain general thought (i.e., the processing of the central systems) is not computational.

Interestingly, in *The Mind Doesn't Work That Way* (herein "MDW") Fodor also seems to be concerned with a different sort of problem, although the problem appears to be conflated with the Relevance Problem.[5] This problem concerns a phenomenon called "globality," which, as noted, arises because a LOT sentence can make contributions to mental processing that do not supervene on the syntax of the sentence. Fodor explains:

> The thought that there will be no wind tomorrow significantly complicates your arrangements if you had intended to sail to Chicago, but not if your plan was to fly, drive or walk there. But, of course the syntax of the mental representation that expresses the thought #no wind tomorrow# is the same whichever plan you add it to. The long and short is: the complexity of a thought is not intrinsic; it depends on

> the context. But the syntax of a representation is one of its essential properties and so doesn't change when the representation is transported from one context to another. So how could the simplicity of a thought supervene on its syntax? As please recall, CTM requires it to do. (MDW, p. 26)

Fodor's rough argument (which I shall call "The Globality Problem") is the following: cognition seems sensitive to global properties. For instance, the addition of a new sentence in LOT frequently complicates an existing plan. But CTM holds that cognition, being computation, is sensitive only to the "syntax" of mental representations, (that is, cognition is sensitive to the type identity of the primitive symbols, the way the symbols are strung together into well-formed sentences, and the algorithms that the brain processes). And syntactic properties are *context insensitive* properties of a mental representation. That is, what a mental representation's syntactic properties are does not depend on what other mental representations in a plan it is combined with: it depends on the type identity of the LOT sentence. But whether a given mental representation has the global properties that it has will typically depend upon the *context* of other representations in a plan (that is, it depends upon the nature of the other LOT sentences in the relevant group). So it seems that cognition then cannot be wholly explained in terms of computations defined over syntactic properties. So CTM is false.[6]

A few lines after the discussion of globality concludes, Fodor moves to the worry that central systems have thoughts that are potentially relevant to any thought in the system's larger database (pp. 37-38). Reading over these passages, it may strike one that MDW seems to present this latter worry as somehow the same as, or at least closely bound up with, the globality worry. In contrast to MDW, in his (1983) Fodor had separated them out. Here, he had observed that central systems have *two* perplexing properties.

> *Being isotropic*: any member of an attitude set is potentially relevant to any other.[7]
>
> *Being Quinean*: certain epistemic properties are defined over a larger set of attitudes.[8]

The Globality Problem appears to be a more precise formulation of the problem that Quinean properties pose for CTM. Isotropy, on the other hand, concerns what I have called the Relevance Problem. Now, the Globality Problem may somehow be closely related to relevance, in a way which requires, for a solution to the Globality Problem, a solution to the Relevance Problem as well. But in light of the difference between the two properties above, we are definitely owed an argument for this putative connection. So until a case is provided for their connection, I will treat the problems as separate. I believe doing so is fruitful; for as we shall see, it allows one to take a divide and conquer approach to the issues.

Now let us continue with tracing the dialectic of MDW, which, at this point, takes an interesting twist. A novel element of the book is that it bridges the above worries about globality and relevance with the massive modularity debate. By "massive modularity" (MM) Fodor means the view that:

> …the cognitive mind is largely modular. This means that there is a more or less encapsulated processor for each kind of problem that it can solve; and, in particular, that there is nothing in the mind that can ask questions about which solution to a

> problem is 'best overall,' that is, best in light of the totality of a creature's beliefs and utilities. (MDW, p. 64)

In essence, MM, as Fodor conceives of it, is the thesis that the mind is almost entirely made up of dedicated, encapsulated modules that have evolved to help our ancestors solve certain problems. Now, to get to the new twist, let us ask: How does the MM issue connect to Fodor's concerns with globality and relevance? Well, if one is persuaded by these concerns, then, according to Fodor, it had better be the case that the extreme version of massive modularity is correct. For MM, as Fodor is conceiving of it, does away with the central system. And if there are no central systems, there is nothing that has the properties of globality and relevance that Fodor is worried about. And, because cognitive science seems to be making strides with respect to the modules, cognitive science is back in business. Unfortunately,

> The bad news is that, since the MM thesis pretty clearly isn't true, we're sooner or later going to have to face up to the dire inadequacies of the only remotely plausible theory of the cognitive mind that we've got so far. (MDW, p. 23)

Fodor offers the following argument against MM. If MM is true then there will be modules that are not merely in the sensory "periphery"; there will be modules that are less perceptual, and which take more sophisticated inputs. Consider, for example, Cosmides and Tooby's Cheater Detection Modules, which is supposed to have evolved to detect cheaters in the context of social exchanges. Obviously some processing is needed to determine what the inputs for the module are, for there are no "transducers" as exist in the case of sensory modules. And, given MM, this processing must be modular too. But it should be less domain specific than the cheater detection module, for it will need a broader range of inputs. And similarly, it appears that the new module would itself need a module, with a broader range of inputs, ad infinitum (MDW, pp. 71-78).

3. Globality, Relevance and Modular Computation

If one shares Fodor's interest in CTM, as I confess to, a natural first reaction to all this is that Fodor has dug us in, and dug us in deep. And it is quite ironic, to say the least, that the person most associated with CTM is doing the digging. Not funny at all.

Indeed, it is important to underscore the significance of Fodor's influence on philosophical thinking about the central systems, and relatedly, to philosophical thinking about the scope and limits of cognitive science. Fodor's worry that cognitive science has little clue how to proceed with respect to the central systems has been regarded as an extremely important line of thinking by many philosophers of mind. As far as I can tell, short of myself, philosophers interested in CTM have not denied the validity of Fodor's concerns about globality and relevance. In addition, many outside of the CTM tradition share Fodor's concern with the relevance problem, as well (e.g., Daniel Dennett, Hubert and Stuart Dreyfus, and John Searle), although their concerns are not primarily motivated by Fodor's work, but by the history of failures of Classical AI.[9] However, I suspect Fodor thinks that short of these philosophers, practically no one in cognitive science pays attention to this stuff; people just press on with their particular areas of research. Indeed, the book is inspired by popular books by Stephen Pinker and Henry Plotkin who, according to Fodor, waxed optimistic about the prospects of cognitive science, paying little attention in their respective books to these problems with the central systems.[10]

However, there are a number of issues with Fodor's discussion that may concern the reader. First, MM, as sketched by Fodor, might look like a bit of a straw man, as he construes MM as holding that there is very little integration in the brain. I know of few in cognitive science or consciousness studies who hold this view, although it may be that such a view is what you get if you took a Rodney Brooks, subsumption architecture, approach, and tried to construct an artificial brain. This strategy may yield simple robots, but this alone cannot build anything like a human brain, which, as we observed, is flexible, creative, and able to integrate seemingly unrelated material. Indeed, it is hard to see how an intelligent system exhibiting MM would even evolve; multimodal integration (i.e., integration of information from multiple sensory modalities) has obvious selective advantage—e.g., a system which can tell how auditory inputs are related to concurrent visual inputs will be better able plan successful responses to predators.[11] In any case, the appeal to a "straw man" view does not, in fact, compromise Fodor's larger argument against the plausibility of explaining the central systems. For Fodor's point is precisely that it is only this extreme view which avoids the tangle of globality and relevance. For if central systems are not amenable to computational explanation, then, as Fodor says, only the extreme form of MM will avoid the mess. And this, as Fodor notes, is hardly a comforting thought, for MM is an implausible view of human cognition. So, the natural question to ask is: do we really need to resort to MM in its extreme form?

I do not believe so, or so I shall proceed to argue in the remainder of this paper. For one thing, MM does not, in fact, avoid the pitfalls of globality and relevance. And further, globality and relevance are not insurmountable problems. To see why modular processes avoid neither globality nor relevance problems, consider a chess playing program. Suppose that a human opponent makes the first move of the game, moving a certain pawn one square forward. Now, the program needs to decide, given the information of what the previous move was, which future move to execute.

The Globality Problem emerges. To keep things simple, let us suppose that there are two game strategies/plans in the program's database, and the program needs to select one, given the first move. Let one plan involve getting the bishop out early in the game, while the other plan involves getting the rook out early in the game. (Where "early" means, say, within three turns). Now, notice that the impact that the addition of the information about what the opponent's first move was on the simplicity of each of the two plans does not appear to supervene on the type identity of the string of symbols encoding the information about the opponent's first move. Instead, the impact of the addition of the string of symbols to the simplicity of each plan depends on the way that the string interacts with the other sentences (i.e., syntactic strings) in the plan. Thus, (the Globality Argument continues) the processing of the chess program is not syntactic, and thus, not computational. So, it appears that a Globality Problem emerges in the context of highly domain specific computing.

The Relevance Problem emerges. Skillful chess playing involves the ability to select a move based on the projected outcome of the move as far into the future of the game as possible. So chess programmers routinely deal with a massive combinatorial explosion. In order to quickly determine the best move, clever heuristics must be used. This is precisely the issue of locating algorithms that best allow for the quick selection of a future move from the greatest possible projection of potential future configurations of the board.[12] And this is just the Relevance Problem, as Fodor and other philosophers have articulated it.

Now, I find it telling that both problems emerge at the level of relatively simple, modular, and uncontroversially computational processes. First, this indicates that Fodor's suggestion that MM is a refuge from problems that plague the central systems is problematic. Second, if both problems can occur in the context of uncontroversially computational processes, the presence of a globality or relevance problem does not entail that the system in question is non-computational.

Because it is implausible to claim that relevance determination is non-computational when the case involved is an uncontroversially algorithmic system, Fodor must say that the Relevance Problem, as it presents itself to the central systems, is somehow distinctive. That is, it is distinctive in a way that suggests that relevance determination in the central systems is non-computational. An obvious point of difference between the cases is that unlike modular processing, central processing is domain general. But this point of difference doesn't seem to warrant the view that the processes in question would be non-computational. For one thing, there are already programs that carry out domain general searches over vast databases. Consider internet search engines. In about 200 ms. one will receive an answer to a search query involving two apparently unrelated words that involved searching a database of over a billion webpages. Second, Fodor's Relevance Problem concerned how to sift through massive amounts of data in real time, and domain generality entails nothing about the size of a database that a relevance search draws from. A database recording the mass of every mass bearing particle in the universe would be topic specific, yet be of a much greater size than a human's memory.

Perhaps there is an alternate construal of the Relevance Problem that Fodor has in mind; one which supplies the needed point of difference. More options for interpreting the Relevance Problem will be canvassed shortly, when I take up the Relevance Problem in more detail. But for now, I shall merely suggest that a very different way to proceed with respect to the Relevance Problem is to assume that the presence of a human relevance problem is not terribly different from relevance problems existing for other *computational* systems. However, in the human case, the 'solution' is a matter of empirical investigation of the underlying brain mechanisms involving human searches. This alternative approach assumes that evolution has provided *homo sapiens* with algorithms that enable quick determination of what is relevant, and further, it is the job of cognitive science to discover the algorithms. On this view, Fodor's injunction that research in cognitive science rest at the modules should be resisted.

4. Tackling Globality and Relevance

Thus far, I have sketched the main argument of MDW, and inserted some doubt that MM indeed is a refuge from globality and relevance problems. The fact that the problems emerge in cases of simple computation may indicate that something is wrong with Fodor's arguments that the problems present obstacles to computational explanation of the central systems. Now I will proceed to locate the underlying flaws in the arguments, beginning with the Globality Problem.

The problem with The Globality Argument is that the same LOT sentence may differ in the effect it has, depending upon the nature of the other sentences in the plan. This is compatible with the requirement that syntax be context insensitive; that is, the requirement that tokens of the same symbol type will make the same syntactic contribution to every belief set that they figure in. The same mental sentence can do so, for all a sentence contributes to a computation is its type identity, and this may have a different impact on different plans/groups of sentences. The impact depends upon the type identity of the added sentence, together with the nature of the algorithms and the type identity of the other sentences in the group.[13] Analogously, consider the case in which one adds a new premise to an existing argument in first-order logic. Put into a different argument, the same premise may have a different impact; for instance, it may now bring about a contradiction. But the difference in impact, while not purely being a matter of the type identity of the premise alone, is syntactic nonetheless. For it depends on the type identity of the premise, together with the type identity of the other sentences in the argument, and the rules.

This response can be stated in terms of Fodor's earlier example. The contribution of a given belief, say, the belief that plane tickets are expensive, will be constant, while what we might call "the interaction effect" is not, since that depends on its interaction with the constant contributions of the other elements in the plan. The constant contribution is the syntax of the belief, while the interaction effect (a global property) depends upon the logical relationships between the belief and the other members of the set. The contribution may differ, while the syntax stays the same. All this is compatible with cognition being computational and symbolic.[14]

Now let me turn to the Relevance Problem. Perhaps the best way to proceed is to state and respond to the various concerns that might motivate the view that relevance determination in the central systems is likely to be non-computational. Since there is much talk of lengthy searches, it is useful to begin with considerations for the view that there are features of the brain which facilitate quick searches, yet somehow are likely to be non-computational.

1. *Human searches are faster because the human brain is faster.* This is due to the fact that the brain is massively parallel, and further, parallel in three dimensions. Humans have on the order of one hundred trillion interneural connections, operating simultaneously.[15] While it is correct that the information processing capacity of the brain vastly outpaces current computers, this does not indicate that human reasoning is non-computational; it just indicates that it may have computational resources that current information technology does not possess.[16]

2. *Worries that the algorithms will be "intractable".* Both Fodor and Carruthers worry about what they call "intractable" computations in the context of the Relevance Problem.[17] This might lead some to suspect that they are concerned that the algorithms that humans use to compute what is relevant are computationally intractable, in the sense detailed by complexity theory. But again, Fodor is not suggesting that relevance is algorithmic. And neither he nor Carruthers have explained why any algorithm that humans use to determine what is relevant would be intractable in the technical sense. And further, even if it turns out that the given algorithm is intractable, computer scientists routinely deal with intractable algorithms all the time, employing heuristics.[18] Nature

would do something like this too, for it is a given that humans determine what is relevant in real time; *if* the processing is indeed algorithmic, the fact that we respond in real time is clear evidence that evolution has obviously found a way to bypass unfavorable complexity results.

3. *The failures of AI indicate that human relevance determination is likely to not be computational.* Let me distinguish some issues that seem to be conflated, making the Relevance Problem seem like a far more serious obstacle to CTM than it really is. The relevance problem that CTM faces can be distinguished from the one that AI faces in some crucial ways. For consider what we might call the "AI challenge":

> AI challenge: build systems that can engage in domain general, commonsense reasoning.

The Relevance Problem, as it relates to AI, is a challenge to engineer such a system. To solve the AI challenge, an actual system has to be built; to solve the CTM challenge, in contrast, one needs to illustrate that domain general thought is likely to be computational.

Prima facie, cognitive science might uncover the principles that humans use to decide what is relevant in fields such as working memory and attention long before machines are built that do these things. We should look to these accounts, rather than to the Classicist AI programs of the 70's and 80's, for the most up to date information about the computational basis of relevance. After all, for the most part, Classicist AI of the past borrowed little from concrete research on the human brain (and indeed, far less information about the brain was available).[19]

A further reason why the AI and CTM problems differ is the following: in AI one might build a system that determines what is relevant that is nothing like a human brain. This would, technically, solve the relevance problem for AI, but not so for CTM. For these reasons, it seems important to keep the AI and CTM issues separate. Indeed, it may make the problem CTM faces seem less overwhelming when it is separated from the AI problem. For the previous failures of AI to build a system that decides what is relevant are daunting. They may prematurely encourage one to conclude that the brain is not computational.[20]

So I assume that for the Relevance Problem to get off the ground, it would not be based on the failure of Classical AI during the 70's and 80's. Nor would it merely be a concern about processing speed or computational complexity. The issue would concern an inability of cognitive science to locate algorithms describing how humans carry out their searches. For Fodor's relevance objection to the central systems to succeed, he needs to establish that this scenario is more likely than not. But is this pessimistic view warranted? I suspect that it is not.

5. Why opt for pessimism?

Contrary to Fodor's sense of the terrain, in which explanation in cognitive science seems to stop at the modules, there are numerous computational explanations of phenomena in the central systems. I will quickly canvass some that strike me as being of particular bearing to the matter of relevance determination. First, the discovery of algorithms for determining relevance is one major task of the fields of working memory and selective attention, which have clearly made great advances over the course of the last decade.[21]

Second, recently, Dahaenne, Changeux and Naccache have been investigating long range neurons that play a role in multimodal or central thought, in hopes of explaining the neural basis for signaling relevance. Such neurons are said to, "break the modularity of the cortex by allowing many different processors to exchange information in a global and flexible manner.[22] They are working within the broad tradition of Global Workspace Theory, (where they call their version of the theory the "Global Neuronal Workspace Theory") originally developed by Bernard Baars, who offers a functional decomposition of the central system, detailing the process by which states go in and out of consciousness. The authors explain:

> The model emphasizes the role of distributed neurons with long-distance connections, particularly dense in prefrontal, cingulate, and parietal regions, which are capable of interconnecting multiple specialized processors and can broadcast signals at the brain scale in a spontaneous and sudden manner. The concept of a 'global neuronal workspace'…builds upon Fodor's distinction between the vertical 'modular faculties' and a distinct 'isotropic central and horizontal system' capable of sharing information across modules.[23]

While it is obvious that much more needs to be done to understand centrality, it seems to me that cognitive science is beginning to grasp computational features of the central systems. If Fodor believes this is not so or that this trend will not continue, then he owes us an explanation this.

In fact, advocates of the Global Workspace Theory have offered the beginnings of a solution to Fodor's Relevance Problem.[24] To see what they have in mind, consider the example that Daniel Dennett has used to illustrate the frame problem, which involves a robot that tries to dismantle a bomb, while slowly walking through of each item in its data set, trying to figure out what is relevant to the task. In the meantime, the bomb detonates, and the robot explodes.[25] As Shanahan and Baars note, such examples seem to rely on the view of a system carrying out a serial computation that walks through a long list of alternatives one by one. If the database is large, as the database in the central system is said to be, then such a model of information flow would be problematic, given the constraints of real time. Indeed, it sometimes seems as if Fodor is using "central" to allude to something like a CPU in the brain, in which all mental operations are sequentially executed. This is something that Steve Pinker and Dan Dennett also suspect that Fodor holds. For instance, in *Consciousness Explained*, Dennett accuses Fodor of upholding a theater model in which there's a CPU in which "it all comes together" and in which there is a "central meaner."[26] When it comes to relevance determination, this leaves us with a rather naive model of information flow (see figure 1).

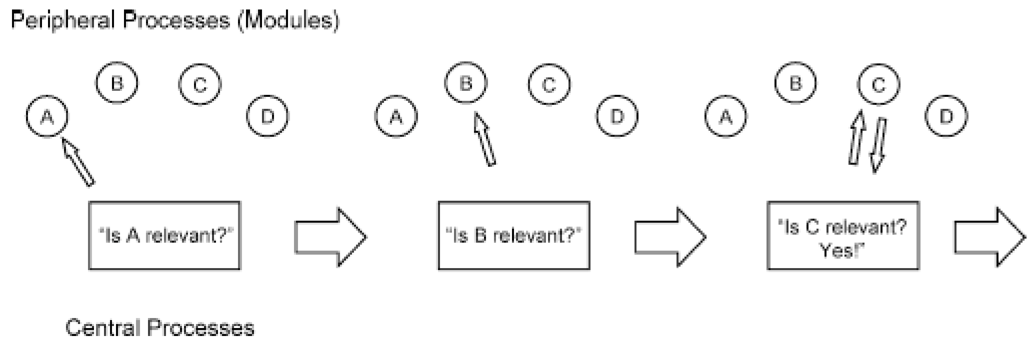

Figure 1: Naïve Model of Information Flow[27]

This view is implausible as a model of human relevance determination, for as is well known, the cognitive brain is massively parallel. But this strong notion of centrality it is certainly not required by CTM.[28] The mark of a central system is informational unencapsulation, not sequential processing.

This being said, let me proceed to sketch the GW solution. The GW theory is an influential scientific theory of the nature of consciousness. Philosophers taking very different positions on the nature of consciousness have nonetheless declared agreement that the GW model, (including the related Global Neuronal Model), are plausible computational theories of consciousness. As Dennett explains:

> Theorists are converging from quite different quarters on a version of the global neuronal workspace model of consciousness, but there are residual confusions to be dissolved. In particular, theorists must resist the temptation to see global accessibility as the *cause* of consciousness (as if consciousness were some other, further condition); rather, it *is* consciousness.[29]

According to the GW view, the role of consciousness is to facilitate information exchange among multiple specialized unconscious processes in the brain. Consciousness is a state of global activation in a "workspace" in which information in consciousness is broadcast back to the rest of the system.[30] At any given time, there are multiple parallel processes going on in the brain which receive the "broadcast." Access to the global workspace is granted by an attentional mechanism; the material in the workspace is then under the "spotlight" of attention and is processed in a serial manner.[31] (And this seems intuitive, as many of our conscious, deliberative, thoughts seem to be serial). As Shanahan explains:

> To my mind, one of the attractions of GWT is that there is no such central processor [CPU]. Rather, we find that high-level cognitive processes are the emergent product of a blend of serial and parallel processing. The serial procession of states in the global workspace can be thought of as the trace of a high-level cognitive process. But this serial procession of states … reflects the combined contributions of massively parallel competing and co-operating processes. (Shanahan, personal communication)

This picture of information flow in the central systems stands in stark contrast to the situation depicted by the naïve information flow (figure 1). On the GW view, when the brain asks what is relevant to a given fact, multiple unconscious processes search and compete for access to the global workspace. The winning competitor has its information broadcast into the GW (and thus, into consciousness). From the first-person perspective, the contents of the workspace seem to unfold serially, but each step is the end result of massive parallel processing. Information may then be broadcast back to the specialist processes for further searching. (Figure 2)

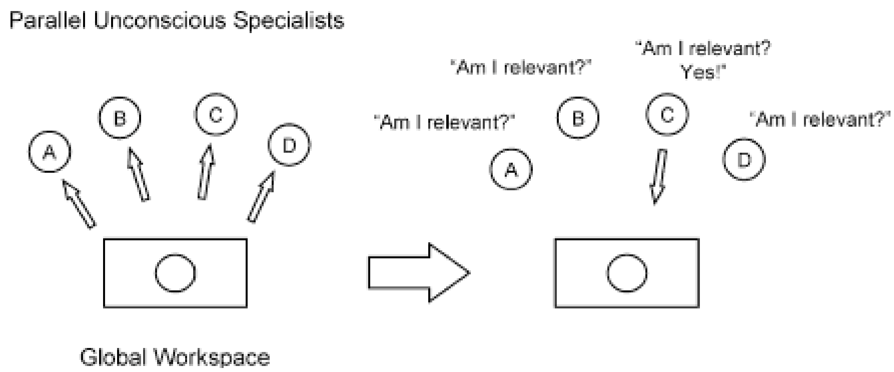

Figure 2: The GW Model of Information Flow[32]

Now, this is clearly intended to be an outline of an answer, rather than a complete account.[33] But bearing in mind the significant empirical support for GW, the suggestion may indeed reflect, albeit in very broad strokes, what the brain is doing.[34] The general point is this: the GW response, together with the body of work on the central systems in cognitive science, suggests that it is far from clear how it is justifiable to adopt the strong claim that such research areas will ultimately fail to progress to the point of locating theories that capture how humans determine relevance. Fodor and other skeptics owe us a justification for this pessimistic view.

Indeed, within the consciousness studies literature many regard information processing problems (such as relevance) as being “easy problems” (Chalmers), because they are thought to have discoverable solutions that await further developments in cognitive science. In contrast, the relevance issue, in the context of critical discussions of computationalism, is supposed to be an in principle reason why CTM, and other computational theories of cognition, cannot succeed. There is clearly some sense of disconnect between these contextually influenced judgments. Perhaps what is needed for perspective on the relevance problem is precisely the hard problem of consciousness, which makes the computational problem look to be merely a question that lacks a complete answer, awaiting more detail from cognitive science.

6. Omissions, Skeletons, etc.

One final thought: Let us suppose that Fodor is correct that globality and relevance are serious worries for CTM. Fodor claims that even if CTM fails to explain the central systems, CTM is still a plausible theory of how part of the mind works. I find this tremendously puzzling. LOT is primarily supposed to be a story about the workings of the central system, for this is the domain that engages in deliberation and reasoning. Why should we believe that if CTM fails to characterize the central systems, that it nonetheless manages to correctly characterize the modules? I throw this question out for further reflection, as I haven't a clue what to say. Good thing the central system is not so mysterious after all.

---

1. Many thanks to Bernie Baars, Tim Bayne, Mark Bickhard, Jerry Fodor, Stan Franklin, and Kirk Ludwig for thought provoking discussions on these issues.

2. Jerry Fodor, *The Modularity of Mind*. Boston: MIT Press, 1983, pp. 47-99. Fodor stops short of saying whether features (2)-(4) are essential. He simply says that a module has "most or all" of these features. (p. 47) However, he intends encapsulation to be an essential property of modules.

3. Fodor, personal correspondence.

4. In his useful overview of the different versions of the Frame Problem, Murray Shanahan distinguishes the original Frame Problem in AI from the philosopher's relevance problem. This latter problem is that of, "how the robot could ever be sure it had sufficiently thought through the consequences of its actions to know that it hadn't missed anything important." The former, in contrast, is: "Using mathematical logic, how is it possible to write formulae that describe the effects of actions without having to write a large number of accompanying formulae that describe the mundane, obvious non-effects of those actions?" Shanahan, Murray, "The Frame Problem", *The Stanford Encyclopedia of Philosophy (Spring 2004 Edition)*, Edward N. Zalta (ed.), URL <http://plato.stanford.edu/archives/spr2004/entries/frame-problem/>. For a discussion of different approaches to the traditional frame problem see Shanahan's (1997), *Solving the Frame Problem: A Mathematical Investigation of the Common Sense Law of Inertia*, Cambridge, MA: MIT Press.. See also Pylyshyn, Z.W. (ed.) (1987), *The Robot's Dilemma: The Frame Problem in Artificial Intelligence*, Norwood, NJ: Ablex.

5. See pp. 37-38 for the conflation.

6. For a longer discussion of this issue and a more formal version of Fodor's argument see Ludwig and Schneider, "Fodor's Challenge to the Classical Computational Theory of Mind", *Mind and Language*, forthcoming.

7. Fodor, *op. cit.,* 1983, p. 105.

8. Fodor, ibid., p. 107.

9. See infra, for further discussion of the challenge presented to CTM by failures of Classical AI.

10. Steven Pinker, (1999), *How the Mind Works,* W. W. Norton & Company; Henry Plotkin, (1997), *Evolution in Mind.* Boston: Harvard University Press. Fodor's original discussion of these issues appeared in the *London Review of Books* Vol 20, No 2, 15 January 1998. He conceived as MDW as a means to expand on this review.

While it may be true that Pinker and Plotkin both paid little attention to globality and relevance problems in their books, Pinker at least is no proponent of MM, as Fodor defines it. Instead, Pinker has a weaker notion of a module, according to which the central systems may be in part modular. (See pp. 30-31 and pp. 314-315, Pinker, op. cit.). So while the central systems may be in part modular, there is plenty of integration between the modules, and he doesn't really ascribe to MM. For more detail, see the interchange between Pinker and Fodor on MDW in the Feb. 2005 issue of *Mind and Language*. Pinker may have more or less ignored the relevance problem in his book, but his reply to Fodor suggests that he finds constraint satisfaction networks promising. (p. 13).

11. For discussion of this issue see Bjorn Merker, "The liabilities of mobility: A selection pressure for the transition to consciousness in animal evolution," *Journal of Consciousness Studies*, 2005. See also Charles Spence and Jon Driver, *Crossmodal Space and Crossmodal Attention*, Oxford: Oxford University Press, 2004.

12. T. Anthony Marsland and Jonathan Schaeffer, eds. *Computers, Chess, and Cognition*. New York: Springer-Verlag, 1990.

13. For further discussion see Kirk Ludwig and Susan Schneider, op. cit.

14. Ludwig and Schneider, op. cit.

15. Ray Kurzweil, (2005), *The Singularity is Near*, NY: Viking, p. 504, note, 27.

16. For a discussion of the development of three dimensional circuitry see ch. 3 of Kurzweil (op. cit.). I will not delve into this because, as I say at infra, x, the issue is not whether a domain general computational system can be built that determines what is relevant; the issue is whether cognitive science will fail to discover computational algorithms describing human relevance determination.

17. Fodor, (op. cit.) 1983, 2000. Peter Carruthers, "Distinctively human thinking: modular precursors and components." In P. Carruthers, S. Laurence and S. Stich (eds.), *The Innate Mind: Structure and Content*. Oxford University Press, 2005. In worrying about intractability Carruthers is considering Fodor's notion of a non-modular, central system.

18. Shanahan, op. cit. p. 9

19. However, I suspect that the future development of AI will borrow heavily from, and proceed roughly in parallel with, empirical discoveries about human and non-human brains. There is currently a massive amount of interest in getting machines to compute algorithms from actual neural processes, and developments follow fairly rapidly from the discovery of the underlying algorithms. E.g., see work in the field of neuromorphic engineering.

20. Relatedly, it is unfair to conflate the LOT, or symbol processing approach, with the crude, logic driven, approaches to AI of the 70's and 80's. If there are "rules" of cognition that an AI system could use, they could be algorithms discovered by cognitive science, not lines of code written at the armchair.

21. For a nice overview of progress in these fields see the chapters on these topics in Michael Gazzaniga, Richard Ivry, and George Mangun, *Cognitive Neuroscience*, (2nd edition), W.W. Norton and Co., 2002.

22. Stanislas Deheane and Jean Pierre Changeux, *Neural Mechanisms for Access to Consciousness*, forthcoming in Michael Gazzaniga, et.al., *Cognitive Neurosciences*, (3rd edition). See also S. Dehaene and L. Naccache. "Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework", *Cognition* 2, 79 (2001).

23. Daheane and Changeux, op. cit., p. 3.

24. Murray Shanahan and Bernie Baars, "Applying Global Workspace Theory to the Frame Problem," *Cognition*, Volume 98, Issue 2, December 2005, pp. 157-176.

25. Daniel Dennett, "Cognitive Wheels, the Frame Problem of AI." In Pylyshyn, op. cit.

26. Daniel Dennett, *Consciousness Explained*, Little, Brown, 1991. Steven Pinker, "So How Does the Mind Work" *Mind and Language*, Vol. 20, No1. Feb. 2005, p. 6. Pinker notes remarks in *The Mind Doesn't Work that Way* in which Fodor associates CTM with the view that the mind is interestingly like a Turing machine. (MDW, p. 30, p. 105, note 3).

27. Baars and Shanahan, op. cit., p. 12.

28. Indeed, in his interchange with Smolensky Fodor points out that LOT is compatible with parallel processing.

29. Dennett, “Are We Explaining Consciousness Yet?” *Cognition*, 2000.

30. Bernie Baars, “The global workspace theory of consciousness”, Max Velmans and Susan Schneider, *The Blackwell Companion to Consciousness*, Oxford: Blackwell Publishers, (in press). Baars, B.J. (1988). *A Cognitive Theory of Consciousness*. New York: Cambridge University Press. Baars, B.J. (1997). *In the Theater of Consciousness: The Workspace of the Mind.* New York: Oxford University Press.

31. S. Dehaene and L. Naccache, op. cit., 2001.

32. This figure is from Shanahan and Baars, op. cit., p. 13.

33. Inter alia, interesting recent work on the GW theory includes the implementation of a GW architecture in artificial systems (Stan Franklin) as well as certain of Dahaene, Naccache and Changeux’ work, which is intended to provide detail as to the structure of the global workspace. As noted, they concentrate on special “workspace” neurons with long distance axons which exchange signals in a fluid fashion, allowing for the manipulation of information across the modules. This allows for the mobilization of various brain areas, bringing their contents to consciousness. (Dehaene and Naccache, op. cit.)

34. For an outline of the empirical support see Baars, op. cit., in press.