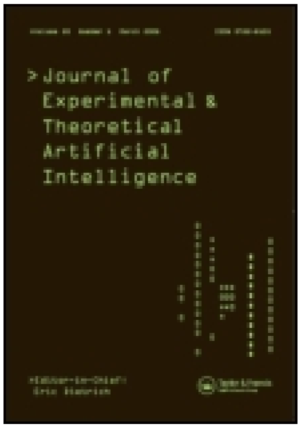

Journal of Experimental & Theoretical Artificial Intelligence
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/teta20

The role of the frame problem in Fodor's modularity thesis: a case study of rationalist cognitive science

ERIC DIETRICH [a] & CHRIS FIELDS [b]

[a] Department of Philosophy, State University of New York, Binghamton, NY, 13902-6000, USA

[b] 412 Apodaca Hill, Santa Fe, NM, 87501, USA

Version of record first published: 30 Jun 2010.

J. EXPT. THEOR. ARTIF. INTELL. 7(1995)279–289

The role of the frame problem in Fodor's modularity thesis: a case study of rationalist cognitive science

ERIC DIETRICH† and CHRIS FIELDS‡

†*Department of Philosophy, State University of New York, Binghamton, NY 13902-6000, USA*
‡*412 Apodaca Hill, Santa Fe, NM 87501, USA*

1. Introduction

Some cognitive scientists view intelligence, especially human intelligence, as approximating ideal rationality. Humans are smarter than chimpanzees, in part, because humans more closely approximate ideally rational creatures. For such scientists, explaining human intelligence requires explaining how we approximate ideal rationality (for example, see Doyle 1989, 1991).

This rationalist view of human intelligence has some influence behind the scenes in all of cognitive science, but it is especially popular in artificial intelligence and philosophy of mind. Jerry Fodor's modularity thesis is a philosophical theory of mind which is squarely in the rationalist camp (Fodor 1983, 1985, 1986, 1987). His thesis cleanly and clearly distinguishes human rationality from the rest of human cognitive behaviour, and has been influential, in part, because it provides an architecture of mind which explains why we *fail* to be ideally rational.

Our main purpose in writing this paper is to question the rationalist view of human intelligence. We think that intelligence does vary across species, and that this fact is scientifically interesting. But, intelligence should *not* be measured in terms of approximating ideal rationality. Human intelligence should not be viewed from the perspective of rationality at all. In fact, rationality should be demoted within the cognitive sciences. Human intelligence should instead be viewed biologically: human intelligence is just one evolutionary response to the problem of surviving in an environment—it is rational to the extent that it works. What makes humans more intelligent than chimpanzees is our ability to solve problems that are quite novel given our background knowledge. This ability cannot be explained in terms of approximating ideal rationality. (What does explain it is an open theoretical question. Our view is that solving strongly novel problems is related to our ability to categorize perceptual experiences.)

The problem with the rationalist agenda is that ideal rationality is incompatible with mechanistic, computational psychology. Any rationalist theory, therefore, is committed to the view that the smarter the organism, the less mechanistic it is, and hence, the less explicable its intelligence is. Fodor embraces this conclusion (1983, pp. 126–129). His modularity thesis is, therefore, an exemplar of rationalist cognitive science, and we will use his theory as a case study of this view. Specifically, we argue that the modularity thesis is incompatible with computationalism. (Computationalism is the hypothesis that cognition is computation. It is the theoretical foundation of modern cognitive science, and we will assume that it is true; Dietrich 1990.) We interpret this incompatibility as symptomatic of

0952-813X/95 $10·00 © 1995 Taylor & Francis Ltd.

the modularity thesis's commitment to the notion of ideal rationality. The incompatibility is shown in its starkest light when one considers the role of the frame problem in Fodor's thesis.

We explain the modularity thesis and the role of the frame problem in it in Sections 2 and 3, respectively. In Section 4, we show the incompatibility between the thesis and computationalism. In Section 5, we suggest a way of preserving what we consider the fundamental insight in Fodor's thesis.

2. The modularity thesis

The modularity thesis is the hypothesis that the human mind comprises a set of informationally encapsulated modules together with a central processor which is not informationally encapsulated. The informationally encapsulated modules ('modules' hereafter) are the perceptual input mechanisms responsible for vision, hearing, tasting etc. They take as input information from the external world, process it, and pass it on eventually to the central processor. There are one or more modules for each perceptual faculty. A module can be construed as a 'special-purpose computer with a proprietary database [where], (a) the operations that [the module] performs have access *only* to the information in its database (together, of course, with specifications of currently impinging proximal stimulations); and (b) at least some information that is available to at least some cognitive process is not available to the module' (Fodor 1985, p. 3; his emphases). The notion of 'information encapsulation' is captured in points (a) and (b), above.

Because they are informationally encapsulated, modules must get by on less than all of the information in the entire system, and, in fact, less than all the information available to the central processor. But more importantly, modules must perform their various input analyses with less than all of the information *relevant* to such analyses (Fodor 1983, 1985, 1987). Moreover, modules typically process only one *type* of information, i.e. modules tend to be 'domain specific, so that—to cite the classic case—the computational systems that deal with the perception of language appear to have not much in common with those that deal with, for example, the analysis of color, or of visual form (or, for that matter, the analysis of nonspeech auditory signals)' (Fodor 1985, p. 4).

The central processor is not informationally encapsulated. It has access to everything the system knows—a sort of epistemic carte blanche. It is, therefore, capable of handling different types of information as well as vast quantities of information. The central processor requires such license because its primary task is 'belief fixation'. The central processor integrates all of the information passed to it by the various modules, proposes new beliefs or *hypotheses* based on the integrated information, attempts to confirm the hypotheses, and updates the system's stored beliefs when a hypothesis is confirmed.

The analogy with confirmation of scientific hypotheses is obvious and deliberate. It is a central tenet of the modularity thesis that the process of belief fixation is *equivalent* to hypothesis confirmation in scientific reasoning: the central processor adds to the belief store just those beliefs which it can inductively confirm (Fodor 1983, esp. pp. 104–110; 1987).

On the modularity thesis, inductive confirmation is rational in the ideal sense: confirmation is rational in this sense if and only if (1) a non-arbitrary sample of all the relevant evidence is considered, and (2) the level of confirmation of any one hypothesis is sensitive to the level of acceptance of any other and to all the hypotheses taken collectively. Property (1) of inductive confirmation is called *isotropy* (Fodor 1983, pp. 104–117, esp. pp. 109–112). Isotropy is the property that 'the facts relevant to the confirmation of a scientific hypothesis may be drawn from anywhere in the field of previously established empirical ... or demonstrative truths' (Fodor 1983, p. 105). In short, 'our botany constrains our astronomy' (Fodor 1983, p. 105). Under the modularity thesis, confirmation is also *Quineian*; this is property (2) (Fodor 1983, pp. 105–117, esp. pp. 107–112). Confirmation is Quineian when 'the degree of confirmation assigned to any given hypothesis is sensitive to [the level of confirmation of all known hypotheses]; (Fodor 1983, p. 107).

Since the central processor operates exactly the way hypothesis confirmation in science is claimed to operate, the operation of the central processor is also both isotropic and Quineian (Fodor 1983, p. 110). This is another way of stating that the central processor is informationally *un*encapsulated (Fodor 1983). The processes carried on by the central processor are processes in which 'a **non**arbitrary sample of the available and relevant evidence' is brought to bear on the problem or task at hand (Fodor 1987, p. 140; Fodor's boldface). And, 'the level of acceptance of any belief is sensitive to the level of acceptance of any other and to the global properties of the field of beliefs [i.e. the belief database] taken collectively' (Fodor 1983, p. 110). According to the modularity thesis, humans owe their success as a species in large part to the fact that their central processors are isotropic and Quineian. It is these two properties that enable central processors to correctly confirm novel hypotheses about our environment while keeping belief stores optimally (or nearly optimally) updated. In short, these two properties are what make us intelligent (Fodor 1983, 1987).

Modules differ from the central processor in other ways, too. Modules are relatively fast, they only perform preliminary analyses of their input, and they are typically associated with fixed neural architectures. The central processor, on the other hand, is relatively slow, analyzes its input in depth, and is not localized in the brain. (For more on these differences, see Fodor 1983, 1985, and 1986.) For our purposes here, we will summarize the modularity thesis as making two major claims: (1) perceptual input mechanisms are modular, and thus only have access 'to less than all of the information at the disposal of the organism' (Fodor 1987, p. 139), and (2) the human mind is not composed completely of modules; the central processor is not a module. Claim (2) is the one incompatible with computationalism.

3. How the frame problem became the most important problem in cognitive science

In the 18 years between 1969, when it was first systematically treated, and 1987, the frame problem went from being a technical problem within artificial intelligence to being one of the most important problems in cognitive science. In 1987, Fodor equated the frame problem with 'the problem of how the cognitive mind works' (Fodor 1987, p. 148) and claimed that understanding how the mind works requires unravelling the nature of inductive relevance and rationality. This was quite a change in status for what initially was a somewhat obscure problem in philosophical logic. Here is what happened during those 18 years.

McCarthy and Hayes's (1969) frame problem was the problem of how to design a logic-based representation of events and an associated inference algorithm so that an intelligent autonomous system such as a robot could, as a consequence of performing any actions A, infer all and only those changes associated with A. Their original frame problem was thus a problem for AI researchers who viewed reasoning as a logical process, and were developing logical representation languages and associated algorithms as models of reasoning. As an example, if a robot had grasped some object, say a toy block, and was moving it from one point to another, then it should have inferred that its hand was occupied and that its hand, arm, and the object were changing positions relative to its centre of gravity.

But the robot need not have inferred that as the toy block moved, neither it nor the room housing the robot changed colour, nor should the robot have inferred that its hand had the same number of fingers after moving the object as it did before moving the object. Simply put, the frame problem was the problem of blocking the vast number of inferences about what had not changed as the result of some action *A* while allowing the small number of inferences about what had changed as a result of *A* (McCarthy and Hayes 1969).

The frame problem is compelling because it has proved to be quite difficult to determine ahead of time which inferences to block and which to allow. To put it another way: it is hard to determine in advance which inferences are going to be relevant to *A* and which are not, because it is hard to know in advance the context in which *A* will occur.

There are two perspectives from which to view the frame problem, and these perspectives are apparent even in the earliest discussions of the problem. From the *external* perspective (from outside the machine), the frame problem is a problem for the designer or developer of an AI system. The 1969 formulation of McCarthy and Hayes discussed above has this perspective; the *designer* is the one who must develop the appropriate logical representation language and associated algorithms. Hayes's (1973) formulation of the frame problem also has an external perspective. He formulated the frame problem as 'the problem of finding adequate collections of laws of motion', where 'laws of motion' are rules that describe the changes that the world undergoes from one situation to the next. The designer is the one who must find the adequate collection of rules.

More recently, McDermott (1987) had discussed an external formulation of the frame problem. he says 'It is widely believed that there is a special problem of getting machines to do as well as people in making all the useful inferences in a situation without any of the useless ones. This is often called the "frame problem" ' (p. 112). *Getting* machines to do as well as humans is obviously the designer's problem. (See also, Hayes 1987, p. 128).

From the *internal* perspective (from inside the machine), the frame problem is a problem for the system itself. When the frame problem is formulated internally, it is stated as the problem of how a system should correctly and efficiently revise its beliefs when performing some task (such as performing some action or solving some problem) or analysing some input. For example, in the paper where Hayes defines laws of motion (Hayes 1973), he also discusses the fact that a child playing with toy blocks knows that moving one toy block has no affect on the other blocks. Hayes is viewing the child as a system that correctly updates its beliefs as it moves a block around. Raphael (1976, p. 147) defines the frame problem as 'the problem of maintaining an appropriate informational context or frame of reference, at each stage during problem-solving [or planning] processes'. This formulation is clearly from the internal perspective: the *system* is the one that needs to maintain the appropriate informational context. Three other formulations from the internal perspective are Dennett's (1978, p. 125)

> ...we must have internal ways of up-dating our beliefs that will fill in the gaps and keep our internal model, the totality of our beliefs, roughly faithful to the world,

Doyle's (1980), p. 232)

> Since we base our actions on what we currently believe, we must continually update our current set of beliefs. The problem of describing and performing this updating efficiently is sometimes called the *frame problem* (Doyle's italics)[1].

and the title of a book on the frame problem: *The Robot's Dilemma: The Frame Problem in Artificial Intelligence* (Pylyshyn 1987).

Fodor also views the frame problem from the internal perspective. He asks 'How, then, does the machine's program determine which beliefs the robot ought to re-evaluate given that it has embarked upon some or other course of action?' (Fodor 1983, p. 114).

The internal and external perspectives complement each other. Unless there were an internal perspective on the frame problem, there would not need to be any external perspective. Somehow, humans don't get bogged down inferring what has remained unchanged in their immediate environments after they have done something. If one wants to *build* an intelligent system, one has to design it so that it also will not get bogged down. Hayes's discussion of the child moving toy blocks clearly exhibits this complimentarity. He asks: 'What assertions could we write down which would capture the knowledge that the child has about the world [i.e., that moving one block leaves the others unchanged]' (Hayes 1973, p. 48).[2]

When viewed internally, the frame problem is equivalent to the problem of inductive confirmation confronting scientists. (Both the scientist and the robot need to reevaluate, revise, and update their beliefs as they confirm various hypotheses. In so doing, both need to decide which information is relevant to confirming a hypothesis, and then which beliefs need revising once a given hypothesis is confirmed.) Since we already know that belief fixation is equivalent to scientific hypothesis confirmation and that we do fix beliefs and confirm scientific hypotheses regularly and with some ease, it follows that belief fixation embodies a solution to the frame problem. Or to put it another way, when fixing a belief, the central processor solves versions of the internal frame problem. (Of course, it also follows that scientific confirmation embodies a solution to the frame problem.) Therefore, figuring out how scientific confirmation works and developing a cognitive theory of belief fixation are equivalent: both require understanding how the (internal) frame problem gets solved (in the one case by scientists, and in the other case by brains) (Fodor 1983, 1987).

This result might gladden some researchers. Developing a theory of higher cognition merely requires figuring out how brains solve their frame problems. Fodor has worked hard to disabuse such comforted souls of this mistaken notion. He says that how brains solve the frame problem will elude us until we have a theory of *global computation*, i.e. a theory of how a processor could be isotropic and Quineian. We are currently completely ignorant of this kind of computation (Fodor 1983, pp. 112ff. and pp. 128–129; 1987). Moreover, Fodor argues that we are likely to remain ignorant about this kind of computation until far into the future, perhaps until we have finished doing science and understand everything we are going to understand (Fodor 1987).[3]

Now we can see why the frame problem is so important. Our central processors solve it routinely, but it is a problem so profound that we ourselves (qua cognitive scientists) will be ignorant of how they accomplish this feat until the end of scientific inquiry. Another way to put this is that our central processors approximate ideal rationality to some interesting extent, but how they do this is now, and will be for a long time, a deep mystery (Fodor 1983, 1987).

4. The central processor is not a computer of any sort

We establish three results in this section.

(1) The central processor is not a computer of any kind, i.e. it cannot be described as any partial recursive function.[4]

(2) A computational theory of global processing, as described by Fodor, is not possible. Phrased in his terms, our result is that no computer could be Quineian and isotropic.

(3) The central processor cannot be a physical system at all since any physical system no matter how complex is simulable to any degree of detail physically possible by

some Turing machine (Fields 1989). This is a strong and novel result. It means that *nothing* could be Quineian and isotropic.

Result 2 entails result 1, so we argue for 2. Result 1 establishes the incompatibility between computationalism and the modularity thesis since the former is the claim that *every* aspect of cognition is completely describable by a partial recursive function.

Before proceeding with our argument, we need to pin down the quarry. There is a perfectly good sense in which ordinary computation is global: the computation of any step in a multi-step algorithm depends on the computations of the other steps. In fact, it is the composition of functions in such a case that defines the algorithm as a single entity, rather than as series of independent one-step algorithms. Beyond this, 'non-ordinary' computation, i.e. connectionist systems, are computational systems where the cognitively interesting information and computation are distributed widely over large subsections of the whole system, sometimes even over the whole system. Fodor means more by 'global' than this.

Fodor gives one particularly strong definition of 'global' when defining isotropy. This definition is given in terms of arbitrariness and relevancy: a global information processing system never considers an arbitrary subset of the information which might be relevant to confirming a hypothesis, nor does it update an arbitrary subset of its beliefs when adding a new belief (Fodor 1987; see also the quote about the non-arbitrariness of isotropic processes in section 2). A local (non-global) information processing system, by constrast, will either consider an arbitrary subset of the information that might be relevant to a hypothesis, or will update an arbitrary subset of the information in its belief store when adding a confirmed hypothesis to it, or both. However, no computational device processes information non-arbitrarily.

The relevance of a piece of information to a particular hypothesis can rarely be determined in advance. If the world were otherwise, we would now know what information is relevant to curing AIDS, stopping the greenhouse effect, and developing warm superconductivity. Instead, the relevance of some piece of information to a hypothesis is seen only after the fact. Who would have thought that Mendelian genetics was relevant to evolution? It's obvious now, but when Mendel's work was rediscovered in the early part of this century it was thought to be an *alternative theory* to Darwinian evolution. And in scientific discovery, relevance is even harder to discern: who would have thought that the structure of the solar system was relevant to the structure of the atom?

Given this fact about relevant information, if a computational system is to consider non-arbitrary subsets of the information which might be relevant to a hypothesis, it will have to consider all the evidence in the universe just to confirm a single hypothesis. This is impossible (it would either take forever, or the lifetime of the universe). Therefore, either the central processor is not a computational system at all, or it prunes the information it must consider.[5] But there is no way to prune out only the irrelevant information because there is no way to determine in advance what information is irrelevant and what is relevant. Indeed, for any hypothesis, scientific or otherwise, information is continually being discovered which is relevant to it; no hypothesis is confirmed or disconfirmed once and for all. Hence, any pruning of the information which might be relevant might in fact prune off relevant information. Any system using pruning, therefore, will likely bypass some relevant information. This is sufficient to make such a system informationally encapsulated and hence a module (i.e. not isotropic). The central processor is *not* a module, so it does not use pruning to limit the information it must consider when confirming a hypothesis. But pruning or not being a computer of any sort are the only alternatives. Hence the central processor must not be a computer of any sort.[6]

Fodor also has a weaker definition of 'global' when defining Quineianness (Fodor 1983). On this definition, a system is a global information processing system when the functions operating on one piece of information in the system can *explicitly* side-effect every other piece of information in the system or can explicitly depend on all of the rest of information in the system (or both). Note that this kind of system *is* restricted to an arbitrary subset of the possibly relevant information: the information in the system. However, this is not sufficient to make the central processor a genuine computational system.

One of the central processor's jobs is keeping the belief store updated (Fodor 1983, pp. 112–117). We have bent over backwards by agreeing with Fodor that there is such a thing as a belief store, i.e. that beliefs are stored in a kind of memory. But there is no such memory (for what kinds of memories current cognitive theory says there are see Tulving and Schacter 1990, and Roediger and Craik 1989). Beliefs are *generated*, not stored. To make this obvious, note that you believe that it is completely dark inside your liver, that there are no horses on the moon, and Fodor is probably not related to either one of us. Even ordinary beliefs are generated, but this happens so quickly and routinely that it is hard to notice.

Given this, how is the central processor to keep its beliefs optimally updated when adopting a newly confirmed hypothesis? It cannot do it by updating the information it uses to generate all of its beliefs because until they are generated, it doesn't know what they are (see the examples just above). So, it has to generate them all and check to see if they are affected by the new belief. How many beliefs can the central processor generate? The number is vast. Consider how many beliefs you generate as you live your life. If you had to check each one against each new belief you adopt, you could spend your entire cognitive resources checking a single belief and still never finish in your lifetime. It follows from this that any computational system that updated its beliefs would update only a small subset of the information it uses to generate its beliefs and hope for the best—a strategy well-known in AI. In Fodor's terms, such a system could only update a fraction of its beliefs. Therefore, such a system is a module, (i.e. not Quineian). Since the central processor is not a module, it must not be a computational system at all.

We have shown that if the central processor is Quineian and isotropic, it cannot be a computer in any current sense of the term. More formally, the central processor is not simulable by any Turing machine because, even on the weak interpretation of 'global', it solves problems in a reasonable amount of time which are either undecidable or intractable. Hence, either the central processor is not executing an algorithm (no undecidable problem is computable by a Turing machine) or it instantiates an extremely unusual and fast architecture (no known architecture—including massively parallel machines—can solve an intractable problem for the large number of inputs the central processor must consider in the lifetime of the universe). So, either the central processor is not a machine at all, or it is a machine we cannot even conceptualize. Computationalism is the thesis that cognition is computation *as we currently understand it*, i.e. as it is expressed in Turing machine theory and the associated theory of computability. Of course, virtually all the details of the algorithms actually executed by the brain remain to be discovered, but discovering that the central processor cannot be simulated by any Turing machine or that our understanding of computation is seriously flawed would be discovering that computationalism is false. It might very well be false, but we have no reason now to believe that it is, and every reason to believe that it is true.

Now for our argument that there is no such physical system as the central processor. Fields (1989) has shown that the theory of measurement which is part of quantum mechanics entails that any measurable behaviour of any system can be fully described computationally. Quantum mechanics (and hence its theory of measurement) is the best supported scientific theory currently in existence. Therefore, any theory which postulates the existence of a noncomputational process, postulates a *non-physical* process, as physics is currently conceived. Given our current understanding of the physical world, therefore, there can be no such device, computational or otherwise, like the central processor.

To summarize, if the central processor solves Fodor's frame problem, i.e. if it is Quineian and isotropic, computationalism is false (and so is quantum mechanics). Since the modularity thesis entails that the central processor can solve Fodor's frame problem, the modularity thesis is incompatible with computationalism. Computationalism, not the modularity thesis, is the more plausible thesis.[7]

5. How to reconcile the modularity thesis and computationalism by using promiscuity

In spite of our criticisms, we think that the modularity thesis is an important theory of mind. Foremost among its attributes, in our opinion, is the role it assigns analogical reasoning and creativity. We know of no comparable theory (with the exception of Hofstadter, 1995) which places as much importance on this crucial aspect of human cognition. Fodor says:

> ...mainstream Cognitive science has managed to get the architecture of the mind *almost exactly backwards*. ... By attempting to understand thinking in terms of a baroque proliferation of scripts, plans, frames, schemata, special-purpose heuristics, expert systems, and other species of domain-specific intellectual automatisms—jumped-up habits, to put it in a nutshell—it missed what is most characteristic, and most puzzling, about the higher cognitive mind: its...creativity, its holism, and its passion for the analogical (p. 4, 1985, Fodor's emphasis).

We agree with Fodor that the central processor has a 'passion for analogical', and that this fact must be explained if we are ever to understand human cognition and construct intelligent artifacts. Our only disagreement with Fodor is over the central processor's alleged unencapsulation. The central processor is not unencapsulated; nothing could be. Nevertheless, the central processor is responsible for creative problem solving which it accomplishes through analogical reasoning. We have our own proposal for how the central processor does this and our proposal reconciles the modularity thesis with computationalism; being a computationalist doesn't mean giving up creativity.[8]

Reconciliation depends on specifying computational mechanisms which are *quasi*-isotropic and *quasi*-Quineian, i.e. mechanisms which approximate isotropy and Quineianness but are nevertheless computable. Such mechanisms would consider a variety of information when attempting to confirm a hypothesis and attempt to keep beliefs (confirmed hypotheses) more or less consistent with one another *only* by executing Turing-computable functions. Besides the general requirement of being computable, each mechanism must meet one further requirement. We take these in turn.

The quasi-isotropic mechanism should not primarily consider information that is *prima facie* relevant to the hypothesis being considered, especially when the hypothesis is novel, as it must be when the organism or system is confronted with a new problem. Since genuine relevance is determined after the fact (as we argued in Section 4), the mechanism will have to consider beliefs that are *prima facie* irrelevant to the hypothesis. Only in this way will it have a reasonable chance of confirming novel hypotheses.

This requirement can be met by making the quasi-isotropic mechanism *promiscuous*. Since it does not know ahead of time which information and which of its other beliefs are relevant to a given hypothesis, it should (at least some of the time) compare beliefs with the hypothesis almost indiscriminately, though not completely indiscriminately because the probability of finding such beliefs is quite low. It can improve this probability enough, we suggest, by considering beliefs which have some simple, *prima facie* relevance, for example, some simple orthographic or syntactic similarity to the given hypothesis. We are suggesting, therefore, that the quasi-isotropic mechanism use slight similarities between beliefs and its current hypothesis to guide it in searching for beliefs that can confirm the hypothesis. This promiscuous strategy strikes a balance between complete random, promiscuity (which in general, would not work) and the pedestrian strategy of considering only those beliefs which fit some relevance criteria fixed ahead of time (which will not work when the hypothesis is novel). Of course, the quasi-isotropic mechanism should not be promiscuous to the exclusion of obvious guides to information which is relevant and will confirm the hypothesis for it. But, we are suggesting that the promiscuous strategy should be part of the repertoire of the quasi-isotropic mechanism.

Now for the quasi-Quineian mechanism. This mechanism must update as many beliefs relevant to a confirmed hypothesis as it can. To do this, it must continually generate beliefs (recall there is no such thing as the belief store) and compare them with one another. However, if it generated beliefs randomly it would never maintain confirmational harmony. On the other hand, it cannot update just those beliefs which are obviously relevant to a particular confirmed hypothesis because these are almost certainly not the only beliefs that are relevant. Again, therefore, a promiscuous strategy is required. Given a confirmed hypothesis (a new belief) the quasi-Quineian mechanism should almost indiscriminately—guided only by orthographic or slight syntactic similarities—search for beliefs to compare with this new belief, updating those that require updating. And again, we are not suggesting that the quasi-Quineian mechanism *only* be promiscuous; it should, of course, update those beliefs that obviously need updating. But if the promiscuous strategy were added to its repertoire, it would update many more beliefs than the merely obvious ones.

To sum up, we claim that a central processor equipped with two promiscuous mechanisms like those just described will (1) confirm novel hypotheses some interesting percentage of the time, and (2) manage to maintain a tolerable degree of confirmational harmony among its beliefs. Furthermore, such a central processor is realizable as a Turing machine or computer. We have, therefore, reconciled computationalism and the modularity thesis.

Finally, notice that on our proposal intelligence is not understood in terms of approximating ideal rationality. Rationality is completely in the background. Promiscuous mechanisms of the kind described are rational if they work, i.e. if they keep the organism alive long enough to produce offspring and get it's genes into the next generation. It is meaningless to say that an ideally rational organism would be more or less cognitively promiscuous than humans.

6. Conclusions

The modularity thesis tries to explain human intelligence as approximating ideal rationality. Any processor which was isotropic and Quineian would be ideally rational, or nearly so (Fodor 1987, p. 140). It would also be virtually omniscient (contra Fodor, 1983, pp. 119–126). It is not conducive to good cognitive science to view human intelligence as failed omniscience. This perspective will encourage us to ask the wrong questions, such as 'What is global information processing?' and it will lead us down unproductive research paths, such as trying to develop a theory of global, nonmechanistic computation.

The proper viewpoint for cognitive science is biology. We should seek an explanation of human intelligence that makes it continuous with cockroach intelligence. Humans are smarter than cockroaches, and any theory of human intelligence should preserve this fact, but our similarities to cockroaches and other fauna (and indeed flora) are as important as our differences. If we want to explain why humans are so intelligent, we would do well by taking these similarities more seriously.

Acknowledgements

The authors thank Imre Balogh, James Fetzer, Ken Ford, Clark Glymour, Rebecca Gomez, Niall Graham, Pat Hayes, Robin Hill, Tony Plate, Roger Schvaneveldt, and Herbert Simon for comments and discussion on the ideas in this paper.

Notes

1. This formulation is particularly interesting because Doyle uses the phrase 'frame problem' to refer both to the scientific problem of describing belief revision and the internal, neurocomputational problem of performing it.
2. Philosophers tend to be interested only in the internal perspective because viewed from this perspective, the frame problem becomes, as Dennett (1979, p. 75) notes, 'an abstract espistemological problem'.
3. In his 1987 paper (pp. 146–147), Fodor suggests that one way to develop theory of global computation would be to develop a principle of 'scientific conservatism'. This principle would provide a formal language for representing all of our scientific theories such that the best theory in every domain would have a canonical property captured completely by the language. Such a principle would, therefore, allow us specify (a) exactly what a minimal change is to any given scientific theory given new data (thus capturing Quineiannes), and (b) which information is relevant to what, relative to any scientific theory (thus capturing isotropy). He then states our chances for developing such a principle:
   "If somebody developed a vocabulary for writing scientific theories which had the property that the shortest description of the world in that vocabulary was always the best theory of the world available, that would mean that the notation gave FORMAL EXPRESSION to our most favored inductive estimate of the world's taxonomic structure. Well, when we have an inductive estimate of the world's taxonomic structure that's good enough to permit formal expression, and a canonical vocabulary to formulate the taxonomy in, most of science will be finished" (Fodor 1987, p. 147).
4. We use the terms 'computer', 'Turing machine', 'partial recursive function', and 'machine' synonymously in this section. Technically, we are going to show that the central processor is not simulable by any Turing machine, or, alternatively, that the central processor does not execute a partial recursive function. For those interested in the formal definitions of these terms see Rogers (1987).
5. Fodor cannot restrict the search to information currently *in* the system, for this is always an arbitrary sample of the evidence. First, at any given time the system has encountered only a very small amount of all the information there is, and these encounters have been governed by caprice and serendipity. Secondly, even if the central processor in fact had all the relevant evidence, it could not know this until all the other facts in the universe were checked. Again therefore, in attempting to confirm a hypothesis, the central processor must consider every fact in the universe.
6. In AI, pruning strategies are called *heuristics*. Heuristics are by no means arbitrary if this word is taken to mean capricious or unreasonable. Nevertheless, because we don't know for sure which information is relevant to any problem of interest, any heuristic may in fact block consideration of some relevant information or evidence. This fact about heuristics is well known to AI researchers.
7. Our argument against Fodor's thesis has been theoretical. There is some experimental evidence that the thesis is false. See Maunsell and Newsome (1987), DeYoe and Van Essen (1986) and Sloman (1989).
8. We are offering only an inchoate hypothesis of the central processor; many details will have to be left to the reader's imagination. Even if our hypothesis is wrong, it demonstrates that computationalism and modularity thesis *can* be reconciled.
9. Some of the material in this paper is reprinted with permission from 'The Wanton Module and the Frame Problem', in *Philosophy and the Computer*, edited by Leslie Burkholder, 1992, Westview Press, Boulder, Colorado.

References

Dennett, D. (1978) *Brainstorms* (Cambridge, MA: Bradford).

DeYoe, E. and Van Essen, D. (1988) Concurrent processing streams in monkey visual cortex. *Trends in Neurosciences*, **11**: 219–226.

Dietrich, E. (1990) Computationalism. *Social Epistemology*, **4**: 135–154.
Doyle, J. (1980) A truth maintenance system. *Artificial Intelligence*, **12**: 231–272.
Doyle, J. (1989). Reasoning, representing, and rational self-government. In Z. Ras (ed.) *Methodologies for Intelligent Systems, 4* (New York: Elsevier), pp. 367–380.
Doyle, J. (1991) The Foundations of Psychology: A logico-computational inquiry into the concept of mind. In R. Cummins and J. Pollack (eds) *Philosophy and AI: Essays at the Interface* (Cambridge, MA: MIT Press) pp. 39–78.
Fields, C. (1989). Consequences of nonclassical measurement for the algorithmic description of continuous dynamical systems. *Journal of Experimental and Theoretical Artificial Intelligence*, **1**: 171–179.
Fodor, J. (1983) The Modularity of Mind (Cambridge, MA: MIT/Bradford).
Fodor, J. (1985) Precis of *The Modularity of Mind. Behavioral and Brain Sciences 8*, pp. 1–42.
Fodor, J. (1986) The Modularity of Mind: Chapter 1 and session 1. In Z. Pylyshyn and W. Demopoulos (eds) *Meaning and Cognitive Structure* (Norwood, NJ: Ablex) pp. 3–29 and pp. 129–167.
Fodor, J. (1987) Modules, frames, fridgeons, sleeping dogs, and the music of the spheres. In Pylyshyn (ed.) *The Robots Dilemma: The Frame Problem in Artificial Intelligence* (Norwood, NJ: Ablex) pp. 139–149.
Hayes, P. J. (1973). The frame problem and related problems in artificial intelligence. In A. Elithorne and D. Jones (eds). *Artificial and Human Thinking* (Amsterdam: Elsevier) pp. 45–59.
Hayes, P. J. (1987) What the frame problem is and isn't. In Pylyshyn (ed.) *The Robot's Dilemma: The Frame Problem in Artificial Intelligence* (Norwood, NJ: Ablex) pp. 123–137.
Hofstadter, D. (1995) *Fluid Concepts and Creative Analogies* (NYC, NY: Basic Books).
Maunsell, J. H. R. and W. T. Newsome (1987) Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience*, **10**: 363–401.
McCarthy, J. and Hayes, P. J. (1969) Some philosophical problems considered from the standpoint of artificial intelligence. In B. Meltzer and D. Michie (eds) *Machine Intelligence 4* (New York: Elsevier) pp. 463–502.
McDermott, D. (1987). We've been framed: Or, Why AI is innocent of the frame problem, in Pylyshyn (ed.) *The Robot's Dilemma: The Frame Problem in Artificial Intelligence* (Norwood, NJ: Ablex) pp. 113–122.
Pylyshyn, Z. (ed.) (1987) *The Robot's Dilemma: The Frame Problem in Artificial Intelligence* (Norwood, NJ: Ablex).
Raphael, B. (1976). The frame problem in problem-solving systems. In N. Findler and B. Meltzer (eds) *Artificial Intelligence and Heuristic Programming* (Edinburgh: Edinburgh University Press) pp. 159–169.
Roediger, H. and Craik, F. (eds) (1989) *Varieties of Memory and Consciousness: Essays in Honor of Endel Tulving* (Hillsdale, NJ: Erlbaum).
Rogers, H. (1987) *Theory of Recursive Functions and Effective Computability* (Cambridge, MA: MIT Press).
Sloman, A. (1989) On designing a visual system. *Journal of Experimental and Theoretical Artificial Intelligence*, **1**: 289–337.
Tulving, E. and Schacter, D. (1990) Priming and human memory systems. *Science* **247**: 301–306.