

Fodor's Challenge to the Classical Computational Theory of Mind

KIRK LUDWIG AND SUSAN SCHNEIDER

Abstract: In *The Mind Doesn't Work That Way*, Jerry Fodor argues that mental representations have context sensitive features relevant to cognition, and that, therefore, the Classical Computational Theory of Mind (CTM) is mistaken. We call this the Globality Argument. This is an in principle argument against CTM. We argue that it is self-defeating. We consider an alternative argument constructed from materials in the discussion, which avoids the pitfalls of the official argument. We argue that it is also unsound and that, while it is an empirical issue whether context sensitive features of mental representations are relevant to cognition, it is empirically implausible.

1. Introduction

In *The Mind Doesn't Work That Way* (TMD), Jerry Fodor, the founding father of the Classical Computational Theory of the Mind (CTM) argues that although computationalism will likely succeed in explaining modular processes, it will fail in explaining what he calls the 'central system'.¹ By a 'central system' Fodor has in mind a non-modular, domain general subsystem in the brain in which information from the different sense modalities is integrated, deliberation occurs, and behavior is planned. According to Fodor, cognitive science, and CTM in particular, will be unable to explain the central system. The main reason is due to a version of the

We wish to thank Jerry Fodor for helpful written comments on a related paper and for many interesting discussions on this issue. Thanks also go to Dominic Murphy for his comments in a symposium on an earlier version of this paper at the 2005 Eastern Division Meetings of the American Philosophical Association, the audience at the 2005 Society for Exact Philosophy meeting at University of California at San Diego in May 2006, and to two anonymous referees for *Mind & Language*. The name order of the authors is alphabetical.

Address for correspondence: Kirk Ludwig, Philosophy Department, University of Florida, Gainesville, FL 32611-8545, USA; Susan Schneider, Philosophy Department, University of Pennsylvania, 423 Logan Hall, 249 S. 36th Street, Philadelphia, PA 19104-6304, USA.

Email: kludwig@phil.ufl.edu; sls@sas.upenn.edu

¹ *The Mind Doesn't Work That Way: the Scope and Limits of Computational Psychology*, London: MIT Press, 2000. All references to page numbers, unless otherwise indicated, are to this book. We will use 'CTM' throughout to refer to the Classical Computational Theory of Mind. This theory holds that the syntactic properties of mental representations that are relevant to cognition are ones accessible to computation on the basis of intrinsic syntactic properties. These are properties that are essential to the representations and are non-relational properties of them. Thus, we use 'CTM' to refer to what Fodor calls 'E(CTM)' in chapter 2 of TMD. Fodor contrasts E(CTM) with M(CTM), which allows relational syntactic properties to be relevant to cognition. As Fodor notes, this is not compatible with a Classical Turing architecture. We return to the contrast in section 6.

Frame Problem, which we shall call ‘The Globality Problem.’ It arises from the thought that beliefs have certain ‘global properties’ that are relevant to cognition.² ‘Global properties’ are properties that particular beliefs (or other attitudes) have that are supposed to be determined by the nature of an attitude set which they are members of.³ For instance, a belief about plane tickets being expensive may complicate a plan to go to China but not a plan to drive to Boston.⁴ So whether a belief tends to complicate planning or not depends on what other attitudes are involved in the plan and how the belief interacts with them. Therefore, its contribution to complexity of planning appears to be a global property of it. If we call what it contributes to the complexity of one’s planning the degree of simplicity/complexity of the belief (‘simplicity’ for short), its degree of simplicity seems to be a global property of it.

Thus, cognition seems sensitive to global properties. But global properties, if relevant to cognition, appear to pose a severe problem for the CTM. For CTM holds that cognition, being computation, is sensitive only to the syntax of mental representations, 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 it is combined with. But whether a given mental representation has the global properties that it has will typically depend upon the context of other representations that it is embedded in (that is, it depends upon the nature of the other attitudes in the relevant group).⁵ If this is right, then such properties are not, and do not supervene on, syntactic properties. Cognition then cannot be wholly explained in terms of computations defined over syntactic properties. If so, then CTM, as a general theory of how the mind works, is untenable.

² A number of problems parade under the banner of ‘The Frame Problem.’ The version Fodor focuses on as central in *TMD* stands out from the usual formulations in focusing on problems global properties raise in relation to the nature of LOT syntax, in particular. Another well-known version is often called ‘The Relevance Problem’. We say more about this distinction in section 6. For more discussion of the frame problem, see Pylyshyn, 1987.

³ Often we will talk of beliefs being a member of a set of beliefs or attitudes. Beliefs are states of thinkers, so by talk of a belief being a member of a set of attitudes *S*, we mean its being had by someone who has the attitudes in *S*.

⁴ Fodor gives examples involving both plans and theories, and treats these as distinct in the text. A plan is a plan for action; a theory is a set of beliefs about how things are. In a footnote he says, ‘[I]acking explicit notice to the contrary, ‘plan’ and ‘theory’ are just cover terms for any collection of one or more propositions. ... For expository convenience, I’ll often use ‘theory’ to cover both theories and plans’ (p. 109). It is not quite clear to us what he has in mind by ‘any collection of one or more propositions’, but we believe the thought is that plans and theories construed as psychologically real are to be taken to consist of the presence of LOT sentence tokens in appropriate functional roles for different sorts of attitudes, desires, intentions, and beliefs, for plans, and just beliefs, for theories.

⁵ Not *all* global properties are context sensitive, however. For example, belonging to a set of attitudes is a global property that every attitude carries with it from set to set. But evidently many global properties will be context sensitive. And these will raise a problem for CTM if Fodor is right that some global properties that are context sensitive are relevant to cognition. It is these context sensitive global properties that we will be specifically concerned with.

We will argue that this important challenge of Fodor's to CTM is unsuccessful. The plan of the paper is as follows. In section 2, we explain the notion of syntax that is relevant to Fodor's argument, and explain why it must be construed as context insensitive. In section 3, we spell out Fodor's argument in detail. This reconstruction follows Fodor in formulating the argument in terms of the example involving the simplicity of a belief. We show that the argument has a false premise, and, in particular, a premise which we can demonstrate to be false using only definitions and assumptions which Fodor himself has introduced and endorsed in the argument. The argument is, therefore, self-defeating. In section 4, we raise the question of whether Fodor has misstated his argument, given that the official version is self-defeating. We consider whether there is another, more successful, argument in the vicinity of the one we criticize, and suggest one possibility. We then show that this version of the argument is also unsound. In section 5, we consider whether the lesson generalizes to properties other than that of simplicity. We argue that in the end it is an empirical question whether there are properties relevant to cognition of the sort which Fodor thinks make trouble for CTM, but also that it is highly plausible that there are not. In section 6, we consider briefly an alternative to the Classical Computational Theory of Mind which Fodor introduces as perhaps a way to save CTM from his criticism, but then criticizes in turn. The alternative allows relational syntactic properties. We head off the charge that our own characterization of how syntax can contribute to computation collapses into the relational version of CTM. In section 7, we distinguish the Globality Problem from a problem with which it is apt to be confused, the Relevance Problem. We conclude in section 8.

It is not part of our project in this paper to vindicate CTM, or to respond to all of the challenges that face it, or to respond to all of the challenges that Fodor himself has raised to it. Rather, it is to show that the Globality Problem, which Fodor makes the centerpiece of TMD, and presents as an in principle objection to the Classical Computational Theory of Mind serving as a general account of cognition, does not succeed, and so does not close the door on the prospects of success for CTM through further empirical research.

2. Syntax in the Language of Thought

Given that Fodor's version of the Frame Problem pays careful attention to the notion of the syntax of a mental representation, it is crucial, at the outset, to explain the relevant sense in which there is supposed to be a Language of Thought (LOT) syntax, and the related sense in which computational properties are said to be 'syntactic.' Speaking of 'LOT syntax' is a bit misleading because the operative notion is not the well-known sense of 'syntax' that involves principles governing well-formed expressions. Although constituents in the language of thought combine to form thoughts in ways that are 'syntactic' in this familiar sense, by

'LOT syntax' philosophers of mind frequently mean something more than this. LOT syntax consists in:

- (i) the class of vocabulary items in the language of thought: that is, the primitive symbols in LOT, and the larger expressions that are built up from the primitives;

and,

- (ii) the rules of composition that apply to the items in the LOT vocabulary.

Correlatively, 'syntactic properties' are:

- (i) grammatical properties of the vocabulary items,

and,

- (ii) other properties that type individuate the vocabulary items.

Any notion of syntactic properties requires (ii) in addition to (i) for the following reason. Consider the case of the type individuation of primitive vocabulary items in LOT. Grammatical properties alone will not suffice for type individuating primitive symbols because any primitive vocabulary item having the same grammatical properties (e.g., being a noun) would be the same symbol. So there would be no way to distinguish between expressions that are grammatically the same, but have different meanings or referents, or are different ways of representing the same referent (e.g. 'dog'/'tree', 'mother'/'water', 'Cicero'/'Tully'). As a result, LOT would have no chance of accounting for cognition computationally. Hence, (i) alone would not yield an adequate theory of primitive expressions.

There are a number of proposals concerning what features, in addition to grammatical category, should type individuate the vocabulary items. However, we will not discuss these here; for deciding between the various views is not really the concern of Fodor's version of the frame problem.⁶ For on any of the existing proposals for individuating LOT syntax, Fodor's problem seems to emerge. This is because at the heart of the Globality Problem is a tension or conflict between a common requirement on any theory of LOT syntax—that syntactic types be context insensitive—and that there are cognitively relevant global properties like simplicity. In Fodor's own words:

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

⁶ For further discussion of these issues see Aydede, 2000a and 2000b; Pessin, 1995; Schneider, 2005, 2008 and forthcoming.

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 (p. 26).

In brief, Fodor's argument seems to be the following: the simplicity of a thought is context dependent; in contrast, the nature of syntax is not. But simplicity is relevant to cognition. So, cognition is not entirely syntactic, contra CTM.

Before developing this rough argument in further detail, it is useful to ask: why must syntax be context insensitive? For if syntax can be context *sensitive*, then there is no obstacle to simplicity supervening on syntax even if it is context sensitive, and the problem does not arise. So let us suppose that a sentence in LOT syntax is type individuated in part by the relations it has to the larger plan that it happens to figure in. Then we face a dilemma. Either mental representations are individuated independently of their syntactic type so that the same mental representations may be parts of different systems of representations, or not. If they are, then generalizations over syntax will not capture generalizations over mental representations. For the syntactic type of a mental representation will change from one system to another, and there will be no generalizations in terms of syntax that capture what we are interested in, namely, what difference a given mental representation makes to mental processes of which it is a part. If mental representations are not individuated independently of their syntactic type, then on this proposal mental representations would not in general be shared between different sets of attitudes. This would have the consequence that generalizations over mental representations, so understood, would not capture the generalizations about mental process we are interested in because those are processes that can involve the same mental states in different systems. For syntax to do the explanatory work it is supposed to do, it must not vary depending on the system of attitudes of which it is a part. It is necessary that the very same sentence of LOT can preserve its syntactic type through changes in a thinker's set of attitudes and be identified across different sets of attitudes. LOT syntax must therefore be context insensitive. More explicitly, the following is a requirement on LOT syntax if it is to do the explanatory work required of it:

Context Insensitivity (CI):

The syntactic properties of a mental representation are invariant across contexts, i.e. the properties that individuate an expression in the language of thought are independent of what set of attitudes the expression is a member of.

3. The Globality Argument

We turn now to a more detailed account of the argument. It is important to grasp the precise way in which simplicity is supposed to present a challenge to CTM in

the light of (CI). First, Fodor defines a notion of simplicity of a belief, for the purposes of the argument roughly sketched above, as follows:

... the simplicity of a [belief is] whatever determines for any given theory that you add it to, how much it complicates (/simplifies) that theory (p. 26).

He then claims that *if* the simplicity of a belief is a context insensitive feature of a mental representation, then, intuitively, it can only contribute to the simplicity of a theory by contributing a ‘constant increment (/decrement)’ to the overall simplicity of the theory (p. 26). So, for a given sentence in the language of thought, the overall contribution made by the simplicity property of the syntactic expression to each belief set must be the same. But, Fodor continues, if this is so, then it seems that adding a belief to one set of attitudes *cannot* make a small difference to the simplicity of an agent’s theory, while adding the same belief (a belief of the same type) to another set of attitudes makes a significant difference to the simplicity of the agent’s theory. And, unfortunately, this sort of thing happens all the time. So it seems that (CI) is violated on the assumption that the simplicity of a belief is a syntactic feature of it or that it supervenes on its syntactic properties

Now let us put the above observations together with some basic facts. We will piece these together into a more formal version of Fodor’s argument, which we shall call the ‘Globality Argument.’ The argument is as follows:

The Globality Argument

1. A belief can contribute to the complexity/simplicity of a theory. [fact]
2. The complexity/simplicity of a theory is relevant to cognition. [fact]
3. What a belief contributes to the complexity/simplicity of a theory is relevant to cognition. [1,2]
4. Def: ‘the simplicity of a [belief is] whatever determines for any given theory that you add it to, how much it complicates (/simplifies) that theory’ (p. 26).
5. The simplicity of a belief is relevant to cognition. [3,4]
6. If the Computational Theory of Mind is true, then the simplicity of a belief must be a context invariant feature of it.
7. If the simplicity of a belief is a context invariant feature of it, then it can only contribute to the complexity of a theory by contributing a ‘constant increment (/decrement) to the overall simplicity of’ the theory (p. 26).⁷

⁷ This is weaker than the claim Fodor makes, which is a biconditional. We need only the left to right direction for the argument.

8. If it can only contribute to the complexity of a theory by contributing a constant increment or decrement, then adding a belief to one set of attitudes *cannot* make a minimal difference to the complexity of an agent's theory, while adding the same belief (a belief of the same type) to another set of attitudes makes a significant difference to the complexity of the agent's theory.
9. However, adding a belief to one set of attitudes *can* make a minimal difference to the complexity of an agent's theory, while adding the same belief to another set of attitudes makes a significant difference to the complexity of an agent's theory.
10. Therefore, CTM is false. [6–9]

Before we evaluate this argument, we want to make a few remarks about its general character. This is what might be called a classic armchair argument. It purports to establish its conclusion on the basis of general considerations available to anyone on reflection -- from the armchair. An armchair argument is not necessary purely a priori, i.e., it does not necessarily rely only on purely a priori premises, though it often involves a priori elements together with what are taken to be obvious facts. In this argument, the underived premises are 1, 2, 4 and 6–9. 1 and 2 are taken to be obvious facts, and indeed may plausibly be claimed to be conceptual truths about beliefs. 4 is definitional in character, but it is a reference fixing definition rather than a concept giving definition. It fixes the denotation of 'the simplicity of a belief' as the property which is the denotation of a certain description. It stipulates what, if anything, 'the simplicity of a belief' is to refer to, and Fodor appears to entertain no doubts about whether it could fail to refer. 6 is a priori because it just draws out a consequence of CTM. 7 is a crucial premise. What is its status? It is introduced in the text as if it is obvious. It is *not* a premise derived from an induction on the failure of CTM, because it is a premise in an argument for the conclusion that CTM must fail. Nor is it presented as derived from a more general principle that has been inductively supported. It appears therefore to be presented to us as a premise that can be known a priori. 8 is clearly a priori, and 9 has the same status as 1 and 2. The argument as a whole, then, is presented as largely a priori. The question whether it is purely a priori hinges on whether 1, 2, and 9 are a priori, and whether 'the simplicity of a belief' can be known a priori to refer given the definition. They are, in any case, clearly taken to be obvious or nearly self-evident truths. It looks to be presented, then, as an in principle challenge to CTM, and that is what we take it to be. It is then, a very significant objection, for it purports to tell us what we cannot do, prior to any further empirical investigation.

Fodor would object to characterizing the argument as aspiring to be a priori or largely a priori because he follows Quine in rejecting the analytic/synthetic distinction, and, along with this, the a priori/a posteriori distinction. However, even for a Quinean, there is a distinction between arguments that draw on assumptions which are very central to the web of belief, like the law of non-contradiction, and ones that do not. Those that do have the character of in principle

arguments, and we might call them ‘in principle’ arguments. It is, then, for a Quinean, presented as an ‘in principle’ argument, as opposed to an in principle argument. However, this will have no practical effect on the style of evaluation that is appropriate, or its practical significance if it is correct.

We have dwelt on the character of the argument because Fodor sometimes presents himself, in his framing remarks, as merely pointing out a pattern of failure in attempts to understand the mind computationally. For example, in the introduction to TMD (p. 6), he writes:

Over the last forty years or so, we’ve been putting questions about cognitive processes to Nature, and Nature has been replying with interpretable indications of the scope and limits of the computational theory of the cognitive mind. The resultant pattern is broadly intelligible; so, at least, I am going to claim.⁸

We are going to evaluate the argument as an in principle (or ‘in principle’—henceforth we omit the qualification) argument against CTM.⁹ If it is meant merely as a kind of induction on the failure of attempts to get a handle on how beliefs contribute to such properties of systems of beliefs as their simplicity, then responding to it as an in principle argument would be mistaken. However, when we get to the details of the Globality Argument, it is clear that it is not a matter of drawing attention to repeated failures to come to grips with how simplicity and complexity of theories figure in cognition. Rather, the observations appealed to can be put in the form of a deductive argument, the premises of which are introduced as if they were obvious independently of any detailed information about the history of attempts to make the CTM work, and, indeed, the central premise is introduced in the fashion in which one normally introduces an assumption which is self-evident, like the law of non-contradiction. Notwithstanding Fodor’s framing remarks, then, the Globality Argument is appropriately taken as an in principle argument against CTM. The framing remarks can be interpreted in the following fashion. First, a Quinean would regard all of these arguments as broadly a posteriori, even if some are relatively more a priori than others, and so any conclusion we reach can be construed as in a sense something that Nature teaches us. Second, though, the framing remarks refer not to the specific, and crucial argument, we focus on in this paper, but to the larger argument of the book, in which Fodor examines a non-classical version of the Computational Theory of Mind CTM (see notes 1 and 14) and considers the problem of relevance in abductive reasoning,

⁸ Also, in his ‘Reply to Steven Pinker ‘So how *does* the mind work?’ (2005) Fodor says ‘TMD offers a diagnosis of a pattern of failures whose self-evidence it takes to be glaring’ (p. 25), suggesting the argument he mounts is inductive in character.

⁹ Recall here our introductory remarks and note 2 in which we say we will use ‘CTM’ for the Classical Computational Theory of Mind, or ‘E(CTM)’ in Fodor’s terminology in chapter 2 of TMD. See also section 6 below for further discussion.

which we will distinguish from the globality problem in section 6, with respect to which the pattern of failure remark would not be inappropriate.

We now turn to evaluating the Globality Argument. The argument begins with some basic, obvious, and perhaps conceptually grounded, facts: (1) Beliefs contribute to the complexity/simplicity of theories; (2) the complexity/simplicity of theories is relevant to cognition. We regard these as obvious. A further, obvious conclusion of (1) and (2) is that what a belief contributes to simplicity of a theory is relevant to cognition. The argument then relies on Fodor's definition of simplicity of belief for the purposes of the argument. We then draw the conclusion that simplicity, understood in this way, is relevant to cognition. So far, so good. Line six says that if CTM is true, simplicity must be context invariant. As argued above, this must be so. So if simplicity is not context insensitive, it cannot supervene on syntactic properties.

But now we arrive at what we believe is a false premise: line seven. Here, Fodor seems to think that a given sentence in LOT *must* contribute a constant increment/decrement, or *the same overall result*, to every belief set. That is, a given sentence must result in the same increase or decrease in the overall simplicity of a plan, no matter what plan is involved. This is incorrect, however, for a constant *result* is not required by context insensitivity. Only a constant *contribution* is. A LOT expression may make the same contribution to any attitude set, where the contribution is simply its syntax. So (CI) is satisfied. But the same mental representation may differ in its *effect*, depending upon the context or nature of the other beliefs in the set. In other words, the contribution may be constant, while the interaction effect is not, since that depends on its interaction with the constant contributions of the other different elements it is combined with.

To see this, let us reconsider premise 7. However, this time, let us substitute in Fodor's definition of 'simplicity':

- 7★. If the property of a belief that determines for any given theory that you add it to, how much it complicates (/simplifies) that theory, is a context invariant feature of it, then it can only contribute to the complexity of a theory by contributing a constant increment or decrement.

When we put in the definition, whatever prima facie appeal the premise had dissipates, and for good reason. For clearly there can be intrinsic features of things which when added to some systems have a small effect on certain global properties, though when added to others have a significant effect. Adding a small amount of mass, say, a gram, to the Moon, will not have much effect on its size. But adding the same amount of mass to a star that is on the brink of collapsing in on itself due to its mass may have an enormous effect on its size. So the size of some systems may be minimally affected by the addition of a quantity of mass, while the same quantity affects significantly the size of others. (Indeed, chaos theory is based on related phenomena: e.g. the falling of a tree would not generally have much effect on weather patterns in a far away region, but it is not out of the question that in certain conditions it would have a significant effect.)

Part of what makes 7 seem plausible, to the extent that it does, is the introduction of the term ‘simplicity of a belief.’ For this suggests that we are thinking of a measure of simplicity of the belief taken by itself. And thinking about that, it may seem puzzling how any intuitively reasonable conception of the simplicity of a belief could contribute in any other way than incrementally to the simplicity or complexity of a theory. It makes it look as if we were discussing a property of a system like mass, and the contributions of the same sort of property of its parts to that global property of it, where the contributions of the masses of the elements is additive. That is why substituting in the definition changes the premise from one that looks at least plausible to one that looks, to say the least, dubious.

Fodor’s Globality Argument employs a definition of the simplicity of a belief as *whatever property of it determines for any given theory how much it complicates that theory*. This is a technical term. It is not supposed to express any intuitive notion of the simplicity of a belief taken *by itself*. Indeed, it is clear that its definition rules this out. The parallel case in our example is defining the *size* of a quantity of mass as *whatever property of it determines how it contributes to the size of any object to which it is added*. Clearly it doesn’t follow that the contribution to the size of an object must be incremental, and clearly the property so defined has nothing to do with the ordinary notion of the size of something.

But even more significantly, if we consider how ‘simplicity’ was defined, we can see that it is guaranteed to be a context invariant property *that also does the right job*. As Fodor writes:

... the simplicity of a [belief is] whatever determines for any given theory that you add it to, how much it complicates (/simplifies) that theory (p. 26).

Rewriting this to make it clearer (and taking a theory to be a set of beliefs):

The simplicity of a belief B = the property P of B such that for any theory T , and any change in complexity R , if R is the change in the complexity of T on adding B to it, then that B has P determines that the change in complexity of T on adding B to it is R .

This is clearly not a property of B which changes depending on what theory it is embedded in: by definition, in fact, it is not, because it is defined in terms of what effect B ’s having P has on the change of complexity of any theory to which it is added. This requires P to be a property B possesses independently of what belief set it is a member of. Thus, the definition of the simplicity of a belief rules out that it is, in Fodor’s terms, a context sensitive property. Thus, in the light of premise 9, we can deduce that 7* and, hence, 7, is false. Thus, Fodor’s own definition of the simplicity of a belief and other premises of his argument, show that the central premise of the argument dealing with the contribution of the simplicity of a belief to the simplicity of a theory to which it is added is false. The argument is therefore self-defeating.

4. The Revised Globality Argument

The problem we have identified hinges on the manner in which Fodor introduces the term 'simplicity of belief'. Given the definition, the simplicity of a belief is guaranteed to be a context insensitive property of it that also does the right job. It is reasonable to ask whether this is really what Fodor intended. That is, does our criticism of Fodor's challenge rest on taking advantage of a careless formulation? The definition itself cannot, we think, be read in a way that does not result in 'the simplicity of a belief' picking out a context insensitive property. For the phrase introduced is a definite description, and it is meant to be read as 'the degree of simplicity of a given belief'. Fixing the belief, there can be only one degree of simplicity for it, absent any further relativization, and none is provided.

We might try to introduce at this point a different definition of simplicity that would avoid this problem. However, it will be more profitable, we think, to consider whether we can construct an alternative argument from considerations in the text leading up to the official argument. We noted above that premise 7 appears more plausible if we suppose that the simplicity of a belief is a determinate of the same determinable as that we apply to a theory or plan in calling it simple or complex, that is, if we assume that relevant property of a belief is being simple to a certain degree in the same sense in which a theory or plan is simple to a certain degree. As we have pointed out, the way that Fodor fixes what property 'the simplicity of a belief' refers to does not at all require that this be so. It might be thought, however, that whatever property of a belief would contribute to the simplicity or complexity of a theory to which it was added would be a determinate of the same determinable that theories or plans have when we call them simple. In the paragraph prior to that in which he gives the definition of 'simplicity of a thought', Fodor entertains a simple model of what an intrinsic syntactic measure a mental representation's simplicity might come to: 'as it might be, the number of [primitive] constituent representations it contains' (p. 26). The same measure might then be applied to a system of representations. On this model, the contribution of a belief to the simplicity of a system to which it belongs would be additive. Fodor points out that this would not account for the fact that the same belief may complicate one plan and simplify another. Thus, number of primitive constituents cannot be the syntactic property that simplicity of a theory or plan supervenes on. Fodor, rightly, does not suggest that this is the only way syntactic properties of a sentence of LOT might determine what they contribute to the simplicity of a system of which they are a part. There is no reason why we should not appeal to the 'logical form' of LOT sentences and what type individuates particular symbols in explaining how a LOT sentence contributes to the simplicity of theories or plans. This is the reason Fodor goes on to give a general argument, which doesn't depend on this model, for the conclusion that the syntax of mental representations cannot account for their contribution to the simplicity of systems to which they are added. Yet, it might be suggested that he is still assuming that

a belief contributes to the simplicity of theories or plans by way of having a degree of simplicity of the same sort as the degree of simplicity of the plans or theories to which it contributes, and that this is what underwrites the assumption that whatever syntactic (or other context insensitive) property that degree of simplicity supposedly supervenes on would have to contribute a constant increment or decrement.

Let us suppose that this is so. Can we use these assumptions to provide a sound argument for the context sensitivity of what a belief contributes to the complexity or simplicity of a theory? If we can, then there will be considerations in the immediate vicinity of the official argument that can be marshaled for the conclusion, and Fodor's charge that CTM cannot serve as a general account of how the mind works will be vindicated. The official argument may then be seen as an unfortunate misstatement of the underlying considerations that make trouble for CTM. The two assumptions around which we need to construct the argument are, first, that beliefs (thoughts) contribute to the simplicity or complexity of theories or plans by way of a degree of simplicity of the same type as the theories or plans have, and, second, if the property of an element and the property of a system to which it is added are determinants of the same determinable, then the contribution is additive. If we combine this with the observation that if CTM is true, the degree of simplicity of a thought must supervene on its syntax, we get the conclusion that beliefs must contribute a constant increment/decrement to any theory to which they are added. More formally:

Revised Globality Argument

1. A belief can contribute to the complexity/simplicity of a theory. [fact]
2. The complexity/simplicity of a theory is relevant to cognition. [fact]
3. What a belief contributes to the complexity/simplicity of a theory is relevant to cognition. [1,2]
4. If a belief contributes to the complexity/simplicity of a theory, then it does so by way of having a degree of simplicity of the same sort as the degree of simplicity that the theory has.
5. If the property of an element and the property of a system to which it is added are determinants of the same determinable, then the contribution of that property of the element to that property of a system to which it is added is additive.
6. What a belief contributes to the complexity/simplicity of a theory is additive. [4,5]
7. If the Computational Theory of Mind is true, then what a belief contributes to the complexity/simplicity of a theory must be a context invariant property of it. [CTM, 3]
8. If the Computational Theory of Mind is true, then what a belief contributes to the simplicity of a theory is always a constant increment/decrement. [6,7]

9. If a belief can only contribute to the complexity of a theory by contributing a constant increment or decrement, then adding a belief to one set of attitudes *cannot* make a minimal difference to the complexity of an agent's theory, while adding the same belief (a belief of the same type) to another set of attitudes makes a significant difference to the complexity of the agent's theory.
10. However, adding a belief to one set of attitudes *can* make a minimal difference to the complexity of an agent's theory, while adding the same belief to another set of attitudes makes a significant difference to the complexity of an agent's theory.
11. Therefore, CTM is false. [8–9]

Unlike the official argument, this argument is not self-defeating. However, we can show that, like the official argument, it is unsound. Our main objection will be to premise 5. But prior to that we wish to raise two questions about premise 4.

The first is whether there is indeed an appropriate notion of simplicity that applies to both beliefs and theories or plans or reasoning indifferently. This is a presupposition of premise 4. If there is not, then in light of premise 1, premise 4 cannot be true. It is not clear that a belief that it is windy, for example, is simple or complex in the same sense in which a plan or a deliberative process to which it is added is simple or complex. What makes for complexity in planning depends both on what one aims at and on what one believes about the world, particularly about means to ends available to one. The measure of complexity might be the number of independent steps that one must take to reach one's goal. There is no analog to this in a belief considered by itself. Beliefs don't involve steps or goals. Perhaps in the case of beliefs and theories (here a set of beliefs) it may be argued that the same notion of simplicity applies, because a set of beliefs in certain propositions is equivalent to a belief in the conjunction of those propositions. However, this is not true. For beliefs in propositions come in different degrees. And this affects how we adjust what we believe when we acquire a new belief. To hold onto a particular belief in the light of a new observation one may have to complicate one's theory considerably. If the degree of belief in it is high, then one's theory must be complicated; if it is low, it may be discarded. So degree of belief matters to the effect new beliefs have. So a system of beliefs is not equivalent to a belief in the conjunction of the propositions believed. And this additional element that plays a role in assessing the simplicity in theories, the relative strengths of beliefs, has no applicability in the case of a single belief. We do not mean to settle the issue whether the same notion applies to beliefs and systems of which they are components. To do this, we would need to provide clear and intuitively adequate accounts of the notions of simplicity that apply to theories, plans, deliberation, reasoning and individual attitudes. Our goal here is to draw attention to the non-obviousness of this presupposition of premise 4, and to what would be needed to establish it. Fodor does not provide any analysis of the target notion of simplicity or complexity. In the absence of that, one cannot establish 4, and there are, as we have indicated, some *prima facie* reasons to doubt that the presupposition is true.

The second question about premise 4 is whether, even if there is a notion that applies to both, we must think of the way a belief contributes to the complexity/simplicity of a theory or plan as going wholly by way of the corresponding measure of the belief. Let us return to our earlier example involving the addition of some mass to an object. The size of the object added may be relevant to the resulting size of the system to which it is added. But as our example shows, other features of it may be relevant to the resulting size as well (namely, mass), even though there is a single determinable that applies both to the added component and to the whole. It does not follow, simply from the fact that the notion of a degree of *F*-ness applies both to a component and a system, that only the degree of *F*-ness of the component can contribute to the degree of *F*-ness of the system. If there is something special about simplicity in this regard, it remains unclear what it would be, and would require a further clarification of what notion is at issue, prior to further argument.

Our main objection, however, is to premise 5, which is demonstratively false. This can be shown with a simple counterexample, illustrated in Figure 1. In this model, elements are squares, and systems are combinations of squares. The simplicity (*S*) of an element or complex is measured by the number of sides it has. The fewer the number of sides the simpler it is. In scenarios (A) and (B), the elements have the same degree of complexity, namely, 4. The system to which the element is added in scenario (A) has a degree of simplicity of 6 prior to the addition of the element to it, and only 4 afterwards. In contrast, in scenario (B), while the element has the same degree of simplicity, the system does not decrease in complexity when the element is added to it, but increases in complexity. This is a function both of the number of sides of the elements and complexes and the way they are combined. This feature of the example may be thought of as analogous to whether a mental representation is tokened in the belief box, or desire box, etc. More complicated dynamical models showing the same thing are possible. Yet this simple example is enough to show that even when the same degree-concept applies to components and the system, the components need not contribute additively to the relevant property of the system. Premise 5 above, which is required to resurrect Fodor's argument along the lines suggested above, is false, and the argument is therefore unsound.

5. General Empirical Considerations Regarding Context Sensitivity of Mental Representations and Cognition

Our discussion thus far has focused on the case of simplicity, as it is the example that Fodor develops throughout his own discussion. But there are other properties that are said to be global.¹⁰ So let us ask: are there reasons to think that any

¹⁰ Fodor briefly discusses 'conservatism', which he characterizes as 'how much it would cost to abandon a belief,' saying that this feature of a mental representation depends on what theory it is embedded in (p. 34). Other examples are the centrality of a given belief to a theory (p. 34), and the property of being relevant to a plan (p. 37).

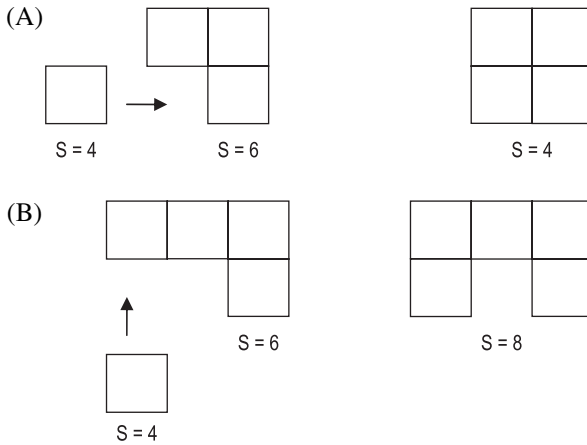


Figure 1 Counterexample to premise 5

candidate for a cognitively relevant global property will turn out, on closer examination, to fare better than simplicity of belief? As a first step toward answering this question, let's consider an example Fodor himself discusses of a putative global property which turns out not to be one. The example is the following:

... say that a text is 'globally odd' if it contains an odd number of words, 'globally even' otherwise; and consider the contribution that the sentence 'John loves Mary' makes to determining whether a text that contains it is globally odd. Query: is this contribution context dependent? Perhaps you're inclined to say, 'sure it is; because if a given text has an odd number of words, then adding 'John loves Mary' makes the resulting text globally even; whereas, if the text has an even number of words, then adding 'John loves Mary' to it makes the resulting text globally odd ... but no. To be sure, the consideration just raised shows that its contributing what it does to the texts that you add it to is a relational property of 'John loves Mary' but it's a context-independent relational property for all that. The sentence makes the same contribution whether the text you add it to is globally odd or globally even; in either case it contributes the number of words it contains. And of course, containing the number of words that it does is a syntactic, hence an essential, property of a sentence, hence not context dependent. *What is context dependent is not what a sentence contributes to determining the global oddity of a text, but rather the result of its contributing what it contributes in determining the global oddity of a text. In some contexts the result of adding three words is a text that's globally odd; in other contexts it's not*¹¹ (p. 27; emphasis added).

¹¹ Fodor credits Paolo Casalegno with the example.

The distinction that Fodor draws here between the result of a contribution and the contribution itself is what we invoked above in showing why the simplicity of a belief (as originally defined by Fodor) is not context dependent, and, hence, not, according to the official definition, a global property of it.¹² In general, in those cases in which it appears that some property of a belief must be a global property, it is plausible to think that we will find that this distinction can be employed to show that there is a property available which is not global and which explains why the result of a constant contribution varies from context to context. If there is some item that we are tracking from context to context, then it must have context invariant properties, which enable us to see it as the same thing in different contexts. If it has different effects depending on different contexts, then the overwhelmingly likely story is that a context invariant property of it (syntactic properties if CTM is right), together with the context invariant properties of the other items in the system, explains why it has different effects in different conditions. For that is the way causation works throughout nature, so far as we know.

It might be thought that we could get a stronger result. For consider the definition of simplicity above. It is introduced as the property of a belief that determines whatever its contributions are to the complexity of various sets it might be a member of. Can we not always introduce similarly a property that is guaranteed to be context invariant to explain the different effects in different contexts? Let the item in question be *I*, let 'S' be restricted to relevant types of systems, and 'E' to relevant types of effects. Here, it might be said, is the context invariant property we can appeal to:

The property *P* of *I* such that for any system *S*, and any effect *E*, if *E* is the effect of adding *I* to *S*, then that *I* has *P* determines that *E* is the effect of adding *I* to *S*.

The property *P* would be context invariant by definition. Unfortunately, it is not so easy to settle the matter. For this succeeds in picking out a property only if the matrix, 'for any system *S*, and any effect *E*, if *E* is the effect of adding *I* to *S*, then that *I* has *x* determines that *E* is the effect on adding *I* to *S*', is uniquely satisfied. Unless there is an a priori reason to think that it must be, we cannot be sure that we have secured a property by this sort of introduction rule. Neither of the forms of argument we have considered above (in sections 3 and 4) succeed in establishing that there is no such property, the first because it presupposes that there is, and the second because it relies on an a priori false assumption. But there also appears to be no a priori reason to think that there must be such a property

¹² Fodor rejects any analogy between simplicity and word number, saying 'Representations contribute the same syntactic structures whatever context you add them to; but thoughts don't contribute the same degree of complexity whatever theory you add them to' (p. 28).

for arbitrary *I* and *E*. That is to say, there does not appear to be any a priori reason to think that CTM is true. CTM appears rather to be an empirical hypothesis.

Whether CTM itself is correct, it does seem to us that, on balance, it is more plausible to believe that context invariant properties of mental representations are responsible for the simplicity or complexity of theories or plans containing them. To appreciate this, we can reflect on what would have to be so for this to be false. There would have to be properties of mental representations that were genuine relational properties of them, properties that they had only in the context of other mental representations, which figure in laws that are not deducible from the intrinsic properties of the elements of the system, their arrangements, and the laws governing them. It would be as if the relational property of *being a planet* were to figure in a brute law. This is not conceptually impossible. But the history of science provides inductive evidence for the overhypothesis that relational properties like this do not figure in brute laws, namely, that we have not so far had to countenance them in any area as ultimate laws. For example, Kepler's laws of planetary motion appeal to the property of being a planet, but have been explained in terms of more fundamental laws that dispense with this higher-level relational property.

6. E(CTM) and M(CTM)

In TMD Fodor distinguishes between what he calls E(CTM) and M(CTM). The former holds that 'only essential properties of a mental representation determine its causal role in a mental life' (p. 24). The latter holds only that the 'role of a mental representation in cognitive processes supervenes on some syntactic facts or other' (p. 29). Essential properties of mental representations are context-insensitive properties, so E(CTM) is incompatible with what beliefs contribute to the simplicity of a theory being context sensitive. Fodor's argument is directed against E(CTM). M(CTM) is supposed to be compatible with the contribution being context sensitive because it allows relational syntactic properties to be relevant to cognition. Fodor holds that the difference between E(CTM) and M(CTM) is important, for Classical computations are defined over the constituent structure of representations, not relational properties of them. This is the basis for Fodor's claim that M(CTM) is not compatible with the view that 'cognitive architecture is Classical Turing architecture' (p. 30).

The Classical Computational Theory of the Mind requires E(CTM), and this has been our concern throughout. We say above that it is very plausible that context invariant features of mental representations, *together with the context invariant features of other mental representations*, accounts for such features of the theories of which they are a part as simplicity. This may sound like a version of M(CTM), however, and so it may seem that although we have objected to Fodor's argument against E(CTM),

we have actually endorsed a version of M(CTM).¹³ But this is not so. The effect of a property of a component of a system on a determinable property of the system as a whole may be mediated by the properties of other elements of the system without the property in question being relational. This is the case with our example involving adding a gram of mass to the moon and to a star on the brink of collapsing. It is the mass together with the mass of the other components that has the effect, but the masses of none of the components are relational properties of them. Similarly, in the example Fodor discusses of a text being globally even or globally odd, it is the contribution of the number of words in a sentence, a context invariant feature of it, *together with the context invariant features of other sentences in the text*, that accounts for the text as a whole being globally odd or globally even. But that does not make the number of words in a sentence a relational property of it. Likewise, the example in figure 1 shows that the simplicity of a system may be determined by the simplicity of its elements and their mode of combination, so that it is the simplicity of an element and how it is combined with the other elements with their simplicities and arrangements which determines the simplicity of the whole, without the simplicity of the elements being a relational property of them.

7. The Globality Problem Distinguished from the Relevance Problem

Before concluding, we want to distinguish the Globality Problem, which we have been discussing, from another problem which also goes under the heading of ‘The Frame Problem’, which it is confused with in some discussions.¹⁴ This will help to locate the bit of the larger debate over the adequacy of CTM with which this paper has been concerned. To do so, we look back to Fodor’s earlier work. In *The Modularity of Mind* (1983), Fodor claimed that central systems had two properties that look to make trouble for CTM, namely:

Being isotropic: any member of an attitude set is potentially relevant to any other (Fodor, 1983, p. 105).

Being Quinean: certain epistemic properties are defined over the entire set of attitudes (Fodor, 1983, p. 107).

¹³ Our thanks to an anonymous referee for *Mind & Language* for the suggestion that further clarification on this point would be useful.

¹⁴ Steven Pinker (2005) treats the two problems we distinguish below as one problem Fodor presents for CTM. As Fodor sometimes talks of both raising difficulties for CTM because of the relevance of global properties of systems of representations to cognition, it is understandable that one could come away with the impression that Fodor regards them as the same problem. Yet the arguments in chapters 2 and 3 of TMD are distinct, the latter being more in the character of an induction on failure, and the former being presented as an in principle argument, as we have noted. They are both ‘Frame Problems’ in the sense that they concern how computational systems can carry out cognitive operations that seem to require access to information that must be extracted from a whole system of representations.

The Globality Problem seems to be a more precise formulation of the problem that Quinean properties are supposed to pose for CTM. The thought is that there are global properties of belief, properties that are defined over the entire set of attitudes, and so which depend on the other beliefs in the set. As we know, Fodor is concerned that such properties do not supervene on the syntax of the belief. Isotropy, on the other hand, concerns what many have called the 'relevance problem.' In general purpose reasoning, any belief (desire, etc.) is potentially related to any other. For example, the thought that it is cold may remind someone of Beethoven's ninth symphony (e.g. it might have been a particularly cold day when one first heard it performed). So isotropy presents the following worry for CTM: how do humans determine what is relevant to a given decision in a computational manner? And further, how can a machine determine what is relevant without needing to walk through virtually every item in its database? In order for a machine to determine which facts are relevant to making any given decision, it seems that the system must examine each of the facts in its database, in order to determine which facts are, in fact, germane. This is an enormous computational task, and it is not clear that human beings do anything like this. The fact that we can determine what is relevant in a timely fashion makes it look as if we have access to the relevance of a fact in a way that is not implemented computationally.¹⁵ Indeed, the Frame Problem of AI is often regarded as being precisely this problem. This is not, however, the same problem as the Globality Problem. The Globality Problem concerns (or at least, aside from inopportune formulations, it was supposed to concern) whether genuinely relational properties of mental representations are relevant to cognition. The Globality Problem is an in principle objection to the success of CTM, for if there are, in the relevant sense, global properties, cognition in general does not supervene on computational processes defined over intrinsic syntax. The relevance problem is not an in principle objection to CTM. The problem is not that properties relevant to cognition might not supervene on the syntax of mental representations. It is that the computational processes that would make use of the relevant properties would be so time consuming that we could not see those as an appropriate model for how human cognition works. The relevance problem remains even if we are confident that the Globality Problem can be put aside.

It is important to keep the two problems distinct. The relevance problem presents itself as an empirical puzzle rather than an in principle or a priori objection. Many would claim that cognitive science has yet to figure out how to solve it.¹⁶ But we don't have an in principle reason to think it can't be solved. We might, indeed, in this case, think that there is some inductive evidence from past failures

¹⁵ Again, see Pylyshyn, 1987.

¹⁶ One of these authors has argued that Fodor's argument that the relevance version of the frame problem is a serious problem for CTM is incorrect. (Schneider, 2007 and 2008.)

that we are unlikely to succeed. But nonetheless it is not the sort of problem, in contrast with the Globality Problem, which would urge us to give up without further effort. Confusion of the two may lead either to seeing the Globality Problem as just an empirical problem, which it is not, or to seeing the relevance problem as an in principle difficulty which makes further empirical work worthless, which it is not.

8. Conclusion

In conclusion, we have argued in this paper that Fodor's official argument in TMD against CTM is unsuccessful, and, in fact, self-defeating. The central assumption, that if the simplicity of a belief is intrinsic, it must always contribute a constant increment or decrement to a theory to which it is added, cannot be true given the definition of 'simplicity of belief' and Fodor's observation, which is also used in the argument, that beliefs do not contribute to the complexity of theories by a constant increment or decrement. We also considered whether an alternative argument for the same conclusion could be constructed from materials in the discussion, which avoided the pitfalls of the official argument. We suggested an argument which rested on the assumptions (i) that the only property of a belief which could contribute to the complexity of a theory would be a degree of simplicity in the same sense as that which applies to the theory as a whole, and (ii) that when there is a property of an element and a property of a system to which it is added that are determinants of the same determinable, the contribution of that property of the element to that property of a system to which it is added is additive. We identified some reasons to think that the first of these assumptions is not true, and argued that any attempt to provide support would have to tackle the job of explaining in detail what notion of simplicity is at work in the argument. We also argued that the second assumption is false by presenting a simple counterexample to it. We considered briefly whether we are guaranteed to be able to find a context invariant property of a mental representation that would explain its contribution to the properties of the whole of a system to which it is added by way of an introduction rule of the sort employed in the definition in section 3 of 'simplicity of belief'. We argued that we have no reason to think this can be ruled out a priori, but that it is an empirical matter whether the relevant description has a denotation. We suggested that there were general grounds for optimism because the alternative must treat higher-level relational properties as appearing in brute laws, and the history of science provides inductive evidence that there are no such brute laws. This is not quite to say that such context invariant properties will turn out to be syntactic in the sense needed for CTM (here E(CTM)), but it leaves it open that they are. Finally, we distinguished the Globality Problem from the Relevance Problem, which remains even if we feel we can safely put the Globality Problem aside, but which is, in contrast to the

Globality Problem, rather to be treated as potentially resolvable on the basis of further empirical work.

Kirk Ludwig
Philosophy Department
University of Florida

Susan Schneider
Philosophy Department
Institute for Research in Cognitive Science
Center for Cognitive Neuroscience
University of Pennsylvania

References

- Aydede, M. 2000a: On the type/token relation of mental representations. *Facta Philosophica: International Journal for Contemporary Philosophy*, 2, 23–49.
- Aydede, M. 2000b: Computation and intentional psychology. *Dialogue*, 39, 365–379.
- Fodor, J. 1983: *The Modularity of Mind*. Cambridge MA: MIT Press.
- Fodor, J. 2000: *The Mind Doesn't Work That Way: The Scope and Limits of Computational Psychology*. Cambridge MA: MIT Press.
- Fodor, J. 2005: Reply to Steven Pinker 'So how does the mind work?'. *Mind & Language*, 25–32.
- Pessin, A. 1995: Mentalese syntax: between a rock and two hard places. *Philosophical Studies*, 78, 33–53.
- Pinker, S. 2005: So how does the mind work? *Mind & Language*, 20, 1–24.
- Pylyshyn, Z. (ed.) 1987: *The Robot's Dilemma: The Frame Problem in Artificial Intelligence*. Norwood, NJ: Ablex Publishing.
- Schneider, S. 2005: Direct reference, Psychological Explanation, and Frege Cases. *Mind & Language*, 20, 423–447.
- Schneider, S. 2007: The language of thought. In P. Calvo and J. Symons, (eds), *Routledge Companion to Philosophy of Psychology*. London: Routledge.
- Schneider, S. 2008: Yes, it does: a diatribe on Jerry Fodor's *The Mind Doesn't Work that Way*. *Psyche*.
- Schneider, S., forthcoming: The nature of symbols in the language of thought: a theory.