Notre Dame Journal of Formal Logic
Volume 27, Number 3, July 1986

Information and Circumstance

JON BARWISE*

I suppose that bronzesmiths in the Bronze Age had a working knowledge of bronze, but not what we would consider a very good theoretic account of bronze. So maybe it should not surprise us to discover that the same holds for information in this Age of Information. For it does. While we process information all the time, personally and with the aid of computers, there is no semblance of agreement as to the basic nature of information and information processing. The basic notions are lacking any commonly accepted philosophical and mathematical foundations. Indeed, there are at least two apparently opposing beginnings toward such foundations, a semantic one that puts the emphasis on "information", and a syntactic version with the emphasis on "processing".

Perhaps coming up with a theory of information and its processing is a bit like building a transcontinental railway. You can start in the east, trying to understand how agents can *process* anything, and head west. Or you can start in the west, with trying to understand what *information* is, and then head east. One hopes that these tracks will meet, but Fodor's paper "Information and Association" [9], in this issue, tries to prove that they won't. Although he begins with a disclaimer that the paper is only cartography, a mapping out of the basic territory of information and information processing to see how various notions might fit into cognitive science, by the end of the paper he has convinced himself, and would convince us, that processing considerations require us to replace any world-oriented semantic account of information content by a syntactic account that has recourse to a formal language of thought.

---

*An earlier version of this paper was read in response to Fodor's "Information and Association" at the APA meeting in San Francisco, March 23, 1985.

Various ideas in this paper grew out of the STASS group at CSLI. Work on this paper was partially supported by NSF Grant DCR-840353 and partially by an award from the System Development Foundation. I would like to thank John Etchemendy, David Israel, John Perry, and Brian Smith for helpful comments on earlier versions of the paper.

Received February 7, 1986

Fodor argues that our ordinary pre-theoretic understanding of information, information processor, and information processing presupposes an underlying internal language, or language of thought, so that any theoretical account of information compatible with these intuitions and a theory of processing will have to incorporate such a formal language. For example, he claims that mental activity, including inference and information processing more generally, corresponds to formal manipulations of formal representations, similar to the formal deductions of proof theory. If Fodor is right about all this, then it undercuts any attempt at developing a world-oriented theory of information content that would eventually square with a theory of information processing, a project dear to the hearts of some of us. But if he is wrong, then seeing where he has gone wrong should help sort out the differences that sometimes seem to separate those interested in using such an account of information content to ground semantics, on the one hand, from those more concerned with processing, cognitive science, or psychology, on the other.

In this paper I want to argue that Fodor is wrong on two counts. In the first section I argue that he has misunderstood a very basic point about the world-oriented account of information, namely its relativity, and so is unduly pessimistic about the fit between such an account and a processing account. In the second section, I will argue that Fodor's conclusion about the formal nature of thought and inference is false. Indeed, I will argue that a certain formal tradition in logic is misguided. Thus, Section 1 is largely a defense of [2] against Fodor's claims, while Section 2 goes beyond it in various ways. Fodor's paper touches on many other issues, but I will devote this reply to these two.

1 The relativity of information Fodor notices something very important about information processing, but he draws entirely the wrong conclusion from it, at least as far as a theory of *information* goes. To explain, let's begin with a bit of cartography of our own. Fodor divides things up as follows. He thinks that there is a view of information, one held by Dretske [4] and put forward by Perry and myself in [2], a view that treats the information content C_s of a situation s as the basic notion, and sees it as being: objective, hence nonperspectival, receiver neutral, intrinsic, and nonintentional.

By contrast, he claims that the notions of the information D_s displayed by a signal S, the information encoded by S, and the information in S available to an agent, are the basic notions needed in cognitive science. He argues that they are: perspectival, receiver relative, nonintrinsic, and intentional. From this he concludes that information is formal, derivative on representation, and subjective.

By contrast with either of these, Perry and I were arguing that the information C_s contained in a situation s, one which is an instance of a signal S, is: objective, independent of representation by the observer, but *essentially relational.* Hence, the information contained in s that happens to be available to an agent is perspectival, and observer relative. Whether it is intentional or not depends on what one means by "intentionality", a matter I will return to below.

1.1 The relational nature of information There is an old habit of supposing that any objective feature of something must be intrinsic to it. It is the basis

of the argument from the perspectival relativity of perception to the subjectivity of the objects of perception, for example. An old habit, as I say, but a bad one. It comes, at least in part, from a general neglect of relations in philosophy, from supposing that relations between things can be reduced to properties of the things related.[1] Take the case of perception. Perception is a relation between perceiver and perceived. How something looks depends on both parties to the transaction, not just on the thing perceived or on the perceiver. Trying to push the object of perception into the head of the perceiver is trying to find something intrinsic to the perceiver to characterize the perceptual event. Perceptual relativity is not a good argument for there being nothing objective to be perceived. Just the contrary. If perception were not of something external, why should the way things look vary in the systematic ways they do?

Many other examples remind us that there are objective relations between things that are not reducible to intrinsic properties of any of the things related. For instance, it is an objective fact about me that I am in California today, but it is not an intrinsic property of me. It has to do with a relation between me and my surroundings. Again, it is an objective feature of the earth that it is the third planet from the sun, but it is not intrinsic to the earth, since it depends on the relation of the earth to the sun and to the other planets.

But is information relational? Surely so. The basic intuition about the information content C_s of a situation s is that it is information *about* something besides s. The information in the rings of a dead tree stump is information about the tree's past. The information in the newspaper about the Vienna Arms talks (if it is information) is about those talks, not about the physical paper. The information contained in the two distinct weather reports (in Fodor's example) that Johnny heard as he was walking down the street was information about the weather in Chicago and Tulsa, whether he knew it or not. In other words, a very basic point about the notion of information content is that it has something to do with relations between different situations: between s and the situation s' that C_s is about. Just what these relations are is part of what a theory of information content must spell out.

The account of the information content C_s of a situation s given by Dretske and that given by Perry and me differ on many points, but they do agree on the relational nature of information. Dretske attempts to characterize information in terms of conditional probabilities (the conditional probability $Pr(s'|s)$ that a situation s' obtains, given that s does, should equal 1), and conditional probabilities are essentially relational. Perry and I build our theory on informational relations R between types of situations: one type of situation involving another type. We call these relations *constraints*. We do not try to reduce these informational relations to probabilities or anything else, and think it is a mistake to do so. There are lots of different sorts of such informational relations, natural laws, necessary relations like those in mathematics, linguistic conventions, to mention just three, many of which cannot be reduced to conditional probabilities. More important though, for the current discussion, is the fact that they cannot be reduced to intrinsic properties of either the situation s containing the information, or the situation s' that C_s is about.

It follows from this account that the information content C_s of a situation s is relative to some informational constraint (or constraints) R. For example,

a tree stump may contain information about the age of a tree relative to one constraint R_1, information about the weather the day before, relative to another R_2, and information about the local ant population, relative to some third constraint R_3. We might indicate the dependence of C_s on this parameter by making it explicit: $C_{s,R}$. Since the information the tree stump contains is relative to constraints, it follows that just what information will be available to an agent who sees the stump will obviously depend on which of these constraints the agent knows.

To recapitulate, informational constraints are essentially relational. They are not reducible to intrinsic properties of the situation containing the information, or of the situation the information is about, or of the agent getting the information. Neither are they reducible to any notion of correlation. The information content $C_{s,R}$ of a given situation s is relative to constraints R, and the information available to an agent will depend on what constraints the agent knows. Consequently, information has just those features of perspectival and receiver relativity that Fodor notices. But that does not make information nonobjective. And it certainly does not make it the least bit like proof theory.

So we all agree that the information in a given situation available to an agent will depend in part on what the agent knows. The way Perry and I put it in [2] was that it depends on what constraints the agent is attuned to. This use of "attunement", borrowed from Gibson and his school, seems to make a cognitive scientist's hair stand on end. Fodor says, for example, that "the notion of attunement is either blatantly behavioristic (and therefore hopeless) or implicitly intentionalistic (and therefore useless)". Given the choice between a hopeless notion and a useless one, one is tempted by the useless, since being useless is itself relative to what you want to use the notion for. And given my understanding of the terms "behavioristic" and "intentionalistic" I think our account is the latter and not the former. So maybe it is useless for the job Fodor has in mind. But I do not think so. Rather, I think Fodor is conflating two quite different things under the confusing banner of intentionality.

Is our account behavioristic? It is in that it sees folk psychology as a theory of the mental that attempts to explain and predict human activity. It is action centered. But it is not behavioristic in the classical sense, the one that has proven hopeless, since it does not attempt to reduce or define the mental in terms of any kind of observable behavior. But for that matter, Perry and I are not inclined to try to reduce or define anything to anything else. So the issue between Fodor and us is not a debate between representationalism and behaviorism, but between a version of representationalism that posits an intrinsically meaningful formal language of thought, and one that doesn't.

Is our account intentionalistic? Well, a behaviorist like Quine would certainly say it is, since it uses properties and relations external to minds to classify both linguistic events and minds. Fodor often says that "intentionality" is just fancy talk for "aboutness". If that is all he means, then our account is intentional to the core, since it takes it as basic that information is information *about*. But all too often talk of intentionality is talk of a mixed bag, where properties, concepts, and meanings get identified, with aboutness lost in the shuffle. If by being intentionalistic Fodor means to suggest that the theory is implicitly relying on some notion of meaning in the head, of the formal language of thought

variety, then he is wrong. I think he must think this, since why else would he think that the notion of attunement is useless in shedding light on mental activity.

By the way, one reason Perry and I talk of attunement to constraints, rather than knowing constraints, is, in fact, to stress that the theory is *not* committed to a view of knowledge that involves people standing in relations to some sort of sentence-like objects in some language of thought. That would be useless. The theory is committed to the existence of mental states, and to their being representational, in that they represent the way the world is, but this view does not commit one to anything like Fodor's language of thought. If Gibson, who had the basic insights about attunement, also thought that talk of the mental could be defined in terms of overt behavior, as Fodor claims, that is not proof that attunement is dependent on this brand of behaviorism.

1.2 The circumstantial nature of information While the relativity of information shows how Fodor has misunderstood the account of information Perry and I were trying to develop in a crucial way, it does not get at some of the other aspects of Fodor's examples, especially the one with Johnny and the radio. Getting this part straight, or straighter at least, is crucial for understanding the attack on formality in the next section. Toward this end, let us factor an *informational situation s* into two parts, the *representation S* and its *embedding circumstances c*. In the case of a dead tree stump, the representation is a pattern of rings, the circumstances are the other features of the situation that are necessary to determine $C_{s,R}$. In the case of a declarative statement, the representation is the sentence S used, the circumstances are the nonlinguistic features of the event s that contribute to its content $C_{s,R}$ relative to the constraints R that make up the conventions of the language used. In the case of a mental situation, the representation is a mental state[2] S of the agent, the circumstances are the features of the mental situation that make the representation have the content $C_{s,R}$ it does, relative to whatever constraints R on agents one is considering. In all cases we write $C_R(S, c)$ for the content, thus displaying three parameters on which information content depends. When the constraints R are held fixed, we will write $C(S, c)$.

The distinction between an informational situation s and its representation S is absolutely central to the view that Perry and I were putting forth in [2]. Representations are what we called "efficient". That is, they have the property of being usable in different situations, getting at quite different contents in these different circumstances. This is in stark contrast to a view of representations as having an intrinsic content, independent of circumstances. Mathematically, it is a question of whether the value of $C(S, c)$ actually depends on the parameter c or not.

In [2], the term "interpretation" was used for what I am here calling the information content of a situation. The term "meaning" was reserved for the property of the representation S whereby it has whatever content it has in various circumstances. That is, we took the meaning $[\![S]\!]$ of S to be the relation between circumstances c and content $C(S, c)$. For Perry and me, the hallmark of meaningful states and sentences is that their content depends crucially on their embedding circumstances.

Fodor's Johnny-example is a case in point. There are two utterances s_1, s_2 involving a single representation, the sentence *It is raining here*, with two vastly different complex embedding circumstances c_1 and c_2, one stretching back to Chicago, the other to Tulsa. The information contained in the two situations is quite different. Of course, Johnny's lack of knowledge of the part of the embedding circumstances makes part of the information contained in the situations inaccessible to him.

Fodor attempts to ground his argument in a notion of "information displayed" by a signal S. Ironically, what Fodor is after is something intrinsic to the signal, and hence something that does not deserve to be called information at all. That is, the information D_s displayed by some representation S is simply the structure of that very representation, nothing more. This is relational only in the sense that the identity relation is a relation, and so it is informational only in a very degenerate sense of information. From "information displayed" Fodor slides into talk of "the display" and from there to a syntactic encoding of the display. How is Fodor led to this view of information as syntax?

Fodor thinks the basic intuition is that

> [information processors] are systems whose behavior in a given situation is determined *by the character of the information that is available to them* in that situation. [9] (The stress is Fodor's.)

The terms "behavior" and "information" are ambiguous as used here. Both have their ordinary worldly reading, as well as solipsistic readings. Suppose we have some mental situation s of an agent B, factored into a state (type) S and circumstances c. The ambiguity in Fodor's use of "information" resides in the question of whether it is the information $C(S, c)$ in s that is at stake, or the information in S alone, whatever that might be. There is a matching ambiguity in the use of "behavior". There is behavior in the normal, everyday sense of the word, where one behaves in certain ways relative to one's family, friends, and environment more generally. And then there is behavior as compatible with methodological solipsism, which I suppose just means mental and bodily movements, disconnected from the environment.

If we are really talking about information, in a way that distinguishes it from misinformation, then it is hard to see what sense to make of the solipsistic reading of the quote. Information is crucially information about the world. And I think that the credibility that the quote seems to have comes from the worldly reading. So let us examine that reading more closely.

There is an obvious sense in which the content $C_R(S, c)$ is available to the agent B, since this is the information that B's cognitive *situation* contains, relative to whatever informational relations R we are using. And it does seem that the reason we think of animals and people as information processors has something to do with the role such information content plays in determining their behavior under ideal conditions. But under the less than ideal conditions in which we live, the behavior of information processors is often not determined by the information available to them. For example, we often act on misinformation. Surely if I falsely believe that I am in California, then that is *misinformation*, not information, but it just as surely affects my behavior as if it were information. This is not just a quibble. As Dretske has stressed, a theory of

information has got to take the difference between information and misinformation seriously and get at the difference, even if information processors can't always tell which they have on their minds. So, on the worldly reading, Fodor is getting at something important, but he gets it wrong, in just the way that Gibsonians get it wrong when they claim, as they sometimes do, that animals never make mistakes.

There is another way, beside error, in which we often fail to act on information that we "encode", to use Fodor's term. Indeed, we frequently encode information that does not control our (worldly) behavior in an appropriate way at all. Suppose, for example, that a reliable German speaker informs you "Jerry Fodor wird jedem, der ihn darum bittet, $1000 geben". Then surely you have encoded the information that Jerry Fodor will give $1000 to anyone who asks for it. You could use an extra $1000. Why don't you ask him for it? Because you don't understand German.

This may seem like an odd example, but it is not unusual in the lives of our fellow information processors—computers. They have all kinds of information encoded in them, but often this information does not affect their behavior at all. Suppose the CSLI computer has a list of all computers at Stanford to be terminated, including itself and that it could, if it wanted, change this list. Why doesn't it? Not because it does not encode the information, in Fodor's sense, but because it is not attuned to the relation between the information and action that makes changing the list in its own best interest. Thus the worldly reading of Fodor's intuition gets at something important about information processors, something that goes some way to explaining the plausibility of Fodor's claim, but it misfires in various ways.

It is often thought (in [5] for example) that the information a computer encodes *never* affects its worldly behavior. This is not true, as shown in [17]. Here is a simple example borrowed from [16]. Consider Turing, the CSLI computer on which this paper was written. It is part of a large network of computers, with a mail service linking the users of these computers. If Turing gets a message earmarked for some other computer, it forwards that message on, but if it gets one earmarked for itself, then it incorporates it into itself, and alerts the user for whom the message was intended. That is, in this case its actions are appropriately related to the information it encodes in the message's address.

2 Inference: situated or formal? What is really behind Fodor's belief in an intrinsically meaningful formal language of thought? He is looking for a solution to a deep problem: How do people think? In particular, how do they do things like deny hypotheses? Or make suppositions or assumptions? Or come to conditional conclusions? Or recognize the kinds of inferential patterns that Fodor discusses in his attack on associative networks? It does not look like tree rings or even frogs can do any of this. How is a worldly account of information going to shed any light on this sort of activity? What sort of mechanism will explain it?

This is a tough problem, and one I don't know the answer to. It might be that *that* kind of mental activity does require language; it does seem to be correlated with the ability to use languages (a point made by David Israel). But even if it does, it does not follow that it requires an innate mentalese, or that this kind

of reasoning bears any similarity to the formal proof theoretic deductions that have captured Fodor's imagination. Nor do I see how positing a language of thought is supposed to explain it, rather than obscure it, though I think I see why Fodor thinks it has a ghost of a chance. What has captured Fodor's imagination is that we logicians have developed formal proof procedures for certain formal languages, procedures that can be used to build inference engines, machines that can carry out formal proofs, even if not very well. Here, Fodor thinks, is hope for a mechanism underlying thinking. Posit an internal, formal language of thought. Call this the "LTH".

2.1 A formality condition I want to suggest that there are good reasons for believing that mental activity in general, and rational inference in particular, are not formal. However, there is a serious question as to just how to cash out the notion of "formal language" being used in Fodor's various formulations of the LTH.[3] I will not try to define it, since Fodor hasn't, but will only assume that such languages at least satisfy the following "Formality Condition":

> In a *formal* language whether or not an inference "From $S_1, \ldots, S_n$ infer S" is legitimate can be determined from the representations $S_1, \ldots, S_n$ and S *without recourse to features of their embedding circumstances.*

By contrast, we have the following partial characterization of what some of us have come to call Situated Language:

> In a *situated* language, whether or not an inference "From $S_1, \ldots, S_n$ infer S" is valid, can depend on the relation between the contents $C(S_1, c_1)$, $\ldots, C(S_n, c_n)$ and $C(S, c)$, and hence can *depend on the embedding circumstances* $c_1, \ldots, c_n$ and c in which the representations occur.

It should be obvious that human languages used for communication are situated, not formal, at least if the representations in question are anything like the surface structures of the language actually used for communication. By contrast, the formal languages of traditional logic are formal, not situated. Indeed, in logic "formal" has always been associated with the doctrine of logical form, which assumes that the validity of an inference can be determined from *form* of the representation. But even the logical languages where semantical considerations are used to define what it means for an inference to be valid satisfy the above formality condition.[4] Indeed, the whole formal tradition in logic has assumed that inference could be determined solely from the "form" of the representations.

This formality condition probably does not capture the essence of, let alone all of, what Fodor wants of a formal language, but it seems clear from Fodor's discussions of formality that it is at least a consequence of what he assumes. But then the situatedness condition does not say all I would want to say about what it is to constitute a situated language, either. However, these two necessary conditions are enough to argue the point I want to make, that ordinary inference is situated, not formal.

Certainly if what we said in the last section is right, then there is a *prima facie* case for inference being situated. And I think that if we could remove our-

selves far enough from the doctrine of logical form, it would be clear which hypothesis best accounts for the observed facts.

For many years certain philosophers and logicians have criticized the standard model-theoretic account of inference as being unfaithful to everyday reasoning, and have attempted to modify it in various ways. This work was inspired by many notorious mismatches between the "laws of logic" and everyday valid reasoning. The mismatches fall into two sorts: (A) inferences that would seem to take the form of a valid logical argument, but are patently unsound, and (B) those that are clearly valid but do not have the form of a valid argument on traditional logical grounds. With work in artificial intelligence, these problems have assumed a new importance, since attempts to use logic to capture anything vaguely approaching rationality are foundering on just such problems.

If everyday human inference takes place in some internal formal language, then these problems with traditional logic, and all attempts at refining it, are hard to understand. However, if inference is situated, then these mismatches are just what you would expect. For a formal language, inference must either ignore circumstance, or hold it fixed, so whenever there is a shift in circumstance, problems of type (A) are likely to occur. For example, if the second hypothesis of an argument shifts the circumstances needed to interpret the first premise, problems of type (A) are bound to occur.[5] More interestingly, if the validity of some inference depends on a shift in circumstances, then problems of type (B) will necessarily occur. Example: "It is raining here", a voice on the radio says; "That guy is not here", Johnny infers, looking at the clear sky. Both of these require the agent to exploit his circumstances in some significant way. If rationality depends on an agent's ability to exploit his circumstances in systematic and significant ways, then the LTH is just wrong; inference is not formal but situated.[6]

Above and beyond the considerations of Section 1.2, I am not going to rehearse yet again all the other reasons that have led many of us to the conclusion that rationality does indeed require situated inference.[7] Rather, I want to take issue with the most commonly cited argument in favor of the LTH, namely, that "it is the only game in town". I will do this by outlining some of the possible mechanisms for embodying situated inference in physical agents, people and computers. Before doing so, however, I need to survey some of the various ways that circumstances can play a role in the determination of content.

2.2 The structure of circumstances Circumstances can help determine the information content C_s of a situation s (consisting of a representation S in circumstances c) in countless ways, each of which can give rise to problems of both sorts (A) and (B). We can partition these contributions into four importantly different kinds by making a couple of informal definitions. First, various features of the representation S are relevant in that they help "articulate" the content. In the case where S is a sentence, these will include things like the words used, but also things like intonation and stress. Call the features of S that help articulate the content the structural features of S. Secondly, we need a notion of "constituent" of the content $C(S, c)$. By this we mean simply any-

thing that the information is about, or is part of what the information is about.[8] Since, as we saw in Section 1.1, the content $C_{s,R}$ of an informational situation s is relative to what informational constraints R one is considering, the notion of what is a constituent of the content is also relative to this same parameter. It follows that the properties of un/articulated non/constituent are also relative to this parameter.

When we speak of the circumstances c of an informational situation s, we have in mind just those circumstances that help determine the content C_s, of course. Our informal definitions allow us to carve these circumstances up into four (possibly overlapping) parts[9]: (i) the part that contributes *articulated constituents* of the content, (ii) the part that contributes *unarticulated constituents*, (iii) the part that constitutes *articulated unconstituents*, and (iv) the part that constitutes *nonarticulated nonconstituents.*[10] Lets look at some natural language examples to get a feel for this four-way division.

Articulated constituents and context: The first part of the circumstances, the part determining the articulated constituents, is frequently called the *context*, in the case of a sentence of natural language. It determines things like the interpretation of "I", "here", "now", "today", and deictic uses of tense, demonstratives, and pronouns.[11] The remainder of c is sometimes called the *background* of the situation.

Unarticulated constituents: In this category we find things like the location where it is claimed that it is raining in an utterance "It is raining", where there is no constituent that contributes the location in the way it does with "It is raining here".[12] If we translate a statement involving pronouns from English into a language like Japanese, which does not contain pronouns, many objects become unarticulated constituents, determined by the background.

Articulated nonconstituents: In this category we find a variety of things typically neglected in semantics. For example, there are the resource situations needed to determine the value-loaded use of a definite description ([2], pp. 146–153) or the situation described by a nonpersistent statement ([2], pp. 159–161, 319). One might argue that this is where one would find the fact that the referent of a use of "John is missing" is named John, or that the person referred to with a use of "she" is female, or the present time in a deictic use of the past progressive.[13]

Unarticulated nonconstituents: Finally, in the fourth category we find a host of standard background assumptions that play a role in the determination of information content. This is what was called "background conditions" in the treatment of conditionals presented in [1]. In particular, it was argued that these are crucial to understanding the logic of conditionals, and the way it seems to differ from traditional formal logic. Consider the faulty inference from "If Claire rubs her eyes, then she is sleepy" to "If Claire rubs her eyes and there is oak pollen present, then she is sleepy". The second conjunct "there is oak pollen present" explicitly changes the background conditions in force in the first statement. That is, the lack of oak pollen is unarticulated and it is not a constituent of the content, but it is crucial to the content being what it is, and to its being information. Another sort of example would be facts which determine whether an utterance of "These eggs need salt" is a statement of fact, or a request to pass the salt.

2.3 Mechanisms for situated inference With this division, we can suggest some possible mechanisms that might allow us to embody situated inference in situated agents, things like people and robots.

1 Exploiting environmental constants Information is almost always, perhaps always, dependent on unarticulated background conditions. That is, the relations that allow one situation s to contain information about another situation s' are conditional on certain background conditions C being met by the environment. In terms of Perry's and my theory, informational relations are usually *conditional* constraints, constraints that only apply under certain conditions, but there is often nothing there in the head that corresponds to these conditions.

For example, a situation with a book released in mid-air contains information about what will come next, a certain falling, as long as we are near earth. That is, as long as we are actually near the earth, we know that if a book is dropped, then it will fall. And so, as long as we are near earth, we can infer a falling from a dropping. Note that we don't need to know anything at all about gravity for this to work, we just have to stay near the earth where gravity works. Take us out of our environment, into space, and this changes, of course, but as long as we are near the earth, it is perfectly sound. So one strategy for building a rational agent is to exploit constancies in its unarticulated background circumstances by allowing it to make inferences which are sound within those circumstances, even if they cannot be seen as formally valid.[14]

2 Circumstantial rules of inference There is a clear sense in which the suggestion made above could be built into a formal language, since it in fact depends on unarticulated nonconstituents that are held constant. The main reason for making it explicit is as a step along the way to the later suggestions. Suppose we design the hardwired inference rules so that their action is sensitive to variable unarticulated background elements. For example, consider an airborne computer which gets time signals from home base, say San Francisco, and automatically adjusts them to the correct time zone, depending on how far the plane has flown in what direction. There is no reason time zones have to be explicitly articulated by anything in the computer, but they are crucial to seeing why the legitimate inferences vary in the way they do. Or consider color constancy, the way we are able to "infer" the color of something from the way it looks in various lighting circumstances. There need be nothing in the representation of the object in question which articulates these lighting circumstances, but they will still play a key role in determining the content of the representation, that is, the color of the object under view as we adjust for them automatically. More generally, what we have in mind are cases where whether or not a given "rule" applies depends not on the form of the representation, or anything else explicitly represented, but on something else, a parameter in the rule itself whose value gets set by the environment. This was the main mechanism suggested in [2] (e.g., pp. 268–272) for situated inference.

3 Making implicit parameters explicit Natural languages have an important feature that suggests another strategy for embodying situated inference, namely the ability to leave parameters implicit in some circumstances, yet make them explicit in others. Compare "It is raining" with "It is raining here" or "It

is raining there", or compare "I am flying" with "I am flying to Canberra, from Sydney, on Sunday, July 1". Natural language allows us great flexibility in making reference to whatever parameters are relevant for the task at hand, but leaving them implicit when they are irrelevant.

It seems clear that our own mental representations have the same feature. Consider the property of being hungry. A small child has little sense of time or planning for her future needs. Hungry is hungry now. But as we mature, we learn to shift back and forth between hunger as a property we either have or don't have, and hunger as a relation having time as a parameter. This ability lets us do things like prepare our meals before we are actually hungry.

In terms of our characterization of circumstances, what is required is that the overall *system* of representations has highly flexible and systematic mechanisms for turning unarticulated constituents into articulated constituents when the need arises. As an implicit parameter is made explicit, the new representational elements will have their values automatically set by the circumstances in force at that time. When it has served its role in inference it can then be "demoted" to an unarticulated constituent. This demotion is important for having computationally tractable simpler representational elements, but it is also important for another reason. It also is necessary if inference is to lead to action. At seven this morning I knew I would be hungry at noon, but that did not cause me to eat at seven, nor even at noon. Rather, it caused me to make a sack lunch. However, now that it is noon, I eat because I am hungry. An explicit element of my thought to represent the time is quite irrelevant.

4 From self-relativity to self-reference An important special case of the move from implicit to explicit (and back) is where an agent moves from a representation S of something to a representation S_I where the agent itself, or some part, aspect, or property of itself, moves from unarticulated to articulated constituent. The moves from being hungry to *I* am hungry, or from x is left of y to x is left of y with respect to *my* orientation, are two simple cases. At a representational level, what we need is the introduction of an extra argument place whose value is set, in a causally connected way, by the relevant aspect of the agent in question. Here the move back and forth between self-relativity and self-reference and its relation to (worldly) behavior becomes very important, as shown in [12].

Three grades of self-referential involvement are surveyed in [16]. The one just mentioned is the simplest, what Smith calls *autonomy*. More sophisticated forms of self-reference are what Smith calls *introspective* and *reflective*. Introspective systems are "... systems with casually connected self-referential mechanisms that render explicit... some of their otherwise implicit *internal* structure" [emphasis added].

The emphasis here is on internal, or mental properties. In order to count as being introspective, the system must have the capability of promoting some such features from implicit to explicit. So, for example, in order to realize that I don't know something, I must somehow survey my own mental terrain and note that it does not answer the question at hand. I refer the reader to [16] for a full discussion of introspective systems, and for the explanation of the fuller leap to reflective systems. Smith not only discusses the importance of these three

mechanisms, but points out examples of computer systems in place that already embody such mechanisms, especially the first two. He argues that the full potential is far from being tapped, however.

2.4 Conclusion It seems to me that Fodor is not really interested in information, or even in information processing. What he really wants to understand are the mechanisms behind thinking. Information processing is interesting to cognitive science to the extent to which it can shed light on thinking. Fodor thinks it can, and I agree. But Fodor assumes that all information processing is formal, and there he is wrong, at least if formal means anything like what logicians have taken it to mean. In fact, there is a view of information processing that is perfectly compatible with a world-oriented view of information, where processing is situated, not formal. Indeed, many current computational systems can be seen as embodying various situated mechanisms.

I have sketched four of these mechanisms for situated inference not because I think they will solve any or all of the problems involved in understanding thinking and rationality in action, or because I think they are new ideas. Indeed, as I have tried to indicate, I think all these mechanisms are used in current computer systems, and such systems are far from rational. Rather, I have listed them because they are simple information processing mechanisms that violate the formality condition. I do not expect these considerations to change Fodor's mind about the nature of thought, but I do hope they will lead to a version of the LTH which recognizes the importance of embedding circumstances on thought and inference. I also hope that logicians will begin developing tools for the study of situated languages. If successful, I suppose the editors of this journal will eventually change its name to *Notre Dame Journal of Situated Logic.*

NOTES

1. For an excellent discussion of the role that the neglect of relations has played in philosophy, see [11].
2. John Perry has suggested calling it a mental county to emphasize its partial nature.
3. For a fuller account of Fodor's views on formality and the LTH, see [7] and [8]. For a much fuller critique of the role of formality in the LTH and in computation than given here, see [17].
4. These languages either ignore embedding circumstances altogether, or treat them as held constant. For in these languages, the content $C(S)$ of a sentence is represented by the class of all models $M(S)$ of S, and an inference of the above form is valid if $M(S_1) \cap \ldots \cap M(S_n) \subseteq M(S)$. Such languages frequently have no complete formal proof theory.
5. In [1] I have argued that many "counter-examples" to the hypothetical syllogism have this feature.
6. Claiming that inference is not formal should not be confused with related claims. For example, it does not follow that inference cannot be studied by using the tools of mathematics that logicians have used. Nor does it follow that there could not be artificial agents like computers, or better, robots, who could embody situated inference. In fact, there already are such things, I would claim, taking issue with Dretske [5], for reasons spelled out in [17].

7. See [1], [6], [10], and [12]–[18].

8. In the discussion that follows I have assumed that we are dealing with content and aboutness in a fairly simple-minded way. In particular, I do not take into account the additional distinction between what an information situation "concerns" and what it is about, a distinction that is important in [13].

9. This division arose in a conversation with John Etchemendy where we were trying to understand the relation between what I was calling the background of an utterance, and the way that term was being used by Perry. It seems that I was using it for (iv) while Perry was using it for (ii).

10. There is, of course, additional relativity about where a given element will fall, since what counts as articulated will depend on one's account of the structural elements of a representation. Some grammatical theories go so far as to posit a "deep logical structure" where all constituents get articulated, by elements that get "deleted" in the surface structure. However, much of the motivation for this analysis simply presupposes that semantic analysis requires an intrinsically contentful representation, which is just what we are questioning.

11. These are basically all of the circumstances that were treated in [2], as part of the discourse situation. It now seems to me that various problems we had were caused by our not distinguishing context from background. For example, the problem about "where" people have the names they have.

12. For more on this important difference, see [13].

13. This marks a departure from [2], where the lack of these distinctions forced us to have the present time as an articulated constituent of the interpretation of such an utterance.

14. This approach to robot design is being explored by Stan Rosenschein and his group at SRI International. See, e.g., [14].

REFERENCES

[1] Barwise, Jon, "The situation in logic II: conditionals and conditional information," *Report CSLI-84-21.* Center for the Study of Language and Information, Stanford University. Also to appear in: Traugott, et al. (eds.), *On Conditionals*, Cambridge University Press.

[2] Barwise, Jon and John Perry, *Situations and Attitudes*, Bradford Books/MIT Press, 1983.

[3] Barwise, Jon and John Perry, "Shifting situations and shaken attitudes," *Report CSLI-84-13.* Center for the Study of Language and Information, Stanford University. Also in *Linguistics and Philosophy*, vol. 8 (1985), pp. 105–161.

[4] Dretske, Fred, *Knowledge and the Flow of Information*, Bradford Books/MIT Press, 1981.

[5] Dretske, Fred, "Machines and the mental," *Proceedings of the American Philosophical Association*, vol. 59 (1985), pp. 23–33.

[6] Etchemendy, John, "The doctrine of logic as form," *Linguistics and Philosophy*, vol. 6 (1983), pp. 319–334.

[7] Fodor, Jerry, *The Language of Thought*, Harvester Press, 1975.

[8] Fodor, Jerry, *Representations*, Bradford Books/MIT Press, 1982.

[9] Fodor, Jerry, "Information and association," *Notre Dame Journal of Formal Logic*, vol. 27 (1986), pp. 307–323.

[10] Kaplan, David, "On the logic of demonstratives," *Journal of Philosophical Logic*, vol. 8 (1978), pp. 81–98.

[11] Olson, Kenneth, *Facts*, Stanford University Ph.D. dissertation, 1985.

[12] Perry, John, "Self-knowledge and self-representation," *Proceedings of IJCAI-1985* (1985), pp. 238–242.

[13] Perry, John, "Thought without representation," *Supplementary Proceedings of the Aristotelian Society*, vol. 60 (1986), to appear.

[14] Rosenschein, Stanley J., "Formal theories of knowledge in AI and robotics," *New Generation Computing*, vol. 3, no. 4 (1985). Ohmsha Ltd.

[15] Rosenschein, Stanley J. and Leslie P. Kaelbling, "The synthesis of machines with provable epistemic properties," *Proceedings of Conference on Theoretical Aspects of Knowledge*, Morgan Kaufmann Publishers.

[16] Smith, Brian Cantwell, "Varieties of self-reference," *Proceedings of Conference on Theoretical Aspects of Knowledge*, Morgan Kaufmann Publishers.

[17] Smith, Brian Cantwell, "Is computation formal?," *Report CSLI-45-85*, Center for the Study of Language and Information, Stanford University.

[18] Suchman, Lucy, "Plans and situated action: The problem of human-machine communication," ISL-6. Xerox-PARC.

Center for the Study of Language and Information, and
Department of Philosophy
Stanford University
Stanford, CA 94305