Searle on Brains as Computers

William J. Rapaport

Department of Computer Science and Engineering,
Department of Philosophy, Department of Linguistics
and Center for Cognitive Science
State University of New York at Buffalo, Buffalo, NY 14260-2000
rapaport@cse.buffalo.edu
http://www.cse.buffalo.edu/∼rapaport/

January 28, 2007

Abstract

Is the brain a digital computer? Searle says that this is meaningless; I say that it is an empirical question. Is the mind a computer program? Searle says no; I say: properly understood, yes. Can the operations of the brain be simulated on a digital computer? Searle says: trivially yes; I say yes, but that it is not trivial.

1 Three Questions

In his 1990 essay, "Is the Brain a Digital Computer?", Searle factors the "slogan … 'the mind is to the brain as the program is to the hardware' " (p. 21) into three questions:

1. Is the brain a digital computer?
2. Is the mind a computer program?
3. Can the operations of the brain be simulated on a digital computer? (Searle 1990: 21.)

Let us consider each of these, beginning with the second.

2 Is the Mind a Computer Program?

What does it mean to say that the mind is a computer program? Surely not that there is a programming language and a program written in it that is being executed on a brain—not for humans, at least. But it could mean that by bottom-up, reverse engineering (neuroscience) together with top-down, cognitive-scientific investigation, we could write a program that would cause a computer to exhibit mental behavior. However, that's question 3, to which Searle gives a different answer.

Possibly question 2 means that the mind plays the same role with respect to the brain that a program does to a computer; call this "Good Old-Fashioned Cartesian Dualism". This may not be much progress over the "slogan" of which question 2 is supposed to be merely a part, but it is worth a small digression.

2.1 Good Old-Fashioned Cartesian Dualism

Computational cognitive science, including what John Haugeland (1985: 112) has termed "good old-fashioned artificial intelligence", is, I believe, good old-fashioned Cartesian dualism. The view that mental states and processes are (or are expressible as) algorithms that are *implemented in* the physical states and processes of physical devices is (a form of) Cartesian dualism: The mental states and processes and the physical states and processes can be thought of as different "substances" that "interact". How might this be?

It should be clear that an algorithm and a computer are different kinds of "substance". If one considers an algorithm as a mathematical abstraction (in the ordinary sense of the term 'abstraction'), then it is an abstract mathematical entity (like numbers, sets, etc.). Alternatively, if one considers an algorithm as a text expressed in some language, then it is, say, ink marks on paper or ASCII characters in a word-processor's file. An algorithm might even be—and indeed ultimately is—"switch settings" (or their electronic counterparts) in a computer. All of these are very different sorts of things from a very physical computer.

How do an algorithm and a computer "interact"? By the latter being a semantic interpretation—a model—of the former. More precisely, the *processes* of the brain/body/computer are semantic interpretations (or models) of the mind/algorithm in the sense of semantics as correspondence. But this is just what we call an implementation. So, an implementation is a kind of semantic interpretation. (For further discussion of this, see Rapaport 1999, 2005.)

Note, by the way, that the mind/algorithm can itself be viewed as a semantic interpretation of the brain/body/computer, since the correspondence can go both ways (Rapaport 2002). How is a mind implemented? Consider a computer program: Ultimately, the program (as text) is implemented as states of a computer (expressed in binary states of certain of its components). That is purely physical, but it is *also* purely syntactic; hence, *it* can have a semantic interpretation. An abstract data type, for instance, can be thought of as the semantic interpretation of an arrangement of bits in the computer (cf. Tenenbaum & Augenstein 1981: 1, 6, 45; see Rapaport 1995, §2.3, and Rapaport 2005). This Janus-faced aspect of the bit arrangements—thought of both as a physical model or implementation of the abstract algorithm and as a syntactic domain interpretable by the algorithm and its data structures—is Marx W. Wartofsky's "model muddle" (Wartofsky 1966, 1979: xiii–xxvi; Rapaport 1995; 1996, Ch. 2; 1999; 2000).

Now, is this really good old-fashioned Cartesian dualism? Is mind-body interaction really semantic interpretation or implementation? Or might this semantic/implementational view be more like some other theory of the mind?

It is not parallelism, since there really is a *causal* interaction: The algorithm (better: the process) *causes* the physical device to behave in certain ways.

So it's not epiphenomenalism, either. Moreover, the device—or its behavior—can produce changes in the program, as in the case of self-modifying programs, or even in the case of a system competent in natural language whose knowledge base (part of the software) changes with each interaction.

Could it be a dual-aspect theory? Perhaps: Certainly, the physical states and processes are one "level of description", and the mental states and processes are another "level of description" *of the same (physical) system*. But talk of levels of description seems to me to be less illuminating than the theory of semantics as correspondence. More to the point, neither "level" is a complete description of the system: The algorithm is not the process, nor can one infer from the algorithm what the future behavior of the process will be: The process can behave in ways not predictable by the programmer (cf. Fetzer 1988, 1991). And even a complete physical description of the system would not tell us *what* it is doing; this is one of the lessons of functionalism.

So dualism is at least plausible. Do the physical states and processes produce mental ones? Here is where the problem of qualia—i.e., subjective qualitative experiences, including pain and physical sensations—enters. (I say more about this in Rapaport 2005, §2.2.)

2.2 Return to Searle

Does question 2 mean that the mind is the *way* the brain behaves? That seems right, but isn't the right analogy: It *doesn't* seem right to say that a program is the way a computer behaves.

"Programs," Searle goes on to say, "are defined purely formally or syntactically" (p. 21). That, I think, is not *quite* right: They require a set of input-output conventions, which would be "links" to the world. In any case, this together with the assertion that "minds have an intrinsic content . . . immediately [implies] that the program by itself cannot constitute the mind" (p. 21). What does 'content' mean?

If it means something internal to the mind (a "container" metaphor; cf. Twardowski 1894, Rapaport 1978), then that minds have intrinsic content could mean that within a mind there are links among mental concepts, some of which play the role of a language of thought and others of which play the role of mental representations of external perceptions (cf. Rapaport 1996, Ch. 3; 2000; 2002; 2006). If so, that would be purely syntactic, as Searle says programs are.

If, on the other hand, 'content' means a relation to an external entity, then why don't programs have that, too (as I noted two paragraphs back)? In any case, programs do take input from the external world: I enter '2' on the keyboard, which results (after a few transductions) in a switch being set in the computer, which the program interprets as the number 2.

So, on either interpretation, the conclusion doesn't follow, since programs can *also* have "intrinsic mental content", whatever that means.

The problem is that question 2 is not the right question. Of course "The formal syntax of the program does not by itself guarantee the presence of mental contents" (p. 26), because the program might never be executed. What Searle should have asked is whether the mind is a computer *process*. And here the answer can be 'yes', since *processes* can have contents. Searle says:

> I showed this [viz., that the formal syntax of a program doesn't guarantee the presence of mental contents] a decade ago in the Chinese Room Argument .... The argument rests on the simple logical truth that syntax is not the same as, nor is it by itself sufficient for, semantics. (Searle 1990: 21.)

This seems to follow from Charles Morris's definitions (1938) of syntax as the study of relations *among* symbols and of semantics as the study of relations *between* symbols and their meanings; thus, syntax $\neq$ semantics. Nor is it the case that semantics can be "derived", "constructed", or "produced" from syntax by Morris's definitions. But the first-person semantic enterprise *is* one of determining correspondences among symbols—between linguistic symbols and internal representations of external objects. Hence, it *is* syntactic even on Morris's definition. The *third*-person semantic enterprise is more like what Morris had in mind. But one person's third-person semantics is another's first-person semantics: If one cognitive agent, Oscar, tries to account for the semantics of another cognitive agent (Cassie) by drawing correspondences between her mental concepts and things in the world, all he can really do is draw correspondences between his representations of her concepts and his representations of things in the world. As with the turtles who hold up the Earth, it's syntax all the way down. (For more about how Cassie and Oscar understand each other, see, e.g., Rapaport 2003.)

3 Can the Operations of the Brain Be Simulated on a Digital Computer?

Let's turn to question 3, the answer to which Searle thinks is trivially—or, at least, uninterestingly—affirmative. "[N]aturally interpreted, the question means: Is there some description of the brain such that under that description you could do a computational simulation of the operations of the brain" (p. 21). Such a description would inevitably be partial (Smith 1985). Hence, so would be the computational simulation. But if it passed the Turing test (i.e., if its effects in the actual world were indistinguishable from those of a human), then what's not in the model is an implementation detail. What might these be? They might include sensations of pain, warm fuzzy feelings associated with categorizing something as "beautiful", etc. (cf. Rapaport 2005). As for pain, don't forget that our *sensation* of it is an internal perception, just like our sensation of an odor (cf. Crane 1998). It might be possible to be in pain and to know that one is in pain without what we normally call a pain sensation, just as it is possible to determine the presence of an object by its odor—by a chemical analysis—without sensing that odor.[1] The "triviality" or "obviousness" of the answer to question 3 stems, according to Searle, from Church's Thesis: "The operations of the brain can be simulated on a digital computer in the same sense in which weather systems, the behavior of the New York Stock market or the pattern of airline flights over Latin America can" (p. 21). And, presumably, since simulated weather isn't weather, simulated brains aren't brains. But the premise is arguable (Rapaport 2005, §3); at least, it does not follow that the behavior of simulated brains isn't *mental*. Brains and brain behavior are special cases.

4 Is the Brain a Digital Computer?

Searle equates question 1 to "Are brain processes computational?" (p. 22). What would it mean to say that the brain was *not* a digital computer? It might mean that the brain is *more* than a digital computer—that only some proper *part* of it *is* a digital computer. What would the rest of it be? Implementation details, perhaps. I am, however, willing to

[1] Angier 1992 reports that "Sperm cells possess the same sort of odor receptors that allow the nose to smell." This does not mean, of course, that sperm cells have the mental capacity to have smell-qualia. Blakeslee 1993 reports that "humans ... may exude ... odorless chemicals called pheromones that send meaningful signals to other humans." She calls this "a cryptic sensory system that exists without conscious awareness ...." And Fountain 2006 discusses a plant that has what might be called a sense of smell, presumably without any smell qualia.

admit that perhaps not all of the brain's processes are computational. Following Philip N. Johnson-Laird (1988: 26–27), I take the task of cognitive science to be to find out *how much* of the brain's processes *are* computational—and surely some of them are (Rapaport 1998). It is, thus, a working hypothesis that brain processes are computational, requiring an empirical answer and not subject to apriori refutation.

On the other hand, to say that the brain is not a digital computer might mean that it's a different kind of entity altogether—that no part of it is a digital computer. But that seems wrong, since it *can* execute programs (we use our brains to hand-simulate computer programs, which, incidentally, is the inverse of Turing's 1936 analysis of computation).

What are brain processes, how do they differ from mental processes, and how do both of these relate to computer processes? A computer process is a program being executed; therefore, it is a physical thing that implements an abstract program. A brain process is also a physical thing, so it might correspond to a computer process. A mental process could be either (i) something abstract yet dynamic or (ii) a brain process. The former (i) makes no sense if programs and minds are viewed as static entities. The latter (ii) would mean that *some* brain processes are mental (others, like raising one's arm, are not). So to ask if brain processes are computational is like asking if a computer process is computational. That question means: Is the current behavior of the computer describable by a recursive function (or is it just a fuse blowing)? So Searle's question 1 is: Is the current (mental) behavior of the brain describable by a recursive function? This is the fundamental question of artificial intelligence as computational philosophy. It is a major research program, not a logical puzzle capable of apriori resolution.

Searle's categorization of the possible positions into "strong AI" ("all there is to having a mind is having a program"), "weak AI" ("brain processes (and mental processes) can be simulated computationally"), and "Cognitivism" ("the brain is a digital computer") is too coarse (p. 22). What about the claim that a computer running the "final" AI program (the one that passes the Turing test, let's say) has mentality? As I argued above, that's not necessarily "just" having a program. But on the *process* interpretation of question 2, Strong AI could be the view that all there is to having a mind is having a *process*, and that's more than having a program. What about the claim that the "final" AI program need not be the one that humans use—i.e., the claim that computational *philosophy* might "succeed", not computational *psychology*? (Computational philosophy seeks to learn which aspects of cognition in general are computable; computational psychology studies *human* cognition using computational techniques. Cf. Shapiro 1992, Rapaport 2003.) This is a distinction that Searle does not seem to make. Finally, Pylyshyn's version of "cognitivism" (1985) does not claim that the brain *is* a digital computer, but that mental processes are computational processes. That seems to me to be compatible with the brain being "more" than a digital computer.

Searle complains that multiple realizability is "disastrous" (p. 26; cf. Rapaport 2005, §4.1). The first reason is that *anything* can be described in terms of 0s and 1s (p. 26), and there might be *lots* of 0-1 encodings of the brain. But the real question, it seems to me, is this: Does the brain *compute* (effectively) some function? What is the input-output description of that function? The answer to the latter question is whatever psychology tells us is intelligence, cognition, etc. For special cases, it's easier to be a bit more specific: For natural-language understanding, the input is some utterance of natural language, and the output is an "appropriate" response (where the measure of "appropriateness" is defined, let's say, sociolinguistically). For vision, the input is some physical object, and the output is, again, some "appropriate" response (say, an utterance identifying the object or some scene, or some behavior to pick up or avoid the object, etc.). Moreover, these two modules (natural-language understanding and vision) must be able to "communicate" with each other. (They might or might not be modular in Fodor's sense (1983), or cognitively impenetrable in Pylyshyn's sense (1985). In any case, solving one of these problems will require a solution to the other; they are "AI-complete" (Shapiro 1992).)

The second allegedly disastrous consequence of multiple realizability is that "syntax is not intrinsic to physics. The ascription of syntactical properties is always relative to an agent or observer who treats certain physical phenomena as syntactical" (p. 26). The observer assigns 0s and 1s to the physical phenomena. But Morris's definition of syntax as relations among symbols (uninterpreted marks) can be extended to relations among components of any system. Surely, physical objects stand in those relationships "intrinsically". And if 0s and 1s *can* be ascribed to a physical object (by an observer), that *fact* exists independently of the agent who *discovers* it.

Searle's claim "that syntax is essentially an observer relative notion" (p. 27) is very odd. One would have expected him to say that about *semantics*, not syntax. Insofar as one can look at a complex system and describe (or discover) relations among its parts (independently of any claims about what it does at any higher level), one is doing *non*–observer-relative syntax. Searle says that "this move is no help. A physical state of a system is a computational state only relative to the assignment to that state of some computational role, function, or interpretation" (p. 27), where, presumably, the assignment is made by an observer. But an assignment is an assignment of meaning; it's an interpretation. So is Searle saying that computation is fundamentally a *semantic* notion? But, for Church, Turing, et al., computation is purely *syntactic*. It's only the input-output coding that *might* constitute an assignment. But such coding is only needed in order to be able to link the syntax with the standard theory of computation in terms of functions from natural numbers to natural numbers. If we're willing to express the theory of computation in terms of functions from physical states to physical states (and why shouldn't we?), then it's not relative.

Searle rejects question 1: "There is no way you could discover that something is intrinsically a digital computer because the characterization of it as a digital computer is always relative to an observer who assigns a *syntactical* interpretation to the purely physical features of the system" (p. 28, my italics). I, too, reject question 1, but for a very different reason: I think the question is really whether *mental processes* are computational. In any event, suppose we *do* find computer programs that exhibit intelligent input-output behavior, i.e., that pass the Turing Test. Computational *philosophy* makes no claim about whether that tells us that the human *brain* is a digital computer. It only tells us that intelligence is a computable function. So at best Searle's arguments are against computational *psychology*. But even that need not imply that the brain *is* a digital computer, only that it behaves as if it were. To discover that something X is intrinsically a digital computer, or a Y, is to have an abstraction Y, and to find correspondences between X and Y.

Perhaps what Searle is saying is that being computational is not a *natural* kind, but an artifactual kind (cf. Churchland & Sejnowski 1992):

> I am not saying there are *a priori* limits on the patterns we could discover in nature. We could no doubt discover a pattern of events in my brain that was isomorphic to the implementation of the vi program on this computer. (Searle 1990: 28.)

This is to admit what I observed two paragraphs back. Searle continues:

> But to say that something is *functioning as* a computational process is to say something more than that a pattern of physical events is occurring. It requires the assignment of a computational interpretation by some agent. (Searle 1990: 28.)

But why? Possibly because to find correspondences between two things (say, a brain and the Abstraction ComputationalProcess—better, the Abstraction Computer) is observer-relative? But if we have *already* established that a certain brain process is an implementation of `vi`, what *extra* "assignment of a computational interpretation by some agent" is needed?

Searle persists:

> Analogously, we might discover in nature objects which had the same sort of shape as chairs and which could therefore be used as chairs; but we could not discover objects in nature which were functioning as chairs, except relative to some agent who regarded them or used them as chairs. (Searle 1990: 2.)

The analogy is clearly with *artifacts*. But the notion of a computational process does not seem to me to be artifactual; it is *mathematical*. So the proper analogy would be something like this: Can we discover in nature objects that were, say, sets, or numbers, or Abelian groups? Here, the answer is, I think, (a qualified) 'yes'. (It is qualified, because sets and numbers are abstract and infinite, while the world is concrete and finite. Groups may be a clearer case.) In any event, is Searle claiming that the implementation of `vi` in my brain isn't `vi` until someone *uses it as* `vi`? If there is an implementation of `vi` on my Macintosh that no one ever uses, it's still `vi`.

Searle accuses computational cognitive scientists of "commit[ing] the homunculus fallacy ... treat[ing] the brain as if there were some agent inside it using it to compute with" (p. 28). But consider Patrick Hayes's (1990, 1997) objection to the Chinese-Room Argument: Computation is a series of switch-settings; it isn't rule-following. (On this view, by the way, the solar system *does* compute certain mathematical functions.) Turing machines do *not* follow rules; they simply change state. There are, however, descriptions—programs—of the state changes, and anything that follows (executes) that program computes the same function computed by the Turing machine. A *universal* Turing machine can also follow that program. But the original, special-purpose Turing machine's program is "hardwired" (an analogy, of course, since everything is abstract here). A universal Turing machine has *its* program similarly hardwired. It is only when the universal Turing machine is fed a program that it follows the rules of that program. But that's what *we* do when *we* consciously follow (hand-simulate) the rules of a program. So it's *Searle* who commits the homuncular fallacy in the Chinese-Room Argument by putting a person in the room. It is not the person in the room who either does or does not understand Chinese; it is the entire system (Rapaport 2000, 2006). Similarly, it is not some *part* of my brain that understands language; it is *I* who understands.

In his discussion of "discharging" the homunculus, Searle says that "All of the higher levels reduce to this bottom level. Only the bottom level really exists; the top levels are all just *as-if*" (p. 29). But *all* levels exist, and *all* levels "do the same thing", albeit in different ways (Rapaport 1990, 2005).

I noted above that systems that don't follow rules can still be said to be computing. My example was the solar system. Searle offers "nails [that] compute the distance they are to travel in the board from the impact of the hammer and the density of the wood" (p. 29) and the human visual system; "neither," according to him, "compute anything" (p. 29). But in fact they both do. (The nail example might not be ideal, but it's a nice example of an *analog* computation.)

> But you do not *understand* hammering by supposing that nails are somehow intrinsically implementing hammering algorithms and you do not *understand* vision by supposing the system is implementing, e.g., the shape from shading algorithm. (Searle 1990: 29; my italics.)

Why not? It gives us a theory about how the system might be performing the task. We can falsify (or test) the theory. What more could *any* (scientific) theory give us? What further kind of understanding could there be? Well, there could be first-person understanding, but I doubt that we could ever know what it is like to be a nail or a solar system. We *do* understand what it is like to be a cognitive agent!

The problem, I think, is that Searle and I are interested in different (but complementary) things:

> ... you cannot explain a physical system such as a typewriter or a brain by identifying a pattern which it shares with its computational simulation, because the existence of the pattern does not explain how the system actually works *as a physical system*. (Searle 1990: 32.)

Of course not. That would be to confuse the implementation with the Abstraction. Searle is interested in the former; he wants to know *how the (human) brain works*. I, however, want to know *what the brain does* and how *anything* could do it. For that, I need an account at the functional/computational level, not a biological (or neuroscientific) theory.

> The mistake is to suppose that in the sense in which computers are used to process information, brains also process information. [Cf. Johnson 1990.] To see that that is a mistake, contrast what goes on in the computer with what goes on in the brain. In the case of the computer, an outside agent encodes some information in a form that can be processed by the circuitry of the computer. That is, he or she provides a syntactical realization of the information that the computer can implement in, for example, different voltage levels. The computer then goes through a series of electrical stages that the outside agent can interpret both syntactically and semantically even though, of course, the hardware has no intrinsic syntax or semantics: It is all in the eye of the beholder. And the physics does not matter provided only that you can get it to implement the algorithm. Finally, an output is produced in the form of physical phenomena which an observer can interpret as symbols with a syntax and a semantics.
>
> But now contrast this with the brain. ... none of the relevant neurobiological processes are observer relative ... and the specificity of the neurophysiology matters desperately. (Searle 1990: 34.)

There is much to disagree with here. First, "an outside agent" need *not* "encode ... information in a form that can be processed by the circuitry of the computer". A computer could be (and typically is) designed to take input directly from the real world and to perform the encoding (better: the transduction) itself, as, e.g., in document-image understanding (cf. Srihari & Rapaport 1989, 1990; Srihari 1991, 1993, 1994). Conversely, abstract concepts are "encoded" in natural language so as to be processable by *human* "circuitry".

Second, although I find the phrase 'syntactical realization' quite congenial (cf. Rapaport 1995), I'm not sure how to parse the rest of the sentence in which it appears. What does the computer "implement in voltage levels": the information? the syntactical realization? I'd say the former, and that the syntactical realization *is* the voltage levels. So there's an issue here of whether the voltage levels are *interpreted as* information, or vice versa.

Third, the output need not be physical phenomena interpreted by an observer as symbols. The output *could* be an action, or more internal data (e.g., as in a vision system),[2] or even natural language to be interpreted by another

[2] Searle seems to think (p. 34) that vision systems yield sentences as output! (See below.)

computer. Indeed, the latter suggests an interesting research project: Set up Cassie and Oscar, two computational cognitive agents implemented in a knowledge-representation, reasoning, and acting system such as SNePS.[3] Let Cassie have a story pre-stored or as the result of "reading" or "conversing". Then let her tell the story to Oscar and ask him questions about it. No *humans* need be involved.

Fourth, neurobiological processes aren't observer-relative, only because we don't care to, or need to, describe them that way. The computer works as it does independently of us, too. Of course, for *us* to understand what the brain is doing—from a third-person point of view—we need a psychological level of description (cf. Chomsky 1968, Fodor 1968, Enç 1982).

Finally, why should "the specificity of the neurophysiology matter desperately"? Does this mean that if the neurophysiology were different, it wouldn't be a human brain? I suppose so, but that's relevant only for the implementation side of the issue, not the Abstraction side, with which I am concerned.

Here is another example of how Searle does not seem to understand what computational cognitive science is about:

> A standard computational model of vision will take in information about the visual array on my retina and eventually print out the sentence, "There is a car coming toward me". But that is not what happens in the actual biology. In the biology a concrete and specific series of electro-chemical reactions are set up by the assault of the photons on the photo receptor cells of my retina, and this entire process eventually results in a concrete visual experience. The biological reality is not that of a bunch of words or symbols being produced by the visual system, rather it is a matter of a concrete specific conscious visual event; this very visual experience. (Searle 1990: 34–35.)

The first sentence is astounding. First, why does he assume that the input to the computational vision system is *information on the retina*, rather than *things in the world*? The former is close to an *internal* symbol representing external information! Second, it is hardly "standard" to have a vision system yield a *sentence* as an output. It might, of course ("Oh, what a pretty red flower."), but, in the case of a car coming at the system, an aversive maneuver would seem to be called for, not a matter-of-fact description. Nonetheless, precisely that input-output interaction *could*, *pace* Searle, be "what happens in the actual biology": I could say that sentence upon appropriate retinal stimulation.

Of course, as the rest of the quotation makes clear, Searle is more concerned with the intervening qualitative experience, which, he seems to think, humans have but computers don't (or can't). Well, could they? Surely, there ought to be an intervening stage in which the retinal image is processed (perhaps stored) before the information thus processed or stored is passed to the natural-language module and interpreted and generated. Does that process have a qualitative feel? Who knows? *How* would you know? Indeed, how do I know (or believe) that *you* have such a qualitative feel? The question is the same for both human and computer. Stuart C. Shapiro has suggested how a pain-feeling computer could be built (Rapaport 2005, §2.3.1); similarly, it's possible that a physical theory of sensation could be constructed. Would it be computational? Perhaps not—but so what? Perhaps some "mental" phenomena are not *really* mental (or computational) after all (Rapaport 2005, §2.3). Or perhaps a computational theory will always be such that there is a role to play for some sensation or other, even though the actual sensation in the event is not computational. That is, every computational theory of pain or vision or what have you will be such that it will refer to a sensation without specifying what the sensation is. (Cf. Gracia's (1990) example of a non-written universal for a written text, discussed in Rapaport 2005, §2.2. See also McDermott 2001.)

Of course, despite my comments about the linguistic output of a vision system, the sentence that Searle talks about could be a "sentence" of one's language of thought. That, however, would fall under the category of being a "concrete specific conscious visual event" and "not . . . a bunch of words or symbols" (Cf. Pylyshyn 1981; Srihari op. cit.; Srihari & Rapaport op. cit.)

Searle's final point about question 1 is this:

> The point is not that the claim "The brain is a digital computer" is false. Rather it does not get up to the level of falsehood. It does not have a clear sense. (Searle 1990: 35.)

This is because "you could not *discover* that the brain *or anything else* was intrinsically a digital computer" (p. 35, my italics). "Or anything else"? Even an IBM PC? Surely not. Possibly he means something like this: Suppose we find an alien physical object and theorize that it is a digital computer. Have we *discovered* that it is? No—we've got an *interpretation* of it *as* a digital computer (cf. "you could assign a computational interpretation to it as you could to

[3] Shapiro 1979, 2000; Shapiro & Rapaport 1987, 1992, 1995; Shapiro et al. 2006. Further information is available online at: [http://www.cse.buffalo.edu/sneps] and at: [http://www.cse.buffalo.edu/∼rapaport/snepskrra.html].

anything else" (p. 35)). But how else *could* we "discover" anything about it? Surely, we could discover that it's made of silicon and has 10^k parts. But that's consistent with his views about *artifacts*. Could we *discover* the topological arrangement of its parts? I'd say 'yes'. Can we *discover* the sequential arrangement of its behaviors? Again, I'd say 'yes'. Now consider this: How do we determine that it's made of silicon? By subjecting it to certain physical or chemical tests and having a theory that says that any substance that behaves thus and so is (made of) silicon. But if anything that *behaves* such and thus is a computer, then so is this machine! So we *can* discover that (or whether) it is a computer. (Better: We can discover whether its processing is computational.)

References

Angier, Natalie (1992), "Odor Receptors Discovered in Sperm Cells", *The New York Times* (30 January 1992): A19.

Blakeslee, Sandra (1993), "Human Nose May Hold An Additional Organ For a Real Sixth Sense", *The New York Times* (7 September 1993): C3.

Chomsky, Noam (1968), *Language and Mind* (New York: Harcourt, Brace & World).

Churchland, Patricia S., & Sejnowski, Terrence J. (1992), *The Computational Brain* (Cambridge, MA: MIT Press).

Crane, Tim (1998), "Intentionality as the Mark of the Mental", in Anthony O'Hear (ed.), *Current Issues in the Philosophy of Mind* (Cambridge, UK: Cambridge University Press): 229–251.

Enç, Berent (1982), "Intentional States of Mechanical Devices", *Mind* 91(362; April): 161–182.

Fetzer, James H. (1988), "Program Verification: The Very Idea", *Communications of the ACM* 31: 1048–1063.

Fetzer, James H. (1991), "Philosophical Aspects of Program Verification", *Minds and Machines* 1: 197–216.

Fodor, Jerry A. (1968), *Psychological Explanation: An Introduction to the Philosophy of Psychology* (New York: Random House).

Fodor, Jerry A. (1983), *The Modularity of Mind: An Essay in Faculty Psychology* (Cambridge, MA: MIT Press).

Fountain, Henry (2006), "This Plant Has the Sense of Smell (Loves Tomatoes, Hates Wheat)", *The New York Times* (3 October 2006): F1.

Gracia, Jorge J.E. (1990), "Texts and Their Interpretation", *Review of Metaphysics* 43: 495–542.

Haugeland, John (1985), *Artificial Intelligence: The Very Idea* (Cambridge, MA: MIT Press).

Hayes, Patrick J. (1990), "Searle's Chinese Room", abstract of presentation to the Society for Philosophy and Psychology Symposium on Searle's Chinese Room and Workshop on Symbol Grounding, University of Maryland (electronic mail posting from Stevan Harnad, 6 June 1990).

Hayes, Patrick J. (1997), "What Is a Computer? An Electronic Discussion", *The Monist* 80(3).

Johnson-Laird, Philip N. (1988), *The Computer and the Mind: An Introduction to Cognitive Science* (Cambridge, MA: Harvard University Press).

McDermott, Drew (2001), *Mind and Mechanism* (Cambridge, MA: MIT Press).

Morris, Charles (1938), *Foundations of the Theory of Signs* (Chicago: University of Chicago Press).

Pylyshyn, Zenon (1981), "The Imagery Debate: Analog Media versus Tacit Knowledge", in Ned Block (ed.), *Imagery* (Cambridge, MA: MIT Press): 151–206.

Pylyshyn, Zenon (1985), *Computation and Cognition: Toward a Foundation for Cognitive Science, 2nd edition* (Cambridge, MA: MIT Press).

Rapaport, William J. (1978), "Meinongian Theories and a Russellian Paradox", *Noûs* 12: 153–180; errata, *Noûs* 13 (1979) 125.

Rapaport, William J. (1990), "Computer Processes and Virtual Persons: Comments on Cole's 'Artificial Intelligence and Personal Identity'," *Technical Report 90-13* (Buffalo: SUNY Buffalo Department of Computer Science, May 1990) [http://www.cse.buffalo.edu/∼rapaport/Papers/cole.tr.17my90.pdf].

Rapaport, William J. (1995), "Understanding Understanding: Syntactic Semantics and Computational Cognition", in James E. Tomberlin (ed.), *Philosophical Perspectives, Vol. 9: AI, Connectionism, and Philosophical Psychology* (Atascadero, CA: Ridgeview): 49–88; reprinted in Toribio, Josefa, & Clark, Andy (eds.) (1998), *Language and Meaning in Cognitive Science: Cognitive Issues and Semantic Theory*, *Artificial Intelligence and Cognitive Science: Conceptual Issues*, Vol. 4 (New York: Garland): 73–88.

Rapaport, William J. (1996), "Understanding Understanding: Semantics, Computation, and Cognition", *Technical Report 96-26* (Buffalo: SUNY Buffalo Department of Computer Science).

Rapaport, William J. (1998), "How Minds Can Be Computational Systems", *Journal of Experimental and Theoretical Artificial Intelligence* 10: 403–419.

Rapaport, William J. (1999), "Implementation is Semantic Interpretation", *The Monist* 82: 109–130.

Rapaport, William J. (2000), "How to Pass a Turing Test: Syntactic Semantics, Natural-Language Understanding, and First-Person Cognition", *Journal of Logic, Language, and Information* 9(4): 467–490; reprinted in James H. Moor (ed.), *The Turing Test: The Elusive Standard of Artificial Intelligence* (Dordrecht: Kluwer, 2003): 161–184.

Rapaport, William J. (2002), "Holism, Conceptual-Role Semantics, and Syntactic Semantics", *Minds and Machines* 12(1): 3–59.

Rapaport, William J. (2003), "What Did You Mean by That? Misunderstanding, Negotiation, and Syntactic Semantics", *Minds and*

Machines 13(3): 397–427.

Rapaport, William J. (2005), "Implementation Is Semantic Interpretation: Further Thoughts", *Journal of Experimental and Theoretical Artificial Intelligence* 17(4; December): 385–417.

Rapaport, William J. (2006), "How Helen Keller Used Syntactic Semantics to Escape from a Chinese Room" [http://www.cse.buffalo.edu/∼rapaport/Papers/helenkeller.pdf].

Searle, John R. (1990), "Is the Brain a Digital Computer?", *Proceedings and Addresses of the American Philosophical Association*, 64(3): 21–37.

Shapiro, Stuart C. (1979), "The SNePS Semantic Network Processing System", in Nicholas Findler (ed.), *Associative Networks: Representation and Use of Knowledge by Computers* (New York: Academic Press): 179–203.

Shapiro, Stuart C. (1992), "Artificial Intelligence", in Stuart C. Shapiro (ed.), *Encyclopedia of Artificial Intelligence, second edition* (New York: John Wiley & Sons): 54–57.

Shapiro, Stuart C. (2000), "SNePS: A Logic for Natural Language Understanding and Commonsense Reasoning", in Łucja M. Iwańska & Stuart C. Shapiro (eds.), *Natural Language Processing and Knowledge Representation: Language for Knowledge and Knowledge for Language* (Menlo Park, CA/Cambridge, MA: AAAI Press/MIT Press): 175–195.

Shapiro, Stuart C., & Rapaport, William J. (1987), "SNePS Considered as a Fully Intensional Propositional Semantic Network", in Nick Cercone & Gordon McCalla (eds.), *The Knowledge Frontier: Essays in the Representation of Knowledge* (New York: Springer-Verlag): 262–315; shorter version appeared in *Proceedings of the 5th National Conference on Artificial Intelligence (AAAI-86, Philadelphia)* (Los Altos, CA: Morgan Kaufmann): 278–283; a revised shorter version appears as "A Fully Intensional Propositional Semantic Network", in Leslie Burkholder (ed.), *Philosophy and the Computer* (Boulder, CO: Westview Press, 1992): 75–91.

tem Shapiro, Stuart C., & Rapaport, William J. (1992), "The SNePS Family", *Computers and Mathematics with Applications* 23: 243–275; reprinted in F. Lehmann (ed.), *Semantic Networks in Artificial Intelligence* (Oxford: Pergamon Press, 1992): 243–275.

Shapiro, Stuart C., & Rapaport, William J. (1995), "An Introduction to a Computational Reader of Narratives", in Judith F. Duchan, Gail A. Bruder, & Lynne E. Hewitt (eds.), *Deixis in Narrative: A Cognitive Science Perspective* (Hillsdale, NJ: Lawrence Erlbaum Associates): 79–105.

Shapiro, Stuart C., & the SNePS Research Group (2006), "SNePS" [http://en.wikipedia.org/wiki/SNePS].

Smith, Brian Cantwell (1985), "Limits of Correctness in Computers", reprinted in Charles Dunlop & Rob Kling (eds.), *Computerization and Controversy* (San Diego: Academic Press, 1991): 632–646.

Srihari, Rohini K. (1991), "PICTION: A System that Uses Captions to Label Human Faces in Newspaper Photographs", *Proceedings of the 9th National Conference on Artificial Intelligence (AAAI-91, Anaheim)* (Menlo Park, CA: AAAI Press/MIT Press): 80–85.

Srihari, Rohini K. (1993), "Intelligent Document Understanding: Understanding Photos with Captions", *Proceedings of the International Conference on Document Analysis and Recognition (ICDAR-93, Tsukuba City, Japan)*.

Srihari, Rohini K. (1994), "Use of Collateral Text in Understanding Photos in Documents", *Proceedings of the Conference on Applied Imagery and Pattern Recognition (AIPR/SPIE, Washington, DC)*.

Srihari, Rohini K., & Rapaport, William J. (1989), "Extracting Visual Information From Text: Using Captions to Label Human Faces in Newspaper Photographs", *Proceedings of the 11th Annual Conference of the Cognitive Science Society (Ann Arbor, MI)* (Hillsdale, NJ: Lawrence Erlbaum Associates): 364–371.

Srihari, Rohini K., & Rapaport, William J. (1990), "Combining Linguistic and Pictorial Information: Using Captions to Interpret Newspaper Photographs", in Deepak Kumar (ed.), *Current Trends in SNePS—Semantic Network Processing System*, (Berlin: Springer-Verlag Lecture Notes in Artificial Intelligence 437): 85–96.

Tenenbaum, Aaron M., & Augenstein, Moshe J. (1981), *Data Structures using Pascal* (Englewood Cliffs, NJ: Prentice-Hall).

Turing, Alan M. (1936), "On Computable Numbers, with an Application to the *Entscheidungsproblem*", *Proceedings of the London Mathematical Society*, Ser. 2, Vol. 42: 230–265; reprinted, with corrections, in Martin Davis (ed.), *The Undecidable: Basic Papers On Undecidable Propositions, Unsolvable Problems And Computable Functions* (New York: Raven Press, 1965): 116–154.

Twardowski, Kasimir (1894), *On the Content and Object of Presentations*, Reinhardt Grossmann (trans.) (The Hague: Nijhoff, 1977).

Wartofsky, Marx W. (1966), "The Model Muddle: Proposals for an Immodest Realism", in *Models: Representation and the Scientific Understanding* (Dordrecht, Holland: D. Reidel, 1979): 1–11.

Wartofsky, Marx W. (1979), "Introduction", in *Models: Representation and the Scientific Understanding* (Dordrecht, Holland: D. Reidel, 1979): xiii–xxvi.