Intentionality, Artificial Intelligence and the Causal Powers of the Brain

Jeffrey M. Whitmer
Northern Illinois University

It seems to be a common belief that in the future, if not in the present, digital computers are going to be capable of cognitive states, experiences, and consciousness equal in every respect to that which exists in human beings.[1] Not everyone, however, is so optimistic. One such skeptic is John Searle and his "Minds, Brains, and Programs"[2] represents a direct confrontation between the skeptic and the proponents of machine intelligence.

In MBP, Searle presents and then attempts to refute the thesis underlying the research of workers in strong AI (Artificial Intelligence). He then presents what can be called his own "positive" view concerning the problem of achieving cognitive states and what sorts of entities can achieve them. The goals of this discussion are to: 1) briefly consider Searle's view on what cognitive states are not (with a focus on understanding), i.e., the refutation of the strong AI thesis, 2) present in as much detail as possible Searle's positive view on cognitive states which turns on the notion of "causal powers of the brain," 3) examine what, if any, relevant differences exist between the position of strong AI and Searle's positive view. Once these three goals are met, I hope to conclude that strong AI has been conclusively refuted, that Searle's positive view is both explicable and plausible, and therefore, that (in light of the refutation of strong AI) it is more reasonable to accept Searle's view than to suspend judgment.

I

THE REFUTATION OF STRONG AI

In his book review "The Myth of the Computer"[3] Searle presents the following argument as a summary of his position against strong AI.

1. Brain processes cause mental phenomena.

2. Mental states are caused by and realized in the structure of the brain.

Auslegung Vol. X, No. 3
ISSN: 0733-4311

So, 3. Any system that produces mental states must have (causal) powers equivalent to those of the human brain.

4. Digital computer programs by themselves are never sufficient to produce mental states.

So, 5. The way the brain produces a mind cannot be by simply instantiating a computer program.

So, 6. If you want to build a machine to produce mental states, then it cannot be designed to do so solely in virtue of its instantiating a certain computer program.

I think that this argument is both valid and sound. Premises one and two seem incontrovertibly true. Searle's arguments in MBP are intended to establish the truth of step four in the above argument and I think they succeed. In order to more fully understand Searle's positive view, however, I will briefly address the salient points of the refutation of strong AI.

The focal point of the debate in question is whether computers can ever achieve cognitive states <u>strictly in virtue of their programs</u>. That is, can a computer achieve any intentional state merely by appropriate programming? The strong AI researcher's claim is yes, and Searle's claim is no. An intentional state of mind is a state that can be described with sentences beginning with "I believe that," "I understand," "I desire that," etc. They are states of mind that are described by Searle as representing objects and states of affairs.[4] In this discussion, as mentioned earlier, the intentional state being considered is that of understanding. The question to be answered is, in virtue of what is the brain the focus of intentionality? Searle and his opponents agree that the brain is the part of the human anatomy that is the source or focus of intentional states, but they disagree about what characteristic of the brain it is that allows it to fulfill this function.

The claim of the researchers in strong AI is that the brain is the focus of intentionality in virtue of the programs it realizes. Therefore, anything that can instantiate a program can achieve intentional states. Consequently, computers, in virtue of their ability to instantiate programs, can achieve cognitive states. It should be noted that for the purposes of this discussion, a computer is defined as any thing or collection of things that is stable enough and complex enough to accurately instantiate a program. This could be an anthill, some toilet paper and stones, an IBM 360, or a collection of beer cans. All of these things (and many others as well) are, or could be made, complex enough to instantiate a variety of computer programs since all that is needed is a structure capable of maintaining certain relationships for an extended period of time. The thesis of strong AI, then, is that: 1) appropriately programmed computers literally have cognitive states, and 2) the programs thereby explain human cognitive states. To use the language of dualism, the strong AI researcher wants to claim that mind is to brain as program is to hardware, i.e., in both relations the former is independent of the latter although the latter is needed to instantiate the former.

Searle refers to one of the examples used in strong AI to clarify this thesis. The example is from Schank and Abelson's Scripts Plans Goals and Understanding.[5] Schank and Abelson develop several programs that fit into the following scenario.

1. The computer is given representative knowledge (a script) intended to be equivalent to what a normal human being would know about the situation in question, e.g., eating in a restaurant.

2. The computer is then given a story about a particular situation, e.g., John ordering a hamburger.

3. The computer is then "asked" questions about the story. These questions commonly refer to things not explicitly stated in the story but which could be derived from the story by anyone (anything) with basic knowledge about restaurants.

4. The computer then answers these questions in a manner we would expect from a human being in similar circumstances.

From this Searle claims (and rightly so) that the supporters of strong AI draw the following conclusions: 1) that the machine can literally be said to understand the story and provide answers to the questions, and 2) that what the machine and its programs do explains the human ability to understand the story and answer questions about it.

From this we can extract a clear and basic thesis indicative of strong AI for the intentional state of, in this case, understanding a natural language. Such a thesis is: S understands P in the case where, given P as input, S realizes a program X which enables S to produce responses which are absolutely indistinguishable from that of a native speaker of the language to which P belongs. Therefore, according to strong AI, the intentional state of understanding all or part of a natural language is achieved when the appropriate program is realized. In other words, S's realizing a program X is sufficient for S to uunderstand P. In this case, the computer program X would have to be such that given P, plus other linguistic data, S could appear to any questioner as indistinguishable from a native speaker of P, i.e., pass the Turing test.[6]

To address these claims, Searle notes that "[o]ne way to test any theory of the mind is to ask oneself what it would be like if my mind actually worked on the principles that the theory says all minds work on."[7] To do this, Searle introduces a series of thought-experiments that are intended to operate on the principles of strong AI, and to show that the claims of strong AI are totally unfounded. Consider the following situation.

> Now suppose that I, who understand no Chinese at all and can't even distinquish Chinese symbols from some other kinds of symbols, am locked in a room with a number of cardboard boxes full of Chinese symbols. Suppose that I am given a book of rules in English that instruct me how to match these Chinese symbols with each other. The rules say such things as that the "squiggle-squiggle" sign is to be followed by the "squoggle-squoggle" sign. Suppose that people outside the room pass in more Chinese symbols and that following the instructions in the book I pass Chinese symbols back to them. Suppose that unknown to me the people who pass me the symbols call them "questions," and the book of instructions that I work from they call the "program"; the symbols I give back to them they call "answers to the questions" and me they call "the computer." Suppose that after a while the programmers get so good at writing the programs and I get so good at manipulating the symbols that my answers are indistinguishable from those of native Chinese speakers. I can pass the Turing test for understanding Chinese. But all the same I still don't understand a word of Chinese and neither does any other digital computer because all the computer has is what I have: a formal program that attaches no meaning, interpretation, or content to any of the symbols.[8]

This thought-experiment has given Searle-in-the-Chinese-room (hereafter referred to as Searle's demon after Haugeland in the commentaries on Searle in BBS) everything that Schank and Abelson's computer has, a script, a story, questions, a program, and Searle's demon does the same thing the computer does--gives back answers in the same language. Furthermore, there are no constraints placed on Searle's demon in this thought-experiment. He can be super-fast, super-intelligent, super-small, whatever is necessary, since these physical characteristics should in no way impare or increase his ability to understand. In this way, Searle's demon can respond as fast or as slow as either a human being or a computer of any design. And since Searle's demon can in every way represent the perspective of the computer when operating on the Chinese symbols, Searle draws the following two conclusions about strong AI.

1) The computer plus program as represented by Searle's demon plus the instruction book does not understand anything. His inputs and outputs are identical to those of a native Chinese speaker but it is clear that Searle's demon does not understand a word of Chinese. Therefore, no computer, however programmed, is capable of any understanding of any stories in any language, solely in virtue of its programming.

2) Since the computer does not actually understand anything, it cannot be an explanation of any human cognitive state. Although it may describe a part of what human cognition is like, the computer plus program does not serve to explain anything, since it understands nothing.

The basic point being made by Searle in these conclusions is that the computer plus program cannot understand anything strictly in virtue of its program, because all this amounts to is giving the computer "syntax" but no "semantics." As Searle notes,

> The computer attaches no meaning, interpretation, or content to the formal symbols, and qua computer it couldn't, because if we tried to give the computer an interpretation of its symbols (semantics) we could only give it more uninterpreted symbols. The computer manipulates formal symbols but attaches no meaning to them . . . .[9]

This is the point made by Searle's demon: he manipulates the formal Chinese symbols, but he doesn't understand them because they have, for Searle's demon, no meaning. And for the supporters of strong AI who wish to claim that even if the Chinese symbols are not understood, the symbolism internal to the machine (the machine language, i.e., Searle's demon's ability to understand the English instruction book) is understood, John Heil has a response in "Does Cognitive Psychology Rest on a Mistake?" that appears consistent with Searle's position.

> It appears, for example, that the sense in which we might want to say that the internal 'machine language' of a digital computer is symbolic--the

> sense, that is, in which it could be said to have meaning (semantics)--is parasitic on its relation to a suitable programming language, and the sense of this language, in turn, dependent on its application by a suitable, language using programmer. The programmer provides an essential link between the states of the machine and the states of affairs in the world to which the former 'refer'.[10]

This relationship between program and programmer can also be used to explain why programmers and researchers in strong AI describe their machines as having intentional states. It is due to the fact that it is obvious to the programmers that the machine has all the necessary information to arrive at the correct answers to their questions. They do not pause to consider that the replies of the computer have no meaning as far as the computer is concerned, but are being interpreted as meaningful by the programmers themselves.

This is the distinction that Searle draws between "intrinsic" intentionality and "observer-relative" intentionality.

> . . . we need to distinguish carefully between cases of what I call intrinsic intentionality, which are cases of actual mental states, and what I call observer-relative ascriptions of intentionality, which are ways people have of speaking about entities figuring in our activities but lacking intrinsic intentionality.[11]

The researchers in strong AI interpret their input into the computer (scripts, stories, and questions) as having meaning, and they ascribe intrinsic intentionality to the computer. But the case of Searle's demon illustrates that all that really obtains is observer-relative intentionality. The computer cannot and does not have intentional states (understanding) strictly in virtue of its program. It has nothing but a bunch of uninterpreted formal symbols and instructions as to how to manipulate symbols, i.e., a syntax but no semantics.

Searle goes on to consider a series of possible replies to his critique of strong AI. For the specifics the reader is referred to MBP. In all of the replies addressed, Searle can accomodate the modified situation into the Searle's demon thought-experiment.[12] In all of the replies, Searle's demon could run the whole operation and not understand anything. We would certainly want to ascribe intentional states to such unified entities that some of the replies suggest, but it would be a case of observer-relative intentionality. As before, we would realize upon a close examination that Searle's demon is merely processing uninterpreted symbols and, once again, he has syntax, but still no semantics.

A close examination of Searle's response to the various replies reveals the following conclusions:

1) The strong AI thesis does not offer the sufficient condition for understanding that it claimed to offer. At best, it is a qualified sufficient condition and it may not even be a necessary condition.

2) Because the thesis of the Combination Reply (as well as some of the other replies) is a brain simulation thesis, the philosophically interesting aspect of the intitial thesis, i.e., mind as independent of brain characteristics, has been sacraficed.

3) Therefore, digital computer programs by themselves are never sufficient to produce mental states (premise four of the main argument).

II

THE CAUSAL POWERS OF THE BRAIN

Since Searle does not come out and explicitly present his own view of intentional states, I will offer a version that Searle could accept based on what he does say. If Searle ever really presents his own thesis, it seems to be "that intentional states processes, and events are precisely that: states, processes, and events. The point is that they are both caused by and realized in the structure of the brain."[13] Also, "[m]ental states and processes, e.g., feeling thirsty or having a visual experience, are both caused by and realized in the neurophysiology of the brain."[14] Finally, "I believe that everything we have learned about human and animal biology suggests that what we call 'mental' phenomena are as much a part of our biological natural history as any other biological phenomena . . . ."[15] The reason that Searle adopts this position (unclear as it may be at this point) is that he cannot understand why anyone would accept that "of all the known types of specifically biological processes, one and only one type is (taken to be) completely independent of the biochemistry of its origins, and that one is cognition."[16] This is a form of the strong AI thesis and Searle has already disposed of this as a viable thesis.

Searle's own view turns on the notion of "causal powers of the brain," and on the intrinsic/observer-relative distinction in the ascription of intentionality. Searle notes that although we may not know <u>how</u> the brain causes or accounts for mental phenomena, we do know that its internal operations are causally sufficient for these mental phenomena. "On my view it it just a plain (testable, empirical) fact about the world that it contains certain biological systems, specifically human and certain animal brains, that are capable of causing mental phenomena with intentional or semantic content."[17] Searle also distinguishes between the internal causes of the brain, and the impact of the external world. We can actually see a tree or we can hallucinate the sight of a tree. Although the external effects on the brain are different, i.e., the former involves the external world while the latter may involve a drug or other neural stimulator, the internal mental states encompass precisely the same intentional state. This is what Searle means when he says "the operation of the brain is causally sufficient for intentionality, and that it is the operation of the brain and not the impact of the outside world that matters for the content of our intentional states, in at least one important sense of 'content'."[18] In other words, even an isolated "brain in a vat" could have intentional states involving trees or whatever, even though the assumed external cause of these states would not exist, i.e., the sight of a tree would not be due to a causal chain starting with an actual tree, proceeding through the eye and optic nerve, and ending in the appropriate part of the brain. Instead, some other stimulation of the brain itself would result in the appearance of a "tree." It is the internal states of the brain (its intrinsic intentionality) that are important, not the impingement of the causally related external world that those of us on the "outside" are aware of (observer-relative intentionality).

From these few claims by Searle, and in what will follow, I will try to show that the following argument is consistent with what Searle does claim about intentionality and the "causal powers of the brain." Because Searle does not specifically endorse this argument, I will refer to it hereafter as the "causal argument" and try to show that Searle's views support this argument.

1. Having the same causal powers as the human brain is sufficient for having intentional states.

2. If B (a non-human brain) exercises its causal powers in the very same way the human brain does then it will have the very same intentional states.

3. For B to have the very same causal powers as the human brain is for its component parts to be capable of functioning in the same mechanistic way (or in a way analogous to) the component parts of the human brain.

4. The component parts of the human brain interact causally via the passage of biochemical electrical currents. This is the mechanism which underlies the causal interactions of the human brain.

So, 5. B will have the same intentional states as a human brain only if its component parts interact electrically in a way analogous to the way the parts of the human brain do.[19]

The first premise is supported by the above comments made by Searle. If the brain, in fact, has intentional states based on certain "causal powers," then if something has these same powers, it has the potential for intentional states. The second premise, in turn, is merely a claim as to the exercising of potential. If anything has the potential for intentional states because it has the same powers as the human brain, then if it exercises these powers in the same way as the human brain, it will have intentional states.

The third premise is perhaps the most controversial. Fortunately, Searle does make some mention of what a claim like this could mean. It should be apparent that Searle is not claiming that carbon-based biochemical entities are the only ones capable of intentional states. That is, "any system that produced mental states would have to have powers equivalent to those of the brain. Such a system might use a different chemistry, but whatever its chemistry it would have to be able to cause what the brain causes."[20] Thus, an entity with a copper-based biochemistry or even an entity with a non-biological structure could possess intentional states, as long as it has the same causal powers as the human brain. This leaves open the possibility of computers having intentional states, which Searle admits. He states, "perhaps its (the computer's or the computer's microchips) electrical properties can reproduce some of the actual causal powers of the electro-chemical features of the brain in producing mental states."[21] To clarify this position, consider the following thought-experiment.

Suppose that the technology is available to separate all the individual neurons of a human brain while maintaining their electro-chemical relations with all the other neurons.[22] In this way, all the causal relations and connections of a normal brain are preserved. Now, suppose that we can create silicon microchips that can exactly re-create the input/output relations that are normal for the synapses of a single neuron.[23] Therefore, if a neuron in the extended brain is damaged, it can be replaced with a microchip that is precisely matched to that particular neuron. The chip will still causally interact with the biochemical neurons because the input/output changes it experiences will result in an alteration of its internal electro-magnetic characteristics. Now, imagine further that every neuron in the extended brain is replaced with a microchip, after each microchip has been turned to match the input/output functions of its respective neuron. I think Searle woul accept this situation as possible, as well as being one that retains the intentional states of the original biochemical brain. This acceptance is based on the fact that the actual causal relations are still intact, even though they are now realized in an electro-magnetic structure instead of an electro-chemical structure. That this might be possible is consistent with the causal argument and with Searle's position since he claims that, "[o]n my account it is a testable empirical claim whether in repairing a damaged brain we could duplicate the electrochemical basis of intentionality using some other substance, say silicon."[24]

Searle would not, however, accept the following argument. Given the manner in which certain microchips are constructed, we can always write a formal program that can stimulate the same formal processes that obtain within the microchip. We can simulate a "gate" within the chip as being "open" or "closed" by the use of '0' and '1' respectively, in our program. Furthermore, a sequence in the chip that takes in three electrical "impulses" of strength $\underline{x}$ and gives out two "impulses" of strength $\underline{y}$ could be represented in a program as an equation of the form $3x \Longrightarrow 2y$. Therefore, we could in principle write a program that would formally represent all of the relations that obtain within the chip. But this is the unacceptable move for the following reasons. First, this move takes the actual, physical states of the chips and describes or represents them using some formal symbolism. Second, Ned Block makes somewhat the same point in a different context[25] and argues that what is happening is that such an approach virtually eliminates the importance of the "primitive processors" in the brain and focuses instead on the formal description or re-presentation of the processes generated by these processors. In either case, we end up with a syntax without semantics. As Searle suggests, "if the simulation of the causes (of intentional states) is at a low enough level to reproduce the causes (as with the microchip or "primitive processor") and not merely describe them (as in a program), the 'simulation' will reproduce the effects."[26] Thus, we can see that the notion of functioning in the same mechanistic way as the human brain is liberal enough to admit non-carbon-based neurophysiology and yet narrow enough to exclude all the things (anthills, bunches of beer cans, etc.) which can only instantiate a digital computer program.

The fourth premise is basically an empirical claim to the effect that the mechanism of the human brain is biological and electro-chemical in nature.

The conclusion of this argument follows from the premises as interpreted above. The notion of having to interact electrically in a way analogous to the human brain has been addressed in the discussion of premise three. However, this conclusion might be more conservative than is necessary. That is, it seems conceivable that an entity could exist whose component mental parts interact magnetically or optically, and we would still want to claim that if they had the appropriate causal powers, they would also have intentionality. It seems apparent, however, that if we can admit electromagnetic interaction to replace electro-chemical interaction, then we could also admit magnetic and optical interaction without any serious harm to the argument. We would still be able to exclude digital computers based on magnetics or optics that only instantiate programs because it would still be an empirical question as to whether such a system could, in fact, reproduce causal powers like those in the human brain. In this way, the above interpretaion of the causal argument based on Searle's various claims admits what it seems reasonable to admit to the class of "things with intentional states" while excluding all those things that were shown to be inadmissable by virtue of Searle's demon.

III

CONCLUSIONS

This discussion has been an attempt to expose the pertinent characteristics of the two views being considered. The difference between the two ultimately turns on how the brain can be the focus of intentional states. For the researchers in strong AI, the brain produces intentional states in virtue of its instantiation of formal programs. For Searle, the brain produces intentionality in virtue of its causal powers and properties. As mentioned above, the difference seems to have its source in the relative priority being placed on the "primitive processors" and the processes as represented by programs.

For the researcher in strong AI, the "primitive processor" (the brain's neurophysiology) is of no importance as can be seen in what can be called a "digital computer." Instead, the researchers are examining the mental processes by questioning and observing human beings in action. From what they observe, they construct a formal program that can take the same symbolic input (words) as the brain qua person, and gives back the same output (words) as the brain qua person. Their claim is that if this is done with adequate attention to detail, the program created and instantiated in ány capable mechanism must achieve the same cognitive states as those observed by the researchers in human subjects. Searle claims this is a gross error because the neurophysiology of the brain is not irrelevant to the analysis of intentional states. Furthermore, since the researchers are (so to speak) on the outside looking in, their ascribing of intentionality to the computer instantiating the program is merely a case of observer-relative intentionality. The causal argument has shown that intrinsic intentionality is exposed by an empirical examination to determine whether the entity in question possesses the appropriate causal powers. But for the researcher in strong AI, the formal relations and structures are constructed through observation and theory, and the data (uninterpreted symbols for the computer, interpreted as words by the researchers) is fitted into this structure. The formal structure precedes the data being related. If you will, the syntax is created by the researchers prior to the semantics, the semantics which never arrives. This is what happens, according to the causal argument, when the formal descriptions of human behavior capable of instantiation by digital computer are given priority over the actual, physical primitive processors that make up the human brain.

According to the causal argument, the perspective and the priority are just the opposite. The intrinsic intentionality of the brain is realized in virtue of the causal relations that exist between the primitive processors of the brain. Whatever the precise causal characteristics of the brain are (however the brain acutally works), they are sufficient to produce intentional states. These characteristics cannot exist soley in the formal, observer-described relations represented in the structure of the brain, so it must be something mechanistically inherent to the brain (but not tied to its particular biochemistry) that accounts for the presence of intentional states. In the causal argument, the primitive processors entering into these formal relations exist prior to the relations. Consequently, we must give priority to these primitive processors rather than to the formal relations we recognize (after the fact). In this case, the semantics exists before the "observed" syntax.

In summary, the relevant difference between the position of strong AI and the causal argument is that the former is based on the formal structure of programs that turn out to be empty of anything to relate and the latter is based on the causal relation and characteristics of the brain which at this point we know to be sufficient for intentional states, but which may at this time be indescribable beyond their being mechanistically grounded. For Searle, the formal relations that obtain and are recognized by strong AI (according to our observer-relative analysis) only serve a purpose in virtue of their ability to describe certain formal mental relationships. They cannot, however, tell us anything about the actual, physical, causal properties of the brain which are the most fundamental source of human intentional states. For this, we must do brain physiology.

In conclusion, there can be little doubt that Searle's positive position is both difficult to present and subtle in the distinctions it draws. However, we have seen that there is a relevant difference between the causal argument and the view of strong AI and that the causal argument can be maintained while the thesis of strong AI is rejected. I think that Searle's argument against strong AI is indubitable. Furthermore, his positive view, as represented in the causal argument, is certainly prima facie plausible. And if we take Searle seriously when he claims that, "[i]f you want to build a machine to produce mental states, then it cannot be designed to do so soley in virtue of its instantiating of a certain computer program, but must have (causal) powers equivalent to those of the brain,"[27] then I think it is more reasonable to embrace Searle's view than to suspend judgment. I think we should accept the causal argument analysis of intentionality.

NOTES

[1]Cf. a) Robert Jastrow, The Enchanted Loom: Mind in the Universe, (New York, NY: Simon and Schuster, 1981). b) Hofstadter/Dennett, The Mind's I, (New York, NY: Basic Books, 1981). c) Schank/Abelson, Scripts Plan Goals and Understanding, (New York, NY: John Wiley and Sons, 1977).

[2]John Searle, "Minds, Brains, and Programs," The Brain and Behavioral Sciences (BBS), (1980) 3, 417-57. (hereafter abbrev. MBP)

[3]John Searle, "The Myth of the Computer," The New York Review of Books, April 29, 1982, 3-6 (hereafter abbrev. MC)

[4]Cf. a) John Searle, "The Intentionality of Intention and Action," Inquiry, 22: 253-80. b) John Searle, "What is an Intentional State?," Mind, 88: 74-92.

[5]Schank, R. C., and Abelson, R. P., Scripts Plans Goals and Understanding, (New York, NY: John Wiley and Sons, 1977).

[6]A. M. Turing, "Computing Machinery and Intelligence," Mind, 59, 236 (1950).

[7]Searle, MBP, p. 417.

[8]Searle, MC, p. 5.

[9]Searle, MC, p. 4.

[10]John Heil, "Does Cognitive Psychology Rest on a Mistake?," Mind, 90: p. 331 (1981).

[11]Searle, MBP, p. 451. It should be noted that this use of intrinsic and observer-relative intentionality may very well seem to open a Pandora's box as far as the question of how we can ever really know that some entity has intrinsic intentionality. However, since the causal argument will show that the notion of "causal powers" is an empirical question, this should not concern us here.

[12]Briefly, these replies are as follows. The Systems Reply claims that Searle's demon is merely a part of a larger system, and that the system as a whole does understand Chinese, even if Searle's demon does not. The Robot Reply asks that we change the program and put the computer in control of a robot such that the robot would receive inputs from various sources and send them to the computer. The computer outputs would, in turn, operate the robot in actions of walking, eating, speaking, etc. Such a robot would be capable of genuine understanding. The Brain Simulator Reply asks us to change the approach to the problem. The program to be developed does not use scripts about the world, but instead it simulates the actual sequence of neural firings in the brain of a Chinese speaker when he has Chinese stories as inputs and gives out Chinese answers. At this level, what could be the difference between the program of the computer and the program of the Chinese brain? The Combination Reply merely asks us to consider in one combined situation, the previous three responses.

[13]Searle, MBP, p. 451.

[14]John Searle, "The Myth of the Computer: An Exchange," The New York Review of Books, June 24, 1982, p. 57. (hereafter abbrev. MCAE)

[15]Searle, MC, p. 4.

[16]Searle, MBP, p. 450.

[17]Searle, MCAE, p. 57.

[18]Searle, MBP, p. 452.

[19]I am indebted to Michael Tye for an earlier formulation of this argument.

[20]Searle, MC, p. 6.

[21]Searle, MC, p. 4.

[22]A part of this thought-experiment is based on Arnold Zuboff's "The Story of a Brain," in The Mind's I, (Hofstadter/Dennett ed.), pp. 202-12.

[23]For the current ideas about the possibility of creating such microchips see: Ernest Kent, The Brains of Men and Machines, (New York, NY: McGraw-Hill, 1981). See especially chapters 1-4.

[24]Searle, MBP, p. 453. It should also be noted that this position seems to draw the distinction between an analog computer and a digital computer. If these microchips were indeed constructed, they would be analog devices because they operate electromagnetically in a manner analogous to the electrochemical operations of the neuron. Thus, Searle can admit that an analog computer of sufficient complexity could have intentional states, since this is precisely what the extended brain described above has become.

[25]Ned Block, "Occasional Paper #22: Mental Pictures and Cognitive Science," Center for Cognitive Science, Massachusetts Institute of Technology, Cambridge, Mass. (1982). In this paper, Block is concerned with the cognitive science interpretation of the pictorial and descriptive analyses of mental imagery. What is relevant here is that Block concludes that even cognitive scientists of the sort who would embrace strong AI must accept that there are primitive processors which cannot be described by representations, but must be explained nomologically (see p. 26). Consequently, the question becomes one of which is more crucial, the representational descriptions or the primitive processors. Block claims that we must place much more emphasis on the primitive processors as analog devices (pp. 39-41); and that is the same point Searle is trying to make, even more fervently, when he says we cannot ignore the neurophysiology of the brain in favor of formal programs alone.

[26]Searle, MBP, p. 453.

[27]Searle, MC, p. 6.

BIBLIOGRAPHY

Arbib, Michael A. The Metaphorical Brain. (New York: John Wiley and Sons, 1972).

Block, Ned. "Occasional Paper #22: Mental Pictures and Cognitive Science." (Cambridge: Center for Cognitive Science, Massachusetts Institute of Technology, 1982).

Boden, Margaret. Artificial Intelligence and Natural Man. (New York: Basic Books, 1977).

Heil, John. "Does Cognitive Psychology Rest on a Mistake?" Mind, 90: 321-42.

Hofstadter, D. R. and Dennett, D. C. The Mind's I. (New York: Basic Books, 1981).

Hunt, Earl B. Artificial Intelligence. (New York: Academic Press, 1975).

Jastrow, Robert. The Enchanted Loom: Mind in the Universe. (New York: Simon and Schuster, 1981).

Kent, Ernest W. The Brains of Men and Machines. (New York: McGraw-Hill, 1981).

Ringle, Martin. (ed) Philosophical Perspectives in Artificial Intelligence. (New York: Humanities Press, 1979).

Searle, John R. "The Intentionality of Intention and Action." Inquiry, 22: 253-80.

______________. "What is an Intentional State?" Mind, 88: 74-92.

______________. "Minds, Brains, and Programs." The Brain and Behavioral Sciences, 3: 417-57.

______________. "The Myth of the Computer." The New York Review of Books, April 29, 1982: 3-6.

______________. "The Myth of the Computer: An Exchange." The New York Review of Books, June 24, 1982, p. 56-57.

Schank, R. C. and Abelson, R. P. Scripts Plans Goals and Understanding. (New York: John Wiley and Sons, 1977).

Sloman, Aaron. The Computer Revolution in Philosophy. (Sussex: The Harvester Press, 1978).

Young, J. Z. Programs of the Brain. (Oxford: The Oxford University Press, 1978).