HOW TO PASS A TURING TEST:
Syntax Suffices for Understanding Natural Language

William J. Rapaport

Department of Computer Science and Engineering,
Department of Philosophy, 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/

June 8, 1999

Abstract

A theory of "syntactic semantics" is advocated as a way of understanding how computers can think (and how the Chinese-Room-Argument objection to the Turing Test can be overcome): (1) Semantics, as the study of relations between symbols and meanings, can be turned into syntax—a study of relations among symbols (including meanings)—and hence syntax can suffice for the semantical enterprise. (2) Semantics, as the process of understanding one domain modeled in terms of another, can be viewed recursively: The base case of semantic understanding—understanding a domain in terms of itself—is syntactic understanding. An internal (or "narrow"), first-person point of view makes an external (or "wide"), third-person point of view otiose for purposes of understanding cognition. The paper also sketches the ramifications of this view with respect to methodological solipsism, conceptual-role semantics, holism, misunderstanding, and implementation, and looks at Helen Keller as inhabitant of a Chinese Room.

This document is *Technical Report 99-1* (Buffalo: SUNY Buffalo Center for Cognitive Science) and *Technical Report 99-??* (Buffalo: SUNY Buffalo Department of Computer Science and Engineering).

1 Introduction

In this paper, I advocate a theory of "syntactic semantics" (Rapaport 1988b, 1995) as a way of understanding how computers can think (and how the Chinese-Room-Argument objection to the Turing Test can be overcome): (1) Semantics, considered as the study of relations *between* symbols and meanings, can be turned into syntax—a study of relations *among* symbols (including meanings)—and hence syntax (i.e., symbol manipulation) can suffice for the semantical enterprise (contra John Searle). (2) Semantics, considered as the process of understanding one domain (by modeling it) in terms of another, can be viewed recursively: The base case of semantic understanding—understanding a domain in terms of itself—is "syntactic understanding" (Rapaport 1986b). An internal (or "narrow"), first-person point of view makes an external (or "wide"), third-person point of view otiose for purposes of understanding cognition. I also sketch the ramifications of this view with respect to methodological solipsism, conceptual-role semantics, holism, misunderstanding, and implementation, and briefly look at Helen Keller as inhabitant of a Chinese Room.

2 The Turing Test

In 1950, Alan M. Turing, already well known for his Turing-machine theory of computation (Turing 1936), published an essay in the philosophy journal *Mind* on the topic of "Computing Machinery and Intelligence". He opened the essay by saying that he would "consider the question, 'Can machines think?' " (Turing 1950: 433). But rather than answer this provocative question directly, he proposed an experiment whose outcome would provide guidance on how to answer it.

He described the experiment by analogy with something he called "the 'imitation game' " (Turing 1950: 433). This parlor game

> … is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart from the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman. He [sic] knows them by labels X and Y, and at the end of the game he says either 'X is A and Y is B' or 'X is B and Y is A'. The interrogator is allowed to put questions to A and B …. It is A's object in the game to try and cause C to make the wrong identification. …
>
> In order that tones of voice may not help the interrogator … [t]he ideal arrangement is to have a teleprinter communicating between the two rooms.[1] … The object of the game for … (B) [the woman] is to help the interrogator. The best strategy for her is probably to give truthful answers. …
>
> We now ask the question, 'What will happen when a machine [specifically, a digital computer; Turing 1950: 436] takes the part of A [the man] in this game?' Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman? These questions replace our original, 'Can machines think?' (Turing 1950: 433–434.)

At this point in his essay, Turing says nothing about what the suitably-programmed computer is supposed to do. Clearly, the computer is supposed to play the role of the man, but the man's task in the original imitation game was to fool the interrogator into thinking that he or she is conversing with the woman. Traditionally, this has been taken to mean that the computer is supposed to fool the interrogator into thinking that it is human *simpliciter*. However, read literally and conservatively, if the computer is supposed to do this by playing the role of the man, then it appears that the computer has a more complex task, namely, to behave like a man who is trying to convince the interrogator that he is a woman! (Colby et al. 1972: 202 makes a similar observation.) Of course, were the computer to be successful in this very much harder task, it would also, *ipso facto*, be successful in convincing the interrogator that it was human *simpliciter*.

Later (p. 442), Turing has us consider "one particular digital computer *C*", and asks:

> Is it true that by modifying this computer to have an adequate storage, suitably increasing the speed of action, and providing it with an appropriate programme, *C* can be made to play satisfactorily the part of A [i.e., the man] in the imitation game, the part of B [i.e., the woman] being taken by a man?

If the part of B is taken by a man, then it follows, from the earlier description that the interrogator's task is to determine which of X and Y is A and B, that B is simply supposed to convince the interrogator that he is the man (or the human) and that the computer's task is to convince the interrogator that *it* is the man (or the human).[2] So it appears that Turing was not overly concerned with the complication discussed in the

[1] Today, of course, we would have them communicate only via a "chat" program over a computer network.

[2] However, this passage *can* be read as follows: A's and B's jobs are still to convince the interrogator that each is a woman; thus, if B is a man, his job is to convince the interrogator that he is a woman, while the computer's job is to convince the interrogator that it is a man who is trying to convince the interrogator that he is a woman. But this is an unlikely interpretation.

previous paragraph (although feminists will note that he apparently thought it important that the human in this human-computer contest be represented by a man, not a woman).

In any case, Turing answered this new question as follows:

> I believe that in about fifty years' time [i.e., by about 2000] it will be possible to programme computers, with a storage capacity of about 10^9, to make them play the imitation game so well that an average interrogator will not have more than 70 per cent. chance of making the right identification after five minutes of questioning. The original question, 'Can machines think?' I believe to be too meaningless to deserve discussion. Nevertheless I believe that at the end of the century *the use of words* and *general educated opinion* will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted. (Turing 1950: 442; my italics.)

3 The Use of Words vs. General Educated Opinion: "Thinking" vs. Thinking

The Turing Test, as the computer version of the imitation game has come to be called, is now generally simplified even further to a 2-player game: Can a human conversing with an unknown interlocutor through a computer "chat" interface determine whether the interlocutor is a human or a suitably programmed computer, or—more simply—can a computer convince an interrogator (who is unaware of who or what he or she is conversing with) that its ability to think, as demonstrated by its ability to converse in natural language, is equivalent to that of a human (modulo the—quite low—70%/5-minute threshold)?

There is an echo of this in Peter Steiner's famous *New Yorker* cartoon (5 July 1993: 61) in which a dog, sitting in front of a computer, observes that "On the Internet, nobody knows you're a dog" (Fig. 1). The success of this cartoon depends on our realization that, in fact—just like the interrogator in a 2-player Turing test—one does *not* know with whom one is communicating over the Internet. This ignorance on our part can have serious real-life implications concerning, e.g., computer security (if I enter my credit-card number on your Web site, have I really bought a book, or have I given my number to a con artist?) and matters of social welfare or personal safety—even life and death (is my daughter chatting with a member of the opposite sex who is about her age, or is she chatting with a potential sex offender?). But note also that, even though many of us are aware of these possibilities, we normally assume that we are *not* talking to a con artist, a sex offender, or even a dog. Or—for that matter—a computer. (My mother did not recognize (or expect) the possibility that she was not talking to a human on the phone, and thus regularly tried to converse with pre-recorded phone messages.) We normally are, in fact, fully prepared to accept our invisible interlocutor as a (normal, ordinary) human with human thinking capacities.

And this, I suggest, was Turing's point.[3] It is, nearly enough, the point of the argument from analogy as a solution to the problem of other minds: I know (or assume) that *I* have a mind and can think, but, when I converse with you face to face, how do I know whether (or can I assume that) *you* have a mind and can think? The argument from analogy answers as follows: You are sufficiently like me in all other visible respects, so I can justifiably infer (or assume) that you are like me in this invisible one. Of course, I could be wrong; such is the nature of inductive inference: You could be a well-designed android, such as *Star*

[3]Although, as my colleague Stuart C. Shapiro pointed out to me, the interrogator, on one reading, knows that he or she is participating in a test. However, another Turing test does not require such knowledge on the interrogator's part: Let the interrogator (unknowingly, of course) begin the conversation with the human; then, at some point, let the computer change places with the human. Can the interrogator tell at what point in the conversation the switch took place? (This is suggested in a passage in Lassègue 1996, §3.2.2. A similar suggestion ("the Extended Turing Test") was made in Abelson 1968: 317–320 and is discussed in Colby et al. 1972: 203–204.)

Figure 1: ©*The New Yorker*, 5 July 1993: 61.

Trek's Commander Data. But we make this inference-to-mindedness—if only unconsciously—on a daily basis, in our everyday interactions. Now, in the case of a Turing test, I (as interrogator) have considerably less analogical information about you; I only have our conversations to go by. But, even in a much weaker case such as this, we do ordinarily infer or assume (and justifiably so) that our interlocutor is human, with human thinking (i.e., cognitive) capabilities.

Is there anything wrong with this? Well, if my interlocutor isn't who (or what) I think he (or she, or it) is, then I was wrong in my inference or assumption. And if my interlocutor was really a suitably programmed computer, then I was certainly wrong about my interlocutor's *biological* humanity. But was I wrong about my interlocutor's (human) cognitive capabilities (independently of the interlocutor's implementation)? That is the question. Turing's answer is: No. Perhaps more cautiously, the lesson of Turing's test is that the answer depends on how you define '(human) cognitive capabilities': One way to define them is in terms of "passing" a Turing test; in that case, of course, any Turing-test-passing interlocutor does think (this is essentially Turing's strategy). Another way is to come up with an antecedently acceptable definition, and ask whether our Turing-test-passing interlocutor's behavior satisfies it. If it does, we have several choices: (1) We could say that, therefore, the interlocutor does think, whether or not it is biologically human (this is, roughly, Turing's strategy, where the antecedently-given definition is something like this: convincing the interrogator of your cognitive capacities with the same degree of accuracy as, in the original game, the man (A) convinces the interrogator that he is the woman (B)); or (2) we could say that there must have been something wrong with our definition if the interlocutor is not biologically human; or (3) we could say that, while the interlocutor is doing something that superficially satisfies the definition, it is not "really" thinking. In case (3), we could go on to say (4) that that is the end of the matter (this is essentially Searle's move in the Chinese-Room Argument) or (5) that the interlocutor is merely "thinking" in some metaphorical or extended sense of that term. Comparison with two other terms will prove enlightening.

3.1 'Fly'

Do birds fly? Of course. Do people fly? Of course not, at least not in the same sense. When I say that I flew to New York City last month, I don't really mean that I flew like a bird ("Didn't your arms get tired?", joke my literalistic friends). What I mean is that I was a passenger on an airplane that flew to New York City. Oh? *Do* airplanes fly? Well, of course; don't they? Isn't that what the history of heavier-than-air flight was all about? Ah, but planes don't fly the way birds do: They don't flap their wings, and they are powered by fossil fuel. So have we, after all, failed in our centuries-old attempt to fly like the birds? No. But how can this be?

There are two ways in which it makes perfectly good sense to say that planes fly: One way is to say that 'fly' is used metaphorically with respect to planes—birds fly; planes only "fly"—but this is one of those metaphors that have become so ingrained in our everyday language that we no longer recognize them as such. (The *locus classicus* of this approach is Lakoff & Johnson 1980.) Turing may have had this in mind when he spoke—in the italicized passage quoted above—about "the use of words" changing. Thus, we can likewise extend 'flying' to cover hot-air balloons (which don't have wings at all), spaceships (which don't travel in air), arrows and missiles (some of which, perhaps more accurately, merely "fall with style", as the film *Toy Story* puts it), and even the movement of penguins under water (more usually called 'swimming'[4]).

The other way in which it makes perfectly good sense to say that planes fly is to note that, in fact, the physics of flight is the same for both birds and planes (e.g., shape of wing, dynamics of airflow, etc.).

[4]"People see films of penguins waddling across the ice, hear them referred to as flightless birds, and naturally assume that penguins can't fly.

"But penguins do indeed fly—they fly in water. Using their wings, which are flat and tapered and have a rounded leading edge, and flapping like any swift or lark, penguins fly through the water to feed and to escape predators" (Ackerman 1989: 45).

What we may have once thought was essential to flying—flapping of wings—turns out to be accidental. Our understanding of what flying really is has changed (has become more general, or more abstract), so that more phenomena come under the rubric of "flying". Turing may have had this option in mind in his remark about "general educated opinion" changing.

The same two options apply to 'thinking': We could say that, insofar as suitably programmed computers pass a Turing test, they do think—extending 'think' metaphorically, but legitimately, just as we have extended 'fly' (which we have always done, even at the very beginnings, centuries ago, of research into human flight). Or we could say that being human is inessential for thinking, the psychological principles of thinking being the same for both humans and suitably programmed computers (or even animals; cf. Griffin 1976, Allen 1997, Allen & Bekoff 1997).

Note that *both* the use of the word 'fly' has changed *and* general educated opinion has changed. Thus, some things (like spaceships and missiles) arguably only "fly", while others (like planes) definitely fly like birds fly. But one can in fact speak of all those things flying "without expecting to be contradicted". Moreover, these two ways need not be exclusive; the common physical or psychological underpinnings of flight or thought might be precisely what allow for the seamless metaphorical extension.

3.2 'Computer'

Another term that has undergone a change of meaning is also instructive and perhaps more to the point: 'computer'.[5] At the time of Turing's 1936 paper on what is now called the Turing machine, a "computer" was primarily a *human* who computed.[6] Turing distinguished between a computing *machine* and a (human) *computer*:

> The behaviour of the *computer* at any moment is determined by the symbols which *he* is observing, and *his* 'state of mind' at that moment. … We may now construct a *machine* to do the work of this computer. To each state of mind of the *computer* corresponds an '*m*-configuration' of the *machine* (Turing 1936 [1965: 136–137]; my italics).

By the time of his 1950 paper on what is now called the Turing test, he posed the question "Can *machines* think?" and spoke of "*digital* computers", "*electronic* computers", and "*human* computers", only rarely using 'computer' unmodified to mean a computing *machine*, as if the modifier 'digital' or 'electronic' still served to warn some readers that *human* computers were not the topic of discussion. Certainly today, 'computer' almost never refers to a human. What happened here? Perhaps first by analogy or metaphorical extension, 'computer' came to be applied to machines. And then, over the years, it has been applied to a large variety of machines: vacuum-tube computers, transistor-based computers, VLSI computers, mainframes, workstations, laptops, "Wintel" machines, Macs, microprocessors embedded in our toasters, etc. What do all these (as well as humans) have in common?—the ability to compute (in, say, the Turing-machine sense).[7] Thus, "general educated opinion" has changed to view 'computer', not so much in terms of an implementing device, but more in terms of functionality—input-output behavior, perhaps together with general algorithmic structure. This change in 'computer' to focus on computational essentials parallels the change in 'fly' to focus on aerodynamic essentials. And it parallels a change in 'think' (and its cognates) to focus on the computational/cognitive essentials. So it is quite possible that Turing was suggesting that the use of 'think'

[5] I am grateful to my colleague Alistair E. Campbell for this example.

[6] In the *OED* (Simpson & Weiner 1989, Vol. III, pp. 640–641), the earliest cited occurrence of 'computer' (1646) refers to humans. The earliest citation for 'computer' referring to machines is 1897, and the next is 1915, both long before the development of modern computers; the bulk of the citations are from 1941ff.

[7] 'Computer' has also been applied to *analog* computers, which, presumably, go beyond the Turing-machine model of computation.

(and its cognates) will undergo a similar conversion from applying only to humans to applying also (albeit not primarily) to machines.

"But," the critic objects, "it isn't *really* thinking; there's more to thinking than passing a Turing test." This is the gut feeling at the heart of Searle's Chinese-Room Argument (Searle 1980, 1982, 1984, 1990, 1993), to which we now turn.

4 The Chinese-Room Argument

The Chinese-Room Argument focuses on language as the expression of thought, and sets up a situation in which an entity passes a Turing test but, by hypothesis, cannot "think"—or, in the envisaged situation, cannot understand language. In this section, I present the argument and two objections.

4.1 The Argument

The situation is this: Searle, who by hypothesis cannot understand written or spoken Chinese, is sealed in a room supplied with paper, pencils, and an instruction book written in English (which he does understand). (1) Through an input slot come pieces of paper with various marks ("squiggles") on them. (2) Searle-in-the-room manipulates the squiggles according to the instructions in the book, and outputs other pieces of paper with squiggles on them that he wrote following the instructions. Steps (1) and (2) are repeated until the experiment stops. From Searle-in-the-room's point of view, that's all he's doing. Unknown to him, however, outside the room (playing the role of interrogator in a Turing test) is a native speaker of Chinese. This native speaker has been inputting to the room pieces of paper with a story (written in Chinese), sufficient background information (written in Chinese) for whoever (or whatever) is in the room to understand the story, and questions (written in Chinese) about the story (perhaps of the sort one finds on an SAT-type test of reading ability). And the native speaker has been receiving, as output from the room (or from whoever or whatever is in it), pieces of paper with excellent answers to the questions, written in fluent Chinese. From the native speaker's point of view, whoever or whatever is in the room understands Chinese (and thus has passed this Turing test). But the native speaker's and Searle-in-the-room's points of view are inconsistent; moreover, Searle-in-the-room's point of view is, by hypothesis, the correct one. Therefore, it is possible for an entity to pass a Turing test without being able to think. More precisely, it is possible to pass a Turing test for understanding natural language without being able to understand natural language. (I return to the differences in point of view in §7.1, below.)

4.2 Two Objections

There have been numerous objections to the Chinese-Room Argument right from the beginning (cf. Searle 1980), but this is not the place to survey them all (a useful compendium may be found in Hauser 1996). I will focus on only two of them. At its core, there are two components to "the" Chinese-Room Argument: an argument from biology and an argument from semantics.

4.2.1 The Argument from Biology

The argument from biology is this:

(B1) Computer programs are non-biological.

(B2) Cognition is biological.

(B3) $\therefore$ No non-biological computer program can exhibit cognition.

I claim that (B2) is wrong: It assumes that cognition (or, in particular, understanding natural language) is not something that can be characterized abstractly and implemented in different (including non-biological) media (cf. Rapaport 1985; 1986a; 1988a; 1996, Ch. 7; and 1999). But if—and I readily admit that this is a big "if"—computational cognitive science succeeds in its goal of developing an algorithmic theory of cognition, then those algorithms will be able to be implemented in a variety of media, including non-biological ones.[8] And any medium that implements those algorithms will exhibit cognition (just as airplanes, as well as birds, do fly). (For a defense of this against two recent objections, see Rapaport 1998.)

4.2.2 The Argument from Semantics

The central concern of the present essay is the argument from semantics:

(S1) Computer programs are purely syntactic.

(S2) Cognition is semantic.

(S3) Syntax alone is not sufficient for semantics.

(S4) ∴ No purely syntactic computer program can exhibit semantic cognition.

I claim that premise (S3) is wrong: Syntax *is* sufficient for semantics. Now, anyone who knows what "syntax" and "semantics" are knows that they are not the same thing—indeed, I spend hours each semester trying to drive home to my students what the differences are. So how can I turn around and say that one suffices for the other?

To begin to see how, consider that what Searle alleges is missing from the Chinese Room is semantic links to the external world, links of the form that such-and-such a squiggle refers to, say, hamburgers: "… 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" (Searle 1982: 5). Note that Searle makes two assumptions: that external links are needed for the program to "attach" meaning to its symbols, *and* a solipsistic assumption that the computer has no links to the external world—that all is internal to it. Now, first, *if* external links *are* needed, then surely a computer could have them as well as—and presumably in the same way that—humans have them (this, I take it, is the thrust of the "robot" reply to the Chinese-Room Argument; Searle 1980: 420). But *are* external links needed? How might we provide Searle-in-the-room with such links? One way would be to give him, say, a hamburger (i.e., to import it from the external world) clearly labeled with the appropriate squiggle (à la Helen Keller; see §9.2.6, below). But now the hamburger is *in* the room; it is no longer part of the external world. Sure—it *came* from the external world, but so did the squiggles. Searle-in-the-room could just as well have been antecedently supplied with a stock of sample objects[9] (and much else besides, for word-*object* links won't

[8] Conceivably, some of the algorithms might be implementation-dependent in some way; see, e.g., Thagard 1986, and cf. Maloney 1987, 1989. But at most this might constrain the nature of the feasible implementing media. It wouldn't necessarily rule out non-biological ones. In any case, the view that an algorithm might be implementation-dependent would seem to go against the grain of the generally accepted view of algorithms as being implementation-*in*dependent.

[9] Cf. *Gulliver's Travels*, where Swift (1726) described the Laputian

> scheme for entirely abolishing all words whatsoever; … [S]ince words are only names for *things*, it would be more convenient for all men to carry about them such *things* as were necessary to express the particular business they are to discourse on. …. **[T]herefore, the room where company meet who practise this art, is full of all *things* ready at hand, requisite to furnish matter for this kind of artificial converse.** Another great advantage proposed by this invention, was that it would serve as an universal language to be understood in all civilised nations, whose goods and utensils are generally of the same kind, or nearly resembling, so that their uses might easily be comprehended. And thus, ambassadors would be qualified to treat with foreign princes or ministers of state, to whose tongues they were utter strangers" (Part III, Ch. 5 [1967: 230–231]; italics in original; boldface added.)

suffice; abstract concepts such as *love*, *number*, etc., will require word-*concept* links).[10] In either case (an imported hamburger delivered from outside or a previously-supplied one stored in the refrigerator at home), the word-meaning links would be internal to the room. As I will argue below, this makes them part of a (larger) syntactic system, and so syntax will *have* to suffice for semantics. To see how, it will help if we review the classical theory of syntax and semantics.

5 Semiotics: Syntax, Semantics, and Pragmatics

Consider some symbol system, i.e., some set of symbols that may or may not be "meaningful". Now, I am stepping on some semiotic toes here when I talk like this, for, in the vocabulary of many (if not most) writers on the subject, symbols are, by definition, meaningful. So, instead, consider a set of "markers" (let us call them) that do not wear any meaning on their sleeves (cf. Fetzer 1994: 14; Rapaport 1998). Think of marks or patterns on paper (or some other medium) that are easily re-identifiable, distinguishable one from another, and relatively unchanging, and do not (necessarily) come already equipped with a semantic interpretation.

According to Charles Morris's classic presentation of semiotics (1938: 6–7; cf. Posner 1992), *syntax* is the study of relations among these markers. Some, for instance, are proper parts of others; certain combinations of them are "legal" (or "grammatical"), others not; and whenever some are in proximity to each other, certain others can be constructed or "derived" from them; etc. (This characterization is intended to cover both the well-formedness and inference rules of complex markers as well as proof-theoretical rules of inference.) Crucially, syntax does not comprise any relations of the markers to any non-markers.

Semantics, according to Morris, is precisely what syntax is not: the study of relations *between* the system of markers and other things. What other things? Traditionally, their "meanings": Traditionally, semantics is the study of the relation of symbols to the things (in the world) that the symbols mean.

Pragmatics will be of less concern to us, but, for the sake of completeness, let me mention that pragmatics is, according to Morris, the study of the relations between markers and their interpreters. Note that this tripartite analysis of semiotics omits a study of the relations between interpreters and symbol-meanings.[11] Perhaps as a consequence, pragmatics is often described as the study of the relations among markers, their meanings, *and* users of the markers. This somewhat more vague study has variously been taken to include the study of indexicals (symbols whose meaning depends on speaker and context), speech acts, discourse phenomena, etc.; it is often characterized as a grab bag of everything not covered by syntax and semantics as above defined.

What is not usually noticed in these definitions is this: If the set of markers is unioned with the set of meanings,[12] and the resulting set considered as a set of (new) markers (i.e., if the "meanings" are made internal to the symbol system), then *what was once semantics*—viz., relations *between* old markers and their meanings—*is now syntax*—viz., relations *among* old and new markers (see §6; these new relations are in *addition* to the old ones that classify markers and provide well-formedness rules). Furthermore, as noted in

Although this may strike one as sounding more like a description of Russellian propositions, if 'things' is replaced by 'concepts', then we have a nice description of a Chinese Room (as well as of interlingual machine translation).

Gulliver's Travels, incidentally, has several other passages potentially relevant to the Turing Test: The King of Brobdingnag determines that Gulliver, who looks to the Brobdingnagians more like a toy, is rational by considering the kinds of answers he gives to various questions (Part II, Ch. 3 [1967: 141–142]). Conversely, Gulliver determines that the Houyhnhnms, who look like horses, are rational by considering their actions, including their language (Part IV, Ch. 1 [1967: 270ff]). (I am indebted to Shapiro for recommending *Gulliver's Travels* as a source of philosophical examples.)

[10] Moreover, as Kearns 1996, 1997 has argued, it is *speech acts*, not expressions, that are the bearers of meaning.

[11] As well as studies of the relations among symbol-meanings (or is that all of science and perhaps some of psychology?) and of the relations among interpreters (or is that part of sociology?).

[12] Taking care in the case of markers that refer to other markers as their meanings, an important special case that I want to ignore for now.

my discussion of pragmatics, it is left open how the symbol-user understands the symbol-meanings. I shall argue that this must be done in a syntactic fashion (see §8). It is in these ways that syntax can suffice for semantics.

But a lot more needs to be said.

6 Syntactic Semantics: I—Turning Semantics into Syntax

One thing that is needed is an argument that the set of (old) markers *can* be unioned with the set of meanings. Insofar as the markers are internal to a mind, we need an argument that the semantic domain can be internalized, so to speak. This *can* happen under certain conditions. In particular, it can happen (and, I claim, does happen) under the conditions obtaining for human language understanding. For how *do* I learn the meaning of a word? Let us, for now, consider only the very simplest case of a word that clearly refers.[13] How do I learn that 'tree' refers to that large brown-and-green thing I see before me? Someone points to it in my presence and says something like "This is called a 'tree' ". Perhaps numerous repetitions of this, with different trees, are needed. I begin to associate[14] two things, but what two things? A tree and the word 'tree'? No; to paraphrase Walker Percy (1975: 43), the tree is not the tree out there, and the word 'tree' is not the sound in the air.[15] Rather, *my internal representation of* the word becomes associated ("linked", or "bound") with *my internal representation of* the tree.[16] Light waves reflected from the tree in the external world enter my eyes, are focused on my retina, and are transduced into electrochemical signals that travel along my optic nerve to my visual cortex. No one knows *exactly* what goes on in visual cortex (or elsewhere) at that point. But surely some nerves are activated that are my internal representation (perhaps permanent, perhaps fleeting) of that tree. Likewise, sound waves emanating from the 'tree'-speaker's vocal tract reach my ears and, via my auditory nerves, ultimately reach my auditory cortical areas, where surely the story is the same: Some nerves are activated that are my internal representation (for the nonce, if not forever) of the word 'tree'. *And these two sets of activated nerves are, somehow, associated, or "bound".*[17] *That* is the semantic relation,

[13]The case of other terms is even more likely to be internal; this is best explored in the context of conceptual-role semantics; cf. §9.2.2.

[14]What constitutes "association"? In this case, simply co-occurrence: When I hear 'tree', I think of trees. Later, it will mean that some kind of "internal" link is forged between the associated things: In the case of Cassie, a computational cognitive agent (introduced later in this section), it will be a semantic-network arc; in the case of a human, it might be some neural sort of "binding" (see Damasio 1989a).

[15]Although apt, this is a slightly misleading paraphrase, since Percy's point is that, in understanding that 'tree' means tree, 'tree' and tree are types, not tokens.

[16]In the case of Cassie, an "internal representation" of a word or object would be a semantic-network node. In the case of a human, it might be a pattern of neuron firings. 'Representation' may not be the happiest term: If there is an external object, then the internal correlate "represents" it in the sense that the internal entity is a proxy for the external one. But if there is no external entity, then it is perhaps inappropriate to speak of 'representation'. See Rapaport 1978, 1981, and Shapiro & Rapaport 1987, 1991 for more on the nature of these Meinongian objects.

[17]For some discussion of this, see Damasio 1989ab and Rapaport 1996, esp. Ch. 3. My Physiology colleague Susan Udin (personal communication, 5 October 1998) informs me that

> nobody really has a clue about how the visual or auditory stimuli actually produce what we experience as the sight or sound. However, we do know that there are many 'association' areas of the cortex in which several modalities, such as vision and audition, activate cells. The primary sensory areas and those association areas will all participate in the immediate sensations triggered by the visual and auditory stimuli. The question of where the learned representations reside is even more obscure, but association areas are likely candidates.

My Biophysical Sciences colleague K. Nicholas Leibovic (personal communication, 23 October 1998) tells me that one shouldn't assume

> that the internal representation is in V1, the primary visual cortex to which the retinal input projects. … [W]e do not know what form the internal representation takes. One form may be an activation of specific neurons. (But it need not be that. It could be a pattern of activity in distributed brain areas, a 'program' of activity. We just don't know

but—taking the activated nerves as the markers (as well as the meanings)—it is a *syntactic* relation. Thus, it is precisely this coordination of multiple modalities that allows syntax to give rise to semantics.[18]

The same holds—or could hold—for a suitably programmed computer. When I converse in English with "Cassie"—a computational cognitive agent implemented in the SNePS knowledge-representation, reasoning, and acting system—she builds internal representations (nodes of a semantic network) of my sentences (Shapiro 1979; Shapiro & Rapaport 1987, 1991, 1992, 1995). If I show her pictures, she builds similar internal representations (more nodes of the same semantic network), and she can associate the nodes from the "linguistic part" of her network with the nodes from the "visual part" (Srihari & Rapaport 1989, Srihari 1991). (The inverse task, of finding—or pointing to—some object in the external world, supplements the nodes with other symbols, as described in detail in Shapiro 1998; roughly, Cassie's internal representation of the object is "aligned" with, again roughly, her internal visual representation of it, and that latter symbol is used to direct her to the corresponding external entity, but in no case does she have direct access to the external entity.)

And the same holds, more generally, whenever we present a formal semantics for some syntactic system: We define the syntactic domain syntactically. Then we present a domain of semantic interpretation. But this domain is *also* presented (or, if it is supposed to be the real world, it is represented) syntactically. And then we relate—syntactically—these two syntactic systems (cf. Smith 1982; Rapaport 1995 and 1996, Ch. 2).

A related argument for an apparently similar conclusion, based on Chomsky's "minimalist" program, has been offered by James McGilvray (1998):

> one should look … to expressions inside the head for meanings …. … [M]eanings are contents intrinsic to expressions … and … they are defined and individuated by syntax, broadly conceived. … these concepts are individuated by internally and innately specified features, not by their relationships to the world, if any. (pp. 225, 228.)

My merger of syntax and semantics into a new syntactic domain whose relation between old (syntactic) markers and new (semantic) markers seems to be echoed by Chomsky's …

> … 'Relation *R*' ('for which read reference', but without the idea that reference relates an LF [logical form] to something 'out there') that stands between elements of an LF and these stipulated semantic values that serve to 'interpret' it. This relation places *both* terms of Relation *R*, LFs and their semantic values, entirely within the domain of syntax, broadly conceived; …. They are in the head. (p. 268.)

7 Points of View

7.1 Whose Point of View Is "Correct"?

The internal-picture sketched above is from the first-person point of view. In studying how a human mind understands language (or, more generally, thinks and cognizes), and in constructing computational models of this (or, more strongly, constructing computational cognitive agents), we must consider, primarily, what

---

what form representation takes in neural terms.)

And my Neurology colleague Alan Lockwood (personal communication, 5 October 1998) informs me that "There are some data about the internal representation process … that show that blind people activate visual cortex when using braille reading (Sadato, *Nature*, vol 380, pp 526–528)". I am also grateful to my Communicative Disorders and Sciences colleague Elaine Stathopoulos for some terminology.

[18] Here, 'syntax', qua "symbol manipulation" (or "marker manipulation"), is to be taken "broadly". For discussion, see Bunn, forthcoming, and Jackendoff, forthcoming.

is going on inside the agent's head, from the agent's point of view. (In Chomsky's and McGilvray's terms (McGilvray 1998: 240–241), we must study an "I-language".) Internally, there are markers that represent or correspond to linguistic entities (words, sentences), markers that represent or correspond to conceptual entities (e.g., propositions and their components), and (perhaps) markers that represent or correspond to entities in the external world.[19] But all of these internal markers are only related to each other. More precisely, the cognitive agent only needs to deal with (i.e., to manipulate) these internal markers; the agent does not need to be concerned with the causal-historical origins of the markers, nor do we need to be concerned with these origins insofar as we are trying to understand how the agent thinks by means of these markers. We need only study the internal relations among them. We do not (at this stage) need to study any external relations between markers and external entities.

The notion of "point of view" is central to the Turing-Test-vs.-Chinese-Room debate, too. As we saw in §4.1, the point of view of the native Chinese speaker differs from the point of view of Searle-in-the-room. Which point of view should dominate? Since all that the Turing Test talks about is the point of view of the interrogator, then—contrary to Searle—what might "really" be going on in the external world—i.e., the point of view of Searle-in-the-room—is irrelevant to the Turing Test.

To get a feeling for why this is, consider the following conversation between Dorothy and the Munchkin Boq, from *The Wizard of Oz* (Baum 1900):

> When Boq saw her silver shoes[20] he said,
> "You must be a great sorceress."
> "Why?" asked the girl.
> "Because you wear silver shoes and have killed the wicked witch. Besides, you have white in your frock, and only witches and sorceresses wear white."
> "My dress is blue and white checked," said Dorothy, smoothing out the wrinkles in it.
> "It is kind of you to wear that," said Boq. "Blue is the color of the Munchkins, and white is the witch color; so we know you are a friendly witch."
> Dorothy did not know what to say to this, *for all the people seemed to think her a witch, and she knew very well she was only an ordinary little girl* who had come by the chance of a cyclone into a strange land. (pp. 34–35; my italics.)

Is Dorothy a witch or not? From her point of view, she is not; but, from the point of view of Boq and the other Munchkins, she is. Dorothy knows herself not to be a witch, no? At least, she *believes* that she is not a witch, as *she* understands that term. But it is certainly possible for her to *believe* that she is not a witch, yet for her to *really be* a witch (in either her terms or the terms of the Munchkins). So, what counts as really being a witch? We must answer this from the point of view of what *Munchkins* take witches to be, for there are many theories of witchcraft, but only the Munchkin theory counts in the present context. The dispute is not about whether Dorothy is "really" a witch in some context-independent sense (from Dorothy's, or the reader's, point of view), but whether she is a witch in the Munchkin sense (from the Munchkin point of view). Boq cites her clothing and actions, which Dorothy admits to. In *Oz*, witches also perform magic, which Dorothy denies having done. But what counts as magic (again from the Munchkin point of view)? Standard magical things like disappearing and transforming one object into another, to be sure, but who is Dorothy (or me, for that matter) to say that, from the Munchkin point of view, her behavior and actions (such as suddenly dropping from the sky) are not included under what they consider to be "magical"? The Munchkin point of

[19]But see Maida & Shapiro 1982, Shapiro & Rapaport 1991, and §7.2, below, for an argument *against* representing external, or "extensional", entities.

[20]A note for those only familiar with the 1939 movie version: The novel has silver shoes, not ruby slippers. And, to those only familiar with the 1939 movie version, shame on you! Baum's *Oz* books are full of wonderful philosophical observations.

view trumps Dorothy's point of view with respect to what it means to be a witch in Munchkinland—they, not Dorothy, are the experts on criteria of their notion of witchcraft.[21]

The Chinese-Room situation is analogous. Does Searle-in-the-room understand Chinese or not? From his point of view, he does not; but from the point of view of the native Chinese speaker, he does. Searle-in-the-room knows himself not to understand Chinese, no? (Certainly, that's what Searle (1980) claims.) At least, he *believes* that he does not understand Chinese, as *he* understands that term. But it is certainly possible for him to *believe* that he does not understand Chinese, yet for him to *really understand* Chinese (see the next paragraph). So, what counts as really understanding Chinese? We must answer this from the point of view of what native Chinese speakers take understanding Chinese to be. For a person might believe that he or she *does* understand Chinese, yet be mistaken; only the native Chinese speaker can ask appropriate questions to determine whether that person really does understand. The native Chinese speaker's point of view trumps Searle-in-the-room's point of view with respect to what it means to understand Chinese—the native Chinese speaker, not Searle-in-the-room, is the expert on criteria of understanding Chinese.

The Chinese-Room case may need a bit more explication, for Searle-in-the-room could legitimately reply to the native Chinese speaker that he, Searle-in-the-room, *still* doesn't believe that he understands Chinese, no matter what the native Chinese speaker says. What I have in mind here is the following sort of situation: As it happens, I understand French to a certain extent; let's say that I believe that I understand 80% of what I hear or read, and that I can express myself with, say, 75% expressiveness: I can carry on a conversation on any topic (even give directions to Parisian taxi drivers), but I always feel that I'm missing something or can't quite generate the right idioms. Suppose, however, that a native French speaker tells me that I am fluent in French. "Ah, if only that were true," I reply. Who's right? Searle (in or out of the room) would say that *I* am—I *don't* (fully) understand French, no matter what the native French speaker tells me.

But Searle-in-the-room is not quite in my situation. He has the advantage of an instruction book (his Chinese natural-language-understanding and -generating program). It is not Chineseless Searle who is conversing with the native Chinese speaker, but Searle-plus-his-instruction-book. This is the "systems" reply to the Chinese-Room Argument (Searle 1980: 419), and I am bringing it up for two reasons. First, it shows that, in the Chinese-Room situation, *unlike* my French situation, Searle by himself cannot insist that, because he (alone) knows no Chinese, his point of view takes precedence—because he is *not* alone: He has his instruction book, and, with its help, he does pass the Chinese-understanding test with flying colors, as judged by the only qualified judge there is. Were Searle-in-the-room, with his book, to be stranded on a desert island and forced to communicate with a Friday who only spoke Chinese, he would be able to do it. The native Chinese speaker is the only person qualified to say, truthfully, "I am conversing with someone (or something) who (or that) understands Chinese". That someone (or something) has *no* right to assert that he (or she, or it) either does *or* does not speak Chinese.[22]

The second point to notice about the systems reply (although it is secondary to my present purpose) is that it is reminiscent of Edwin Hutchins's theory of "cognition in the wild" (Hutchins 1995ab). The extended cognitive system that navigates a ship, consisting of the crew *plus* various instruments, is a real-life counterpart of Searle-in-the-room plus his instruction book. Hutchins argues that it is not any individual crew member who navigates the ship, but the crew plus instruments that are external to the crew's minds: "systems that are larger than an individual may have cognitive properties in their own right that cannot be reduced to the cognitive properties of individual persons" (Hutchins 1995b: 266). Similarly, I argue with the systems reply that Searle plus the instruction book that is external to his mind has the cognitive property of

[21] Given that they were also taken in (perhaps) by the Great Oz himself, arguably they are not experts, but one can easily imagine a slightly different situation in which they would be. On the other hand, who's to say that, *from their point of view*, Oz was *not* a wizard?

[22] Cf. my Korean-Room Argument and my example of a student who doesn't understand what greatest common divisors are but who can compute them, in Rapaport 1988b, §§4–5.

understanding Chinese and that this is not (therefore) a cognitive property of Searle-in-the-room by himself (which is consistent with Searle-in-the-room's protestations that he (alone) still doesn't understand Chinese).

7.2 No Direct Access

To return to an earlier point, external links of the sort that Searle believes necessary are not needed, because the cognitive agent has no direct access to external entities. Those are fighting words, so what do I mean by them? I mean, simply, that if I want to say that 'tree' refers to that tree over there, I can only do so by associating my internal word 'tree' with my internal representative of that tree over there. Let me spell this out in more detail: I see a tree over there, and—while pointing to it—I say, "That's what 'tree' refers to" (or, more simply, "That's a tree"; but cf. Percy 1975: 258–264 on the dangers of this formulation). But what do I see? I am directly aware of the following visual image: my hand pointing to a tree. The visual image of the pointing hand and the visual image of the pointed-to tree are all internal. I go up and touch the tree (how much closer to the external world could I get?). But now all I have is an internal tactile image of the tree. It's all internal. I only indirectly access the external tree. (As Percy (1975: 33) puts it, "… it is not really the world which is known but the idea or symbol …, while that which it symbolizes, the great wide world, gradually vanishes into Kant's unknowable noumenon.")

Why do I believe that visual (and other sensory) images are internal, that I have no direct access to the external world, or, better, that my access to the external world—for I do believe that we have such access!—is always mediated by internal representatives of it? I am convinced by the following simple experiments (versions of the argument from illusion; cf. Ayer 1956: 87–95): Look at some distant object, such as a small light source about 10 feet away. Close your left eye; you still see the light. Now open your left eye and close your right; you still see it. But are you seeing the same thing you were in the two previous cases? In one sense, presumably, the answer is "Yes": You are seeing the same distal object—but only indirectly and as mediated by an intensional representative. In another sense—the one I am interested in—the answer is "No": The two (intentional) objects directly seen by your two eyes are slightly different (different locations relative to other entities in your visual field; different shapes; in my own case, at times, slightly different colors).[23] And how do I know that there are two objects? Because, by crossing my eyes, I can see both at once (and, in so doing, I can compare their different colors)! Since, by hypothesis, there are not *two* of them in the external world, the internal images and the external object are even numerically distinct. (There is even a third object: the apparently 3-dimensional one constructed by stereoscopic vision (cf. Julesz 1971), which differs in shape and location from the other two. All are internal visual images—representations of the external object. And the stereoscopically constructed image is not identical with the external object, precisely because it is constructed by the "mind's eye".)

I am not a pure solipsist, merely a representative realist. There *is* an external world, and my internal images are directly caused by external objects. But *I* have (perhaps better: my mind has) no (direct) access to the external objects. Does anyone? Surely, you say, *you* could have access to both worlds. From this third-person point of view, you could have access to my brain and to the external world, and—in the golden age of neuroscience—will be able to associate certain nerve firings with specific external objects. Similarly, I—as Cassie's programmer—can associate nodes of her semantic-network "mind" with things in the external world. Or consider again the situation in which I point to a tree and say 'tree'. From your point of view, you see both the tree and me pointing to it—both of which are, apparently, in the external world. Aren't we both looking at the same tree?

Not really. For suppose I associate Cassie's node `B1` (which, let's say, she lexicalizes as 'tree') with

[23] I can only speculate on the source of the different colors: Perhaps it is caused by different lighting conditions and the slightly different angles from which my two eyes see the distal object, or perhaps it is caused by different amounts of light entering each eye (different *f*-stops?).

that tree over there. What am I really doing? I'm associating *my internal representation of Cassie's node* with *my internal representation of the tree*. And this is all internal to *me*. In the case of my pointing to the tree, all *you* are seeing is the following internal image: my hand pointing to a tree. We can only *assume* that there is an external tree causally responsible for our two internal-to-ourselves tree-images. This is what the third-person point of view really amounts to. (As Bernard Williams (1998: 40) puts it, "Kant was rightly impressed by the thought that if we ask whether we have a correct conception of the world, we cannot step entirely outside our actual conceptions and theories to as to compare them with a world that is not conceptualized at all, a bare 'whatever there is.' ")

There is a similar problem with knowledge vs. belief, taking knowledge as justified true belief (modulo Gettier 1963). Cassie has some beliefs. Does she have any knowledge? Suppose that she believes that p, that p is true, and that she is justified in her belief. Does she know that p? Well, *she* cannot legitimately claim to *know* that p is true; she can only *believe* that it's true (and that may or may not be the same as her mere belief that p). She may in fact be justified in her belief, but she need not *know* (or believe) that she is thus justified. From the first-person point of view, there is no knowledge, only belief. From the third-person point of view—my view of Cassie, say—I can judge that she believes that p, I can judge that p is true, and I can judge that she is justified in her belief. So *I* can say that she knows that p. But, on the third-person point of view properly treated (see above), what is really going on is that *I believe* that Cassie believes that p, *I believe* that p (is true), and *I believe* that Cassie is justified in her belief. So *I believe* that Cassie knows that p. But that's just my belief. (For details and a computational implementation of this, see Rapaport et al. 1997; cf. Rapaport 1998, §3.)

So, by merging internalized semantic markers with (internal) syntactic markers, the semantic enterprise of mapping meanings to symbols can be handled by syntactic symbol (or marker) manipulation, and, thus, syntax *can* suffice for the (first-person) semantic enterprise.

8 Syntactic Semantics: II—A Recursive Theory of Semantic Understanding

There is a second way to approach syntactic semantics. Semantics is concerned with two domains and one binary relation: (1) the domain of the syntactic markers, characterized by (syntactic) formation or inference rules—call this the syntactic domain; (2) the domain of the semantic interpretation, the domain of the entities that are the meanings (or semantic interpretations) of the syntactic entities—call this the semantic domain; and (3) a mapping between the syntactic and semantic domains—the semantic interpretation.

What is the purpose of a semantic interpretation of a syntactic domain? Typically, we use the semantic domain to understand the syntactic domain. If we understand one thing in terms of another, ideally that other must already be understood. The semantic domain, therefore, must ideally be antecedently understood. How? There are two ways to understand the semantic domain: We could turn around and treat *it* like a syntactic domain—like a domain of (uninterpreted) markers characterized syntactically—and then find some *third* domain to play the role of semantic interpretation for *it*. And so on, in what Brian Cantwell Smith (1987) has called a "correspondence continuum". At some point, this process must stop. Our understanding of the last domain in the sequence must be in terms of the domain itself.[24] And the only way to understand

[24] For the sake of clarity, let me provide a couple of examples. But I hasten to warn that they are more complex than they seem. For a fuller discussion of the issues involved, see Rapaport 1995.

Example 1: We understand *numerals* ('1', '2', etc., or 'I', 'II', etc.) in terms of *numbers* (1, 2, etc.). "But what are numbers?", asks the philosopher of mathematics. One answer that has been given is that they are certain *sets* (e.g., $\{\emptyset\}$, $\{\{\emptyset\}\}$, etc.). And what are sets? Whatever answer is given, at some point we must stop (and be satisfied) with something that we feel we understand by itself. One complication here is that all of our discourse about numbers is expressed in terms of numerals ('1', 'I', 'one', '/w∧n/' (i.e., the spoken word), or even '$\{\emptyset\}$', etc.); thus, it could be said that we understand numbers in terms of certain numerical markers. This is the "model muddle" alluded to in §9.2.5. (I am grateful to my Linguistics colleague Jeri Jaeger for phonemic assistance.)

Example 2: When presenting a formal language, we semantically interpret terms, n-place predicates, and wffs by, say, individuals, a domain in terms of itself is syntactically; i.e., we understand it by being conversant with manipulating its markers: That is what syntactic understanding amounts to (cf. Rapaport 1986b). To give the most obvious example, we understand a deductive system syntactically when we understand it proof-theoretically. On this recursive picture of understanding, semantic understanding is, in the final analysis—the base case of the recursion—syntactic understanding. (It is also possible that the correspondence continuum ends in a circle of domains, each of which is understood in terms of the next one in the cycle. In this case, our understanding of any domain in the circle must always be relative to our understanding of the other domains. In fact, we would be better off considering the cycle of domains as a single, large domain, understood syntactically. For details and further discussion, see Rapaport 1995.)

I understand the internal symbols of my own Mentalese language of thought syntactically. One could say that "mental terms" don't *mean*; they just *are* (shades of Gertrude Stein?). More precisely, they *interact*: I manipulate them according to certain (no doubt unconscious) rules. Cassie does the same with her nodes. The meaning of any node in her semantic network consists, essentially, of its relations to all the other nodes in the entire network, or, as it is often put, its meaning is its location in the network (cf. Carnap 1928; Quine 1951; Quillian 1967, 1968; Rapaport 1988b). For some purposes, this may be too much and would need to be constrained to some suitable subnetwork (cf. Hill 1994, 1995; in this way, we can come to learn dictionary-like meanings of new words from context, without any recourse to external sources—cf. Ehrlich 1995; Ehrlich & Rapaport 1997; Rapaport & Ehrlich 1999).

How does Searle-in-the-room understand the native Chinese speakers? In the same way that I understand you: By mapping internal representations of your utterances, considered as syntactic entities, to my internal symbols (which, as we have seen, will include internal representations of external objects), and then doing symbol manipulation—syntax—on them. This is what Searle-in-the-room does: He maps internal representations of the native Chinese speaker's utterances (i.e., he maps the squiggle-input) to his internal symbols (as specified in the instruction manual, which must—although Searle did not specify it—contain a knowledge-representation and reasoning system), and then he manipulates the symbols. (See Rapaport 1988b, §3.5.)

Here is where the two approaches to syntactic semantics merge: Some mental terms (namely, the "old" markers) *do* have meanings, but their meanings are just *other* mental terms (namely, the internalized representatives of the external semantic domain), and these latter "just are".[25]

9 Ramifications of Syntactic Semantics

9.1 The Mind as Syntactic System

What is required of a cognitive agent for it to be able to understand and generate language in this syntactic fashion? A lot. As I have urged before (Rapaport 1988b, 1995), it will need to be able to take discourse (not just individual sentences) as input; it must understand all input, grammatical or not; it must perform inference and revise beliefs; it must make plans (including planning speech acts for natural-language generation, planning for asking and answering questions, and planning to initiate conversations); it must understand plans (including the speech-act plans of interlocutors); it must construct a "user model" of its interlocutor; it must

n-place relations, and states of affairs. "But what are individuals, relations, and states of affairs?", asks the metaphysician. Various ontological theories can be provided; perhaps n-place relations are sets of ordered n-tuples of individuals. What, then, are sets, n-tuples, and individuals? Again, whatever answer is given, at some point we must stop (and be satisfied) with something that we feel we understand by itself.

[25] Where do these latter mental terms—these initial concepts—come from? Each heard word is accompanied by a "bare-particular" concept (see Shapiro & Rapaport 1995), whose only "content" is that it is that which is expressed by that word (cf. the semantics of the SNePS "lex" arc, Shapiro & Rapaport 1987). Connections to other concepts give it more detail. Thus, all such information is "assertional", not "structural", to use Woods's (1975) distinction.

learn (about the world and about language); it must have lots of knowledge (background knowledge; world knowledge; commonsense knowledge, perhaps à la CYC—see Lenat & Feigenbaum 1991; and practical, "how-to", knowledge—see Erion, forthcoming); and it must remember what it heard before, what it learns, what it infers, and what beliefs it revised ("Oh yes, I used to believe that, but I don't any more"). And it must have effector organs to be able to *generate* language. In short, it must have a mind.[26] But note that the necessary mind, thus characterized, will be a purely syntactic system: a system of markers (perhaps semantic-network nodes, perhaps a neural network) and algorithms for manipulating them.

Such algorithms and markers are sometimes called 'rules and representations', but I dislike that phrase. First, 'rules' suggests rigid, unbreakable, antecedently-set-up laws. But the algorithms for manipulating the markers need not be lawlike (they would probably need to be non-monotonic "default" or "defeasible" "rules"), and they could be created on the fly (the system has to be able to learn).[27] Second as I urged in §5, the markers should not be thought of as symbols *representing* something external to the system; although they *can* be related to other things by a third person, the only relations needed by the cognitive agent are all internal. Finally, 'rules and representations' is usually taken as a euphemism for what John Haugeland (1985) called "GOFAI": good old-fashioned, classical, symbolic AI (and often for a particular subspecies of GOFAI: production systems). But 'markers and algorithms' applies equally well to connectionist, artificial neural networks, which disdain rules and representations as being too inflexible or too high-level, and everything that I have said about syntactic semantics applies to connectionist, artificial neural networks, taking the nodes of an artificial neural network as the markers.

9.2 The Nature of Syntactic Semantics

In this section, I offer promissory notes to cover various implications of my theory. All are cashed out elsewhere (mostly in Rapaport 1996).

9.2.1 Methodological Solipsism

The theory I have sketched here has solipsistic overtones, but they are "methodologically" solipsistic, to use Putnam's and Fodor's expression (Putnam 1975, Fodor 1980). From the first-person point of view, a cognitive agent can only *assume* that the external world exists, and it can represent it and its denizens—as well as *non*-existents!—internally; any needed external objects can be internalized. Methodological solipsism, or what is sometimes called "narrow functionalism" (e.g., Harman 1988), is all that is needed to understand how a cognitive agent understands its own behavior. It is all that is needed to *construct* a (model of a) *computational* cognitive agent. *Wide* functionalism—the rest of the story, so to speak—at best tells *external* observers something more about what is going on. It gives them the *de re* information that the cognitive agent itself lacks. But the cognitive agent itself does not need this information for its own purposes. And the external observer only has it indirectly and internally, as well: by having an internal mental model of both the cognitive agent being studied and the context in which it is situated, both of which are "seen" from the third-person observer's own first-person point of view. And, just in case you forgot, the first-person semantic enterprise is one of correspondences among symbols; hence, it is purely syntactic. (For a defense

[26] As Shapiro has pointed out to me, without the effectors, it might have a mind, but not one that would be detectable via a (purely linguistic) Turing test. Cf. the comments in Shapiro 1995: 521–522 concerning the cognitive abilities of humans with physical disabilities. (See also Maloney 1987: 352–353; 1989: 146–149.)

[27] Shapiro (1997) prefers the term 'procedure' to 'algorithm' because, on the standard introduction-to-computer-science-level definition of 'algorithm', algorithms halt and are correct, but many interactive computational procedures (e.g., those for natural-language understanding and generation, or even an airline reservation system) do neither. See also Rapaport 1998 for further discussion of what an algorithm is.

and elaboration, see Rapaport 1996, Ch. 6; cf. Rapaport, et al. 1997, §7.4.2; and see Minsky 1965 for a similar claim.)

9.2.2 Conceptual-Role Semantics

Such a methodologically solipsistic, internal, first-person, syntactic semantics is, in essence, what has been called a conceptual-role semantics (e.g., Sellars 1963; Harman 1982, 1987; Loar 1982; Loewer 1982; Maloney 1998): The meaning of any expression is the role that it plays in the complete system of expressions (cf. the meaning of a node, above; I think it is also important to distinguish between *conceptual*-role semantics and—what is often taken to be synonymous; cf., e.g., Fodor & Lepore 1991—*inferential*-role semantics, for I think that there is more to meaning and understanding than *inference*: "mere" association is also important). Some have argued for the primacy of external, or "wide", semantics (Putnam 1975, Burge 1979), while others have argued for a two-factor analysis (e.g., Harman 1987, 1988; Lepore & Loewer 1987). But, although two factors can be specified—one internal and first-person, the other only specifiable in a third-person way—only the internal, first-person one is needed for understanding how someone understands. Conceptual-role semantics is the appropriate sort of semantics to account (from the first-person perspective) for how a cognitive agent understands language. A truth-conditional semantics can still be provided, but only from a third-person perspective. (For a defense and elaboration, see Rapaport 1996, Ch. 4.)

9.2.3 Holism

If the meaning of any part of the system depends on the entire rest of the system, as it does in conceptual-role semantics, then meaning is holistic. Some find this to be problematic (notably Fodor & Lepore 1992); I do not. One of the problems that Fodor and Lepore raise has to do with the apparent incommensurability of different systems of roles. (Cf. McGilvray 1998: 241: "virtually no one has exactly the same repertoire of SEMs in their operating lexicons as another, hence—again—their I-languages will differ. And one's repertoire changes over time." SEMs are the minimalist program's semantic interfaces, which, with its LFs, "are syntactically specified, period"; McGilvray 1998: 228n3.) Perhaps, Fodor and Lepore suggest pessimistically, one will have to be reconciled to a theory of *similarity* of meaning, rather than of identity of meaning. There are indeed cases where roles can't be cleanly compared. The clearest cases come from language translation. The role of the French preposition '*à*' is simply not played by any one preposition in English, nor is the role of the English preposition 'in' played by any one preposition in French. Yet this prevents neither translation nor mutual comprehension. Nor do cases of dissimilar roles among nouns prevent everyday translation or comprehension, though they wreak havoc with literary and poetic translation, not to mention puns and even everyday associations or connotations. So be it. One can always convey the foreign meaning by a suitable, if prosaic and pedantic, gloss (Jennings 1985; Rapaport 1988b, §3.5). The same holds within one's one language:

> … how does one guarantee that one's neighbor shares the same meanings for terms? The answer is that one doesn't, at least not formally or exhaustively. Rather, in practice, one defeasibly assumes by default that everyone agrees, but one keeps in reserve the ubiquitous conflict resolution method that says "one may call into question whether they and their neighbor are simply disagreeing over the meaning of some terms". (Lenat & Feigenbaum 1991: 236.)

(For elaboration, also see Rapaport 1996, Ch. 4.)

9.2.4 Misunderstanding and Negotiation

A methodologically solipsistic, holistic, conceptual-role semantics would seem to be isolated—insulated, if you will—from the rest of the world of language and communication, and this, surely, cannot be. The external world *does*, of course, impinge on the internal syntactic system. Not only are our internal representations (except possibly for hallucinations) caused by the external world, but we *do* communicate with others. And we avoid *real* first-person solipsism in this way. What is required for two cognitive agents to communicate *successfully* about an external thing is the detection of misunderstandings and their correction by negotiation. When we communicate, we attempt to convey our internal meanings to an audience (interlocutor or reader) by means of a public communication language. I can't have direct access to your thoughts, only to your speech acts (to your language, including gestures). My interpretation is not of what you are thinking, but of your language and gestures. It is, indeed, a conclusion; understanding involves inference, albeit *defeasible* inference. In so doing, however, we almost always fail. Yet we almost always nearly succeed: The resulting misunderstandings, if near enough, can be ignored. Even if we can't ignore them, we can still minimize them through negotiation[28]—learning what our audience meant or thought that we meant. (Bruner 1983 presents psychological evidence that children do just that, and McRoy & Hirst 1993 and Hirst et al. 1994 present a computational theory of correction by negotiation. See Rapaport 1996, Ch. 5, for further discussion.)

9.2.5 Is a Model Syntactic or Semantic?

I said that, typically, the semantic domain is used in order to understand the syntactic domain. But often it is the other way around: A syntactic domain is devised in order to talk about (and thereby linguistically understand) a semantic domain. Marx W. Wartofsky (1966) called this "the model muddle", and I have explored it elsewhere (Rapaport 1995; 1996, Ch. 2). The model muddle is one aspect of the correspondence continuum; another aspect is the computer-science notion of implementation. In Rapaport 1999, I argue that implementation is semantic interpretation.

9.2.6 Helen Keller

I also said that a hamburger labeled with a squiggle would suffice to teach Searle the meaning of that squiggle,[29] much as Helen Keller (allegedly; see Leiber 1996) learned that the finger-spelling 'w-a-t-e-r' meant the cool liquid that was flowing over her hand at the same time. But what really went on at Helen's well house? Does her story tell us anything about what it's like to be in a Chinese Room? I think it does:

> The morning after my teacher came she … gave me a doll. … When I had played with it a little while, Miss Sullivan slowly spelled into my hand the word "d-o-l-l." I was at once interested in this finger play and tried to imitate it. When I finally succeeded in making the letters correctly I was flushed with childish pleasure and pride. Running downstairs to my mother I held up my hand and made the letters for doll. *I did not know that I was spelling a word or even that words existed; I was simply making my fingers go in monkey-like imitation.* (Keller 1905: 35; my italics.)

At the beginning of this passage, one expects that the antecedently-played-with doll would be associated with the finger-spelled word 'd-o-l-l'. But as can be seen from Helen's later claim of ignorance ("I did not know

[28] Or, perhaps, by self-organization: "Given a system in which there is natural variation through local fluctuations [read: individual differences in meanings], global coherence … may emerge provided certain kinds of positive feedback loops [read: negotiation?] are in place" (Steels 1998: 138).

[29] Subject, perhaps, to Quinean (1960), 'gavagai'-like, radical-translation problems; but I suspect that either these can be overcome by negotiation or else ultimately don't matter.

…"), her statement that she "made the letters *for doll*" (my italics) must be taken *de re* (they were letters-for-dolls), since, clearly, Helen did not know that she was "making … *letters*" (my italics) or that they were "for doll". The last sentence of the quotation is, I think, a significant passage. It is a wonderful description of pure syntax. Searle would be pleased. Annie Sullivan (Helen's teacher), on the other hand, no doubt would have had reason to believe that Helen *did* know what she was doing: Annie plays native-Chinese speaker to Helen's Searle-in-the-room. (I explore this throughout Rapaport 1996, esp. Ch. 8.)

9.2.7 Naming

And what is the nature of naming—the nature of that squiggle-label "naming" a hamburger, or of 'w-a-t-e-r' "naming" water? Herbert S. Terrace (1985) has argued that *naming* is a very special activity, restricted to humans:

> … the main function of … [the "use of a symbol as a name"] in the use of human language [is] the transmission of information from one individual to another *for its own sake*. (Terrace 1985: 1016–1017; my italics.)

Terrace claims that non-human primates are unable to name in this sense. Is that true? If so, is it also true for *computerized* cognitive agents? I think the answer, in both cases, is 'No' (see Rapaport 1996, Ch. 9).

10 Who Can Pass a Turing Test?

I believe that a suitably programmed computer could pass a Turing test. I don't think that this has happened yet, examples such as Eliza, Parry, or the Loebner competitions notwithstanding.[30] Nor do I think that it is going to happen in the near future. As I write, 2001 is close upon us, but HAL is not (cf. Stork 1997), and I won't venture to make any more precise predictions: Both Turing (who, in 1950, predicted 2000) and Herbert A. Simon and Allen Newell (who, in 1957, predicted 1967 for the chess version of a Turing test, missing by 30 years; see Simon & Newell 1958) were way off, and I couldn't hope to compete with the likes of them.[31]

But I believe that a suitably programmed computer *will*, eventually, pass a Turing test. And, more importantly, I believe that such a Turing-test-passing computer *will* "really" think, for the reasons adumbrated above, namely, syntax suffices for semantic understanding. More cautiously, I believe that it is a worthy research program to try to build such a computer (i.e., to write such programs) and that such an attempt is the *only* way to find out whether such a computer can be built (cf. Rapaport 1998).

But there is another reason that a Turing test will eventually be passed. It is less interesting from a computational point of view, more so from a sociological point of view. It is simply that—to return to the earlier discussion of the Internet dog—for whatever reasons (and what these are is worth exploring), humans tend to treat other entities with which they interact as if they were human:

> As [software] agents are better able to create the illusion of artificial life, the social bond formed between agents, and the humans interacting with them, will grow stronger. New ethical questions arise. Each time we inspire an agent with one or more lifelike qualities, we muddy the

[30] On Eliza, see Weizenbaum 1966, 1967, 1976. On Parry, see Colby, Weber, & Hilf 1971; Colby et al. 1972; Colby 1981. And on the Loebner competitions, see Loebner 1994, 1998; Rees et al. 1994; Shieber 1994ab.

[31] Although Simon says that "it had nothing to do with the Turing Test" and that "(a) I regard the predictions as a highly successful exercise in futurology, and (b) placed in the equivalent position today, I would make them again, and for the same reasons. (Some people never seem to learn.)" (personal communication, 24 September 1998). At the end of the next millennium, no doubt, historians looking back will find the 40-year distance between the time of Newell and Simon's prediction and the time of Kasparov's defeat to have been insignificant.

> distinction between users being amused, or assisted, by an unusual piece of software and users creating an emotional attachment of some kind with the embodied image that the lifeless agent projects. (Elliott & Brzezinski 1998: 15.)

Call this ‘anthropomorphism’ if you wish. Call it Daniel C. Dennett’s “intentional stance”, if you prefer (Dennett 1971). We have already witnessed tendencies along these lines with Eliza, the winners of the Loebner competitions, and even Garry Kasparov’s attitude toward Deep Blue.[32]

What will happen when we accept a computer as having passed a Turing test? Surely, I predict, we will accept it as a thinking thing. If that means, to paraphrase Turing, that the use of the word ‘think’ will have *altered* (or been metaphorically extended) “so much that one will be able to speak of machines thinking without expecting to be contradicted”, so be it. But my main point in this paper has been to show that no such change is needed. “General educated opinion” will come to see that syntax suffices for *real* thinking.

Acknowledgments

I am grateful to Stuart C. Shapiro and the other members of the SNePS Research Group and to John Kearns, Justin Leiber, Ausonio Marras, James McGilvray, and Paul E. Ritty, and to the attendees of the SUNY Binghamton PACCS Master Class for comments on earlier versions of this essay.

References

Abelson, Robert P. (1968), “Simulation of Social Behavior”, in Gardner Lindzey and Elliot Aronson (eds.), *The Handbook of Social Psychology, 2nd edition* (Reading, MA: Addison-Wesley), Vol. 2, Ch. 12: 274–356.

Ackerman, Diane (1989), “Penguins”, *The New Yorker* (10 July 1989): 38–67.

Allen, Colin (1997), “Consciousness, Animal”, in Edward N. Zalta, ed., *Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/entries/consciousness-animal/>.

Allen, Colin, & Bekoff, Marc (eds.) (1997), *Species of Mind: The Philosophy and Biology of Cognitive Ethology* (Cambridge, MA: MIT Press).

Associated Press (1997), “Opponent bytes at offer for draw with Kasparov”, *Buffalo News*, 7 May 1997: A7.

Ayer, A. J. (1956), *The Problem of Knowledge* (Baltimore: Penguin).

Baum, L. Frank (1900), *The Wonderful Wizard of Oz* (New York: Dover, 1966 reprint).

Bunn, James H. (forthcoming), “Universal Grammar or Common Syntax? A Critical Study of Jackendoff’s *Patterns in the Mind*”, *Minds and Machines*.

Bruner, Jerome (1983), *Child’s Talk: Learning to Use Language* (New York: W. W. Norton).

Burge, Tyler (1979), “Individualism and the Mental”, in Peter A. French, Theodore E. Uehling, Jr., & Howard K. Wettstein (eds.), *Studies in Metaphysics*, Midwest Studies in Philosophy, Vol. 4 (Minneapolis: University of Minnesota Press): 73–122.

[32]Kasparov has tended to speak of Deep Blue (and its ancestor, Deep Thought) on occasion using “intentional-stance” terminology:

1. “Computers have their psychology too,” he [Kasparov] said. “If you know a computer well, you can anticipate its moves.” (Schonberg 1989: B2.)
2. “I knew enough to put the computer in misery,” boasts Kasparov. (Levy 1997: 54.)
3. “After beating the computer last year [1996], he [Kasparov] said it exhibited the stirrings of genuine thought: ‘I believe signs of intelligence can be found in the net result, not in the way the result is achieved,’ he said before this week’s rematch. ‘I don’t care how the machine gets there. It feels like thinking.’ ” (Johnson 1997.)
4. “Today I didn’t play well, but the computer was a computer and eventually it *knew* enough not to lose the game” (Associated Press 1997; my italics).

Carnap, Rudolf (1928), *The Logical Structure of the World*, Rolf A. George (trans.), (Berkeley: University of California Press, 1967).

Colby, Kenneth Mark (1981), "Modeling a Paranoid Mind", *Behavioral and Brain Sciences* 4: 515–560.

Colby, Kenneth Mark; Weber, Sylvia; & Hilf, Franklin Dennis (1971), "Artificial Paranoia", *Artificial Intelligence* 2: 1–25.

Colby, Kenneth Mark; Hilf, Franklin Dennis; Weber, Sylvia; & Kraemer, Helena C. (1972), "Turing-Like Indistinguishability Tests for the Validation of a Computer Simulation of Paranoid Processes", *Artificial Intelligence* 3: 199–221.

Damasio, Antonio R. (1989a), "Time-Locked Multiregional Retroactivation: A Systems-Level Proposal for the Neural Substrates of Recall and Recognition", *Cognition* 33: 25–62.

Damasio, Antonio R. (1989b), "Concepts in the Brain," in "Forum: What is a Concept?", *Mind and Language* 4: 24–27.

Dennett, Daniel C. (1971), "Intentional Systems", *Journal of Philosophy* 68: 87–106; reprinted in Daniel C. Dennett, *Brainstorms* (Montgomery, VT: Bradford Books): 3–22.

Ehrlich, Karen (1995), "Automatic Vocabulary Expansion through Narrative Context", *Technical Report 95-09* (Buffalo: SUNY Buffalo Department of Computer Science).

Ehrlich, Karen, & Rapaport, William J. (1997), "A Computational Theory of Vocabulary Expansion", *Proceedings of the 19th Annual Conference of the Cognitive Science Society (Stanford University)* (Mahwah, NJ: Lawrence Erlbaum Associates): 205–210.

Elliott, Clark, & Brzezinski, Jacek (1998), "Autonomous Agents as Synthetic Characters", *AI Magazine* 19(2): 13–30.

Erion, Gerald J. (forthcoming), *Ontology, Theory, and Ethics for Common Sense* (tentative title), Ph.D. dissertation, Department of Philosophy, SUNY Buffalo.

Fetzer, James H. (1994), "Mental Algorithms: Are Minds Computational Systems?", *Pragmatics and Cognition* 2: 1–29.

Fodor, Jerry A. (1980), "Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology", *Behavioral and Brain Sciences* 3: 63–109.

Fodor, Jerry, & Lepore, Ernest (1991), "Why Meaning (Probably) Isn't Conceptual Role," *Mind and Language* 6: 328–343.

Fodor, Jerry, & Lepore, Ernest (1992), *Holism: A Shopper's Guide* (Cambridge, MA: Basil Blackwell).

Gettier, Edmund L. (1963), "Is Justified True Belief Knowledge?", *Analysis* 23: 121–23; reprinted in A. Phillips Griffiths (ed.), *Knowledge and Belief* (Oxford: Oxford University Press, 1967).

Griffin, Donald R. (1976), *The Question of Animal Awareness: Evolutionary Continuity of Mental Experience*, New York: Rockefeller University Press.

Harman, Gilbert (1982), "Conceptual Role Semantics", *Notre Dame Journal of Formal Logic* 23: 242–256.

Harman, Gilbert (1987), "(Nonsolipsistic) Conceptual Role Semantics", in Ernest Lepore (ed.), *New Directions in Semantics* (London: Academic Press): 55–81.

Harman, Gilbert (1988), "Wide Functionalism," in Stephen Schiffer & Susan Steele (eds.), *Cognition and Representation* (Boulder, CO: Westview Press): 11–18.

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

Hauser, Larry (1996), "The Chinese Room Argument", in James Fieser (ed.), *The Internet Encyclopedia of Philosophy*, <http://www.utm.edu/research/iep/c/chineser.htm>.

Hill, Robin K. (1994), "Issues of Semantics in a Semantic-Network Representation of Belief", *Technical Report 94-11* (Buffalo: SUNY Buffalo Department of Computer Science).

Hill, Robin K. (1995), "Non-Well-Founded Set Theory and the Circular Semantics of Semantic Networks", in E. A. Yfantis (ed.), *Intelligent Systems: 3rd Golden West International Conference* (Dordrecht, The Netherlands: Kluwer Academic Publishers): 375–386.

Hirst, Graeme; McRoy, Susan; Heeman, Peter; Edmonds, Philip; & Horton, Diane (1994), "Repairing Conversational Misunderstandings and Non-Understandings", *Speech Communication* 15: 213–229.

Hutchins, Edwin (1995a), *Cognition in the Wild* (Cambridge, MA: MIT Press).

Hutchins, Edwin (1995b), "How a Cockpit Remembers Its Speeds", *Cognitive Science* 19: 265–288.

Jackendoff, Ray (forthcoming), "Bringing *Patterns* into Focus: A Response to Bunn", *Minds and Machines*.

Jennings, Richard C. (1985), "Translation, Interpretation and Understanding", paper read at the American Philosophical Association Eastern Division (Washington, DC); abstract, *Proceedings and Addresses of the American Philosophical Association* 59 (1985): 345–346.

Johnson, George (1997), "Ghost in the Chess Machine: Brain or Box? Think About It", *The New York Times* (9 May 1997): A1, B4
Julesz, Bela (1971), *Foundations of Cyclopean Perception* (Chicago: University of Chicago Press).
Kearns, John (1996), *Reconceiving Experience: A Solution to a Problem Inherited from Descartes* (Albany: SUNY Press).
Kearns, John (1997), "Propositional Logic of Supposition and Assertion", *Notre Dame Journal of Formal Logic* 38: 325–349.
Lakoff, George, & Johnson, Mark (1980), *Metaphors We Live By* (Chicago: University of Chicago Press).
Lassègue, Jean (1996), 'What Kind of Turing Test Did Turing Have in Mind?', *Tekhnema: Journal of Philosophy and Technology* 3: <http://www.gold.ac.uk/tekhnema/3/lassegue/read01.html>.
Leiber, Justin (1996), "Helen Keller as Cognitive Scientist", *Philosophical Psychology* 9: 419–439.
Lenat, Douglas B., & Feigenbaum, Edward A. (1991), "On the Thresholds of Knowledge," *Artificial Intelligence* 47: 185–250.
Lepore, Ernest, & Loewer, Barry (1987), "Dual Aspect Semantics", in Ernest Lepore (ed.), *New Directions in Semantics* (Orlando, FL: Academic Press): 83–112.
Levy, Steven (1997), "Man v. Machine", *Newsweek* (5 May 1997): 51–56.
Loar, Brian (1982), "Conceptual Role and Truth-Conditions", *Notre Dame Journal of Formal Logic* 23: 272–283.
Loebner, Hugh Gene (1994), "In Response [to Shieber 1994a]", *Communications of the ACM* 37(6): 79–82.
Loebner, Hugh Gene (1998), "Home Page of the Loebner Prize—'The First Turing Test' ", <http://www.loebner.net/Prizef/loebner-prize.html>.
Loewer, Barry (1982), "The Role of 'Conceptual Role Semantics' ", *Notre Dame Journal of Formal Logic* 23: 305–315.
Maida, Anthony S., & Shapiro, Stuart C. (1982), "Intensional Concepts in Propositional Semantic Networks", *Cognitive Science* 6: 291–330; reprinted in Ronald J. Brachman & Hector J. Levesque (eds.), *Readings in Knowledge Representation* (Los Altos, CA: Morgan Kaufmann, 1985): 169–189.
Maloney, J. Christopher (1987), "The Right Stuff", *Synthese* 70: 349–372.
Maloney, J. Christopher (1989), *The Mundane Matter of the Mental Language* (Cambridge, U.K.: Cambridge University Press).
Maloney, J. Christopher (1998), "A Role for Conceptual Role Semantics", in Francesco Orilia & William J. Rapaport (eds.), *Thought, Language, and Ontology: Essays in Memory of Hector-Neri Castañeda* (Dordrecht: Kluwer Academic Publishers, 1998): 169–180.
McGilvray, James (1998), "Meanings Are Syntactically Individuated and Found in the Head", *Mind and Language* 13: 225–280.
McRoy, Susan, & Hirst, Graeme (1993), "Misunderstanding and the Negotiation of Meaning", *AAAI Fall Symposium on Human-Computer Collaboration (Raleigh, NC)* (Menlo Park, CA: AAAI): 57–62.
Minsky, Marvin L. (1965), "Matter, Mind, and Models", in Marvin L. Minsky (ed.), *Semantic Information Processing* (Cambridge, MA: MIT Press, 1968): 425–432.
Morris, Charles (1938), *Foundations of the Theory of Signs* (Chicago: University of Chicago Press).
Percy, Walker (1975a), "The Delta Factor: How I Discovered the Delta Factor Sitting at My Desk One Summer Day in Louisiana in the 1950's Thinking about an Event in the Life of Helen Keller on Another Summer Day in Alabama in 1887", in Walker Percy, *The Message in the Bottle: How Queer Man Is, How Queer Language Is, and What One Has to Do with the Other* (New York: Farrar, Straus and Giroux), Ch. 1.
Percy, Walker (1975b), "Semiotic and a Theory of Knowledge", in Walker Percy, *The Message in the Bottle: How Queer Man Is, How Queer Language Is, and What One Has to Do with the Other* (New York: Farrar, Straus and Giroux), Ch. 11.
Posner, Roland (1992), "Origins and Development of Contemporary Syntactics", *Languages of Design* 1: 37–50.
Putnam, Hilary (1975), "The Meaning of 'Meaning' ", reprinted in *Mind, Language and Reality* (Cambridge, UK: Cambridge University Press): 215–271.
Quillian, M. Ross (1967), "Word Concepts: A Theory and Simulation of Some Basic Semantic Capabilities", *Behavioral Science* 12: 410–430; reprinted in Ronald J. Brachman & Hector J. Levesque (eds.), *Readings in Knowledge Representation* (Los Altos, CA: Morgan Kaufmann, 1985): 97–118.
Quillian, M. Ross (1968), "Semantic Memory", in Marvin Minsky (ed.), *Semantic Information Processing* (Cambridge, MA: MIT Press): 227–270.

Quine, Willard Van Orman (1951), "Two Dogmas of Empiricism", reprinted in *From a Logical Point of View, 2nd Edition, Revised* (Cambridge, MA: Harvard University Press, 1980): 20–46.

Quine, Willard Van Orman (1960), *Word and Object* (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. (1981), "How to Make the World Fit Our Language: An Essay in Meinongian Semantics", *Grazer Philosophische Studien* 14: 1–21.

Rapaport, William J. (1985), "Machine Understanding and Data Abstraction in Searle's Chinese Room", *Proceedings of the 7th Annual Conference of the Cognitive Science Society (University of California at Irvine)* (Hillsdale, NJ: Lawrence Erlbaum Associates): 341–345.

Rapaport, William J. (1986a), "Philosophy, Artificial Intelligence, and the Chinese-Room Argument", *Abacus* 3 (Summer 1986) 6–17; correspondence, *Abacus* 4 (Winter 1987) 6–7, *Abacus* 4 (Spring 1987) 5–7.

Rapaport, William J. (1986b), "Searle's Experiments with Thought," *Philosophy of Science* 53: 271–279; preprinted as *Technical Report 216* (Buffalo: SUNY Buffalo Department of Computer Science, 1984).

Rapaport, William J. (1988a), "To Think or Not to Think", *Noûs* 22: 585–609.

Rapaport, William J. (1988b), "Syntactic Semantics: Foundations of Computational Natural-Language Understanding", in James H. Fetzer (ed.), *Aspects of Artificial Intelligence* (Dordrecht, The Netherlands: Kluwer Academic Publishers): 81–131; reprinted (with numerous errors) in Eric Dietrich (ed.), *Thinking Computers and Virtual Persons: Essays on the Intentionality of Machines* (San Diego: Academic Press, 1994): 225–273.

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

Rapaport, William J. (1996), *Understanding Understanding: Semantics, Computation, and Cognition*, *Technical Report 96-26* (Buffalo: SUNY Buffalo Department of Computer Science); <ftp://ftp.cse.buffalo.edu/pub/tech-reports/96-26.ps.Z>; <http://www.cse.buffalo.edu/~rapaport/Papers/book.ps>

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; longer version preprinted as *Technical Report 97-15* (Buffalo: SUNY Buffalo Department of Computer Science) and *Technical Report 97-5* (Buffalo: SUNY Buffalo Center for Cognitive Science): <ftp://ftp.cse.buffalo.edu/pub/tech-reports/97-15.ps.Z>, <http://www.cse.buffalo.edu/~rapaport/Papers/implementation.ps|pdf>

Rapaport, William J., & Ehrlich, Karen (1999), "A Computational Theory of Vocabulary Acquisition", in Łucja 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); pre-printed as *Technical Report* 98-05 (Buffalo: SUNY Buffalo Department of Computer Science) and *Technical Report* 98-1 (Buffalo: SUNY Buffalo Center for Cognitive Science); <ftp://ftp.cse.buffalo.edu/pub/tech-reports/98-05.ps.Z>; <http://www.cse.buffalo.edu/~rapaport/Papers/vocab.cogsci.tr.ps>

Rapaport, William J.; Shapiro, Stuart C.; & Wiebe, Janyce M. (1997), "Quasi-Indexicals and Knowledge Reports", *Cognitive Science* 21: 63–107; reprinted in Francesco Orilia & William J. Rapaport (eds.), *Thought, Language, and Ontology: Essays in Memory of Hector-Neri Castañeda* (Dordrecht: Kluwer Academic Publishers, 1998): 235–294.

Rees, Roderick; Crawford, Charles; Bowyer, Adrian; Schubert, Len; & Loebner, Hugh (1994), "Forum: Notes on the Turing Test", *Communications of the ACM* 37(9): 13–15.

Schonberg, Harold C. (1989), "Kasparov Beats Chess Computer (for Now)", *New York Times* (23 October 1989): A1, B2.

Searle, John R. (1980), "Minds, Brains, and Programs", *Behavioral and Brain Sciences* 3: 417–457.

Searle, John R. (1982), "The Myth of the Computer", *New York Review of Books* (29 April 1982): 3–6; cf. correspondence, same journal (24 June 1982): 56–57.

Searle, John R. (1984), *Minds, Brains and Science* (Cambridge, MA: Harvard University Press).

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

Searle, John R. (1993), "The Failures of Computationalism", *Think* (Tilburg, The Netherlands: Tilburg University Institute for Language Technology and Artificial Intelligence) 2 (June 1993): 68–71.

Sellars, Wilfrid (1963), *Science, Perception and Reality* (London: Routledge & Kegan Paul).

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. (1995), "Computationalism", *Minds and Machines* 5: 517–524.

Shapiro, Stuart C. (1997), "What Is Computer Science?", <http://www.cse.buffalo.edu/∼shapiro/Papers/whatiscs.pdf>.

Shapiro, Stuart C. (1998), "Embodied Cassie", *Working Notes of the AAAI 1998 Fall Symposium on Cognitive Robotics, Orlando, FL*; <http://www.cse.buffalo.edu/∼shapiro/Papers/embodiedcassie.ps>.

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, 1986): 278–283; revised version of the 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.

Shapiro, Stuart C., & Rapaport, William J. (1991), "Models and Minds: Knowledge Representation for Natural-Language Competence", in Robert Cummins & John Pollock (eds.), *Philosophy and AI: Essays at the Interface* (Cambridge, MA: MIT Press): 215–259.

Shapiro, Stuart C., & Rapaport, William J. (1992), "The SNePS Family", *Computers and Mathematics with Applications* 23: 243–275; reprinted in Fritz 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 Narrative", 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.

Shieber, Stuart M. (1994a), "Lessons from a Restricted Turing Test", *Communications of the ACM* 37(6): 70–78.

Shieber, Stuart M. (1994b), "On Loebner's Lessons", *Communications of the ACM* 37(6): 83–84.

Simon, Herbert A., & Newell, Allen (1958), "Heuristic Problem Solving: The Next Advance in Operations Research", *Operations Research* 6: 1–10; reprinted in Herbert A. Simon, *Models of Bounded Rationality* (Cambridge, MA: MIT Press, 1982), Vol. 1, pp. 380–389.

Simpson, J. A., & Weiner, E. S. C. (preparers) (1989), *The Oxford English Dictionary, 2nd edition* (Oxford: Clarendon Press).

Smith, Brian Cantwell (1982), "Linguistic and Computational Semantics", *Proceedings of the 20th Annual Meeting of the Association for Computational Linguistics (University of Toronto)* (Morristown, NJ: Association for Computational Linguistics): 9–15.

Smith, Brian Cantwell (1987), "The Correspondence Continuum", *Report CSLI-87-71* (Stanford, CA: Center for the Study of Language and Information).

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., & 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.

Steels, Luc (1998), "The Origins of Syntax in Visually Grounded Robotic Agents", *Artificial Intelligence* 103: 133–156.

Stork, David G. (1997), *HAL's Legacy: 2001's Computer as Dream and Reality* (Cambridge, MA: MIT Press).

Terrace, Herbert S. (1985), "In the Beginning Was the 'Name'," *American Psychologist* 40: 1011–1028.

Thagard, Paul (1986), "Parallel Computation and the Mind-Body Problem", *Cognitive Science* 10: 301–318.

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.

Turing, Alan M. (1950), "Computing Machinery and Intelligence", *Mind* 59: 433–460; reprinted in Alan Ross Anderson (ed.), *Minds and Machines* (Englewood Cliffs, NJ: Prentice-Hall, 1964): 4–30, and in Allan Collins & Edward E. Smith (eds.), *Readings in Cognitive Science* (San Mateo, CA: Morgan Kaufmann, 1988): 6–19.

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

Weizenbaum, Joseph (1966), "ELIZA—A Computer Program for the Study of Natural Language Communication between Man and Machine", *Communications of the ACM* 9: 36–45; reprinted, with correspondence from *CACM* 17 (1974): 543, in *Communications of the ACM* 26(1) (1983): 22–28.

Weizenbaum, Joseph (1967), "Contextual Understanding by Computers", *Communications of the ACM* 10: 474–480.

Weizenbaum, Joseph (1976), *Computer Power and Human Reason: From Judgment to Calculation* (New York: W. H. Freeman).

Williams, Bernard (1998), "The End of Explanation?", *The New York Review of Books* 45(18) (19 November): 40–44.

Woods, William A. (1975), "What's in a Link: Foundations for Semantic Networks," in Daniel G. Bobrow & Allan M. Collins (eds.), *Representation and Understanding: Studies in Cognitive Science* (New York: Academic Press): 35–82; reprinted in Ronald J. Brachman & Hector J. Levesque (eds.) *Readings in Knowledge Representation* (Los Altos, CA: Morgan Kaufmann, 1985): 217–241.