Semantics as Syntax

William J. Rapaport

Department of Computer Science and Engineering,
Department of Philosophy, Department of Linguistics,
and Center for Cognitive Science
University at Buffalo, The State University of New York
Buffalo, NY 14260-2500
rapaport@buffalo.edu
http://www.cse.buffalo.edu/~rapaport/

August 16, 2016

Abstract

Let $S$, $T$, be non-empty sets. The syntax of $S$ (or $T$) is the set of properties of, and relations among, the members of $S$ (or $T$). The ontology of $T$ (or $S$) is its syntax. The semantic interpretation of $S$ by $T$ is a set of relations between $S$ and $T$. Semantics is the study of such relations between $S$ and $T$. Let $U = S \cup T$. Then the syntax of $U$ provides the semantics of $S$ in terms of $T$. Hence, semantics is syntax.
The Chinese room shows what we knew all along: syntax by itself is not sufficient for semantics. (Does anyone actually deny this point, I mean straight out? Is anyone actually willing to say, straight out, that they think that syntax, in the sense of formal symbols, is really the same as semantic content, in the sense of meanings, thought contents, understanding, etc.?)
[Searle, 1993, p. 68]

My thesis is that (suitable) purely syntactic symbol-manipulation of a computational natural-language-understanding system’s knowledge base suffices for it to understand natural language.
[Rapaport, 1988, pp. 85–86]

Does that make any sense? Yes: Everything makes sense. The question is: What sense does it make? Stuart C. Shapiro (in conversation, 19 April 1994)

1 Syntax vs. Semantics

Does syntax suffice for semantics? John Searle famously says that it does not. I have argued that it does. More precisely, I have argued that semantics is nothing but syntax. These slogans need to be cashed out.

1.1 Syntax

Let’s begin with syntax. The word ‘syntax’ has at least two meanings: a narrow or specific one, and a wide or general one. On the narrow (and perhaps more usual) meaning, the syntax of a language is its grammar, and the syntax of a logic is its proof theory.

The wide meaning, which is the one I want to focus on, includes both narrow meanings, but goes beyond them. It is roughly synonymous with Charles Morris’s ‘syntactics’: “the formal relation of signs to one another . . . in abstraction from the relations of signs to objects or to interpreters” [Morris, 1938, pp. 6, 13]. (The former relations are those of semantics, the latter are those of pragmatics.)

On the wide view, syntax is the study of the properties of the “symbols” of an (uninterpreted) “symbol system” and the relations among them, including how the symbols can be “manipulated”. But ‘symbol’ is a charged term, used by some to mean an interpreted “sign”: a sign together with its meaning. Worse, ‘sign’ is yet another charged term, because signs are supposed to be signs of something. I want to focus on the “sign” or “symbol” itself, devoid of any meaning, so I will use the more neutral terms ‘mark’, ‘mark system’, and ‘mark manipulation’ instead of the more familiar ‘sign’, ‘symbol’, ‘symbol system’, and ‘symbol manipulation’.

This is a kind of very “pure” syntax. Dale Jacquette does not believe in the existence of such “pure syntax … entirely divorced from semantics”. But all that he says in defense is that such marks “lack even derivative meaning or intentionality” [Jacquette, 1990, p. 294]. (‘Intentionality’, by the way, seems to have two different, albeit related, meanings in the literature. In a technical sense deriving from Brentano, it means “directedness to an object”; in the sense in which Jacquette, Searle, and others use it in the context of the Chinese Room Argument, it seems to be roughly synonymous with ‘cognition’, ‘understanding’, or even ‘consciousness’.) Jacquette goes on to say that even purely syntactic “computer programs … are always
externally interpreted” (p. 295, my italics). I agree with the latter comment, but I still think that there is such a thing as pure syntax in the sense that I am using it here. But this debate would take us too far astray.4

The (purely) syntactic properties of marks include their “shape” (what a mark looks like and how it differs from other marks in the system), an inventory of the marks of the system (an “alphabet” or “vocabulary” of “primitive”, or “basic”, marks), and relations spelling out how marks may be combined to form more complex ones from simpler ones (usually given by recursive rules that take primitive, or given, or “atomic” marks as the base case, and show how to combine them to produce “molecular” marks, or “well-formed formulas” [wffs]). This much would normally be called the “grammar” of the system if the system were to be thought of as a language.

As I noted above, some mark systems, especially those that really are languages (formal or otherwise) might have other syntactic properties and relations, in addition to shape, inventory, and grammatical combinatory relations. For instance, some molecular marks—well-formed formulas—might be taken as axioms. And some sets of molecular marks might stand in certain relations such that, whenever some of them have been collected together, others—called ‘theorems’—might then be “legally” allowed to be added to the collection. Here what I have in mind are transformation rules or rules of inference. Thus, in addition to “grammatical” syntax, a set might also have a “logical” or “proof-theoretic” syntax. The production of molecular wffs and theorems is usually what is meant by ‘symbol manipulation’ (i.e., mark manipulation). Note, however, that I am not requiring the transformation rules to be “truth preserving”, because I take “truth” to be a semantic property. (I take correspondence theories of truth to be semantic. Coherence theories seem to be more syntactic, or holistic. We’ll come back to holism in §3.1.)

But I want to be even more general than Morris: I see no reason not to include systems that might not normally be considered to be languages. On my view, any set of objects has a syntax if the objects have properties and stand in relations to each other. On this view, even biological neural networks have syntax [Rapaport, 2012, p. 38], as does the world itself.

1.2 Semantics

Semantics, of course, is the study of meaning and truth. The meaning of some piece of language (a word, phrase, sentence, what have you) might be its referent (if there is one), i.e., something else, in the world, rather than in the language (with the exception, of course, of things like names of words, whose referents are other words in the language—after all, language is part of the world). Or the meaning of some piece of language might be its sense (Fregean or otherwise)—again, something else, outside of the language, though not, perhaps, “in the world”. Or it might be a Meinongian object or an idea in a mind.5 But, in any case, the meaning of a piece of language is not typically thought of as being part of the language; rather, it is something else that the piece of language stands in relation to. One exception is conceptual-role or holistic theories of meaning, but we’ll come back to that in §3.1.

So, whereas syntax only requires one domain (call it the “syntactic domain”), semantics requires two: a syntactic domain and a “semantic domain”. The syntactic domain is the thing that needs to be understood, to be interpreted. The semantic domain is the thing that provides the understanding, that provides the interpretation.

Following [Morris, 1938, p. 21], then, I take semantics to be the study of the relations between the marks of two systems. Because syntax is the study of the properties and relations of a single system, it would seem

---

4I discuss these issues in the context of the relation of computer programs to the world in [Rapaport, 2015].
that, indeed, syntax does not suffice for semantics. Yet I argue that it does. Let’s look into this more closely.\(^6\)

## 2 Two Syntactic Systems

### 2.1 The Syntax of \(L\)

On the standard view, a syntactic domain is usually some (formal or formalized) language \(L\), which is described syntactically—that is, in terms of its marks and rules for manipulating them. Thus, for instance, \(L\) might be described as having terms, perhaps of two (simple, or atomic) kinds: individual constants \(a, b, \ldots\) (e.g., proper names or other nouns) and individual variables \(u, v, \ldots\) (e.g., pronouns). “New” (complex, or molecular) terms (e.g., noun phrases) can be constructed from previously given or previously constructed (“old”) ones (whether atomic or molecular) by means of function symbols of various arities, \(f, g, \ldots, f_i, \ldots\) (e.g., ‘the father of . . .’, ‘the average of . . . and . . .’), together with “grammar” rules specifying the “legal” structure (or “spellings”) of such molecular terms (say, if \(t_1, \ldots, t_n\) are terms, and \(f^n\) is an \(n\)-place function symbol, then \(\llbracket f^n(t_1, \ldots, t_n) \rrbracket\) is a term).

In addition, \(L\) will have predicate symbols of various arities: \(A, \ldots, Z, A_i, \ldots\) (e.g., verb phrases); connectives and quantifiers: \(\neg, \lor, \forall, \exists\ldots\) (e.g., ‘it is not the case that . . .’, ‘. . . or . . .’, ‘for all . . ., it is the case that . . .’); and more “grammar” rules specifying the “legal” structure of well-formed formulas (or sentences): If \(t_1, \ldots, t_n\) are terms, and \(P^n\) is an \(n\)-place predicate symbol, then \(\llbracket P^n(t_1, \ldots, t_n) \rrbracket\) is a well-formed formula (wff); if \(\varphi\) and \(\psi\) are wffs, and \(v\) is an individual variable, then \(\llbracket \neg \varphi \rrbracket, \llbracket (\varphi \lor \psi) \rrbracket, \llbracket \forall v [\varphi] \rrbracket\) are wffs.

Note that \(L\) is a language. Sometimes \(L\) is augmented with a logic: Certain wffs of \(L\) are distinguished as axioms (or “primitive theorems”), and rules of inference are provided that specify how to produce “new” theorems from “old” ones. For instance, if \(\varphi\) and \(\llbracket (\varphi \rightarrow \psi) \rrbracket\) are theorems, then so is \(\psi\). A proof of a wff \(\psi\) (from a set of wffs \(\Sigma\)) is a sequence of wffs ending with \(\psi\) such that every wff in the sequence is either an axiom (or a member of \(\Sigma\)) or follows from previous wffs in the sequence by one of the rules of inference.

And so on. I will assume that the reader is familiar with the general pattern.\(^7\) The point is that all we have so far are marks and rules for manipulating them either linguistically (to form wffs) or logically (to form theorems). All we have so far is syntax in Morris’s sense.

Actually, in my desire to make the example perspicuous, I may have given you a misleading impression by talking of “language” and “logic”, of “nouns” and “verb phrases”, etc. For such talk tends to make people think either that I was talking, albeit in a very strange way, about language and nouns and verbs—good old familiar languages like English with nouns and verbs like ‘dog’ and ‘run’—or that I had that in the back of my mind as an intended interpretation of the marks and rules. But marks are merely (perhaps) physical inscriptions or sounds that have only some very minimal features such as having distinguished, relatively unchanging shapes capable of being recognized when encountered again.

\(^{6}\)Much of what follows is a detailed elaboration of comments I made in [Rapaport, 1995, §2.2].

\(^{7}\)See, e.g., [Rapaport, 1992b], [Rapaport, 1992a] for more details.

\(^{8}\)Or ‘dog’ (plural: ‘dogs’) and ‘(to) dog’: ‘Dogs dogs dog dogs’ is syntactically correct and semantically meaningful (if not pragmatically acceptable) as a sentence of English, meaning “Dogs—whom other dogs follow—follow other dogs.” Or ‘buffalo’ (plural: ‘buffalo’(!)) and ‘(to) buffalo’ (meaning “(to) intimidate”): ‘Buffalo buffalo buffalo buffalo buffalo’ is likewise syntactically correct and semantically meaningful (but only pragmatically acceptable in Buffalo, NY); see http://www.cse.buffalo.edu/~rapaport/BuffaloBuffalo/buffalobuffalo.html
2.2 The Syntax of $L'$

So, let me offer a somewhat less familiar syntactic domain $L'$, which I will call, this time, not a “language”, but merely a “mark system”. First, I need to show you the marks of $L'$. To really make my point, these should be quite arbitrary: say, boxes, circles, or squiggles of various kinds. But I will make life a bit easier for the reader and the typesetter by using letters and numerals.

$L'$ consists of the following marks:

\[
A_1, \ldots, A_i, \ldots; \\
F_0, F_1, F_2, F_3; \\
(, ), .. ; \\
\text{[i.e., a left-parenthesis, a right-parenthesis, a comma, and a semi-colon]}
\]

I want to show you a certain class $K$ of marks of $L'$. To talk about them, I’ll need another set of marks that are not part of $L'$, so we’ll let ‘$A$’, ‘$B$’, ‘$C$’, ‘$B_1$’, ‘$B_2$’, . . . be variables ranging over the members of $K$. Now, here are the members of $K$:

1. $A_1, \ldots, A_i, \ldots \in K$
2. If $A, B \in K$, then $[F_0(A)]$, $[F_1(A, B)]$, $[F_2(A, B)]$, $[F_3(A, B)] \in K$.
3. Nothing else is in $K$.

We could ask questions of this formal mark system. For instance, which molecular marks are in $K$? By suitable mark manipulation, following (1)–(3), we can ascertain that $A_1, A_{100}, F_0(A_{100}), F_0(F_0(A_{100})), F_3(F_0(F_0(A_{100})), F_2(A_1, A_{100})) \in K$, but that $F_0(F_6), B \notin K$.

Now, let’s make $L'$ a bit more interesting. Let $H \subseteq K$; let $A, B \in K$; and let’s say that an $(H, A)$-sequence is a sequence of members of $K$ such that $A$ is the last item in the sequence, and, if $B$ is in the sequence, then either $B \in H$ or there is a set $\{B_1, \ldots, B_n \mid (\forall 1 \leq i \leq n)[B_i \in K]\}$ such that $\{R(B_1, \ldots, B_n; B)\} \in \mathcal{R}$, where $\mathcal{R}$ is defined as follows (remember that ‘$R$’ is a mark of $L'$; I am defining $\mathcal{R}$ as consisting of certain sequences of marks beginning with ‘$R$’):

\begin{align*}
\mathcal{R}_1. & \quad R(A; F_1(A, B)) \in \mathcal{R} \\
\mathcal{R}_2. & \quad R(B; F_1(A, B)) \in \mathcal{R} \\
\mathcal{R}_3. & \quad R(F_1(A, B), F_0(A); B) \in \mathcal{R} \\
\mathcal{R}_4. & \quad R(F_1(A, B), F_0(B); A) \in \mathcal{R} \\
\mathcal{R}_5. & \quad R(F_2(A, B); A) \in \mathcal{R} \\
\mathcal{R}_6. & \quad R(F_2(A, B); B) \in \mathcal{R} \\
\mathcal{R}_7. & \quad R(A, B; F_2(A, B)) \in \mathcal{R} \\
\mathcal{R}_8. & \quad R(F_3(A, B), A; B) \in \mathcal{R} \\
\mathcal{R}_9. & \quad \text{If there is an $(H, B)$-sequence whose first item is $A$,} \\
& \text{then $R(; F_3(A, B)) \in \mathcal{R}$} \\
& \text{[Note: There is no symbol between ‘$(‘ and ‘$;‘].}
\end{align*}
\begin{itemize}
    \item If there is an \((H, F)\)-sequence whose first item is \(A\),
    \(\mathcal{R} \ni [R(;F_0(A))] \equiv [R(A)] \in \mathcal{R}\).
    \item If there is an \((H, F)\)-sequence whose first item is \(F_0(A)\),
    \(\mathcal{R} \ni [R(A)] \equiv [R(\mathcal{A})] \in \mathcal{R}\).
    \item Nothing else is in \(\mathcal{R}\).
\end{itemize}

We can now ask more questions of our system; e.g., which marks \(A\) are such that \([R(A)] \in \mathcal{R}\)? By suitable mark manipulations, following \(\mathcal{R} 1–\mathcal{R} 12\), we can ascertain that, e.g., \([R(F_3(A_0,A_0))] \in \mathcal{R}\) (this is actually fairly trivial, since \(A_0\) is an \((A_0,A_0)\)-sequence whose first item is \(A_0\)).

Hard to read, isn’t it! You feel the strong desire to try to understand these squiggles, don’t you? (Are you, perhaps, beginning to feel like Searle-in-the-Chinese-Room?) You would probably feel better if I showed you some other domain—a semantic domain—with which you were more comfortable, more familiar, into which you could map these squiggles. I will. But not yet.

Of course, I could be sadistic and suggest that you “get used to” \(\mathcal{L}'\) by manipulating its symbols and learning more about the members of \(\mathcal{K}\) and \(\mathcal{R}\). After all, as John von Neumann allegedly said, “in mathematics you don’t understand things. You just get used to them.”

To understand a syntactic domain \(S\) is either:

1. to “get used to” \(S\), or else
2. to understand \(S\) in terms of a semantic domain \(T\).

The latter is semantic understanding in Morris’s sense: understanding one thing in terms of something else. The former is what I have called “syntactic understanding” [Rapaport, 1986c]: understanding something in terms of itself. And in the case of semantic understanding, how do you understand the semantic domain \(T\)? Normally, \(T\) is assumed to be antecedently understood. But that has to mean that it is understood syntactically—you have gotten used to it. If \(T\) is not antecedently understood, then it has to be considered as a syntactic domain in its own right and understood in terms of yet another semantic domain \(T'\). And so on. Ultimately, I claim, all understanding is syntactic understanding (the base case of the recursion).

But “syntactic understanding”—the sort of thing that you come to have by getting used to the syntactic domain—does not seem, on the surface, to be any kind of “real” understanding. This is the intuition underlying Searle’s Chinese Room Argument and its earlier incarnation in the guise of Leibniz’s mill.

Where is the meaning or understanding (or “intentionality” or “consciousness”) in this kind of “meaningless” (yet rule-based, or regulated) mark manipulation? As I read him, [Jacquette, 1990, p. 293] suggests that it can generate understanding, in turn suggesting that what we have here is a clash of

---

10See [Rapaport, 1995] for further discussion.
11“Imagine there were a machine whose structure produced thought, feeling, and perception; we can conceive of its being enlarged while maintaining the same relative proportions among its parts, so that we could walk into it as we can walk into a mill. Suppose we do walk into it; all we would find there are cogs and levers and so on pushing one another, and never anything to account for a perception.” [Leibniz, 1714, §17]
12“... as we move closer to an exact microlevel decentralized ... input-output isomorphism with the neurological activity of natural intelligence, the intuition that we do not thereby also precisely duplicate the brain’s causal powers by which it produces intentionality begins to fade and lose its grip on our pre-theoretical beliefs. The imaginable causal efficacy of microlevel input-output functionalities raises difficulties about the adequacy of the Chinese Room example to support Searle’s thesis that a functioning program in which physical syntax tokens causally interact with themselves and a machine environment at the proper level of design could not produce intentionality just as effectively as the natural system it simulates.”
fundamental intuitions: Some (e.g., Searle) say that such mark systems and mark manipulation cannot suffice for understanding; others (perhaps Jacquette, and certainly I) say that it can.

In any case, you could just try to get used to $L'$ by doing mark manipulation, and I believe that you would thereby come to understand it. But I won’t be that mean. First, we need to move away from pure syntax and find out what semantics consists of.

2.3 A Semantic Interpretation of $L$

Given some syntactic domain—some (formal) mark system—one can ask two sorts of questions about it. The first sort is exemplified by those we asked above: What are the members of $K$? Of $R$? These are purely “internal”, syntactic, questions. The second sort is, in short: What’s the meaning of all this? What do the marks mean (if anything)? What, for example, is so special about the members of $K$ or the marks of the form $R(A)$? To answer this sort of question, we must go outside the syntactic domain: We must provide “external” entities that the marks mean (that they can be understood in terms of), and we must show the mappings—the associations, the correspondences—between the two domains.

But, as I have said elsewhere,

Now a curious thing happens: I need to show you the semantic domain. If I’m very lucky, I can just point it out to you—we can look at it together, and I can describe the correspondences (“The symbol $A_{37}$ means that red thing over there.”). But, more often, I have to describe the semantic domain to you in . . . symbols [i.e., marks], and hope that the meaning of those symbols will be obvious to you.13

[Rapaport, 1995, §2.2.2]

So, let’s provide a semantic interpretation of our first formal mark system, $L$. Since $L$ had individual terms, function marks, and predicate marks14—which could be combined in various (but not arbitrary) ways—I need to provide meanings for each such mark as well as for their legal combinations. So, we’ll need a non-empty set $D$ of things that the terms will mean—a Domain of interpretation (sometimes called a Domain, or universe, of discourse)—and sets $F$ and $R$ of things that the function and relation symbols will mean, respectively. These three sets can be collectively called $M$ (for Model). What’s in $D$? Well, anything you want to talk or think about. What are in $F$ and $R$? Functions and relations on $D$ of various arities—i.e., anything you want to be able to say about the things in $D$. That’s our ontology, what there is.

Let’s pause a moment here for an important point: $D$ has members; the members of $D$ have properties; and the members of $D$ stand in various relations to each other. The study of such objects, their properties, and the relations among them is ontology. But I have defined the study of objects, their properties, and the relations among them to be syntax. Thus, ontology is simply the syntax of the semantic domain.

Now for the correspondences. To say what a mark of $L$ means in $M$ (what the meaning, located in $M$, of a mark of $L$ is), we can define an interpretation function $I: L \rightarrow M$ that will assign to each mark of $L$ something in $M$ (or it might be an interpretation relation if we wish to allow for ambiguity), as follows:

1. If $t$ is an individual term of $L$, then $I(t) \in D$.

(Which element of $D$? Whichever you want, or, if we spell out $L$ and $D$ in more detail, I’ll tell you; for example, perhaps $I(‘Barack Obama’) =$ the 44th President of the U.S., if ‘Barack Obama’ is an individual constant of $L$, and $D$ is the set of humans.)

13For further discussion of this problem, see [Smith, 1987], [Rapaport, 1995, §2.5].
14I.e., what are normally called ‘function symbols’ and ‘predicate symbols’.
2. If \( f \) is a function symbol of \( \mathcal{L} \), then \( I(f) \in \mathbf{F} \).

3. If \( f(t_1, \ldots, t_n) \) is a (molecular) term of \( \mathcal{L} \), then \( I(f(t_1, \ldots, t_n)) = I(f)(I(t_1), \ldots, I(t_n)) \in \mathbf{D} \).

(I.e., the interpretation of \( f(t_1, \ldots, t_n) \) will be the result of applying (a) the function that is the interpretation of \( f \) to (b) the elements of \( \mathbf{D} \) that are the interpretations of the \( t_i \); and the result will be an element of \( \mathbf{D} \).)

4. If \( P \) is a predicate symbol of \( \mathcal{L} \), then \( I(P) \in \mathbf{R} \).

So far, so good. Now, what do wffs mean? Those philosophers and logicians who take \( \text{wff} \) would be a kind of syntactic understanding \[ [\text{Chang and Keisler, 1973, pp. 4ff}], \] but that’s not very interesting or useful for \( \mathbf{M} \). In this ideal situation, when one studies, not isolated or made-up sentences, but just one item in \( \mathbf{M} \) is an isomorphism—a 1–1 and onto homomorphism; that is, every item in \( \mathbf{M} \) is the meaning of just one symbol of \( \mathcal{L} \). (Being “onto” is tantamount to \( \mathcal{L} \)’s being “complete”.) Perhaps isomorphism is less than ideal, at least for the case of natural languages. David P. Wilkins has observed that when one studies, not isolated or made-up sentences, but ...

... real, contextualised utterances ... it is often the case that all the elements that one would want to propose as belonging to semantic structure have no overt manifestations in syntactic structure. ... [The degree of isomorphism between semantic and syntactic structure is mediated by pragmatic and functional concerns....] [Wilkins, 1995, p. 381]

In this ideal situation, \( \mathbf{M} \) is a virtual duplicate or mirror image of \( \mathcal{L} \). (Indeed, \( \mathbf{M} \) could be \( \mathcal{L} \) itself [Chang and Keisler, 1973, pp. 4ff], but that’s not very interesting or useful for semantic understanding of \( \mathcal{L} \); rather, it would be a kind of syntactic understanding!)

In less ideal circumstances, there might be marks of \( \mathcal{L} \) that are not interpretable in \( \mathbf{M} \); in that case, \( I \) would be a partial function. Such is the case when \( \mathcal{L} \) is English and \( \mathbf{M} \) is the world (‘unicorn’ is an English word, but unicorns don’t exist), though if we “enlarge” or “extend” \( \mathbf{M} \) in some way—e.g., if we take \( \mathbf{M} \) to be Meinong’s \textit{Ausserein} instead of the actual world—then we can make \( I \) total [Rapaport, 1981].
In another less ideal circumstance, “Hamlet’s Law” might hold: There are more things in M than in L; i.e., there are elements of M not expressible in L: I is not onto. And, as noted earlier, I might be a relation, not a function, so L would be ambiguous. There is another, more global, sense in which L could be ambiguous: By choosing a different M (and a different I), we could give the marks of L entirely distinct meanings. Worse, the two Ms need not be isomorphic. (This can happen in at least two ways. First, the cardinalities of the two Ds could differ. Second, suppose L is a language for expressing mathematical group theory. Then M₁ could be an infinite cyclic group (e.g., the integers under addition), and M₂ could be M₁ × M₁, which, unlike M₁, has two disjoint subgroups (except for the identity).)

2.4 A Semantic Interpretation of L’

Let’s consider an example in detail; I’ll tell you what the marks of L’ mean. First, I need to show you M. To do that, I need to show you D: D will include the marks: ϕ₁, . . . , ϕᵢ, . . . (so, I’m explaining one set of marks in terms of another set of marks; be patient). D will also include these marks: ¬, ∨, ∧, →. Now I can tell you about K (in what follows, let Aᵢ be the i-th atomic marks of K, let ϕᵢ be the i-th atomic marks of D, and let A, B ∈ K):

\[
I(Aᵢ) = ϕᵢ \\
I(F₀) = ¬ \\
I(F₁) = ∨ \\
I(F₂) = ∧ \\
I(F₃) = →
\]

I(⌜F₀(A)⌝) = ¬I(A) 
I(⌜F₁(A, B)⌝) = I(A) ∨ I(B) 
I(⌜F₂(A, B)⌝) = I(A) ∧ I(B) 
I(⌜F₃(A, B)⌝) = I(A) → I(B)

I assume, of course, that you know what ‘¬’, ‘(I(A) → I(B))’, etc., are (namely, the negation sign, a material conditional wff, etc.). So, the elements of K are just wffs of propositional logic (as if you didn’t know)!

What about R? Well: I(R) = ⊢ (where ⊢ ∈ R and where R, of course, is part of M); i.e., R means the deducibility relation on wffs of propositional logic. So, the elements of R are rules of inference:

\[
I(⌜R(A; F₁(A, B))⌝) = A \vdash I(A ∨ B) \quad \text{(i.e., } ∨\text{-introduction)} \\
I(⌜R(B; F₁(A, B))⌝) = B \vdash I(A ∨ B) \quad \text{(i.e., } ∨\text{-introduction)} \\
I(⌜R(F₁(A, B); F₀(A); B)⌝) = (A ∨ B) \vdash ¬A \vdash B \quad \text{(i.e., } ∨\text{-elimination)} \\
I(⌜R(F₁(A, B); F₀(B); A)⌝) = (A ∨ (B)) \vdash ¬B \vdash A \quad \text{(i.e., } ∨\text{-elimination)} \\
I(⌜R(F₂(A, B); A)⌝) = (A ∧ B) \vdash A \quad \text{(i.e., } ∧\text{-elimination)} \\
I(⌜R(F₂(A, B); B)⌝) = (A ∧ B) \vdash B \quad \text{(i.e., } ∧\text{-elimination)} \\
I(⌜R(A; B; F₂(A, B))⌝) = A, B \vdash (A ∧ B) \quad \text{(i.e., } ∧\text{-introduction)} \\
I(⌜R(F₃(A, B); A; B)⌝) = (A → B), A \vdash B \quad \text{(i.e., } →\text{-elimination, or Modus Ponens)}
\]

Before we can finish interpreting R, I need to tell you what an (H, A)-sequence means: It is a proof of I(A) from hypotheses I(H) (where, to be absolutely precise, I should specify that, where H = {A, B, . . . } ⊆ K, I(H) = {I(A), I(B), . . . }). So:

15"There are more things in heaven and earth, Horatio, / Than are dreamt of in your philosophy." Hamlet (I, 5, ll. 167–168). This can also be interpreted as a summary of G"odel’s Incompleteness Theorem (where ‘dreamt of’ means “provable”).

16I am grateful to Nicolas Goodman for this example.
\( I(\mathcal{R}9) \) is:
if there is a proof of \( I(B) \in D \) from a set of hypotheses \( I(H) \) whose first line is \( I(A) \),
then \( \vdash [I(A) \rightarrow I(B)] \)  
(i.e., \( \rightarrow \)-introduction, or Conditional Proof)

\( I(\mathcal{R}10) \) is:
if there is a proof of \( [I(B) \land \neg I(B)] \) \( \triangledown \) from a set of hypotheses \( I(H) \) whose first line is \( I(A) \),
then \( \vdash \neg I(A) \)  
(i.e., \( \neg \)-introduction)

\( I(\mathcal{R}11) \) is:
if there is a proof of \( [I(B) \land \neg I(B)] \) \( \triangledown \) from a set of hypotheses \( I(H) \) whose first line is \( \neg I(A) \),
then \( \vdash I(A) \)  
(i.e., \( \neg \)-elimination)

So, now you know: \( \mathcal{L}' \) is just ordinary propositional logic in a weird notation. Of course, I could have told you what the marks of \( \mathcal{L}' \) mean in terms of a different model \( \mathcal{M}' \), where \( D' \) consists of states of affairs and Boolean operations on them. In that case, \( \mathcal{L}' \) just is ordinary propositional logic. That is, \( \mathcal{M} \) is itself a syntactic formal mark system (namely, \( \mathcal{L} \)) whose meaning can be given in terms of \( \mathcal{M}' \), but \( \mathcal{L}' \)'s meaning can be given either in terms of \( \mathcal{M} \) or in terms of \( \mathcal{M}' \).

There are several lessons to be learned from this. First, \( \mathcal{L}' \) is not a very “natural” mark system. Usually, when one presents the syntax of a formal mark system, one already has a semantic interpretation in mind, and one designs the syntax to “capture” that semantics: The syntax is a model—an implementation—of the semantics.\(^{17}\)

Second, it is possible and occasionally even useful to allow one formal syntactic system to be the semantic interpretation of another syntactic system. Of course, this is only useful if the interpreting syntactic system is antecedently understood. How? In terms of another domain with which we are antecedently familiar! So, in our example, the unfamiliar \( \mathcal{L}' \) was interpreted in terms of the more familiar \( \mathcal{M} \) (i.e., \( \mathcal{L} \)), which, in turn, was interpreted in terms of \( \mathcal{M}' \). And how is it that we understand what states of affairs in the world are? Well . . . we’ve just gotten used to them. (We’ll come back to this in §4.)

Finally, note that \( \mathcal{M} \) in our example is a sort of “swing” domain: It serves as the semantic domain relative to \( \mathcal{L}' \) and as the syntactic domain relative to \( \mathcal{M}' \). We can have a “chain” of domains, each of which except the first is a semantic domain for the one before it, and each of which except for the last is a syntactic domain for the one following it. To understand any domain in the chain, we must be able to understand the “next” one. How do we understand the last one? Syntactically.\(^{18}\)

### 3 Syntax Suffices for Semantics

#### 3.1 Syntactic Understanding

Let’s take stock. Given any (non-empty) set \( S \) of objects of any kind, the specification of the properties of \( S \)'s members and of the relations that they stand in to each other is the syntax of \( S \). These properties and relations may be of different kinds, so we might be able to identify a “grammatical” syntax of \( S \) as well as a “logical” or “proof-theoretic” syntax of \( S \). We can understand \( S \) in terms of its syntax by “getting used to” manipulating its members according to these properties and relations. This is syntactic understanding.\(^{19}\)

\(^{17}\)On the nature of implementation, see [Rapaport, 1999], [Rapaport, 2005].

\(^{18}\)For more on these “chains” and their possible components, see [Rapaport, 1995, §§2.3ff].

\(^{19}\)I explore this, with an example from elementary algebra, in [Rapaport, 1986c].
Syntactic understanding is holistic in the following way: The syntax of $S$ can be represented by a graph whose vertices are the members of $S$ and whose edges represent its properties and relations. Such a graph is often called a ‘semantic network’, and such networks have rightly been criticized as really being “syntactic” networks. (The Semantic Web is really a syntactic web.)\textsuperscript{20} In such a network, the “meaning” of any vertex is its location in the network—its relations to all other vertices in the network. This is conceptual-role semantics or semantic holism. Semantic holism is just more syntax.\textsuperscript{21}

### 3.2 Semantic Understanding

But there is another kind of understanding: semantic understanding. Here, we need another set, $T$, in terms of which we understand $S$. When we ask what $s \in S$ means, our answer is some $t \in T$. But $T$ will have its own syntax. As I noted earlier, I see no difference between the syntax of $T$, thus understood, and the ontology of $T$, though we tend to reserve the former term for languages and logics, and the latter term for the realms that those languages and logics describe or are “about”. Thus, if we are understanding $S$ in terms of $T$, we would speak of the syntax of $S$ and the ontology of $T$, but that is merely a manner of speaking.

Semantic understanding requires relations between $S$ and $T$: relations of meaning, reference, etc. But these relations are not among the (internal) relations of $S$’s syntax or $T$’s ontology. They connect $S$ and $T$, but are external to both.

### 3.3 Syntax Is Semantics

So, how can we talk about those semantic relations? We cannot use either $S$ or $T$ by themselves to talk about them, because the semantic interpretation function is not part of $S$ alone or of $T$ alone. But we can talk about them by taking the union of $S$ and $T$; call it $U$. (In earlier writings about computational theories of cognition, I have called this the “internalization” of the semantic domain into the syntactic domain [Rapaport, 2012, §3]. See §4.2, below.) What is the syntax of $U$? It consists, in part, of the inventory of properties of members of $U$. This includes all of the properties of members of $S$ and all of the properties of the members of $T$. It also consists, in part, of the inventory of relations among the members of $S$ and the relations among the members of $T$. But it also includes the semantic relations between the members of $S$ and the members of $T$.

So, it is the syntax of $U$ that enables us to talk about the semantics of $S$. Semantics is, thus, just more syntax—the syntax of the union of a syntactic domain and its semantic domain. QED

### 4 Implications

I will close with brief comments on two philosophical issues that can be illuminated by this theory.

#### 4.1 Twin Earth

Hilary Putnam has argued that, not only can “two terms . . . have the same extension and yet differ in intension”, but that “two terms can . . . differ in extension and have the same intension” [Putnam, 1973, p. 700, and, more famously, in [Putnam, 1975]]. The latter claim is intended to be surprising, because it is

\footnote{[Berners-Lee and Fischetti, 1999, p. 12; cf. pp. 184ff]. [Ceusters, 1975]; for discussion relevant to the present essay, see, esp., [Rapaport, 2006, p. 393] and also [Rapaport, 2012, pp. 41–42, 48].}

\footnote{I explore conceptual-role semantics and holism in more detail, and respond to [Fodor and Lepore, 1992]’s objections, in [Rapaport, 2002].}
typically held that intensions determine extensions. Putnam offers his Twin Earth thought experiment as a counterexample.

I do not want to rehearse these arguments here, but merely point out some similar issues and see what my semantics-as-syntax theory might have to say.

First, note that the fact that the intensions of ‘water’ (on Earth) and ‘water’ (on Twin Earth) might be identical yet their extensions (H$_2$O and XYZ, respectively) be different parallels a situation with computer programs: By way of an “intuition pump”, recall that intensions (and Fregean senses) are sometimes modeled as (computable) functions or algorithms. Now, it can be the case that a single algorithm with a given input can have different outputs depending on the context in which those algorithms are executed. For one example from the literature, an algorithm (recipe) for producing hollandaise sauce when executed on Earth will likely produce something quite different when executed on the Moon [Cleland, 1993]. (Strictly speaking, perhaps, the context should be taken as part of the input, so the algorithms will, in fact, have the same outputs if given exactly the same inputs. I discuss this in greater detail in [Rapaport, 2015].)

Second, the relation of a word to its intension is simply one kind of meaning. The relation of a word to its extension is another kind of meaning. (There is no such thing as “the” meaning of a word; to claim that “meanings ain’t in the head” [Putnam, 1973, p. 704] is highly misleading, because some of them are!) What I want to point out is that the relation of an intension to an extension is yet another kind of meaning: All three of these relations are semantic in my sense, because they are relations between two domains.

4.2 Computational Cognition

If semantics is nothing but syntax (albeit syntax writ large), how do we understand language? How might a computer understand language? Searle says that it can’t, and I suggested the intuition behind this in §2.2. But I also mentioned a way out in §3: via “internalization”. I have cashed this out in the series of essays cited in footnote 2, but I will summarize my position here.

What seems to be missing in the Chinese Room is “real” semantics—links to the external-world referents of the words and sentences of language. How does Searle-in-the-room (actually, Searle-in-the-room together with the instruction book!) know that the word ‘hamburger’ means “hamburger” (i.e., refers to hamburgers), or that a certain (Chinese) “squiggle” does?

According to the theory I presented above in §3, it would seem that an actual hamburger would somehow need to be “imported” into Searle-in-the-room’s instruction book (the computer program for natural-language understanding and generation). Instead, a representative of an actual hamburger is thus imported into Searle-in-the-room’s “mind” (or “semantic network”). The hamburger is “internalized”. In the case of a real human being, this representative is the end result of, say, the visual process of seeing a hamburger (or the olfactory process of smelling one, etc.), resulting in a “mental image” of a hamburger. (To speak with Kant, it is an “intuition” or concept of a hamburger, not the hamburger-in-itself.) More precisely, the biological neural network in the human’s brain has neurons whose firings represent the word ‘hamburger’, and it has neurons whose firings represent the actual hamburger. Both of these sets of neuron firings are in the same “language”—the same syntactic system. Call it “U”. As in §3.3, $U = S \cup T$, where $S$ is the neuron firings of language, and $T$ is the neuron firings of perceptual images. $U$ is the “language of thought” [Fodor, 1975]. Yes, $T$ is just “more symbols” (as [Searle, 1980, p. 423] has objected; more precisely, $T$ is just more neuron firings—the “marks” of $T$ as a syntactic system). But that’s how semantics works. The same thing happens (or can happen) for computers; though the language of thought won’t be a biological neural network (it might be a computational semantic network such as SNePS or an artificial |22Again, see the essays in footnote 2, as well as [Shapiro and Rapaport, 1987], [Shapiro and Rapaport, 1992].

12
neural network). Thus, a combination of the “robot reply” (for internalization) and the “systems reply” (because it is never Searle-in-the-room alone) show us how to escape the Chinese Room.\footnote{23More details can be found in [Rapaport, 2000] and [Rapaport, 2006].}

References


23 More details can be found in [Rapaport, 2000] and [Rapaport, 2006].


