Chapter 3

Language Analysis and Understanding

(Following section is taken from Chapter 3 “Language Analysis and Understanding”) of the book: “Survey of the state of the art in human language technology”)

3.5 Semantics

Stephen G. Pulman
SRI International, Cambridge, UK
and University of Cambridge Computer Laboratory, Cambridge, UK

3.5.1 Basic Notions of Semantics

A perennial problem in semantics is the delineation of its subject matter. The term meaning can be used in a variety of ways, and only some of these correspond to the usual understanding of the scope of linguistic or computational semantics. We shall take the scope of semantics to be restricted to the literal interpretations of sentences in a context, ignoring phenomena like irony, metaphor, or conversational implicature (Grice, 1975; Levinson, 1983).

A standard assumption in computationally oriented semantics is that knowledge of the meaning of a sentence can be equated with knowledge of its truth conditions: that is, knowledge of what the world would be like if the sentence were true. This is not the same as knowing whether a sentence is true, which is (usually) an empirical matter, but knowledge of truth conditions is a prerequisite for such verification to be possible. Meaning as truth conditions needs to be generalized somewhat for the case of imperatives or questions, but is a common ground among all contemporary theories, in one form or another, and has an extensive philosophical justification, e.g., Davidson (1969); Davidson (1973).

1This survey draws in part on material prepared for the European Commission LRE Project 62-051, FraCaS: A Framework for Computational Semantics. I am grateful to the other members of the project for their comments and contributions.
A semantic description of a language is some finitely stated mechanism that allows us to say, for each sentence of the language, what its truth conditions are. Just as for grammatical description, a semantic theory will characterize complex and novel sentences on the basis of their constituents: their meanings, and the manner in which they are put together. The basic constituents will ultimately be the meanings of words and morphemes. The modes of combination of constituents are largely determined by the syntactic structure of the language. In general, to each syntactic rule combining some sequence of child constituents into a parent constituent, there will correspond some semantic operation combining the meanings of the children to produce the meaning of the parent.

A corollary of knowledge of the truth conditions of a sentence is knowledge of what inferences can be legitimately drawn from it. Valid inference is traditionally within the province of logic (as is truth) and mathematical logic has provided the basic tools for the development of semantic theories. One particular logical system, first order predicate calculus (FOPC), has played a special role in semantics (as it has in many areas of computer science and artificial intelligence). FOPC can be seen as a small model of how to develop a rigorous semantic treatment for a language, in this case an artificial one developed for the unambiguous expression of some aspects of mathematics. The set of sentences or well formed formulae of FOPC are specified by a grammar, and a rule of semantic interpretation is associated with each syntactic construct permitted by this grammar. The interpretations of constituents are given by associating them with set-theoretic constructions (their denotation) from a set of basic elements in some universe of discourse. Thus, for any of the infinitely large set of FOPC sentences we can give a precise description of its truth conditions, with respect to that universe of discourse. Furthermore, we can give a precise account of the set of valid inferences to be drawn from some sentence or set of sentences, given these truth conditions, or (equivalently, in the case of FOPC) given a set of rules of inference for the logic.

### 3.5.2 Practical Applications of Semantics

Some natural language processing tasks (e.g., message routing, textual information retrieval, translation) can be carried out quite well using statistical or pattern matching techniques that do not involve semantics in the sense assumed above. However, performance on some of these tasks improves if semantic processing is involved. (Not enough progress has been made to see whether this is true for all of the tasks).

Some tasks, however, cannot be carried out at all without semantic processing of some form. One important example application is that of database query, of the type chosen for the Air Travel Information Service (ATIS) task (DARPA, 1989). For example, if a user asks, “Does every flight from London to San Francisco stop over in Reykjavik?” then the system needs to be able to deal with some simple semantic facts. Relational databases do not store propositions of the form every X has property P and so a logical inference from the meaning of the sentence is required. In this case, every X has property P is equivalent to there is no X that does not have property P and a system that knows this will also therefore know that the answer to the question is no if a non-stopping flight is found and yes otherwise.

Any kind of generation of natural language output (e.g., summaries of financial data, traces of KBS system operations) usually requires semantic processing. Generation requires the construction of an appropriate meaning representation, and then the production of a sentence or sequence of sentences which express the same content in a way that is natural for a reader to comprehend, e.g., McKeown, Kukich, et al. (1994). To illustrate, if a database lists a 10 a.m. flight from London to Warsaw on the 1st–14th, and 16th–30th of November, then it is more helpful to answer the question What days does that flight go? by Every day except the 15th instead of a list of 30 days of the month. But to do this the system needs to know that the semantic representations of the two propositions are equivalent.
3.5 Semantics

3.5.3 Development of Semantic Theory

It is instructive, though not historically accurate, to see the development of contemporary semantic theories as motivated by the deficiencies that are uncovered when one tries to take the FOPC example further as a model for how to do natural language semantics. For example, the technique of associating set theoretic denotations directly with syntactic units is clear and straightforward for the artificial FOPC example. But when a similar programme is attempted for a natural language like English, whose syntax is vastly more complicated, the statement of the interpretation clauses becomes in practice extremely baroque and unwieldy, especially so when sentences that are semantically but not syntactically ambiguous are considered (Cooper, 1983). For this reason, in most semantic theories, and in all computer implementations, the interpretation of sentences is given indirectly. A syntactically disambiguated sentence is first translated into an expression of some artificial logical language, where this expression in its turn is given an interpretation by rules analogous to the interpretation rules of FOPC. This process factors out the two sources of complexity whose product makes direct interpretation cumbersome: reducing syntactic variation to a set of common semantic constructs; and building the appropriate set-theoretical objects to serve as interpretations.

The first large scale semantic description of this type was developed by Montague (1973). Montague made a further departure from the model provided by FOPC in using a more powerful logic (intensional logic) as an intermediate representation language. All later approaches to semantics follow Montague in using more powerful logical languages: while FOPC captures an important range of inferences (involving, among others, words like every, and some as in the example above), the range of valid inference patterns in natural languages is far wider. Some of the constructs that motivate the use of richer logics are sentences involving concepts like necessity or possibility and propositional attitude verbs like believe or know, as well as the inference patterns associated with other English quantifying expressions like most or more than half, which cannot be fully captured within FOPC (Barwise & Cooper, 1981).

For Montague, and others working in frameworks descended from that tradition (among others, Partee, e.g., Partee, 1986, Krifka, e.g., Krifka, 1989, and Groenendijk and Stokhof, e.g., Groenendijk & Stokhof, 1984; Groenendijk & Stokhof, 1991a), the intermediate logical language was merely a matter of convenience which could, in principle, always be dispensed with provided the principle of compositionality was observed. (I.e., The meaning of a sentence is a function of the meanings of its constituents, attributed to Frege, (Frege, 1892)). For other approaches, (e.g., Discourse Representation Theory, Kamp, 1981) an intermediate level of representation is a necessary component of the theory, justified on psychological grounds, or in terms of the necessity for explicit reference to representations in order to capture the meanings of, for example, pronouns or other referentially dependent items, elliptical sentences or sentences ascribing mental states (beliefs, hopes, intentions). In the case of computational implementations, of course, the issue of the dispensability of representations does not arise: for practical purposes, some kind of meaning representation is a sine qua non for any kind of computing.

3.5.4 Discourse Representation Theory

Discourse Representation Theory (DRT) (Kamp, 1981; Kamp & Reyle, 1993), as the name implies, has taken the notion of an intermediate representation as an indispensable theoretical construct, and, as also implied, sees the main unit of description as being a discourse rather than sentences in isolation. One of the things that makes a sequence of sentences constitute a discourse is their connectivity with each other, as expressed through the use of pronouns and ellipsis or similar devices. This connectivity is mediated through the intermediate representation, however, and cannot be expressed without it. The kind of example that is typically used to illustrate this is the following:

A computer developed a fault.
A simplified first order representation of the meaning of this sentence might be:
exists(X,computer(X) and develop_a fault(X))

*There is a computer X and X developed a fault.* This is logically equivalent to:
not(forall(X,not(computer(X) and develop_a fault(X))))

*It isn’t the case that every computer didn’t develop a fault.* However, whereas the first sentence can be continued thus:

A computer developed a fault.
It was quickly repaired.

—its logically equivalent one cannot be:

It isn’t the case that every computer didn’t develop a fault.
It was quickly repaired.

Thus, the form of the representation has linguistic consequences. DRT has developed an extensive formal description of a variety of phenomena such as this, while also paying careful attention to the logical and computational interpretation of the intermediate representations proposed. Kamp and Reyle (1993) contains detailed analyses of aspects of noun phrase reference, propositional attitudes, tense and aspect, and many other phenomena.

### 3.5.5 Dynamic Semantics

Dynamic semantics (e.g., Groenendijk & Stokhof, 1991a; Groenendijk & Stokhof, 1991b) takes the view that the standard truth-conditional view of sentence meaning deriving from the paradigm of FOPC does not do sufficient justice to the fact that uttering a sentence changes the context it was uttered in. Deriving inspiration in part from work on the semantics of programming languages, dynamic semantic theories have developed several variations on the idea that the meaning of a sentence is to be equated with the changes it makes to a context.

*Update semantics* (e.g., Veltman, 1985; van Eijck & de Vries, 1992) approaches have been developed to model the effect of asserting a sequence of sentences in a particular context. In general, the order of such a sequence has its own significance. A sequence like:

Someone’s at the door. Perhaps it’s John. It’s Mary!

is coherent, but not all permutations of it would be:

Someone’s at the door. It’s Mary. Perhaps it’s John.

Recent strands of this work make connections with the artificial intelligence literature on truth maintenance and belief revision (e.g, Gärdensfors, 1990).

*Dynamic predicate logic* (Groenendijk & Stokhof, 1991a; Groenendijk & Stokhof, 1990) extends the interpretation clauses for FOPC (or richer logics) by allowing assignments of denotations to subexpressions to carry over from one sentence to its successors in a sequence. This means that dependencies that are difficult to capture in FOPC or other non-dynamic logics, such as that between *someone* and *it* in:

Someone’s at the door. It’s Mary.

can be correctly modeled, without sacrificing any of the other advantages that traditional logics offer.
3.5 Semantics

3.5.6 Situation Semantics and Property Theory

One of the assumptions of most semantic theories descended from Montague is that information is total, in the sense that in every situation, a proposition is either true or it is not. This enables propositions to be identified with the set of situations (or possible worlds) in which they are true. This has many technical conveniences, but is descriptively incorrect, for it means that any proposition conjoined with a tautology (a logical truth) will remain the same proposition according to the technical definition. But this is clearly wrong: all cats are cats is a tautology, but The computer crashed, and The computer crashed and all cats are cats are clearly different propositions (reporting the first is not the same as reporting the second, for example).

Situation theory (Barwise & Perry, 1983) has attempted to rework the whole logical foundation underlying the more traditional semantic theories in order to arrive at a satisfactory formulation of the notion of a partial state of the world or situation, and in turn, a more satisfactory notion of proposition. This reformulation has also attempted to generalize the logical underpinnings away from previously accepted restrictions (for example, restrictions prohibiting sets containing themselves, and other apparently paradoxical notions) in order to be able to explore the ability of language to refer to itself in ways that have previously resisted a coherent formal description (Barwise & Etchemendy, 1987).

Property theory (Turner, 1988; Turner, 1992) has also been concerned to rework the logical foundations presupposed by semantic theory, motivated by similar phenomena.

In general, it is fair to say that, with a few exceptions, the contribution of dynamic semantics, situation theory, and property theory has so far been less in the analysis of new semantic phenomena than in the exploration of more cognitively and computationally plausible ways of expressing insights originating within Montague-derived approaches. However, these new frameworks are now making it possible to address data that resisted any formal account by more traditional theories.

3.5.7 Implementations

Whereas there are beginning to be quite a number of systems displaying wide syntactic coverage, there are very few that are able to provide corresponding semantic coverage. Almost all current large scale implementations of systems with a semantic component are inspired to a greater or lesser extent by the work of Montague (e.g., Bates, Bobrow, et al., 1994; Allen, Schubert, et al., 1995; Alshawi, 1992). This reflects the fact that the majority of descriptive work by linguists is expressed within some form of this framework, and also the fact that its computational properties are better understood.

However, Montague’s own work gave only a cursory treatment of a few context-dependent phenomena like pronouns, and none at all of phenomena like ellipsis. In real applications, such constructs are very common and all contemporary systems supplement the representations made available by the base logic with constructs for representing the meaning of these context-dependent constructions. It is computationally important to be able to carry out at least some types of processing directly with these underspecified representations: i.e., representations in which the contextual contribution to meaning has not yet been made explicit, in order to avoid a combinatorial explosion of potential ambiguities. One striking motivation for underspecification is the case of quantifying noun phrases, for these can give rise to a high degree of ambiguity if treated in Montague’s fashion. For example, every keyboard is connected to a computer is interpretable as involving either a single computer or a possibly different one for each keyboard, in the absence of a context to determine which is the plausible reading: sentences do not need to be much more complex for a large number of possibilities to arise.

One of the most highly developed of the implemented approaches addressing these issues is the quasi-logical
form developed in the Core Language Engine (CLE) (Alshawi, 1990; Alshawi, 1992) a representation which allows for meanings to be of varying degrees of independence of a context. This makes it possible for the same representation to be used in applications like translation, which can often be carried out without reference to context, as well as in database query, where the context-dependent elements must be resolved in order to know exactly which query to submit to the database. The ability to operate with underspecified representations of this type is essential for computational tractability, since the task of spelling out all of the possible alternative fully specified interpretations for a sentence and then selecting between them would be computationally intensive even if it were always possible in practice.

3.5.8 Future Directions

Currently, the most pressing needs for semantic theory are to find ways of achieving wider and more robust coverage of real data. This will involve progress in several directions: (i) Further exploration of the use of underspecified representations so that some level of semantic processing can be achieved even where complete meaning representations cannot be constructed (either because of lack of coverage or inability to carry out contextual resolution). (ii) Closer cooperation with work in lexicon construction. The tradition in semantics has been to assume that word meanings can by and large simply be plugged in to semantic structures. This is a convenient and largely correct assumption when dealing with structures like every X is P, but becomes less tenable as more complex phenomena are examined. However, the relevant semantic properties of individual words or groups of words are seldom to be found in conventional dictionaries and closer cooperation between semanticists and computationally aware lexicographers is required. (iii) More integration between sentence or utterance level semantics and theories of text or dialogue structure. Recent work in semantics has shifted emphasis away from the purely sentence-based approach, but the extent to which the interpretations of individual sentences can depend on dialogue or text settings, or on the goals of speakers, is much greater than had been suspected.

3.6 Chapter References


3.6 Chapter References


Chapter 3: Language Analysis and Understanding


3.6 Chapter References


Chapter 3: Language Analysis and Understanding


3.6 Chapter References


3.6 Chapter References


Chapter 3: Language Analysis and Understanding


3.6 Chapter References


Chapter 3: Language Analysis and Understanding


3.6 Chapter References


