The Syntax, Semantics, and Inference Mechanism of Natural Language

Francis Y. Lin
Computational Linguistics Unit
School of English and Philosophy
University of Wales Cardiff

Abstract

It is both desirable and plausible to treat natural language itself as a knowledge representation (KR) formalism. Every KR formalism has syntax and support certain inferences. The syntax of a KR formalism specifies the form in which knowledge must be encoded, and its inference mechanism depends on its syntax. If natural language is a formalism for representing knowledge, then its syntax provides specifications of well-formed formulae. In addition, reasoning must be done directly on the basis of syntax, as in any artificial KR scheme. In this respect, existing syntactic theories are inadequate, because syntax in these theories does not support reasoning. In this paper I present a new conception of the syntax, semantics, and inference mechanism of natural language. The central ideas are (1) that word use determines the forms of sentences, and (2) that these forms are used to express meaning and perform inference. Syntax, semantics, and inference are seen as an integrated whole, rather than separate and autonomous fields as in many existing studies of natural language.

Introduction

Existing knowledge representation (KR) systems do not use natural language directly or fully. The common practice is to translate natural language into an artificial formalism, on which subsequent operations are based. There have been attempts to make KR formalisms close to natural language to gain wider linguistic coverage or to obtain more efficient reasoning (Barwise & Cooper, 1981; Kayser, 1987; Alshawi, 1992; Iwanska, 1992). But, since people use natural language to exchange information, to gain new knowledge, to express thought, to perform reasoning, and so on, it is not only desirable but also plausible to treat natural language itself as a KR formalism. Every KR formalism has syntax and supports certain inferences. The syntax of a KR formalism specifies the form in which knowledge must be encoded, and its inference mechanism depends on its syntax. Some KR formalisms also have a semantic component, which determines what a well-formed formula means and how meaning is ascribed to it. If natural language is a formalism for representing knowledge, then its syntax provides specifications of well-formed formulae. In addition, reasoning must be performed directly on the basis of syntax, as in any artificial KR scheme. This is a point that has been neglected so far. If we apply this requirement to existing accounts for syntax, we will easily find they are inadequate, that is, existing syntactic rules are not designed to support inference and it is doubtful that they can be used directly to perform reasoning. What we need then is a new conception of syntax of natural language, and a corresponding new conception of semantics and reasoning.

Syntax and Word Use

Ordinary speakers can potentially recognise and produce an infinitely large number of sentences. This phenomenon is called the *competence* of the language speaker (Chomsky, 1957). As with Chomsky, this phenomenon can only be accounted for by assuming that the language speaker utilises a finite set of grammar rules which are capable of generating an infinite set of sentences. One major goal of generative grammar is to account for this phenomenon. A generative grammar typically consists of a set of abstract rules and principles, from which all sentences in a language can be derived. These rules and principles are so abstract that they can hardly be said to be known to the ordinary speaker in any real sense. So they are assumed to be 'tacitly' known to the speakers, or even 'innate' in their genes (Chomsky, 1988).

Chomsky's conception of grammar rules is in direct contrast to that of Wittgenstein. According to the later Wittgenstein, language consists of language games which are played according to rules (Wittgenstein, 1972). Rules must be able to be expressed and they must be transparent to participants in a rule-

governed practice (Baker & Hacker, 1985, pp. 62-3). In this section I outline an account of grammar whose rules are *actually* known to the speakers. I shall start with the notion of sentence structure and try to explicate the notion of grammar via some examples.

Consider the following sentence:

(1) John kissed Mary in front of Peter on purpose, because he wanted to annoy him.

The structure of the sentence, in my analysis, is:

(2) something$_x$ happened because something$_y$ happened[1]

where *something$_x$ happened* refers to:

(3) John kissed Mary in front of Peter on purpose.[2]

This sentence in turn has a structure, which is:

(4) somebody did something on purpose.

where *somebody* is John, *somebody did something* is:

(5) John kissed Mary in front of Peter.

The structure of sentence (5) is:

(6) somebody$_x$ did something in front of somebody$_y$.

where *somebody$_x$ did something* refers to:

(7) John kissed Mary.

whose structure is:

(8) somebody$_x$ kissed somebody$_y$.

It can be seen from the above example that what I call structures are expressions of word use. For example, (2, 4, 6, and 8) show how the words (and phrases) 'because', 'on purpose', 'in front of', and 'kiss' are used; and one only needs a little reflection to confirm that this is the case. I shall therefore call a structure like (2) a **usax**,[3] and say that a word is used grammatically if it conforms to its usax (or one of its usaxes if it has more than one). We can now have a better idea about the notions of sentence structure and can appreciate the equivalence relationship between the following propositions:

(9) A string of words is a sentence.

(10) The string of words has a sentence structure.

(11) The sentence is grammatical.

(12) All the words have been used grammatically.

A few other examples may help to clarify the notions of structure, usax, and grammaticality. (13) is a grammatical sentence, it conforms to the structure (15), which is a usax for 'kind'. (14) is ungrammatical, because there is no such usax as (16):

(13) It's kind of you to come to see me.

(14) *It's hungry of you to come to see me.

(15) It's kind of somebody to do something.

(16) *It's hungry of somebody to do something.

It is easy to see that usaxes, as those illustrated, are general ways of how words are used and that they are finite in number. Usaxes are grammar rules, according to which we recognise sentences and determine their grammaticality. We have learned them, and they are known to us. In the rest of this section I shall show how this finite number of explicit grammar rules can generate an infinite number of sentences.

Usaxes are concrete, in the sense that they are expressible in natural language. On the other hand they are also general. The more general a usax is the more sentences it can cover, that is, it can generate. For example, there exists an increasing order of generality among the following sentences:

(5) < (6) < (17) < (18) < (19)

where (17-19) are:

(17) Somebody did something.

(18) Something happened.

(19) Something is the case.

(19) covers events, states, properties, and whatever you have. In fact (19) is the most general form of sentences (or usaxes).

The process of *generalisation* starts, as illustrated above, from any sentence and ends in the most general form, which is (19). What is the converse process? It is an easy matter to see that it is *generation*, in the Chomskyan sense. For example, one can begin with (19), through (18), (17), and (6), and end with (5).[4] Notice that each general usax can generate a large number of usaxes which are less general, and the latter are capable of generating a great many even less general usaxes. Note also that each least general usax can still generate a large number of concrete sentences. Through this process a grammar can generate very many sentences. To ensure that a grammar can generate an infinite number of sentences, recursion will have to be build into it. But I shall say no more about this here.

To sum up, usaxes are grammar rules, they are finite and are known to the speakers. Through the process of generalisation, one gets more and more general usaxes and through the converse process, which is generation, one produces (or can produce) all the sentences in a language.

[1] I use somebody$_x$ to mean that there is an entity x which is referred to by the word 'somebody'. The referents will be left out when there are no ambiguities present.

[2] For simplicity the analysis of the second half of (1), which is 'he wanted to annoy him', will be omitted here.

[3] This notion is employed to stress the relationship between the use of words and the syntax of sentences.

[4] But one does not need to always start from (19) to generate a sentence. One can start from anywhere in the chain of generation. Note also that one does not always need to reach the end of the generation process either. One can stop in the midway, to arrive at certain sentences, e.g. (17).

Semantics and Inference

Semantics is the study of meaning. My aims in this section are to present my view on meaning, and to provide justification for it. I shall show that the theory developed here subsumes several major semantic theories.

Wittgenstein rejects the idea that meaning is something abstract, something like a 'shadow' of words and sentences (Wittgenstein, 1972, PP. 117&120). He says that to know the meaning of a word/sentence we must look at how it is used and in what situation it is used (Wittgenstein, 1972, PP. 43&592). The meaning of a word/sentence is the explanation we give (in each of the concrete situation in which it is used). In Wittgenstein's words, 'meaning is what an explanation of meaning explains' (Wittgenstein, 1974, p. 69). Now, how is explanation done? As one cannot simply memorise all sentences (in order to recognise or produce them), one cannot simply memorise the meaning of all sentences (in order to explain the meaning). To know the meaning of a sentence, which is one of a potentially infinite set, we must perform some inferences.

I submit that the meaning of a word/sentence is the set of sentences which can be inferred from the word/sentence, and I shall also call the latter the *meaning-bearing* sentences of the former. By inference (or reasoning) I mean not only formal inference (e.g. deduction), but also informal or practical reasoning (e.g. speech acts), and inference of other sorts. Thus semantics in the present framework is compatible with Lyons' *linguistic semantics* (Lyons, 1995). According to this conception, there is no meaning in isolation: meaning arises from interconnections (Strawson, 1992; Quine & Ulian, 1978).

One influential semantic theory is based on semantic primitives (Katz, 1972; Jackendoff, 1976). There are well-known problems with such an approach, such as that it is difficult, or even impossible, to identify the primitive elements, and that in the vast majority of cases such reductions are inadequate as full specification of meaning. As Lyons (Lyons, 1995) points out, componential analysis based on semantic primitives can be replaced by the use of meaning-postulates. In general, the semantic primitives of a word/sentence are at best equivalent to a subset of its meaning-bearing sentences.

It is also common practice to specify the meaning of a sentence by assigning cases or participant roles for the objects mentioned in the sentence, e.g. (Fillmore, 1968; Halliday, 1994). For example, in 'John kissed Mary' 'John' is said to be the *agent*, 'Mary' the *patient*. There are often difficulties in determining how many roles there are and what they are, and in assigning roles to objects mentioned in a sentence. Another problem is that roles are usually not capable of capturing enough, let alone complete, meaning. Consider 'John dried the shirt' for instance, the roles can not capture the meaning such as 'the shirt was wet and it is now dry'. Viewed from the vantage point of the current theory, roles are merely short-form for some (a subset) of the meaning-bearing sentences, e.g. 'John did the kissing', 'Mary was kissed'. Thus, in a certain sense role-based theories can be seen as a simplified version of the present theory, in which only certain types of inference are in operation (see next section).

A third major semantic theory utilises the notion of truth conditions, the most important version being Davidson's. Davidson explicates the meaning of a sentence in terms of its truth conditions, conditions under which the sentence is true (Davidson, 1967, 1970). As Dummett (1976) points out, the theory does not explain how the speaker can know those truth conditions. It also implies that the ordinary speaker masters the predicate-calculus and the translation rules which paraphrases the ordinary sentences into their Davidsonian replacements, which is both 'unrealistic' and 'unnecessary' (Strawson, 1992, p. 103). It is not difficult to see that, like the roles-based theories, the truth conditions of a sentence are a subset of its meaning-bearing sentences.

Moravcsik (1990) contends that 'the meaning of a word "w" is that in virtue of which an element of reality counts as a w'. For example, the meaning of 'water' is that it is a clear, odourless liquid, and is used for drinking, and irrigation, etc. The present account largely agrees with Moravcsik's view of meaning. What he regards as the meaning of a word is effectively the set of sentences we can infer from the word. These meaning-bearing sentences are what we know about (or infer from) the word, and hence are its meaning.

Moravcsik's theory is only about words, and he has not yet touched upon the meaning of sentences, which is an equally, and probably more important part of a theory of meaning. The account presented here has just done this. It is a natural extension of Moravcsik's work, for either the meaning of a word or that of a sentence is the set of its meaning-bearing sentences, thus they can be, as they are in the present framework, treated in the same way.

So far I have examined four major semantic theories. These theories appear to be diverse, but from the present point of view they are actually closely related, and they can all be subsumed in the present theory. The meaning of a word/sentence in these theories is part of what it is in the present framework. These theorists attempt to impose certain constraints on in-

ference, though they may be unaware of this. In other words, the theories I have just studied only employ certain types of inference.

Inference in Natural Language

Having established that the meaning of a sentence is the set of sentences which can be inferred from it, the task of semantics then becomes the task of explaining how inference is done. In parallel to the conception of grammar presented in Section 2 I shall provide an account for semantics based on a finite number of inference rules which are known to the speakers

Inference of meaning can take many forms (or types), four of which will be discussed here. I shall first explicate what the four types of inference are by giving examples, and then briefly compare my treatment of them with some other approaches. The most common type of meaning inference is performed through the meaning of words (and phrases), which speakers are aware of and are able to explain. Many new words are learned also in this way, i.e. through their explanation in terms of other words (Johnson-Laird, 1983, ch. 10). For example, we know:

(20) If you do something on purpose, you do it deliberately.

(21) If you do something in front of somebody, you do it when he is present.

(22) If you kiss somebody, you touch him with your lips.

I shall call sentences like (20-22) *meaning elucidations*, following (Higginbotham, 1989).[5] It is readily seen that meaning elucidations also have a structure (e.g. If x then y), which consists of structures of the constituent sentences (compare (20-22) with (4, 6, 8)). If we apply the meaning elucidations (20-22) to sentences (4, 6, 8) respectively, then we should be able to infer the corresponding sentences. I shall use the term *elucidation* to refer to this type of inference.

Another common way of explaining the meaning of a sentence is to use its equivalent or near-equivalent sentences. For example, (23) is equivalent to (24), and (25) means the same as (26):

(23) You gave something to somebody.

(24) You gave somebody something.

(25) There is something$_x$ in something$_y$

(26) Something$_y$ has something$_x$ in it.

I shall call equivalent sentences such as (23-26) *meaning equivalences*. This type of inference will be called *equivalisation*.

The third type of commonly used inference is *deletion*, which basically means that we can infer from some sentences some of their parts. The following sentence is a deletion rule:

(27) If you did something at a certain time, in a certain place, in a certain way, or for a certain reason. etc., then you did it.

The fourth type of inference is what I shall call *exposition*. It is used to expose the semantic roles (sometimes called case roles or participant roles) in sentences. Below are a few examples of exposition rules:

(28) If x did y, then what x did is to do y.

(29) If x did y, then what (or the person who) did y is x.

(30) If x did y to z, then what (or /the person whom) x did y to is z.

It is easy to see that all these four types of inference are based on usaxes: Given any sentence, we can abstract its usax. This usax will match one (or more) semantic rules, and corresponding inferences will be obtained. It is here that grammar elucidated in Section 2 and semantics discussed in this section come to contact with each other, and the interaction is very smooth.

It is also here that we see more clearly the relationships between my account of meaning and other semantic theories. By using elucidation and equivalisation we can replace theories based on semantic primitives. By using exposition we can replace semantic roles. And by using deletion we can replace at least the part of Davidson's theory which accounts for the derivation of (32) from (31):

(31) John kissed Mary in front of Peter.

(32) John kissed Mary.

In doing so, we also avoid many fundamental problems inherent in those theories, which were mentioned in the previous section.

Conclusions and Discussions

It is both desirable and plausible to use natural language as a means of representing knowledge and performing reasoning. This requires that the syntax of natural language should specify the format in which knowledge is encoded, and that reasoning should operate directly on the basis of syntax as in any artificial KR scheme. Existing accounts of syntax are inadequate, because the syntax developed so far is separate from and cannot support reasoning. A new conception

[5] This is not to advocate a 'dictionary' view of meaning, which implies that meaning is fixed. According to Wittgenstein (Wittgenstein, 1972, P. 456), if somebody just said 'I kissed Mary' and I now simply mimic his voice and also say 'I kissed Mary', then I didn't really kiss Mary (i.e. touch her with my lips). But this is due to the meaning of the word 'mimic'. It is true that the meaning of words/sentences depends on the situations in which they are used, but we must know the meaning of the words/sentences which describe such situations and perform corresponding inference.

of syntax, semantics, and reasoning is needed.

In this paper I presented a new way of looking at the syntax, semantics and inference mechanism of natural language. The basis of investigation is word use. In learning a language we learn not only words, but also how they are used and what they mean. The rules of grammar, meaning, and inference are closely related to word use and they are known to ordinary speakers. In order to understand these notions I have taken necessary excursions into linguistics and philosophy. What was achieved was a unified account for these notions. As I argued, usaxes are grammar rules, they determine sentence recognition and generation. They are the forms in which meaning is expressed, and reasoning is performed. Thus, syntax, semantics, and reasoning form an integrated whole, rather than separate and autonomous components of natural language.

Further work is still needed to develop the present account of language into a satisfying theory. One direction is to explain within this framework more syntactic phenomena, e.g. passivisation, reflexives, etc, as dealt with in Government and Binding Theory (Haegeman, 1994). Another direction is to explore further the inference process based on usaxes (see Lin (1996)).

Acknowledgements

I am greatful to Sir Peter Strawson for discussions, comments, and encouragement. All errors are mine.

Reference

Alshawi, H. (Ed.). (1992). *The Core English Engine.* MIT Press.

Baker, G. P., & Hacker, P. M. S. (1985). *Wittgenstein: Rules, Grammar and Necessity.* Blackwell, Oxford.

Barwise, J., & Cooper, R. (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy, 4,* 159–219.

Chomsky, N. (1957). *Syntactic Structures.* Mouton, The Hague.

Chomsky, N. (1988). *Language and Problems of Knowledge.* The MIT Press, London.

Davidson, D. (1967). Truth and meaning. Reprinted in his Inquiries into Truth and Interpretation, Clarendon Press, Oxford, 1984.

Davidson, D. (1970). Semantics for natural languages. Reprinted in his Inquiries into Truth and Interpretation, Clarendon Press, Oxford, 1984.

Dummett, M. (1976). What is a theory of meaning? (ii). Reprinted in his The Seas of Language, Clarendon Press, Oxford, 1993.

Fillmore, C. (1968). The case for case. In Bach, E., & Harms, R. T. (Eds.), *Universals in Linguistic Theory.* Rinehart and Winston, New York.

Haegeman, L. (1994). *Introduction to Government and Binding Theory* (2nd edition). Blackwell, Oxford.

Halliday, M. A. K. (1994). *Introduction of Functional Grammar* (revised edition). Edward Arnold, London.

Higginbotham, J. (1989). Elucidations of meaning. *Linguistics and Philosophy, 12,* 465–517.

Iwanska, L. (1992). A general semantic model of negation in natural languages: representation and inference. In *Proceedings of the Third Conference on the Principles of Knowledge Representation and Reasoning.*

Jackendoff, R. (1976). Towards an explanatory semantic representation. *Linguistic Inquiry, 7,* 89–150.

Johnson-Laird, P. N. (1983). *Mental Models.* Cambridge University Press, Cambridge.

Katz, J. (1972). *Semantic Theory.* Harper & Row, New York.

Kayser, D. e. a. (1987). A strategy for reasoning in natural language. *Applied Artificial Intelligence, 1,* 205–231.

Lin, F. Y. (1996). *On the notion of logical form.* Computational Linguistics Unit, School of English and Philosophy, University of Wales Cardiff. Manuscript.

Lyons, J. (1995). *Linguistic Semantics.* Cambrige University Press, Cambridge.

Moravcsik, J. M. E. (1990). *Thought and Language.* Routledge, London.

Quine, W. V. O., & Ulian, J. (1978). *The Web of Belief* (rev. edition). Random House, New York.

Strawson, P. F. (1992). *Analysis and Metaphysics: An Introduction to Philosophy.* Oxford University Press, Oxford.

Wittgenstein, L. (1972). *Philosophical Investigations* (3rd edition). Blackwell, Oxford.

Wittgenstein, L. (1974). *Philosophical Grammar.* Blackwell, Oxford. Rhees, R. (ed.).