Semantics, Translation, and Anaphora

Rebecca Root

Although most machine translation systems (proposed or implemented) make some use of a semantic level of description, this is typically something quite different in interpretation, if not in form, than what is found in other areas of natural language processing (NLP). This state of affairs might seem surprising at first to formal semanticists, since meaning would appear to be as central to translation as it is to query understanding, for example. But it is not so surprising on closer inspection. The primary function of the semantics component in a formal description is to account for the relationship between a syntactic structure and objects in the world (or model of the world). This is often done by laying down rules by which linguistic structures can be mapped to expressions of some intermediate language, for example, first order predicate calculus, for which some conventional mapping to a model of the world has been defined. It is obvious why this sort of approach to natural language semantics is attractive for NL interfaces to databases. If the database is considered to be the model of the world, and the database query language is used as the intermediate language, the NL semantic component directly provides the mapping from natural language to formal query language that is necessary. The same convenience holds for other applications in which the linguistic input must be mapped to an external, non-linguistic data structure or application program. This is not the case for MT. Here the task is to map between linguistic objects. There is no a priori requirement for a component which deals with a model of the world (but see Boitet and Gerber(84) for discussion which suggests that a truth conditional approach might be useful). The sort of semantic information that has proven most useful in this context is functional structure of predicates and semantic selectional restrictions.

If all a formal semantics supplied was truth conditions, then there would be no reason to question the status quo. And, indeed, I don't want to imply that a change is necessarily called for. However, truth conditional semantics can bring more to an overall theory of grammar than simply providing the truth conditions. It can also act as a "semantic filter" to rule out unwanted interpretations as a side effect. When this happens, the overall grammar is simplified because a need for specific syntactic constraints is eliminated. On a lexical level, the reliance on semantic filtering is uncontroversial. To take a well-known example, no one would attempt to explain the unacceptability of "Colorless green ideas sleep furiously" in terms of syntactic constraints. Rather, it is attributed to the fact that the normal interpretation of the words involved rule out a true interpretation of the sentence in any model corresponding to the real world. An exactly parallel situation can arise with respect to the interpretation of larger constituents. For an example see Ladusaw(82) or Hoeksema(85). Here the authors independently provide accounts of the English partitive construction, exemplified here:

(1)

a. all of the boys
b. * all of some boys

which do not rely on the explicit statement of the partitive constraint, i.e., that the noun phrase following "of" must be definite. Instead, the interpretations assigned to the partitive construction and to indefinite noun phrases rule out a meaningful interpretation
when the two are combined, thereby simplifying the statement of the partitive construction.

Of course, in the two examples mentioned here, the complication to the syntax would be minor and localized. In each case, simple feature passing and checking could be used to cover most examples. But as semantic analysis continues to be approached with the same rigor that syntactic analysis has been, it is quite possible that for certain phenomena, a semantic account could be replaced by a syntactic one only if it were quite ad hoc and cumbersome. For a system which has done without a formal semantic component for reasons of economy, this could pose an interesting dilemma. In the rest of this paper, I will discuss the problems associated with accounting for the interpretation and distribution of anaphora in multisentence discourse, and will present an analysis which overcomes most of the difficulties. This discussion is interesting in the present context because the solution offered is quite simple, but it relies crucially on an appeal to a truth conditional semantic component, and so is beyond the scope of most MT systems.

1. Discourse Anaphora

One of the most studied topics in both theoretical and computational linguistics is anaphora resolution. It has been approached from a wide variety of angles, but most work addresses one of two questions: what anaphoric links are possible and what links are intended. The present work addresses the first of these. Discourse anaphora is particularly interesting because, on the one hand, it is complex, and on the other hand, it is common enough to warrant both a theoretically and a computationally interesting solution.

The term anaphor is used here in the general sense of a word which points back to some previous linguistic or semantic object. This differs from the usage within Government and Binding Theory in which the term is used more narrowly to refer to linguistic elements, including empty categories, which are necessarily bound by an antecedent according to strict syntactic criteria. Since this work is concerned almost exclusively with relationships across sentence boundaries, there should be little confusion in terminology.

Syntactic descriptions have primarily characterized what anaphoric relationships are impossible by means of a variety of principles of disjoint reference. At the discourse level, a syntactic solution does not look promising because of the wide variety of surface configurations which are possible, especially when plural anaphors are involved. One finds straightforward links to antecedents in virtually any noun phrase position.

(2)

   He liked it.

b. John asked Mary about her problem.
   He knew it was bothering her.

But in addition to these, there are links from anaphors to discontinuous antecedents, as in (3):

(3)

a. John found a piano teacher.

b. They are both fond of Mozart.

And there are links from plural anaphors to singular antecedents in distributive contexts, as in (4):

(4)

a. Every girl brought a cake.

b. They were chocolate.
   * It was chocolate.

And to see that proper resolution of discourse anaphora can be necessary for proper translation, consider the following German to English translation: