The Place of Semantics in MT Systems

WINFIELD S. BENNETT
Siemens Information Systems Inc., Austin, Texas, USA

Abstract

Semantics has been an issue in machine translation research from its inception. The basic assumption is that semantics is necessary for any MT system, with some uncertainty about the extent necessary for that semantics. This paper examines the general issue of MT semantics in an attempt to put the questions in perspective. From a practical standpoint, semantics is not necessary for an MT system, although highly desirable. Further, something less than a full-blown semantics can contribute to an MT system.

1. Introduction

Semantics has been an issue from the beginning of machine translation research. Warren Weaver (1949) discusses semantics in his famous memorandum, although his views of how to deal with semantic problems were quite naïve. Bar-Hillel (1960*a*) and Yngve (1964) both clearly stated their belief in the necessity for semantics in MT. For some time, then, a basic assumption about machine translation systems is that semantics is necessary to any system. For example, Lytinen states: 'It has long been realized by MT researchers that semantics must be used to resolve many of the lexical and structural ambiguities that occur in natural language' (Lytinen, 1987*a*), Lytinen's statement certainly represents the general view espoused by many MT researchers.

Any discussion of MT semantics naturally includes the question of what form such semantics might take. Stated another way: to what extent should semantics in MT go? As Lehrberger and Bourbeau put it: The need for semantic analysis is generally recognized; the big question is how to do it' (Lehrberger and Bourheau, 1988).

The discussion below puts both of these questions into perspective.

2. Semantics and machine translation

It is probably uncontroversial to state that semantics would generally enhance any machine translation system. This statement, of course, assumes that the semantics makes some sort of contribution to the overall effectiveness of the system without decreasing its efficiency. Here I take the position that efficiency in terms of speed and/or cost are important factors in any machine translation system. The issue, as I see it, is not whether semantics would enhance an MT system; rather, it is whether semantics is *necessary* for an MT system, as the consensus would have it.

In the decades since ALPAC, researchers have revised their expectations of MT to reflect a greater understanding of the limitations and potentials of translation systems (Hutchins, 1986*b*). In spite of more realistic expectations, discussions about MT semantics still approach the issue from the point of view that machine translation should translate any form of natural language. This is certainly the ultimate goal, but, pragmatically, not an immediate one. Consider, for example, the following pair of sentences, which Lytinen (1987*b*) presents as evidence that MT syntactic systems need semantics for disambiguation.

> The cleaners dry-cleaned the coat that Mary found at the rummage sale for $10.
> The cleaners dry-cleaned the coat that Mary found in the garbage for $10.

The ambiguity in the first sentence is indisputable; however, the difficulty with such subtle ambiguity is that is does not occur regularly in the types of texts for which most machine translation systems are designed and used. The same may be said for Carbonell's (1987) example:

> While driving down route 72, John swerved and hit a tree.

While these examples are excellent illustrations of problems in disambiguation, one should be cautious before assuming that they are evidence for the necessity for semantics in MT. In providing instances of ambiguity it is always possible to find examples which are ambiguous enough to confuse human translators as well as computers. Consider the simple sentence:

> The dinosaur is in the pen.

On the surface this looks like a variation on the well-known Bar-Hillel (1960*b*) example:

> The box is in the pen.

Bar-Hillel's example is unambiguous for native speakers of English, but the same may not be said for my example, since one can purchase writing instruments in the barrel of which a small plastic dinosaur is embedded. Subtle ambiguities are not in themselves grounds to declare the necessity for semantics in MT systems: even subtler ambiguities are always possible.

If the basis for arguments for the necessity of semantics in machine translation systems is to be ambiguities such as those presented in Lytinen's and Carbonell's examples, then such ambiguities must occur frequently enough in the texts to reduce the overall accuracy of the machine translation system.

I contend that the ambiguity which MT systems must handle is much less complex semantically than that found in the examples presented as justification for the necessity of semantics in MT systems. Based on my

Correspondence: Winfield S. Bennett. Siemens Information Systems, Inc., Linguistics Research Center, PO Box 7247, Austin, Texas 78713-7247, USA

Literary and Linguistic Computing, Vol. 4, No. 3, 1989

research experience the sorts of ambiguity actually encountered in MT texts are more like the following:

The computer outputs the data in the file.
The system uses the information to print the rule lines in the footer.
Write the name on the piece of paper.

All of these illustrate rather obvious problems with prepositional phrase attachment; however, the possibility for the resolution of these kinds of ambiguity is not particularly enhanced with even the most sophisticated semantics. Each of the sentences can certainly be semantically analyzed with the prepositional phrase attached to the object noun phrase as well as to the verb. The issue thus becomes one of pragmatics, requiring either a pragmatics for the system or heuristics to choose one interpretation over the other.

None of this discussion is meant to dismiss the realities of ambiguity as a major problem in machine translation; certainly semantics will provide solutions to some of the ambiguity problems. I do not question in the least the importance of research to create systems capable of semantic invariance, defined as: 'preserving invariant the meaning of the source text as it is transformed into the target text' (Carbonell and Tomita, 1987). Ultimately, the issue is not whether semantics *should* be a part of an MT system, but whether it *must* be a part. In the best of all possible worlds, semantics would be a part of any MT system, but, failing that, it is not absolutely necessary for most MT. It is time to recognize that semantics in MT is highly desirable, but not necessary.

3. The extent of semantics

Given that semantics is desirable (or necessary) for any machine translation system, the question then is: to what extent should MT semantics go?

There is rather significant diversity among researchers, in answering this question. Recognizing the dangers of overgeneralization, I will simply characterize the prevailing ideas in this issue as ranging from the use of semantic features to augment syntactic analysis to full-blown semantic analysis, with any number of positions in between. I will not go into the various definitions for different kinds of semantics (for a detailed discussion of such issues, see Hutchins, 1986a).

If semantics is to be a part of a machine translation system, the extent of the semantics should be determined by the needs of the system. Clearly a full-blown semantics should, in theory, meet all the needs of a system, but its cost may be greater than its ability to fulfill the needs. While it can be correctly argued in principle that the fullest possible semantics is needed for any given MT system, the reality is that an MT system may not need the full power of such a system most of the time.

To illustrate I will use one example from my experience. German prepositions do not have a one-to-one correspondence with English prepositions. For example, the German preposition *nach* may be translated in most instances as one of three English prepositions: *after, according to, to,* e.g.

nach dem Konzert = after the concert
nach der Formel = according to the formula
nach Austin = to Austin

Similarly German *zu, unter*, and many other prepositions have one-to-many correspondences with English. This is hardly surprising, given the nature of prepositions in human languages, but these linguistic realities pose something of a problem for an MT system involving German analysis. The solution to the problem may take a number of forms, but any solution should take into account the relevant information from the preposition and its object and do the necessary work to ensure that *nach* is translated as *after* or *according to* or *to,* as appropriate. Presumably a full-blown semantics would handle this, but is such power necessary for the disambiguation needed at this level? In fact a semantic feature system is sufficient to ensure the correct translation for the German prepositions in most instances. Using semantic features may not be the ideal solution, theoretically, but is is certainly effective and efficient, in this example.

Other examples of the usefulness of semantics in machine translation can be cited. As one such example one can note, with Hutchins (1988), that a number of systems now use case roles as part of the analysis of clauses, but without using a full-blown semantic analysis. Another example is tense and aspect (see, for exampie, Meya and Vidal, 1988).

As desirable, in theory, as a full-blown system might be, it is generally true that highly accurate machine translation can be produced by a system with less computational cost using less than a full-blown system Use of lexical semantic features should not be dismissed out of hand, since this approach may be the quickest and cheapest way to achieve some semantic control in a giver system,

Once again the issue comes down to desirability versus necessity. As desirable as full-blown systems are, an MT system may be able to do extremely well using something less. Full-blown semantics in MT are highly desirable but certainly not necessary.

4. Conclusion

In the course of a few pages I have sought to examine the two closely related issues: Is semantics *necessary* for MT? and To what extent should MT semantics go?

I take a practical approach in answering both questions from the standpoint of necessity, i.e. what is practically necessary for an MT system? The answers, it seems to me are clear. Semantics in a machine translation system is highly desirable, but not necessary. The extent of the semantics is as minimal as is necessary for accuracy in real test applications.

Acknowledgements

This paper is the result of frequent discussions with colleagues both at the Linguistics Research Center and elsewhere. In particular, 1 wish to thank Ana Santistehan, Katherine Hoyl, and Joseph Liro for their assistance in this paper. The positions presented in this paper are, of course, my own.

References

Bar-Hillel, Y. (1960*a*). The Present Status of Automatic Translation of Languages', in F. L. Alt (ed.), *Advances in Computers,* 1, 91-163. New York: Academic Press.

---(1960*b*). Ibid. 158.

Carbonell, J. G. (1987). 'Knowledge-based Machine Translation'. Presented as a Tutorial at the 25th Annual Meeting of the Association for Computational Linguistics, 6 July 1987. Stanford University, Handout.

---and Tomita, M. (1987). 'Knowledge-based Machine Translation, the CMU Approach', in: S. Nirenburg (ed.), *Machine Translation,* 68-89. London: Cambridge University Press.

Hutchins, W. J. (1986*a*). *Machine Translation: Past, Present, Future*. Chichester, England: Ellis Horwood Ltd.

-- (1986*b*). Ibid. 175.

-- (1988). 'Recent Developments in Machine Translation. A Review of the Last Five Years', in D. Maxwell, K. Schubert, and T. Witkam (eds.), *New Directions in Machine Translation,* 15. Dordrecht, Holland: Foris Publications.

Lehrberger, J., and Bourbeau, L. (1988). *Machine Translation, Linguistic Characteristics of MT Systems and General Methodology of Evaluation. Studies in French and German Linguistics,* 15, 103, Amsterdam: John Benjamins.

Lytinen, S. L. (1987*a*). 'Integrating Syntax and Semantics', In S. Nirenburg (ed.), *Machine Translation. Theoretical and Methodological Issues.* 302. London: Cambridge University Press.

--- (1987*b*). Ibid. 303.

Meya, M., and Vidal, J. (1988). 'An Integrated Model for the Treatment of Time in MT-systems', *Proceedings of the 12th International Conference on Computational Linguistics,* 2, 437-41. Budapest: John von Neumann Society for Computing Sciences.

Weaver, W. (1949). 'Translation', reprinted in W. N. Locke and A. D. Booth (eds.), *Machine Translation of Languages,* 15-23. New York: John Wiley and Sons, 1955.

Yngve, V. H. (1964). 'Implications of Mechanical Translation Research', *Proceedings of the American Philosophical Society,* **108**, 279-81.