

[From: IFTT'89, April 26-28, Oiso Prince Hotel, Japan]

Man Machine Interactions in MT
Jun-ichi Tsujii
Centre for Computational Linguistics, UMIST
P.O.Box 88, Manchester M60 1QD, U.K.
(tsujii@ccl.umist.ac.uk)

1. MT - An Underconstrained Problem

MT or natural language translation in general is typical of "underconstrained" problems which we often encounter in the field of artificial intelligence. That is to say, the same "messages" can and should be translated differently depending on the surrounding context, and on the speakers' intention etc.

It is all too often the case that this information, which is necessary for the selection of the appropriate target expressions or text structure, is not made explicit in source texts prepared for translation. The author of the source text naturally follows the "rules" of the source language in preparation of source texts and assumes that the factors which will affect the selection of target expressions are self-evident.

2. Understanding or Interactions

MT systems developed so far or being developed have been trying to compensate this genuine property of language translation by extending the units of translation from sentences to texts or by introducing "understanding" based on "domain specific knowledge".

This course of research would be inevitable, if we were to confine ourselves to translation of prepared texts, such as computer manuals, which already exist before translation. In such cases, we have to recover from text itself or by using extra "knowledge", such as implicit information which is necessary for formulating target expressions. However, we can imagine a quite different course of research for developing a different type of MT system, i.e. an "expert" system which can play the role of an "intelligent secretary with knowledge of the foreign language". Such a system would not usually require the user (the writer) to prepare full source texts in advance. It would start from rough sketches of what the writer wants to say and would gather the information necessary for formulating target texts by asking the writer questions, because writers are the people who really intend to communicate and have a clear idea about what they want to say.

This sort of system is different from the widely promoted "Translator's Workbench" idea, the main aims of which are to help TRANSLATORS to translate texts. In this scenario, both the system and the user have knowledge about both source and target language, and it is sometimes difficult to see where the most appropriate division of labour should occur: indeed, there is sometimes a conflict between what the system offers the translator-user, and what the user already knows, or between the extent to which the system or

the user should take the initiative, which might differ from occasion to occasion. On the other hand, in the proposed expert system scenario, the partition of knowledge is clear: the system knows mainly about translation, the writer knows only about the desired communicative content of the message. There is no conflict between what the system assumes to be the extent of the writer's (the user's) knowledge, nor in the writer's expectations.

3. Translation as Information Retrieval

Notice that the concept of "source text" in the above context of "intelligent secretary" is quite different from that in the normal context of MT. That is, we don't have a source text to translate as such, but instead, the user has his communicative goals and the translation system can help to formulate the most appropriate target linguistic forms by gathering information necessary to accomplish these goals through interactions.

MT systems so far have been developed based on the implicit assumption that source texts contain all (or almost all) information to be translated. But this assumption is simply not true, especially when we consider such a language pair as English and Japanese.

In the above scenario, the system tries to gather information necessary for formulating target texts through interactions. This means the system formulates target texts by adding information to "source texts" (in a conventional sense). We can extend this idea further. In the extreme cases, we can imagine a system which have stereotypical target texts in certain restricted domains (e.g. business correspondences in specific areas), retrieves appropriate texts through dialogues with users and reformulates them to fulfil the specific requirements expressed by users. In this scenario, MT systems become a kind of information retrieval systems and add a lot of information not contained in "source texts" at all.

4. Conclusions

Considering the genuine property of language translation, we cannot expect to have full automatic, high-quality MT. We have to develop systems which allow flexible and effective human interventions. Though the ideas in this paper are very speculative at this moment, my contention is that we have to explore diversified approaches to interactive MT and an interactive system for monolingual users is worth for exploration.

Acknowledgement

I have been inspired very much by the discussions with colleagues of UMIST during the preparation of this paper. Especially, the discussions with Dr.H.Somers, Mr.J.McNaught, Mr.D.Jones and Mr.J.J.Carroll have been extremely useful for the preparation of the paper.