The Development and Use of Machine Translation Systems

Makoto Nagao
Department of Electrical Engineering
Kyoto University

1. Introduction

The development of machine translation systems has reached the stage where some commercial systems have become available. These are the products of their manufacturers' creativity, along with technology transferred from universities/research institutes that have been studying the process for more than a quarter of a century. Some users of machine translation systems say they can make more profit by using the systems than without them. Generally speaking, however, the systems are quite imperfect and are not easily used by ordinary persons.

In the face of these imperfect machine translation systems, there are several factors we must consider. First, we must improve the systems. We must discern the directions in the near future that will be most fruitful for practical machine translation applications. Second, we must become accustomed to the use of present-day machine translation systems. Then we may be able to find the proper application areas where we can earn a profit by utilizing the system. The third factor is somewhat related to the second. We must realize that machine translation systems are not biological systems, but engineering systems. This means that the system will work properly only within the range of the design specifications. Therefore, we must be very careful about what kinds of sentences and texts are fed into the system. The sentential structures and the text written must be clear, so that anyone can follow the context without deep inferential procedures.

The final point to be stressed is, of course, intensive research for machine translation systems based on completely new paradigms. These paradigms must have human-like flexibility in interpretation and understanding, and also human-like learning capability. More than ten years will be required for such research to become incorporated into practical systems.

We cannot, however, idly wait for more than ten years for the completion of such an ideal system. We must make all-out efforts to facilitate items one to three mentioned above, and we must answer the needs of the rapidly growing translation market under both present and future international circumstances. And, in the early period of the twenty-first century, when a new, human-like machine translation system has been realized, gradual replacement will take place from the systems based on the extension of the present technology to the remarkable system we anticipate.

2. Utilization of Existing Systems

Pre-editing the input text and post-editing the machine-translated materials will be unavoidable throughout the near future. There are many text materials that are incomprehensible even to human beings. If the sentences are

sufficiently simple and clear to represent ideas, pre-editing will not be required. Unfortunately, there are very few such cases. Raw translation by humans must be checked by a career translator. When we see the details of Japanese-to-English translation in Japanese translation companies, we are astonished by the difference between the raw translation by a Japanese translator and the post-edited text by a native speaker of English. We can imagine how painful and time-consuming post-editing is to the post-editor. We sometimes cannot discover similar phrases between the raw translation and the final post-edited text.

Such is the reality of human translation in Japan, so that post-editing is a routine matter in machine translation.We have, however, some exceptional cases. For example, when the translated materials are (a) perused by topically limited persons who read merely to acquire new information as quickly as possible, (b) throw the material away after reading, or (c) when necessary will read the articles in detail. In such a case, no editing will be required, and the present-day machine translation systems are quite useful.

Machine translation systems must be evaluated according to a wider environment. This means from the time when the texts to be translated are handed over, to the time when the required translation is returned to the customer. Whether manual translation is better, or a process including machine translation is more profitable, must be judged not only by the total cost for the entire process, but also by the speed and the quality. This includes whether special terms are translated properly throughout the entire text.

These are judged according to profitable points such as speed, translation of special terms by the preparation of terminology dictionaries, etc. One factor that we must keep in mind is that no machine translation system is usable immediately after its introduction—unlike a washing machine that works perfectly anytime and anywhere. The translation system must be tuned and improved very carefully to accommodate the properties of the text we wish to translate. Therefore, companies that intend to introduce a translation system must appoint persons who are dedicated to the system's operation and improvement. One or two years of effort will be necessary to draw out the abilities of the system. Without such efforts, the introduction of a machine translation system will result in complete failure.

3. Looking Toward the Future
It is evident that translation quantity will increase dramatically as international communication becomes more frequent through the continuous expansion of international air traffic and telecommunications networks. Machine translation systems are the only way we can think of to cope with such an expected increase of translation quantity.

As for the translation of technical manuals and similar documents, the improvement of present-day machine translation systems—all of which are based on the compositionality principle—will be the best way for creating dictionaries, grammatical rules, and guidelines for writing technical documents precisely, and for widely promoting such technical writing methods throughout society. As for business letters, news articles, and related texts, the best method is to establish prototypical phrases, sentential styles, and text structures, and to prepare typical translations for them.

The flexible utilization of such prototypical examples is quite practical for the translation of these document categories. We must, however, develop a fanciful "translation and editor" system that will not be difficult to use.

Dialog sentences and actual utterances have varieties of ungrammatically, and are quite difficult to translate by machine. We must conduct very basic research on analytical methods for imperfect input, on knowledge information-processing, and on natural-language understanding—as well as on basic linguistic theory studies. On the other hand, we must not forget such domain-specific machine translation systems as TAUM METEO, which appear to be the shortest avenues to success.

Multilingual machine translation is becoming more and more important, but is expected to be far more difficult than a machine translation system of fixed language pairs. International cooperation is essential. Dictionaries, grammatical rules, and software systems for machine translation—all of these are to be developed by mutual agreement among different countries. We must make efforts to find proper international cooperative systems for this purpose.

Reference: Makoto Nagao -- How Far Can It Go? Oxford University Press, 1989 (Translated from the Japanese Original, Iwanami, 1987)