TWO YEARS AFTER THE MT SUMMIT (MT Summit Keynote Speech)

Makoto NAGAO

Department of Electrical Engineering
Kyoto University
Kyoto, Japan

1. THE ITEMS CLARIFIED IN THE FIRST MT SUMMIT

The purpose of the First MT Summit, which was held in Hakone, Japan in the autumn of 1987, was to discuss the problems of machine translation from the following viewpoints: (a) to make clear the basic research topics in machine translation; (b) to consider ways for users to accept machine translation after eliciting the opinions of users of the machine translation system; and (c) how governments should support the development, introduction, and utilization of machine translation systems.

Our expectations for practical results from the First MT Summit were not extravagant because this was the first conference held from such viewpoints. But in this summit it was possible to grasp the state of the art of the research supported by the governments in each country, the opinions and the outlines of MT technology, and the ways to use MT systems. These were great achievements.

As a result, the significance of the conference was recognized and the holding of a Second MT summit in Germany two years later was suggested by Mr. Czermak of BMFT. His suggestion was actually facilitated, resulting in this wonderful conference in Munchen. I would like to thank Mr. Czermak and Professor C. Rohrer—the chairman of the conference committee—and the other participants for the success of this event.

In the First MT Summit, the following matters were clarified. In Japan, the Mu-2 Japanese-to-English machine translation system for technical documents began at JICST. CICC began to construct machine translation systems between Japanese and the languages of neighboring countries—specifically China, Thailand, Malaysia, and Indonesia. The Japan Electronic Dictionary Research Institute Ltd. was established, and the compilation of large-scale dictionaries was begun for use in machine translation and computerized natural-language-processing. Many machine transla-

tion systems have been developed and marketed by private corporations. These systems are currently beginning to be used on a large scale by some corporations.

In Europe, in addition to SYSTRAN, Siemens was continuing efforts to put the system to practical use. EUROTRA and other projects were also being developed. In Canada, although there were various attempts after the success of TAUM METEO, it seems that the development of a technical-term dictionary was emphasized from a practical approach. In the United States, there were a few commercial systems such as SYSTRAN, LOGOS, and ALPS, but there was practically no development except at Carnegie-Mellon University. The United States Department of Commerce is the major force in considering the introduction of a Japanese-to-English translation system for the facile translation of Japanese technical documents. There were doubts, however, as to whether or not usable systems exist. In Southeast Asia, some studies have commenced, but the work is still in a preliminary stage.

The changes and developments in this field since the First MT Summit are not outstanding. But it is significant that steady progress is being made. For instance, in the past, major MT systems in Japan could only be used by large computers. At present, however, all the systems can be used by personal computers. Cost performance has been improved substantially and more efforts are under way for improvement of the quality of translation. Also, systems have been developed for pre-editing the sentences to be translated and for post-editing sentences after translation. By attaching an OCR (Optical Character Reader) to an MT system, enhanced efficiency is being achieved. Printed Japanese texts can be recognized automatically with an accuracy of more than 99 percent. The pre-editing and post-editing systems have been tested for practical use with the high-quality printout system. The number of users of MT systems have been increasing gradually, and there are some cases where a cost saving of approximately 50 percent was realized using an MT system.

With the progress of MT technology and the intensification of user interest, an International Forum on Translation Technology—*Harmonizing Human Beings and Computers in Translation*—was held in Oiso in April of this year. A primary goal of the forum was to interface users and makers, and to enhance understanding among users who had doubts or lacked interest in MT systems. In this conference, we took up the ALPAC report published in the United States in 1966, and discussed how the report can be evaluated in the present situation. The conclusions of the discussions will be reported in the other session of this conference.

The following points were clarified. The viewpoints and evaluations of today's machine translation have been radically changed by recent developments. The old and simple-minded view

that "machine translation is impossible; therefore, the study and development of an MT system are meaningless as clarified by the ALPAC report," is no longer valid. We should recognize that the MT system is usable, and that we must contribute to the worldwide circulation of information by developing machine translation, because that goal is unachievable without machine translation.

Other objectives of the conference were to investigate and estimate the modes of the MT system's usage and Japan's translation market, and to discuss future demand for machine translation. By making users aware of the actual condition of present MT systems—which are quite imperfect—they understood the effectiveness of the system, depending on the use. When provided with proper advice on methods of use, users can make an intelligent decision about the introduction of an MT system.

Also, by listening to the experiences and opinions of actual users, the makers of the system understood the factors required for an improved version. Four hundred persons participated in the conference, about half of whom were MT users. In conjunction with the conference, fourteen companies presented exhibitions of MT systems and related technology, enabling the participants to obtain heightened recognition of the present MT systems.

Items clarified at the conference included the following.

(1) Although the present MT systems are imperfect, they are economical, depending on the application, when compared with human translation.

(2) To use MT systems practically, it is important effectively to execute pre-editing and post-editing. When a mass of documents are translated, division of labor in such editing is particularly effective.

(3) The compilation of a technical-term dictionary has been under way, but additional term dictionaries are required in many specialized fields. Also, there are often cases where the party ordering the translation specifies the translation words to be used, so that the system must be applicable in such a situation.

(4) We should develop a system whereby the user can input the experiences of machine translation post-editing, a system that could be improved not only by makers but also by users.

(5) The experiences of the users of MT systems should be studied carefully. Case studies must be conducted from the standpoints of (a) how long the trial usage continued after the introduction of an MT system before obtaining a profit through normal operation; (b) how extensive were the efforts of dictionary tuning to the texts to be translated; (c) how many people are engaged in the system's maintenance; (d) what is the speed of machine translation, pre-editing, post-editing; and so on. Such information is invaluable for prospective users of MT systems.

(6) More compact and portable MT systems that include word-processing capabilities will comprise the next-generation system.In such a system, interactive machine translation will be adopted and the system will be used not only by translation companies, but also by private users.

(7) In the past, many good technologies for machine translation were developed at research sections. Many of these have already been transferred to the industry and have been incorporated into todays MT systems. At universities and research institutes, however, we must devote much more study to the basic mechanisms of human translation. Such study includes basic research on theoretical linguistics, cognitive science studies of translators' behavior, and a computational model of language understanding. As themes for the study, the following should be considered: the use of knowledge; the use of artificial intelligence technology; the addition of a learning function; the automatic use of post-editing records for improving grammatical rules; the use of such semantic information as a thesaurus and bilingual-example sentences; elucidation of the mechanism of natural-language understanding by the use of contextual information; the study of parallel processing hardware and software; the development and use of a large-scale dictionary; etc.

(8) To translate properly and correctly is the paramount concern not only for machines, but also for humans. This subject depends on whether the translator properly reads and understands the original texts. This is the most important factor in translation. Accuracy requires that the original sentences should be written without vagueness and with no possibility of misunderstanding. We must establish a method of technical writing that gives necessary and sufficient information to readers. Or, taking a step forward, limits should be given to the linguistic expression to achieve correct translation—especially by machine. As a result of such restrictions, MT systems would become completely practical. A psychological study concerning such linguistic limitations is vital, because writers must be able to compose sentences easily within the predefined limits.

For the future development of MT systems and the proliferation of MT users, the participants in the IFTT considered that users and makers should exchange information and ideas, and create a situation where everyone can benefit. Accordingly, we agreed upon the necessity of creat-

ing an entity such as an International Association of Translation Technology, or an International Association of Machine Translation, where we can exchange information and opinions about the practical use of MT systems.

Considering the progress of the past two years, the items to be discussed, to be clarified, and to be agreed upon in this Second Summit should include the following.

(1) To understand sufficiently the state of the art of MT systems in 1989. It is quite important to know what companies sell what kind of MT systems, what are the specific areas fitted to these systems for translation, what is the mode of uses of these systems, how perfect or imperfect are the systems, and so on.

(2) To discuss ways for a wider circulation of the MT system. For what kind of users is the introduction of an MT system effective, what kind of organizational systems should such users build in introducing and using MT systems, and how many years and what kind of efforts will be considered for extending MT use and its effectiveness after its introduction? For each of these topics, we should make case studies based of varieties of MT users. Then we will be able to achieve the smooth introduction of MT systems and their effective uses.

(3) Examining the compliance of users with the MT system, we should make clear the parts that should be improved. Also, it is necessary to consider changing the systems into ones that users can improve for themselves.

(4) We should make clear the faults of the present MT system and discuss technical methods for overcoming the faults. Especially needed is a controlled language in which to express the original sentences properly, as well as efforts to popularize the language.

(5) To make a basic study concerning natural language for translations with a better quality than is achieved by the present MT system. The direction of the study and its theme should be made clear, and cooperative study should be conducted among computational linguists, theoretical linguists, and translators.

(6) To make a wider-range basic study about natural-language-processing. The study of machine translation is important not only for itself, but also for the information society of the twenty-first century. This is because the study provides basic technologies for man-machine interfaces, information retrieval, interactive systems, etc.

(7) International cooperation is necessary for machine translation. Each country must cooperate and promote rapport between MT users and MT-system makers. Accordingly, the establishment should be considered of an International Association of Translation Technology, or an International Association for the Promotion of Machine Translation.

REFERENCES

1. *Kikai Honyaku Summit (*Machine Translation Summit) by Makoto Nagao (ed.), Ohmsha, 1989.

2. *Proceedings of the International Forum on Translation Technology,* Oiso, April 1989.

3. The JEIDA, (ed.): *A Japanese view of machine translation in light of the considerations and recommendations reported by ALPAC,* U.S.A., April 1989.

4. *Kikai Honyaku wa Dokomade Kanouka* by Makoto Nagao, Iwanami Shuppan, 1986.2. (English Translation: *Machine Translation: How far can it go?* Oxford University Press, 1989.)