

## 7 • 1 Report on Machine Translation Summit

Makoto Nagao  
Kyoto University, Japan

The Machine Translation Summit was held for three days from September 17th to 19th, 1987, at the Hakone Prince Hotel. The Summit was sponsored by JEIDA (Japan Electronic Industry Development Association) and supported by MITI, ICOT (Institute for New Generation Computer Technology), CICC (Center of the International Cooperation for Computerization) and EDR (Japan Electronic Dictionary Research Institute). A total of 200 participants, of which 85 were from 15 foreign countries and 115 were Japanese, took part in the Summit. Although the Summit was little advertised both domestically and abroad, a number of requests for participation from throughout the world were received even shortly before the scheduled dates. Nevertheless, because of the limited capacity (200 maximum) of the hall as well as of the accommodations, some of the requests had to be turned down, with much regret, received only a few days prior to the opening of the Summit.

The meeting was called "Summit" since one of its goals was to examine machine translation from a truly international point of view, inviting a number of participants from U.S., Canada, Europe, China, South Korea, India and many other southeast Asian countries as well as Japan. Furthermore, the "Summit" was intended as a place where participants from a variety of fields, such as, technological engineers, users, and policy decision-makers at the governmental level can discuss the present and the future of the machine translation.

### History of Machine Translation

This section briefly describes the background

which led to this Summit and traces the history of researches and developments conducted in the area of machine translation. In Japan, researches in machine translation began as early as in 1957 at Denki Shikenjo (present Electro-technical Laboratory) and universities. The attempt however, was largely limited to basic research without any practical application. In 1977, a success of the Canadian meteorological center in providing a 24-hour English-French translation service was reported. Moreover, around the same time, the European Economic Community started considering a large-scale project for development of multi-language machine translation. Prompted by such active involvement in machine translation abroad, the importance of conducting project-oriented researches beyond basic level, as well as of positive involvement of private firms, were greatly stressed in Japan. Under such circumstances, the Machine Translation Investigation Committee which was established within JEIDA in 1981, began to examine technologies of machine translation both domestically and abroad, discuss various problems inherent in R & D and market development and report the results of their achievements. The Committee's active engagement in such tasks continues up to present.

Concrete examples of machine translation systems include a system developed in 1979 which was capable of translating, to a significant extent, Japanese to English and vice versa, the titles of scientific and technological papers. As a result of this development, the Japanese government recognized the need and reached a decision to develop, under the leadership of Science and Technology Agency and with cooperation from MITI and Ministry of Education, an English-to-Japanese and Japanese-to-English translation system for abstracts of scientific papers. This government-

supported research, conducted for four years since 1982, had proven the potential practicality of the system. Presently, the earlier system is being modified for practical application. Stimulated by such a government endeavor, a number of private firms launched on similar projects, resulting in the recent development of a series of commercial translation systems publicized by different companies.

The member countries of EC joined so-called EUROTRA project which was organized by EC for the development of multilingual machine translation system. The project which started its preliminary consideration in 1978, was faced with serious difficulty once it launched on its attempt to develop a system capable of mutually translating 7 languages to start with, and later 9 languages, after the new membership of Spain and Portugal. To make matters worse, most of those from different European countries who participated in the project, were new to the field and therefore required education in basic principles of machine translation. The largest difficulty, however, was generated from the disagreement among the member countries as to the basic structure of the system which they were to develop in cooperation. Indeed, almost the entire preliminary period which started in 1978 was devoted to the discussion of this subject alone. The disagreement on details which continued even during the period of the First EUROTRA Plan (1982 -), caused some of the most competent but frustrated researchers to resign from the project, and as a result, the project was significantly reduced in its scale. The experiment conducted at the end of the First period involved merely a toy-like system of a mini scale. Presently, as of September 1987, the achievements of the First period are reviewed and evaluated so as to reach some kind of conclusion by October regarding the question of whether the project should be continued on to its Second period, and if so, what the goals of the Second period should be.

In the United States, on the other hand, the ALPAC report submitted in 1965, stated the impracticability of such a system, which resulted in suspension of any government funding to university researches in the field. Consequently, to date, there has been no specific researches in machine translation conducted in

the U.S. higher institutions. However, those who previously worked in universities moved into private sector and succeeded in developing a commercial system called SYSTRAN which has been adopted at several places since 1973 with additional improvements made afterwards. Since then, although commercial systems such as LOGOS, SMART and ALPS were developed, most of their applications have been limited to particular group of users. Nevertheless, stimulated by the efforts made in Japan and in Europe in the areas of machine translation, the Machine Translation Center was recently established at Carnegie-Mellon University for the purpose of showing possibilities of multilingual translation within specific domains. A similar attempt was also started at Georgetown University, the center of machine translation researches, and the fact that M. Zarechnak who used to be a prominent figure in the field in 1960's has resumed his researches deserves much attention.

### **Current Situation of Machine Translation System**

Let's consider machine translation from the viewpoint of a person (or a company) wishing to have one's papers machine translated. Given a tremendous volume of papers requiring translation, manual translation, as opposed to machine translation, resulted in a number of inevitable problems, such as, low speed, inconsistency both in style and word usage and mistranslation due to translation performed by a number of different translators who do not (or cannot) sufficiently understand the content of material. Indeed, there has been a number of cases where the users trying to work by reading a manually translated manual, complained of the poor quality of the translation which was found incomprehensible, and returned the manual to the supplier. Therefore it is quite natural for those companies who had been victimized by erroneous manual translation to hope to solve the past problems through the development of machine translation. Despite such high expectations, available machine translation systems are not capable of fulfilling such needs. First of all, it is very difficult for a newly introduced system to immediately translate such documents. That is, for a

system to actually launch on practical translation, it is necessary to collect technical terms used in a given document area, give each of them an appropriate translation and register the translated terminology in the dictionary. It takes tremendous time and money to make a good dictionary covering several tens of thousands of words. Furthermore, while general terminology is already registered within a system, the terminology has to be changed according to different subject areas of the materials to be translated.

Similar problems are also inherent in the syntax of a system. While average syntax is readily available within a system, any syntactic diversions found in the material must be reincorporated through addition of new syntactic rules as well as modification of what is already available. Therefore, for the first couples of years since the introduction of a system, the user is obliged to engage in those above-mentioned tasks in cooperation with its manufacturer. The performance of the system after completing such tasks, however, remains to be seen. In the case of SYSTRAN French-English translation system introduced by EC, after several years of initial adjustment period, the system performance was proven fairly satisfactory, which was partly due to the similarity of the two languages, and provided that the materials for translation are carefully screened.

In the recent years, there has been a considerable progress in R & D of machine translation in Japan, with a number of private firms investing a significant amount of money and labor for the purpose. Already 10 or so Japanese companies have announced their machine translation system and at least another 10 or so are engaged in the development of related systems. The prospective users, however, appear to have a mixed feeling about possibilities of machine translation. On one hand, they generally have high expectations and high regard for the possibility of a machine performing an intellectual activity similar to that of a human. However, on the other hand, those prospective users having an assumption that such a marvelous system has been already perfected to the level equivalent to human intelligence, are largely disappointed with the poor quality of the state-of-the-art performance of

currently available systems. In this sense, general public is regarding machine translation systems with a kind of suspicion, and this fact puts a serious constraint on the potential machine translation market.

### **Goals of the Summit**

Under such circumstances, prior to the organization of the Summit, following goals were considered essential for the future of machine translation:

- (1) To elucidate present technological level of machine translation systems and relevant technological problems to be solved in the coming years.
- (2) To have potential users fully recognize the present level of the system and understand the fact that apparent imperfections are surmountable through carefully-devised application and that the system can be highly practical from the cost-effective point of view.
- (3) To clearly state the indispensability of governmental and other research assistances for a wide range of R & D to be conducted in the areas of linguistic information processing so as to improve on currently available commercial systems which are, in effect, far from being perfect.

It was therefore considered particularly important in Japan to have the above issues widely understood by individuals in different fields in order to realize full-fledged development of machine translation market. With regard to European countries, we hoped to have them recognize the level of machine translation technology available in Japan, while at the same time, have them fully realize the necessity of further advancement as well as development of the EUROTRA project without being hampered even in the face of another negative report similar to ALPAC. For the United States and Canada, we also hoped to have the two countries recognize the available Japanese technology in machine translation and have the Summit serve as an incentive to resume and promote similar researches back home.

## **Circumstances leading to the organization of the Summit**

Included in various remarkable operations conducted by the Machine Translation Investigation Committee of JEIDA since its establishment, was annual investigation of overseas machine translation technology and reporting the results to the registered and unregistered general companies to provide them with information on international technological standards in the areas of machine translation. Particularly in recent years, much of the accomplishments made by such an operation centered around research-based exchanges with international experts of the field. Below presents a list of cooperative research and information exchange programs organized abroad by the Committee.

1982: A 2-day workshop, by inviting a research group of EUROTRA, was held in Brussels, Belgium. We presented R & D in machine translation that were being conducted by Japanese universities, research institutes and private firms at the time, and learned the outline of the EUROTRA project in exchange. Very meaningful discussions took place at the workshop, with the presence of more than 10 central figures of the EUROTRA project and as many Japanese researchers.

1984: A 2-day workshop was held among researchers from the U.S., Europe and Japan at the Stanford University. The workshop ended in particular success with the presence of approximately 50 major researchers in the field of machine translation who took part in the International Conference of Computational Linguistics which was held just prior to the workshop. The participants discussed in detail and exchanged their opinions on available basic technologies for machine translation concerning the syntax, dictionary, software, system configuration, etc.

1985: Another 2-day workshop by inviting a research group of EUROTRA was held in Geneva, Switzerland. We presented practical machine translation systems developed by different Japanese companies to illustrate the situation in Japan and the EUROTRA group reported on the results of their 3 years of research in machine translation. While the workshop

proved beneficial for both groups, the presentation of the situation in Japan appeared to have made a particular impact on the EUROTRA researchers whose research attempts were often subject to disagreements among the participating countries.

As a result of such activities organized by the Committee, researchers of machine translation throughout the world started to show much interest in the accomplishments made in Japan. In fact, we received a number of requests from foreign researchers to hold an international conference of machine translation in Japan, on the basis of the results of the workshops.

## **The Nature of the Summit**

Under such circumstances, the Machine Translation Summit was organized under the sponsorship of JEIDA. The major objectives of the Summit were not confined to mere discussion of various technological problems in machine translation. Instead, the Summit was intended as an opportunity to examine and discuss a wide range of issues concerning the future of machine translation from a variety of points of view, with an active participation by users of the systems as well as government offices willing to supply financial assistance for future researches. With much understanding of such intended purposes, MITI kindly provided us with generous financial contribution for the organization of the Summit. In addition, we were able to have the Director of Industrial Electronics Division of MITI, Mr. Honda give an opening address on the manners in which Japan as a country engages in the researches of machine translation and Deputy Director, Mr. Kubota of Industrial Electronics Division of MITI participates in the panel discussion. All of such cooperation we received from different sectors contributed in making a strong impression on the world as to Japan's positive attitude towards the promotion of international exchanges in machine translation.

Another machine translation project involving Japanese and other Asian languages was started this year. CICC, the organization in charge of the project.

also largely contributed to the Summit with their presentation of specific plans of the same project. Owing much to this CICC project, we were very pleased to have so many participants from other Asian countries as well, making the Summit "international" in the true sense of the word. Other accomplishments made at the Summit included; ICOT director, Dr. Fuchi's explanation of intended purposes of developing inference machine for the fifth generation computers as well as a detailed description of activities conducted by EDR for further development of machine translation systems in Japan. All of such presentations given at the Summit have made a large impact on the participating European and U.S. researchers involved in similar fields.

Although the Summit was rather compact in scale with 16 invited presentations, 2 panel discussions (with 18 panelists) and 3 reports on final conclusions during its three-day holding, each of the program has been proven truly meaningful. In particular, one of the initial presentations made by Hutchins of Great Britain was unquestionably rich in content covering almost all of the problems involved in machine translation from a wide range of perspectives. The presentation should serve as an important point of reference for anyone conducting the researched in the future.

One of the very unique characteristics of the Summit lied in its organization of a special session where a presentation of a system by a manufacturer was followed by that of a user who had actually worked with the system. The session which successfully illustrated how some of the representative Japanese machine translation systems are being introduced and adopted in user companies, and what specific requests such companies have for the manufacturers, was able to provoke much interest on the part of the audience. Such a format, tried for the first time in Japan and with few precedents abroad, proved especially significant and is well worthy of being regarded as one of the most remarkable successes of the Summit.

Another new attempt made during the Summit was the presentation of the ways in which governments of different countries engage themselves in the researches in machine translation. In addition to detailed reports on the present condition of EUROTRA

and the nature of the Japanese ODA plan, government involvement in similar researches in the Southeast Asia and other foreign countries, their future direction and the types of expectations they have of Japan were discussed at the Summit. We believe that such an opportunity enabled governments of different countries to gain additional understanding of their expected roles in the development of machine translation. The importance of such an information exchange was also demonstrated by active exchange of opinions and many questions posed by the participants of this particular panel discussion.

Among a number of reports presented on the technological standards of machine translation in different countries, the Japanese presentation which demonstrated sophistication and thoroughness of its state-of-the-art technology, was proven especially successfully in making a strong impression on other participants from abroad.

On the final day of the Summit, three reports were made on the results and conclusions attained through the Summit. Dr. M. Kay of the United States concluded by stressing that although he admitted the fact that machine translation was by far still a technology of the future, with a development of technology to support well-devised optimal combination of human and artificial intelligence, the system could be realized despite the limited standard of technology that is available today. Mr. L. Rolling of EC summarized, on the basis of his experience with the SYSTRAN over the years, the factors that are essential in determining the future development of machine translation systems. Finally, I was given an opportunity to state that the machine translation system which comprises merely one facet of a gigantic system requires overall planning which is appropriate for the development of syntax, dictionaries, etc. for the purpose of controlling general system development and that there are number of research projects to be conducted to have a better understanding of linguistic characteristics. At the end of my final report, I stressed that while the commercial systems might have to face a number of difficulties in the next 2 or 3 years to accomplish additional improvements, as is illustrated in the diagram covering the past, present and future of the machine translation, the

market is apt to gradually expand and that we already see light at the end of the tunnel.

### Results of the Summit

At the Summit 9 Japanese companies and 1 university gave a demonstration of their respective translation machine. It was particularly effective to have not only the foreign participant but also the Japanese participants representing the user side, actually see different machines translate various sentences of their own choice, in a large single room.

Some words should be also said about the site of the Summit. As is mentioned earlier, the Summit was held by reserving almost all of Hakone Prince Hotel. The superb atmosphere of the Hotel enabled enjoyable stay for all of the participants. Although unfortunately we had much rain during the Summit because of a typhoon approaching the nearby regions, it stopped raining on the very evening of the scheduled banquet. We held a cocktail party on a boat sailing around the Ashinoko Lake, which helped to enhance the marvelous atmosphere of the dinner which was held afterwards.

One priceless harvest of the Summit lies in its favorable effect on those Japanese working in private companies who had been engaged in the development of machine translation system over a number of years.

The Summit provided those Japanese engineers who were often faced with many difficulties and hardships, a chance to regain confidence in what they had been doing as well as to see some light in the future of their endeavor despite the severe competition. The Summit made it clear that Japan is the country to lead the world in the field of machine translation. It gave another hope to the participating users by indicating various technological and market potentials through new applications of machine translation systems. As for the participants from abroad, as it has been repeatedly stated in this paper, the Summit has served as an opportunity for them to fully recognize the actual technological standards of Japan, while at the same time, obtain much stimulation for enhancing their R & D efforts back home. These many positive results of the Summit are substantiated by the closing remarks given Mr. Czermak, a representative of Ministry of Science and Technology, F.R.G. In his remarks, Mr. Czermak expressed his gratitude to Japan as well as his strong determination to make best effort to realize in 2 years, a similar Machine Translation Summit in West Germany.

Finally, the author would like to close this paper by expressing his sincere gratitude to everyone who contributed to the success of the Summit, particularly, Mr. Suzuki, Executive Director of JEIDA and Mr. Honda of MITI and many others. Furthermore, the author would like to deeply thank the members of the Committee on Machine Translation Systems of JEIDA for all their time and effort in preparing for the Summit for a period of over one year.

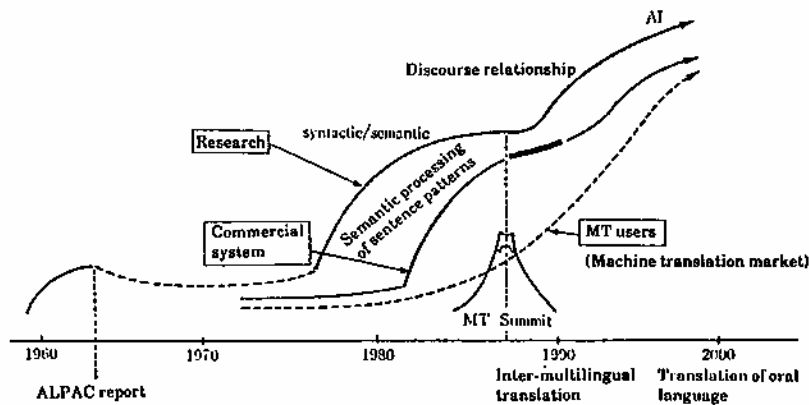


Fig. 7-1 Development of machine translation