

MT SUMMIT VI: SPECIAL REPORT

Contributors:

Alan Melby and Muriel Vasconcellos



Alan Melby

Machine Translation Summit VI, held at the Catamaran Resort Hotel in San Diego, California, from October 19 through November 1, 1997, was a gala affair celebrating 50 years of machine translation.

The Summit was organized this year by the Association for Machine Translation in the Americas (AMTA) on behalf of the International Association for Machine Translation (IAMT). The Information Sciences Institute of the University of Southern California served

as institutional host, and the conference had a total of 21 sponsors, including eight US Government agencies, three associations, and six corporations.

The biennial MT Summits are held in rotation in Asia, Europe, and the Americas. Six years ago, at MT Summit III in Washington, DC, the series acquired formal structure through creation of the IAMT and the three regional associations that would be responsible for organizing the conference when it came their turn. Summit VI brought this initiative full circle. The IAMT chose to mark this occasion by honoring its founder, Professor Makoto Nagao, with the first IAMT Award.

MT PAST: TALES FROM THE PIONEERS

One of the highlights of the conference was the panel "The Pioneers of Machine Translation," which featured eight of the early figures in the field. The session opened with a video greeting from Andrew Booth, whose involvement in machine translation dates back to the very beginning, half a century ago. Most of the pioneers who were present (including the panel chair, Muriel Vasconcellos) had associations with MT that went back at least 40 years.

Victor Yngve's assignment as an MT researcher at the Massachusetts Institute of Technology in 1953 made him the first person to be employed full-time in the field. His table-driven syntactic parsing scheme published in 1955 is considered to be the first published translation scheme that was completely explicit and programmable. During his 12 years at MIT he founded the journal *Machine Translation*, which published over 1,000 pages of contributed papers on MT from 1955 to 1968, and co-founded the Association for Machine Translation and Computational Linguistics (later the Association for Computational Linguistics). Yngve reviewed highlights from the International Conference on Machine Translation held at MIT in June 1952 and showed photographs of the participants and other early MT-ites. He indicated that he is still attempting to work out a suitable theoretical basis for machine translation and other tasks in natural-language processing.

Georgetown University was represented by Christine Montgomery, who began her lifelong career in machine translation working with Paul Garvin, the linguist responsible for the famous Georgetown-IBM experiment in 1954; Peter Toma, who worked on the General Analysis Technique (GAT) with Michael Zarechnak (see *MT Summit VI Proceedings*, p. 31) and a decade later invented Systran, the first commercial MT system; and A.E.R. (Tony) Brown, who worked independently at first and later reprogrammed the entire GAT system. Also from the Georgetown project was Roger Heller, head of the data-entry center in Frankfurt, Germany, where text and dictionary entries were keypunched onto cards that were subsequently airlifted back to Georgetown.

MT PRESENT

Of course, the bulk of the presentations were aimed at giving a view of machine translation today. The overwhelming majority of invited talks and refereed papers addressed the practical aspects of using MT in real settings. The titles included "MT from the Production Perspective" (Mary Flanagan of CompuServe and Frederike Bruckert of Logos), "Managing Distributed MT Projects" (Jennifer Brundage et al.), "Translating Scientific Texts Using MT and MAT Systems" (Vladimir Petrov for Olga Bezhanova), "Sharable Formats and Their Supporting Environments for Exchanging User Dictionaries among Different MT Systems" (Shin-Ichiro Kamei, et al.), "Production MT in the Public Agencies" (Dimitri Theologitis of the European Commission and Joel Ross of the US Foreign Broadcast Information Service), "MT as a Commercial Service" (Terence Lewis et al.), "Using MT in a Corporate Setting" (Lou Cremers of Océ and Christine Kamprath of Caterpillar), and "The MT Market Perspective" (L. Chris Miller of Multilingual Computing Solutions, Colin Brace of Language Industry Monitor, and Yuko Shimamoto-Caleo of Fidelity Management & Research).

MT FUTURE

The future trajectory of machine translation begins, of course, with the research and development going on today. In keeping with Summit tradition, reviews of current R&D provided an overview of the work that is being undertaken—and the paths being charted—around the world: "MT R&D in Europe" (Peter Quartier of Lotus and Roger Havenith of the European Commission), "MT R&D in Asia" (Hozumi Tanaka of the Tokyo Institute of Technology and Keh-Yin Su of BCG), and "MT R&D in North America" (Eduard Hovy of the University of Southern California Information Sciences Institute and Elliott Macklovitch of the University of Montreal), while the session on

"R&D for Commercial MT" (Gregor Thurmair of Gesellschaft Multilinguale Systeme [GMS] and Laurie Gerber of Systran) gave special attention to the commercial aspect.

Other ongoing research was reported under the headings "Machine Translation through Language Understanding" (Makoto Nagao of Kyoto University), "JEIDA's Bilingual Corpus and Other Corpora for NLP Research in Japan" (Hitoshi Isahara of JEIDA), "Multilingual Spoken Dialogue Translation System Using Translator-Drive Machine Translation" (Hideki Mima et al. of ATR Interpreting Telecommunications Research Laboratories), and "User-Friendly Machine Translation: Alternate Translations based on Differing Beliefs" (David Farwell and Stephen Helmrreich of New Mexico State University).

MT users were asked what developments they would like to see in the future, and their answers were heard in the session "What Do Users Need from MT Systems?" (Edith Westfall of Multilingual Computing Solutions and Michael Marubio of Logos). Bentz Maegaard's "Whither MT?" provided an excellent summary of future trends that can be expected, based on what she sees as the "pull" of the market.

CONCLUSIONS

Conspicuously absent was the announcement of any fundamental breakthroughs in how to improve the quality of machine translation of dynamic general language. But somehow this does not seem to matter as much anymore. The big news items in the field are the effect of the Internet, the integration of various techniques, and price reductions.

The growth of Web sites in languages other than English has produced a sudden demand for "gisting" tools that produce "indicative" translations of Web pages or blocks of text from a site.

Publication quality is not essential, but instant turnaround is all-important, so machine translation fits well with this task. Another application of "down-and-dirty" machine translation is the US Army Research Laboratory's new FALCon system, shown at the conference as a tabletop demonstration. FALCon converts foreign-language documents into an approximate English translation, allowing US forces to assess the military significance of documents captured in the field and determine whether to pass them to a linguist for full translation and analysis.

The Carnegie-Mellon system for translating broadcast captions is a good example of integration of various techniques. If a knowledge-based translation of a segment of text fails and the segment is not found in the translation-memory database, an example-based translation may be used if there is a similar segment in the database. As a last resort, a word-for-word dictionary lookup translation is supplied. All this happens in real time.

Prices have dropped across the board. For example, the METAL system sold by Siemens used to cost a small fortune, but now it has been repurposed by GMS as a Langenscheidt consumer product. Systran, long an IBM mainframe program, now has a Windows version for PCs. Very sophisticated machine translation systems are available for under US\$1,000, and some less ambitious systems are available for under US\$200.00.

Machine translation is finding its niches. There is detente between human translators and the machine. Will the next half-century be as full of amazing developments as the last 50 years? Only time will tell.

FOR FURTHER INFORMATION

For copies of the US\$10 NSTC videotape of the "MT Pioneers" session, contact MurielVasconcellos@compuserve.com