

3 • 1 Machine Translation Projects

Shun Ishizaki

Electrotechnical Laboratory, Japan

[1] EUROTRA project

The EC membership is comprised so far of 12 countries where nine languages are officially spoken. Started out in 1982, EUROTRA is a project designed to develop a system to translate into these languages an increasing number of minutes of EC legislature and economical activity reports and business related documents. Now in its second phase, the project is to develop such prototype system as having 2,500 words toward the second half of 1988. The vocabulary will amount to 20,000 items in the two-year third phase to consummate the entire project. The project will be evaluated on an overall basis upon completion in that phase. The translation system will eventually be less of transfer type and generally more of interlingua type.

[2] ALVEY MT project

This is part of an English government project ALVEY for information technology research and development. The five-year-long research efforts for machine translation system development, which lasted into 1987, gave rise to two systems: NTRAN developed by the University of Manchester Institute of Science and Technology (UMIST) and AIDTRAN by the University of Sheffield. NTRAN is an English-Japanese machine translation system designed exclusively for writing computer manuals in Japanese and AIDTRAN a Japanese-English system employing Dr. Jelinek's Japanese parser.

[3] Asian project

Asian nations — Thailand, Malaysia, Indonesia, China and Japan — have been working on a multilingual machine translation system which adopts an interlingua approach to use internationally. The research items of the six-year project, which was launched in 1987, include analysis, generation, electronic dictionary, I/O system, translation support, and an overall system as well as interlingua. Japan's Center of the International Cooperation for Computerization is joined by Japanese computer firms, and research organizations of other countries join the project. Electrotechnical Laboratory (ETL) carries out the fundamental research of its own for interlingua, and provides consultative service to the international project.

[4] GETA project

France's University of Grenoble has been engaged since 1971 in GETA, a project intended for French, Russian and a number of Asian as well as other European Languages. In 1978 the project produced ARIANE-78, which was later repeatedly improved. Thanks to the improvements, the system now operates with a personal computer. It features transfer approach, heuristic linguistic processing, and multi-level structural description.

[5] Mu project

A Japan Science and Technology Agency-funded machine translation project, Mu (the first phase)

completed in March 1986 both Japanese-English and English-Japanese translation of the entries of electrical engineering papers at JICST. Kyoto University, Electrotechnical Laboratory, JICST and many firms joined the Project. The achievement, which took four years, features transfer approach, vocabulary drive-processing approach, and use of grammatical descriptive language GRADE. The second phase Mu started in April 1986 and aims to modify the first one and develop an machine translation system to serve the JICST needs.

[6] CMU project

The Carnegie-Mellon University established an machine translation laboratory in 1986 to carry out an machine translation system research employing both computational linguistics and AI approaches. The 30-member research project is concerned with universal

parser and generator, multilingual translation, and multi-domains. A few Japanese businesses take part in the project.

[7] Electronic dictionary project

In 1986, Japanese government and 8 Japanese computer firms began a nine-year project to publish a large-scale electronic dictionary. The dictionary would be contributing to natural language processing and knowledge inference technology. Japan Electronic Dictionary Research Institute leads the project to develop, among others, master dictionaries, concept dictionaries, data management systems, and corroborative evaluation systems. The master dictionaries will contain both English and Japanese basic vocabulary, 200,000 words each. The concept dictionary will be made up of 30,000 conceptual descriptions and 400,000 conceptual hierarchies.