[From: *Zápisník slovenského jazykovedca* 10 (1/2), 1991, 13-16]

MACHINE TRANSLATION IN THE EIGHTIES

No more than fifteen years ago it appeared to most observers that machine translation (MT) had been one of the "great failures" of modern research. In 1966 a committee set up by the major US sponsors of MT research (ALPAC) had concluded that MT was slower, less accurate and twice as expensive as human translation and that "there is no immediate or predictable prospect of useful machine translation". The report brought a virtual end to MT research in the United States for over a decade and had a profound impact on research sponsorship elsewhere. However, by the end of the 1970s a revival had begun which accelerated during the 1980s until now there is probably more activity in this field than ever before.

By the late 1970s it had become the generally accepted view that the 'direct' translation approach of 'first generation' MT systems (systems designed in all details specifically for one particular pair of languages) is inherently incapable of producing good quality output. The 'indirect' approaches of most later systems analyse source language (SL) texts independently of potential target language (TL) forms, and employ intermediary representations to carry information from source text to target text. Some systems aspire to 'interlingual' (IL) representations (ideally universal, or at least language-independent), with a two-stage translation process: from SL to IL and from IL to TL. Others prefer a three-stage 'transfer' approach: from SL to an abstract interface (a SL-oriented representation), then to a TL-oriented representation, and from the latter to a final TL text.

The 1980s have seen developments on many fronts. Firstly, there has been the research, development and in some cases implementation and commercial exploitation of transfer systems of various kinds. Secondly, there have been implementations and improvements of earlier (basically 'direct') designs. Thirdly, there has been continuing research on artificial intelligence approaches and techniques and their application to MT system design. And fourthly, there have been explorations of alternative models, including not only innovative interlingua approaches but also applications of statistical techniques, the beginnings of experiments on speech ranslation and investigations of systems designed for non-translators.

In 1981 and 1982 appeared the first commercial systems (ALPS, Weidner system, Japanese systems, e.g. Oki's PENSEE, Mitsubishi's MELTRAN, and Sanyo's SWP-7800 'Translation Word Processor', Fujitsu's ATLAS/I, Sharp's DUET and more sophisticated systems, e.g. the Logos systems (initially for German-English, later other languages), the Fujitsu ATLAS/II system for Japanese-English translation, the NEC PIVOT system for bi-directional English and Japanese translation, and above all the METAL systems for German-English and English-German (based on the advanced research undertaken for many years at the University of Texas at Austin and since 1979 supported by the Siemens Company at Munich), with research proceeding on systems combining English and German with Spanish, French and Dutch.

All these commercial systems, however 'advanced', generally require a considerable initial commitment by purchasers to adapt programs and dictionaries to local needs, often in collaboration with vendors. In this sense, nearly all MT installations are tailored to particular environments, they are in effect 'in-house' systems which could not, without further adaptation, be transferred to another operational context.

Many MT systems have been explicitly designed for particular clients and using a controlled language which it is known the computer program can tackle with little or no need for external assistance or subsequent expensive post-editing (METEO system designed by the Montreal MT group TAUM for translating weather reports from English into French, SPANAM (1980) and ENGSPAN (1985) developed at the Pan American Health Organization in Washington for translating medical and public health literature). In 1978 the European Community decided to support research on a multi-national multilingual system based on the latest advances in computational linguistics. This was the Eurotra project, which involved teams of researchers in

each member nation of the Community and intended to produce a MT system for translating from and to all nine Community languages. It has undoubtedly been the most ambitious project of the 1980s. Its general design owed much to the most advanced transfer systems of the time, the GETA-Ariane and the SUSY systems of Grenoble and Saarbrücken respectively. Each project has made substantial contributions to MT theory and practice, and in addition Eurotra has inspired innovative theoretical linguistic and computational-linguistic research, particularly in the Netherlands, Belgium, Denmark, Germany and Great Britain.

In addition to these 'transfer' systems the 1980s have seen the revival of interlingua systems - at the present time there are two interlingua systems under development, both in the Netherlands: the DLT (Distributed Language Translation) project at the BSO company in Utrecht, based on Esperanto, and the Rosetta project at the Philips company in Eindhoven, exploring the use of Montague grammar in interlingual representations.

For many observers of MT development it has appeared that the most likely source of techniques for improving MT quality is the application of research on natural language processing within the context of Artificial Intelligence (AI). The main US centre is at Carnegie-Mellon University; other projects are found in California, New York, New Mexico, Texas, British Columbia, Montreal, Toronto, etc. In Europe, AI approaches are found at Saarbrücken, Stuttgart, Berlin and Bielefeld, at Cap Sogeti Innovation in France. In Japan there is much activity in this area too.

Until the late 1970s most MT research activity was undertaken in academic environments with relatively little regard for immediate or even long-term potential applications. During the 1980s there has been a remarkable shift: some research is still based in universities and institutes, but most is now undertaken by independent companies for short- or long-term commercial interests. This has had an impact not only on the pace of development but also on the range of languages covered. Until the mid-1960s, systems were developed primarily for the use of scientists to keep abreast of technological activity. Research concentrated on translation from Russian, or - in the case of Soviet MT research - from English. In the 1970s systems were designed for the pressing administrative and legislative needs of bilingual Canada and the multilingual European Community. In the 1980s the demand has been for systems covering the major commercial languages of the world: English, French, German, Spanish, and Japanese, to which have been added in recent years Arabic, Chinese and Korean.

With the variety and wealth of activity outlined briefly it is difficult to summarize the changes during the last decade. Major features would have to include: the shift from large-scale batch systems to microcomputer-based interactive systems, the impact of artificial intelligence, the development of controlled language and of sublanguage systems, the revival of interlingua and statistical approaches, the commercialisation of MT development, the growing internationalisation of MT activity, and above all the emergence of Japanese systems on a field previously dominated by American and European systems. The rapid growth of MT in the last decade and particularly in the last five years means that there are almost certainly more researchers and developers active in the field than there were at the height of the 1960s before the ALPAC report appeared. The demand for translation is growing at a pace well beyond the capacity of the translation profession. Machine translation is no longer a slightly suspect academic pursuit, it has established itself as a technology-based industry of international dimensions.

W.John Hutchins
(8.X.1990 - Bratislava)

Bibliografia

Machine translation: past, present, future. Chichester: Ellis Horwood, 1986.

Recent developments in machine translation: a review of the last five years. In: *New directions in machine translation*: conference proceedings, Budapest 18-19 August, 1988. Ed: D. Maxwell et al. Dordrecht: Foris, 1988.