The history of machine translation in a nutshell

John Hutchins
[revised January 2014]

§1. Before the computer

It is possible to trace ideas about mechanizing translation processes back to the seventeenth century, but realistic possibilities came only in the 20th century. In the mid 1930s, a French-Armenian Georges Artsrouni and a Russian Petr Troyanskii applied for patents for 'translating machines'. Of the two, Troyanskii's was the more significant, proposing not only a method for an automatic bilingual dictionary, but also a scheme for coding interlingual grammatical roles (based on Esperanto) and an outline of how analysis and synthesis might work. However, Troyanskii's ideas were not known until the end of the 1950s. By then, the computer had been born.

§2. The pioneers, 1947-1954

Soon after the first appearance of 'electronic calculators' research began on using computers as aids for translating natural languages. The beginning may be dated to a letter in March 1947 from Warren Weaver of the Rockefeller Foundation to cyberneticist Norbert Wiener. Two years later, Weaver wrote a memorandum (July 1949), putting forward various proposals, based on the wartime successes in code breaking, the developments by Claude Shannon in information theory and speculations about universal principles underlying natural languages. Within a few years research on machine translation (MT) had begun at many US universities, and in 1954 the first public demonstration of the feasibility of machine translation was given (a collaboration by IBM and Georgetown University). Although using a very restricted vocabulary and grammar it was sufficiently impressive to stimulate massive funding of MT in the United States and to inspire the establishment of MT projects throughout the world.

§3. The decade of optimism. 1954-1966

The earliest systems consisted primarily of large bilingual dictionaries where entries for words of the source language gave one or more equivalents in the target language, and some rules for producing the correct word order in the output. It was soon recognised that specific dictionary-driven rules for syntactic ordering were too complex and increasingly ad hoc, and the need for more systematic methods of syntactic analysis became evident. A number of projects were inspired by contemporary developments in linguistics, particularly in models of formal grammar (generative-transformational, dependency, and stratificational), and they seemed to offer the prospect of greatly improved translation.

Optimism remained at a high level for the first decade of research, with many predictions of imminent "breakthroughs". However, disillusion grew as researchers encountered "semantic barriers" for which they saw no straightforward solutions. There were some operational systems – the Mark II system (developed by IBM and Washington University) installed at the USAF Foreign Technology Division, and the Georgetown University system at the US Atomic Energy Authority and at Euratom in Italy – but the quality of output was disappointing (although satisfying many recipients' needs for rapidly produced information). By 1964, the US government sponsors had become increasingly concerned at the lack of progress; they set up the Automatic Language Processing Advisory Committee (ALPAC), which concluded in a famous 1966 report 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." It saw no need for further investment in MT research; and instead it recommended the development of machine aids for translators, such as automatic dictionaries, and the continued support of basic research in computational linguistics.

§4. The aftermath of the ALPAC report, 1966-1980

Although widely condemned as biased and short-sighted, the ALPAC report brought a virtual end to MT research in the United States for over a decade and it had great impact elsewhere in the Soviet Union and in Europe. However, research did continue in Canada, in France and in Germany. Within a few years the Systran system was installed for use by the USAF (1970), and shortly afterwards by the Commission of the European Communities for translating its rapidly growing volumes of documentation (1976). In the same year, another successful operational system appeared in Canada, the Meteo system for translating weather reports, developed at Montreal University.

In the 1960s in the US and the Soviet Union MT activity had concentrated on Russian-English and English-Russian translation of scientific and technical documents for a relatively small number of potential users, who would accept the crude unrevised output for the sake of rapid access to information. From the mid-1970s onwards the demand for MT came from quite different sources with different needs and different languages. The administrative and commercial demands of multilingual communities and multinational trade stimulated the demand for translation in Europe, Canada and Japan beyond the capacity of the traditional translation services. The demand was now for cost-effective machine-aided translation systems that could deal with commercial and technical documentation in the principal languages of international commerce.

§5. The 1980s.

The 1980s witnessed the emergence of a wide variety of MT system types, and from a widening number of countries. First there were a number of mainframe systems, whose use continued almost to the present day. Apart from Systran, now operating in many pairs of languages, there was Logos (German-English and English-French); the internally developed systems at the Pan American Health Organization (Spanish-English and English-Spanish); the Metal system (German-English); and major systems for English-Japanese and Japanese-English translation from Japanese computer companies.

The wide availability of microcomputers and of text-processing software created a market for cheaper MT systems, exploited in North America and Europe by companies such as ALPS, Weidner, Linguistic Products, and Globalink, and by many Japanese companies, e.g. Sharp, NEC, Oki, Mitsubishi, Sanyo. Other microcomputer-based systems appeared from China, Taiwan, Korea, Eastern Europe, the Soviet Union, etc.

Throughout the 1980s research on more advanced methods and techniques continued. For most of the decade, the dominant strategy was that of 'indirect' translation via intermediary representations, sometimes interlingual in nature, involving semantic as well as morphological and syntactic analysis and sometimes non-linguistic 'knowledge bases'. The most notable projects of the period were the GETA-Ariane (Grenoble), SUSY (Saarbrücken), Mu (Kyoto), DLT (Utrecht), Rosetta (Eindhoven), the knowledge-based project at Carnegie-Mellon University (Pittsburgh), and two international multilingual projects: Eurotra, supported by the European Communities, and the Japanese CICC project with participants in China, Indonesia and Thailand.

§6. The early 1990s

The end of the decade was a major turning point. Firstly, a group from IBM published the results of experiments on a system (Candide) based purely on statistical methods. Secondly, certain Japanese groups began to use methods based on corpora of translation examples, i.e. using the approach now called 'example-based' translation. In both approaches the distinctive feature was that no syntactic or semantic rules are used in the analysis of texts or in the selection of lexical equivalents; both approaches differed from earlier 'rule-based' methods in the exploitation of large text corpora. However, traditional rule-based projects continued, e.g. the Catalyst project at Carnegie-Mellon University, the project at the University of Maryland, and the ARPA-funded research (Pangloss) at three US universities. Thirdly, the first translation memory systems came to the market (Trados, etc.), enabling translators easy access to previously translated texts.

A fourth innovation was the start of research on speech translation, involving the integration of speech recognition, speech synthesis, and

translation modules – the latter mixing rule-based and corpus-based approaches. The major projects have been at ATR (Nara, Japan), the collaborative JANUS project (ATR, Carnegie-Mellon University and the University of Karlsruhe), and in Germany the government-funded Verbmobil project.

Another feature of the early 1990s was an increase of MT activity to research on practical applications, to the development of translator workstations for professional translators, to work on controlled language and domain-restricted systems, and to the application of translation components in multilingual information systems.

§7. The late 1990s.

These trends continued into the later 1990s. In particular, the use of MT and translation aids (translator workstations and translation memories) by large corporations grew rapidly – a particularly impressive increase was seen in the area of software localization (i.e. the adaptation and translation of documentation for new markets). There was a huge growth in sales of MT software for personal computers (primarily for use by non-translators). The demand was met not just by new systems but also by 'downsized' and improved versions of previous mainframe systems. While in these applications, the need has been be for reasonably good quality translation (particularly if the results were intended for publication), there was an even more rapid growth of automatic translation for direct Internet applications (electronic mail, Web pages, etc.), where the need has been for fast real-time response with less importance attached to quality. Significant for the future was the introduction of MT online services (first Babelfish and later Google Translate). With these developments, MT became a mass-market product.

§8. Since 2000

Statistical machine translation (SMT) has now become the dominant framework of MT research; there are now very few researchers continuing with exclusively rule-based methods. The main reasons are: i) the availability of large monolingual and bilingual corpora, ii) the open-source availability of software for performing basic SMT processes (alignment, filtering, reordering), such as Moses, GIZA, etc.; iii) the availability of widely accepted metrics for evaluating systems (BLEU, and successors). Statistical methods do not require researchers to know the languages involved in systems (or, at least, to have in-depth knowledge) and do not demand complex large-scale acquisition of rules and lexical data. While SMT approaches dominate, there are still many aspects of MT for which rule-based approaches have continuing relevance, e.g. syntactic analysis to improve reordering of sentences between typologically different languages (e.g. English and Japanese), the treatment of morphologically rich languages (such as Russian, Finnish and agglutinative languages), the problems of

transliterated names (particularly for Chinese), and the problems of discourse relations (e.g. treatment of pronouns). As a consequence, many researchers adopt 'hybrid' approaches combining SMT and rule-based approaches.

The use of MT continues to increase in volume and to spread to new areas of application (such as patents, television subtitles, website localization, restaurant menus, social networking, humanitarian aid, company information services, public lectures, and more). Large-scale use is found in an increasing number of global companies and translation services, often in conjunction with the pre-processing of input (e.g. by controlled language and terminology checking) and with the post-editing of outputs (including now statistics-based methods and crowdsourcing). The earlier antagonism of professional translators appears to be much reduced: translators are using MT as well as translation memories as aids in the production of draft translations. As for the general public, access to MT is now almost exclusively through free on-line services (such as Google Translate) and no longer through software on PC systems. There is a perception that translation quality is improving both in general-purpose on-line MT as well as in research systems. Nevertheless, the development of evaluation metrics remains an area of major importance for both general users and researchers, in particular because human assessments of translation quality often differ widely from statistical measures.

§9. Further information

This outline includes only the major and most significant developments. For more detail see my publications on the history of MT at: http://www.hutchinsweb.me.uk, and the resources available in the *Machine Translation Archive* (http://www.mt-archive.info).