[Paper presented at the first MT Summit, September 1987, Hakone, Japan]
[From: *Machine translation summit*, ed. M. Nagao et al. (Tokyo: Ohmsha Ltd.), pp.7-12]

Prospects in machine translation

W. John Hutchins
University of East Anglia, England

Research on machine translation began in the 1950s and has largely remained to this day an activity which combines an intellectual challenge, a worthy motive and an eminently practical objective. The challenge is to produce translations as good as those made by human translators. The motive is the removal of language barriers which hinder scientific communication and international understanding. The practical objective is the development of economically viable systems to satisfy a growing demand for translations which cannot be met by traditional means. However, the realization has been disappointing in many respects; and, although recently optimism has been growing, operational machine translation systems are still far from satisfactory.

On the one hand, there are many who have unbounded faith in the power of modern computers, and do not understood why after years of intensive research it is not always possible to translate simple sentences with perfect accuracy; the full complexity of natural language is not grasped. On the other hand, there are those who are fully aware of its complexities and subtleties and are convinced that any kind of machine translation is inherently impossible. The research and development is not identified as a worthwhile activity due to such unfulfilled expectations and sceptical prejudices, and it has already taken twenty years since the ALPAC report to demonstrate that there is practical value in systems capable of producing translations which are less than perfect. It is only recently that the results and progress of machine translation research are beginning to be recognised by translators, administrators and the general business community. Nevertheless, there is a great deal more to be achieved before it can be said that machine translation research has delivered what it set out to do.

No current operational machine translation systems can produce good quality output without either placing restrictions on input texts or involving human assistance before, during or after translation processes. Present machine translation systems make ‘simple’ grammatical errors which no human translator would make. All have difficulties in the selection of pronouns, prepositions, definite and indefinite articles-particularly when translating from languages such as Russian which do not have articles. The revision of machine translation output typically involves a great deal of low-level correcting, and it is not surprising that translators have a generally poor opinion of machine translation systems. Revisers of machine translation output are prepared to accept difficulties with technical vocabulary and with homographs and polysemes, but the repetitive correction of the same mistranslation is irritating, both for post-editors and for operators of interactive machine translation systems. Much can be done to simplify the editing facilities, but it would be preferable if the mistakes were not there in the first places.

Many of the difficulties can be avoided by limiting systems to particular subjects or styles of language. This can reduce problems of lexical homography and syntactic ambiguity, and at the same time enable higher levels of comprehensiveness and completeness in dictionaries. Here we may identify three basic options: restriction to a particular ‘sublanguage’ or ‘subworld’, restriction to particular text types, and restrictions on the vocabulary and grammar of input texts. The ‘sublanguage’ approach has the advantage of minimising problems of analysis, disambiguation and generation of target language, but in practice few texts are limited to single subject fields: the weather reports of METEO may be the exception rather than the rule. Restriction to particular text types, such as titles (e.g. TITRAN) or abstracts reduces grammatical complexity to some extent but with little impact on solution of lexical problems, unless systems are further limited in subject range. The practice of 'pre-editing' is now rarely adopted, primarily for economic reasons, since output inevitably requires revision as well. The alternative option of accepting only texts composed in a 'restricted language', e.g. the Multinational Customized English of Xerox, becomes economically feasible if there is to be multilingual output.

Post-editing, pre-editing, sublanguages, restricted languages – all these are means of accommodation to the limitations of existing operational systems. But they are not solutions. Ultimately, significant improvements in machine translation can come only from fundamental research on fully automatic translation. Human intervention will be essential in some form in any future system that can be envisaged, but the aim must be to reduce human revision to matters of expertise and of stylistic smoothness.

The dominant framework for most machine translation systems under current development is based essentially on the syntax-oriented approach of computational linguistics (the successor of earlier lexicon-based approaches of 'direct translation' systems). There are two basic 'indirect' strategies: interlingual and transfer. The initial projects based on the interlingua approach proved too ambitious, and in the last two decades there has been a general agreement within the machine translation community on the basic transfer system design: three stages of analysis, transfer and synthesis, with relatively language-independent tree-transduction procedures and abstract language-specific intermediary representations incorporating information from all levels of analysis. Current examples are the GETA-ARIANE system at Grenoble, the SUSY and ASCOF systems at Saarbrücken, the multinational EUROTRA project, and the Japanese Mu project at Kyoto University and elsewhere.

The approach has its major strengths built up during many years of experience with large-scale complex systems: a solid body of well-tested and efficient methods of morphological and syntactic analysis (context free parsers, ATN parsers, tree transducers, charts, etc.), now supported by theoretical developments in 'unification grammars' (e.g. Lexical Functional Grammar, Generalized Phrase Structure Grammar and Definite Clause Grammar), and the modularity and flexibility of system architectures permitting progressive incorporation of newer techniques, including recent artificial intelligence (AI) methods.

Its weaker features include the familiar limitation of analysis to sentences (with problems of pronouns, intersentential relationships, and elliptical constructions), and the constraints of the syntactic framework, however abstract and language-independent, for semantic representations. In semantic analysis there has been successful treatment of homography and syntactic ambiguity; and there have been successful implementations of case frames, of semantic features, and recently of Montague semantics; but, nevertheless, the profounder problems of interlingual semantic analysis have continued to prove elusive.

The abstractness of intermediary representations and therefore the 'depth' of semantic analysis depends on whether the system is intended to be bilingual or multilingual. Bilingual systems can operate with direct equivalences of lexical items and syntactic structures, and can ignore interlingual or 'universal' considerations. By contrast, multilingual systems inevitably tend towards interlingual semantic representations (as the EUROTRA researchers have found). As a consequence, problems of semantic analysis are exacerbated. Fully automatic multilingual systems based on the transfer approach are perhaps still too ambitious in the present status of linguistic and computational knowledge.

The 'universal language' which would serve as an interlingua for machine translation is as elusive as ever. The belief that machine translation 'pivot languages' can be constructed from scratch, based on logical formula and semantic primitives, retains its attractions, despite many past failures. More realistic may be efforts to establish semantic elements common to a family of historically related languages, such as the Romance languages, or possibly the common core of the European languages. An alternative framework is to modify and expand an already constructed ‘artificial’ international language, such as Esperanto in the DLT project at Utrecht -or, even more speculatively, to build upon a supposedly ‘logical’ natural language, as the South American language Aymara is claimed to be.

Whatever the approach to language-independent or interlingual representations, the major problems for all machine translation systems will always be found in the processing of the illogical and vague expressions of natural languages and in the extraction of the precise meaning from insufficiently explicit, context-bound source-language texts – texts which human readers can interpret with greater or lesser facility but which remain intractable to computational analysis.

While there is undoubtedly still scope for considerable improvements in advanced transfer approaches, greater expectations are focused on AI approaches. In this framework, translation involves the ‘understanding’ of source texts and the expression of the (language-independent) ‘meaning’ in a target language text. No rigid distinction is made between linguistic and non-linguistic information, nor between syntax and semantics; text processing is not limited to sentences; understanding and text interpretation refer to knowledge databases and inference mechanisms of specific subworlds.

In all these aspects AI offers new directions for machine translation research. At present there are no large-scale purely AI-based translation projects which can test the feasibility of methods developed in small-scale experiments. However, many ‘linguistics’-based machine translation research systems are now introducing AI techniques in various forms: greater integration of syntactic and semantic information, the incorporation of ‘subworld’ knowledge databases, and the embodiment of translators’ expertise in disambiguation and transfer components (‘expert systems’). But there is still scepticism about their general applicability. There are perhaps three main issues. The first is the huge size of AI knowledge bases for fragments of natural language vocabulary much smaller than those required in machine translation systems even when restricted to specific sublanguages; AI researchers claim that knowledge bases are expandable without performance degradations, but this has yet to be demonstrated in a machine translation environment. The second issue is the doubt whether ‘comprehension’ to the depth assumed by AI researchers is really necessary for translation purposes; many translators and many operational machine translation systems produce fully satisfactory versions without going any further than ‘superficial’ lexical equivalences; this is an issue bound closely to that of translation quality – there are no clearcut answers, and the debate will continue for the foreseeable future. A third issue is the abstractness of ‘content’ representations resulting in losses of information about ‘surface’ structures of texts; versions produced by AI methods are not translations but rather paraphrases of source texts.

What is generally agreed is the future of machine translation as one component of integrated man-machine communication systems which include authoring, summarization, electronic mail, word processing, publishing facilities, access to remote databases, and so forth. One sign already is the interest of many machine translation projects in translation workstations which would provide translators with facilities ranging from straightforward dictionary consultation (in-house and remote) to full-scale automatic translation. To this picture must be added the obvious attractions of spoken language input and output, i.e. translation systems for telephone communication and ultimately simultaneous interpretation. The prospects are still distant, however: while speech generation is fairly well advanced, the problems of speech recognition are much more intractable.

At the most fundamental level, machine translation systems will have to be designed as ‘learning’ systems. Both AI and computational linguistics focus on the formalization and modelling of language. Much can and has been achieved, but it will never be the complete answer. Natural language resists rationalistic classifications, logical analyses and efforts to extract underlying uniformities. Language use is dynamic and perpetually innovative; knowledge is constantly changing and expanding. At present, systems inevitably represent states of knowledge and language at particular moments; they can be updated only by direct human input. Future systems must be capable of ‘learning’ indirectly from the expertise of translators, from the texts they process and from the corrections made by revisers of machine translation output. This must be a major objective in future machine translation research.

In the immediate future it would appear that the best prospects for fully automatic translation systems will be those which combine the virtues of the traditional ‘linguistics’ approaches and the newer ‘knowledge-based’ approaches. The most likely option is the integration of AI techniques within the well-founded ‘computational linguistics’ framework of powerful and flexible transfer systems. This option is already being explored in many existing projects. Alternatively, AI-based systems might incorporate techniques of computational linguistics. Here, too, there are signs that this is happening. Advances in microcomputer technology have made possible a proliferation of small-scale machine translation projects, and there will be many AI-inspired projects, but major advances will probably come primarily from the larger projects (Grenoble, EUROTRA, Kyoto, etc.) Only these will have the capacity to test new linguistic and AI techniques on a sufficiently large scale.

Experimental projects can be innovative, but systems designed to be operational within a fixed time-span must be founded on techniques that ensure reasonable success. Machine translation systems have long gestation periods; the commercial risks are considerable, with high investment in system development and in after-sales maintenance; and so there is a natural reluctance to invest in untried experimental methods. Basic machine translation research will continue to be primarily pursued at academic institutions, with increasing collaboration of commercial and industrial interests. International coordination is highly desirable. In certain aspects of system design there are large areas of agreement: morphological analysis and dictionary structure and lexical information. Collaboration would seem obvious, but it is still rare; there is too much duplication of machine translation research effort.

The complexity of machine translation computational processes demands close collaboration between the research fields of linguistics, artificial intelligence and computer programming. In the past, it was desirable to separate the tasks in machine translation projects between linguists and programmers, to divide strictly between linguistic data and computer algorithms. The distinctions are still valid, but the separation of tasks has become less meaningful in recent developments. Theoretical linguistics is moving closer to computational formalisms – particularly in lexical-functional grammar and the unification grammars – and the higher level languages such as LISP and Prolog can be applied almost directly to the formalizations of linguists. The increasingly close links between theoretical linguistics and computer science will be even more vital when the promised era of parallel processing comes.

In the immediate future, operational machine translation systems designed for the production of quality translations will continue to have four options: the use of ‘regularized’ or pre-edited input, the restriction of systems to specific sublanguages, the involvement of human interaction during machine translation processes, and the acceptance of the need for more or less extensive post-editing of texts. The options are not mutually exclusive since post-editing may be needed for restricted language systems, and interactive systems may be limited to particular sublanguages.

For the foreseeable future, the ‘regularized input’ option will be economically viable only in multilingual situations, where the same text is to be translated into a number of languages. This does not mean that the machine translation system itself has to be multilingual; there can be (as in the Xerox case) a series of bilingual systems. Interactive machine translation systems are in principle feasible for simultaneous multilingual output, and this is the intention of the Utrecht DLT project. For the more traditional ‘batch processed’ machine translation with post-editing there will continue to be bilingual systems, with the newer transfer systems promising better quality than those operating at present. There is an undeniable need for multilingual systems, but it seems likely that the bilingual design will still be popular because multilingual systems have an inherently greater complexity and therefore greater risk of failure.

The attitude of translators to machine translation is unlikely to change dramatically. Some will be prepared to work as post-editors and others will operate interactive systems. There would be more if 'raw' machine translation output could be rid of those irritating incorrect prepositions, articles and pronouns; translators do not want to be 'slaves' to machines which produce unsatisfactory results. They are professionally committed to versions of high quality. Consequently, most will prefer sophisticated machine aids, accessed at a single terminal, e.g. a translator's workstation, combining aids which are already becoming commonplace: on-line text editing, in-house specialized glossaries, on-line access to remote term banks, transmission of texts by telecommunication links, optical character recognition, and generation of high-quality printed output. In this way they can achieve higher rates of productivity and greater levels of consistency while maintaining their traditional high standards, particularly in scientific, technical and administrative translation. There is still room for much improvement; irritating equipment incompatibilities, inadequate provision of non-English characters and non-Roman alphabets, etc., but there are solutions within existing technologies and more advanced facilities will be established using the dictionary construction method in both machine translation and AI, and semantic research. The growth of machine translation has in fact increased the options for translators; they will be able to concentrate on translations which need their creative flair, and they will leave the tedious, repetitive and mundane to the machine.

The area of translation with greatest potential for machine translation systems is the largely untapped demand for rough translations. The information needs of scientists, technicians, and administrators can often be satisfied by the unedited output of existing machine translation systems. The recipients must have the subject knowledge enabling them to understand the intentions of authors even if the message is obscured by lexical and grammatical errors and by stylistic idiosyncrasies. The toleration of errors varies, however, according to the nature of the material translated; greater accuracy is probably required for titles and abstracts than for full texts, where the larger context provides more clues for interpretation. Even greater tolerance may be found if recipients also have some knowledge of the source language; for these virtually word-for-word 'pidgin' versions may be comprehensible. In general, however, the aim must be much higher quality than in present systems, since as unedited output improves so demand will increase and even higher expectations will be voiced. But what is not necessarily required is an AI-type 'understanding' component, since rough translations are intended for those who can call upon their own far richer knowledge resources. It is clear that in the immediate future, the best prospects lie in systems designed for particular sublanguages. Most demand will come from the applied sciences, engineering, and social sciences, where there tends to be less knowledge of foreign languages than in the pure sciences.

The most innovative direction for future machine translation development exploits the recognition that AI systems do not produce strict translations of a text but rather paraphrases in another language. The idea is that 'paraphrase translations' should be produced by systems which combine composition in the user's own language and simultaneous translation into another. In this case, the most immediate application is clearly conventional business correspondence, where what is most important is that the basic 'message' is conveyed in the foreign language. If speech recognition and spoken output were also to be developed within the same restricted area, the commercial prospects would look very attractive and there would be no lack of financial support.

Paraphrase translation in the other direction, into the mother tongue of the recipient, would also be a feasible interactive machine translation model. There are undoubtedly a sufficient number of specialists with some basic knowledge of a foreign language who would be capable of translating a text in their subject field with minimal computer assistance. Developments in this direction might eventually, with progressive miniaturization, lead to the long predicted pocket ‘personal translators’ for travellers and businessmen.

In the more distant future there must be the prospect of systems combining translation and summarization. The prospect of producing summaries of foreign language reports or articles for administrators, businessmen and scientists in their own language is almost certainly more attractive than paraphrase or rough translations of full texts. AI researchers and others have begun experiments on small-scale summarization programs in restricted subworlds, but it has already become apparent that the complexities of the task are at least equal to those of machine translation itself.

A major factor to be considered in the development of any future machine translation system will be its capacity to be integrated easily into many kinds of computerized office environments. This entails not just equipment compatibility, input and output facilities (OCR, network links, printing and publishing, access to remote databanks, etc.), facilities for creating in-house glossaries, easy system maintenance, but above all absolute reliability of hardware and software.

Reliability and robustness are essential. Systems must not fail either because of the unforeseen effects of dictionary changes or because of ungrammatical or illogical input texts. The levels of tolerance which may be observed at present will decrease as the quality of machine translation output improves and as machine translation is integrated more closely with other computer and telecommunication systems. Least tolerance is to be expected from businessmen using ‘para-translation’ systems and from translation agencies operating ‘batch processing’ and interactive machine translation systems.

Whatever the machine translation system, the quality of dictionaries is crucial. Good dictionaries require many years of painstaking work; arguably the present success of Systran is attributable primarily to the amount of effort which has been devoted to its large and detailed dictionaries. It is vital that operators who are not familiar with computers and programming should be able to amend and update dictionaries easily. It is also obviously desirable that machine translation dictionaries should be more compatible with each other, so that lexical information from one can be used in another.

The progressive improvement of machine translation systems, desirable for many reasons (not least to reduce post-editing costs), argues for the maximum flexibility and modularity of system design and in particular for good communication between designers, users and translators. The Systran implementation at the European Communities has benefited considerably from practical suggestions by experienced translators. A failure of past machine translation research has been the neglect of the expertise of professional translators, lexicographers and terminologists. Machine translation systems must be built to respond to the actual experience of users, to provide facilities which users perceive as necessary for improving performance.

Predictions in the machine translation field have in the past been notorious for excessive optimism or for disillusioned pessimism. The safest prediction is that there will be a long-term future for nearly all varieties of systems, machine aids, workstations, and interactive machine translation will be preferred by professional translators. Post-edited machine translation will continue in large translation services and in translation bureaux. Restricted input machine translation will remain the option for some multinational companies, unless ‘raw’ machine translation output improves greatly. Unrevised machine translation and ‘para-translation’ systems will be adopted increasingly for information-only translations. Automatic telephone translation seems unlikely in the near future, but universal access on demand via public networks to many different kinds of automatic translation facilities, whether interactive or ‘para-translation’ or unrevised ‘batch’ systems, does not now seem to be in the least Utopian. As always, it is the responsibility of the machine translation community to ensure that the general public is not disappointed by unrealistic promises.