[From: *Physics Bulletin*, vol.20, 1969]

Machine translation at the NPL

A. J. SZANSER, MSc, FIL, *National Physical Laboratory, Teddington, Middlesex*

A research project in Machine Translation (MT) from Russian into English was started at the National Physical Laboratory in 1960 and carried on until its conclusion in 1966. Over these years the project evolved from a simple 'improved word for word translation' into a more advanced stage of translation of sentences and phrases with a certain amount of syntactical analysis, and included also idiomatic translation of many word combinations. Lastly, an attempt to provide for dictionary deficiencies by quasi-translation procedures was made. The computer used throughout was the experimental machine ACE, built at the NPL in 1959.

The project was concluded because it was felt that the stage achieved (which, incidentally, proved that the idea was not only plausible but actually worked) was sufficient for a pilot experiment. Finally, an evaluation of the achieved results was carried out.

THE NPL MT SYSTEM

The NPL Russian-English translation project was limited to technical and scientific texts, more specifically to the field of electronics and allied subjects. Regarding the quality of translation, the aim was to provide a *usable* translation for an expert in the given field.

Any computer operated processing of natural languages requires first encoding of the text into a machine-acceptable form (a binary representation of alphanumerical characters). For translation, a word list, i.e. dictionary, containing the equivalents in both languages, must also be prepared in advance, in a similar encoded form. The dictionary entries should contain all grammatical information pertaining to the individual words. The translation itself consists of matching the text words with the dictionary entries, considering the grammatical compatibility and making choices, where applicable, on this ground. The output equivalents in the 'target' language would form a crude word for word translation. This would be unsatisfactory, however, because of the multiple equivalents still remaining, especially in a highly inflected language like Russian. For example, one Russian word may mean either 'physicists' or 'of physics'. Therefore, an analysis *within the context* (syntactic analysis) has to be done. Finally, the English words should be inflected and put into a grammatically acceptable order (syntactic synthesis) and then printed in an ordinary form.

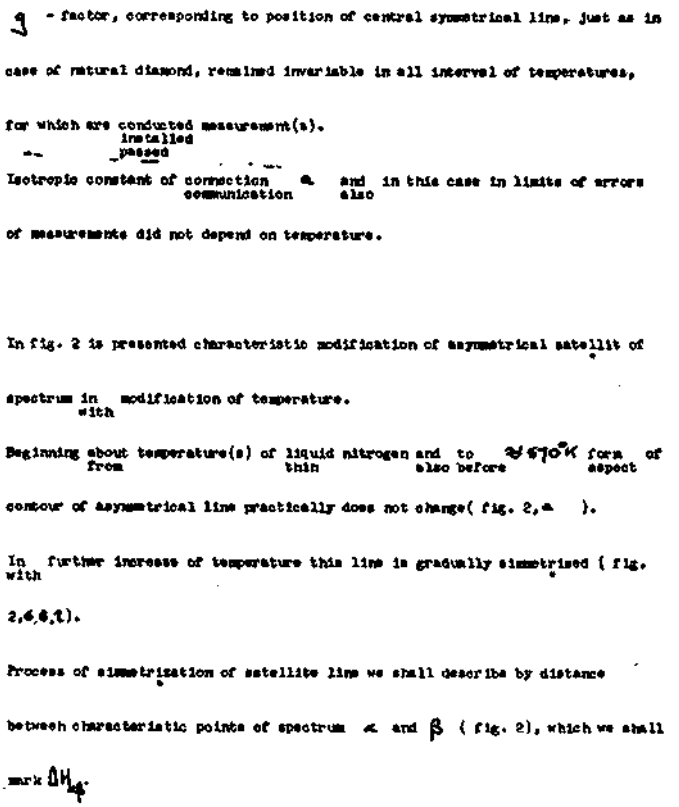
Our basic dictionary was originally obtained from Harvard University, and was then revised and adapted to our needs. It ultimately contained about 15000 words represented by 18000 entries. The difference represented irregular forms, while the regular ones were provided by storing the word stems only, after splitting off the affixes (inflections). Each entry comprised the Russian stem (the full word if irregular) and the English equivalents, together with the respective grammatical information. The dictionary was stored on magnetic tape.

The text was input punched on computer cards and stored on magnetic tape. From then on, all further processing was fully automatic. In the preliminary processing the text words were searched for affixes which, if found, were replaced by coded information (e.g. if English were the source language, the word 'extended' would be split into stem 'extend' and affix 'ed', which would be coded to mean 'past tense or past passive participle of a verb'). The text was then sorted into alphabetical order of (Russian) words, adding to each entry the serial number for a later reverse operation.

This text was matched against the dictionary by running the two tapes simultaneously (the socalled 'serial access' to the dictionary in contrast to 'random access', which is possible with more developed equipment). The most economical length of the text processed depended on both the size of the dictionary and the sorting and re-sorting time and was found to be about 3000 words (an average article). Whenever a match between the text word and a dictionary entry occurred, the whole information (which includes some grammatical information) from the latter was transferred to the former and then the text was re-sorted into the original order. This stage, the dictionary look-up, included also *idiom recognition* in which one of a list of selected idioms (up to five words in length) was identified, first by its 'key words' and then by the other words in context. If recognized, the new translation overruled that of the separate words.

The text prepared in this way was ready for syntactic analysis, which was necessary to elucidate the meaning of phrases and sentences and to resolve the ambiguities remaining after the dictionary look-up. In this respect the theoretical work reached a fairly sophisticated stage, but the implementation of the successive routines into a sequence of programmes operating on the dictionary proceeded more slowly, so that at the termination of the project, when the evaluation experiment was about to commence, only a few fairly simple procedures (such as the integration of nominal and predicative blocks and simpler cases of coordinate groups[1]) were included. More refined procedures such as verb government (in English, for example, direct and indirect objects; in Russian the situation is much more involved) or the resolution of personal pronouns (the respective Russian and English sets are by no means a one to one mapping) were ready and tested automatically on a simulated text, but there

[1] Instead of definitions, some examples (in English) will serve to illustrate these terms. *Noun group:* 'two large, brick buildings': *predicative group:* 'will be carefully assessed'; *coordinate group:* 'electrons and some other kinds of particles'. Any one of these and similar groups forms a single syntactic unit and, therefore, their recognition is vital as a step towards uncovering the proper meaning of the sentence or of particular words and, further, for subsequent correct synthesis in the target language.



- factor, corresponding to position of central symmetrical line, just as in case of natural diamond, remained invariable in all interval of temperatures, for which are conducted measurement(s). installed passed

Isotropic constant of connection communication and also in this case in limits of errors of measurements did not depend on temperature.

In fig. 2 is presented characteristic modification of asymmetrical satellit of spectrum in with modification of temperature.

Beginning about temperature(s) of liquid nitrogen from thin and to also before form aspect contour of asymmetrical line practically does not change(fig. 2,).

In further increase of temperature this line is gradually simmetrised (fig. with 2,6,8,1).

Process of simmetrization of satellite line we shall describe by distance between characteristic points of spectrum and (fig. 2), which we shall mark

Figure 1 A sample page of the NPL Russian-English translation (The words marked by an asterisk are transliterated, possible 'anglicized' words. Alternative equivalents are printed one under another.)

was no time for their inclusion in the system.

As the syntactic complexity increased it was found more and more difficult for the procedures to be incorporated directly into the final system. A theoretical model was therefore evolved according to which all the procedures were included in one sequence of analysis of the source language, gradually building a 'list structure' of the sentence in a logical tree form which at the same time provided the necessary elements for the synthesis into the target language. The complete operations of analysis and synthesis thus became separated from each other.

Two routines were written to improve the straight transliteration (from Cyrillic into Latin alphabet) of

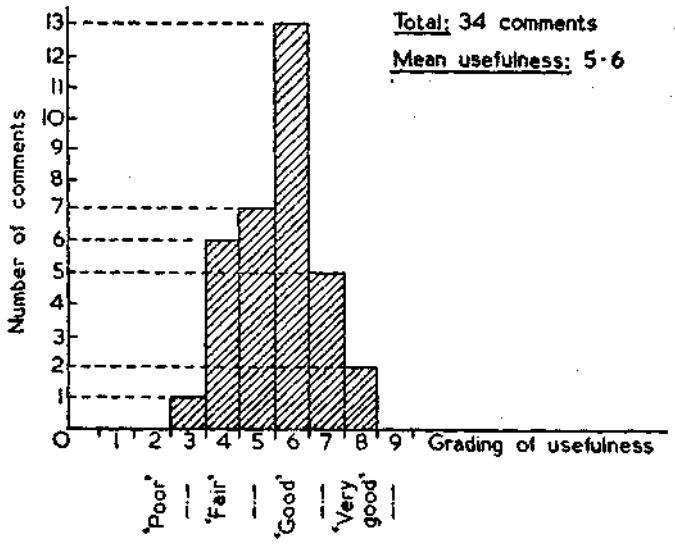


Figure 2 Assessment of usefulness of NPL machine translation output (reproduced by permission of *The Incorporated Linguist*)

words not found in the dictionary. First of these was the prefix splitting routine, which would attempt to find and split off possibly occurring Russian prefixes (such as in English would be 'radio' in 'radiolocation' or 'multi-' in 'multi-electrode'). A list of some 300 such prefixes was prepared and the remainders of the compound words could often be found in the dictionary. The other routine analysed the Russian ending and attached its English equivalent to the unaltered stem, in the hope that it would be (as indeed it often was) international, and therefore understandable. For example, 'dinam-ic' is recognized as 'dynamic' (the first part being rendered phonetically, as were likewise all the words not found in the dictionary), whereas the transliteration could be, for example, 'dinamicheskogo'.

Finally, the English equivalents from the text were chosen, following reordering and inflection, and the results were output on paper tape. This, in turn, being fed into an electric printer, produced the final copy of the translation. A sample page of a translation is reproduced in figure 1.

More detailed accounts of the NPL machine translation system will be found by interested readers in McDaniel *et al.* (1967) and Szanser (1966, 1967).

THE EVALUATION EXPERIMENT

The NPL system brought up to this stage was then submitted to an operational test. In May 1966, a number of industrial and university research establishments were invited to submit Russian scientific texts for translation. In response to this request, 40 articles were received. Some of these were rejected for formal reasons; the remainder were processed automatically and the resulting translations were sent to 45 institutions, including the original ones. Of 39 comments received, a few were rejected as too vague and the remaining 34 became a basis for evaluation.

The method of quantizing the comments was our own[2]. The main criterion, according to the professed aim of our research, was *usefulness to an expert reader.* The scale adopted was from 2 (poor) to 8 (fully adequate) expressed in even integers (the odd ones were used for intermediate steps and the extremes, 0 and 10, were improbable). The readers were *not* invited to apply the scale, to avoid too elastic an interpretation. A few evaluators at the NPL were asked instead to read the verbal comments (some of them lengthy) and translate these into the gradings. As a justification for this method we may observe that the differences between individual assessments were no more than one or, exceptionally, two points. Figure 2 shows the distribution of the comments received over this scale of usefulness.

Most of the comments received were informative and to the point. The prevalent criticism was the inadequacy of English equivalents offered. This, of course, is a matter of dictionary compilation and is not directly relevant to the MT processing system itself. The same applies to the many remaining unidiomatic renderings of word combinations. We used a list of about 500 idioms, whereas 1500 probably would be adequate.

Next, the ambiguity of choice could in many cases be eliminated, or reduced, if further syntactic procedures, already prepared, could have been incorporated into the system. Among other inadequacies which would require special treatment (not necessarily difficult) were multiple ambiguities arising from the many meanings of some common conjunctions or prepositions. Finally, more

[2] The various existing evaluation procedures were unsuitable for our purpose *(see* Szanser 1967).

attention could have been given to abbreviations, especially acronymic ones, common in Russian. On the other hand, a problem quite intractable at the present state of computer technology is the existence of choices arising from semantic ambiguities, the simplest example being homonymy. A possible approach to this problem has been discussed by the author elsewhere (Szanser 1966, 1967). The complete solution would require an associative memory system of a capacity not yet available. There seems to be, however, no 'logical bar' against it.

MACHINE TRANSLATION ELSEWHERE

Apart from theoretical studies, the NPL project was, as far as the writer is aware, the only *complete* one done in the field of machine translation in this country.

Abroad, most of the work undoubtedly has been done in the USA. The research in computational linguistics, as in other fields, is carried on in the United States, both at the universities and by industrial corporations, under Government sponsorship. Restricting this brief account to practical applications, i.e. MT-proper, there were several groups working in this field, mostly on Russian-English translation. Some of these achieved fairly high standards, for example, Georgetown University, IBM and the Bunker-Ramo Corporation.

In terms of basic techniques (such as dictionary look-up, which is highly dependent on the hardware used), some American results outpaced ours, as far as speed is concerned. The best example is the IBM 'photostore' memory, which makes random-access look-up *and* a fully inflected dictionary possible, thus eliminating many preliminary routines. As regards syntactic procedures, however, there seems to exist no great superiority, if any, in the results achieved by the US practical MT systems, compared with ours. In assessing results, incidentally, one has to be aware of a certain degree of elasticity in programming which can produce much smoother, but not necessarily more faithful, translation by introducing a large amount of socalled preferential choice.

In 1966 a special committee formed by the National Academy of Sciences published a report on the whole problem of research in computational linguistics, including MT (Automatic Language Processing Committee 1966). The main conclusion of the report was that financial support for MT projects on the existing scale was unjustified by the results achieved. The Committee recognized instead a (limited) need for pure research in computational linguistics and for the development of machine-aided human translation, such as automatized dictionaries. Since the publication of this report, many MT groups in the USA have either stopped their work, or changed its subject.

In the USSR, the only country where the MT research has been carried out on a scale comparable with that in the USA, the research centres are grouped in Government institutes, of which the largest is VINITI ('All-Union Institute for Scientific and Technical Information') in Moscow, or at the universities (notably Leningrad). The work, in contrast to that done in the Western countries (again apart from the theoretical study) is chiefly oriented towards multilingual translation, especially from and into the languages of the Union, although English-Russian translation is also represented. As regards the achievements of the latter, the only one in which a comparison can be made, and insofar as Russian results are available, these seem to be no more advanced than the work described in this article. It must be said in fairness, however, that the syntactic analysis of English is considerably more difficult than that of Russian.

In continental Europe outside Russia, the most important groups are attached to universities, for example Grenoble in France or Milan in Italy. There are also national societies concerned with MT and applied linguistics. Some research centres are associated with international bodies, such as the European Coal and Steel Community, whose translation bureau in cooperation with the Free University of Brussels elaborated a very promising system of machine-aided translation (automatic dictionary, operating in the context of whole sentences), and Euratom with its research centre at Ispra, Italy.

CONCLUSION

Machine translation quite understandably is now going through a phase of unpopularity as an aftermath of the brisk progress in the late 1950s (during which exaggerated hopes were entertained, especially in the less responsible sector of scientific journalism) and the following near-stagnation in the 1960s. It is perhaps characteristic that, during the 1968 International Federation for Information Processing congress in Edinburgh, out of more than 200 papers there was only one (by B. Vauquois of Grenoble) connected with MT.

It seems that the practical ceiling, conditioned by hardware available in the present MT systems, is now within sight. It is of course possible to increase the sophistication of the systems and so to improve the quality of translation, but the gain probably would be disproportionate to the effort in terms of cost and man-hours.

A further breakthrough can come in two ways: either by a radical improvement in computer storage combined with a low access time, as a result of which enormous quantities of associations could be stored and searched, in analogy with, perhaps, the working of the human brain; or, as some would see it, by deepening and organizing the knowledge of language to such an extent that it would become possible to both analyse and synthesize the discourse at a lesser cost in programming work than hitherto. The writer is sceptical as to the sufficiency of the latter method alone. Both will most probably be needed. But reiterating the opinion expressed earlier, there is no inherent impossibility in achieving a good quality machine translation.

The work described in this article was carried out at the National Physical Laboratory.

FURTHER READING

AUTOMATIC LANGUAGE PROCESSING COMMITTEE 1966 *Language and Machines— Computers in Translation and Linguistics* NAS-NRC publication 1416 (Washington, DC: National Academy of Sciences-National Research Council).

MCDANIEL, J., DAY, A. M., PRICE, W. L., SZANSER, A. J., WHEI.AN, S. and YATES, D. M. 1967 *Translation of Russian Scientific Texts into English by Computer — a Final Report* (Teddington, Middx: National Physical Laboratory), Auto 35.

SZANSER, A. J. 1966 'Machine translation research at the National Physical Laboratory, Teddington' *The Incorporated Linguist* **5**, 102-9.

----- 1967 'Machine translation — the evaluation of an experiment' *The Incorporated Linguist* **6**, 90-6.