

Cybernetics and the Science of Language

RECENTLY a translation of the English sentence "The fundamental problem in the theory of differential equations is that of deducing the properties of the solutions of a given differential equation from the analytic form of the equation" was made at the Steklov Mathematics Institute in Moscow with the aid of an electronic brain. The only mistake made in the translation, which does not prevent anyone from getting the full meaning of the text, was that the machine gave the word "property" instead of "properties."

The machine processing of texts is thus far only of a preliminary character, limited by theoretical problems and initial experiments. The rules for machine translation from French into Russian have been tested in the USSR more fully. In testing these rules not only separate sentences were translated but also full fragments from French mathematical texts. The results of the experiments are quite satisfactory but we have still far to go for the solution of practical problems of translation by machine.

The main difficulty in finding a final solution of practical problems of machine translation from one language into another is that languages have been insufficiently studied from the point of view of cybernetics.

The translation should transmit the initial meaning of the text even though the method of expressing this meaning fully changes. Therefore, it may prove more convenient to make the translation by means of an intermediary language consisting of a complete set of all the meanings which may be met in the translated text of various languages. Thus, at first the translation would be made from one language into this intermediary language specially worked out for machine translation, and then from the intermediary language into the other language. Several synonymic words or combinations of words of any other language can correspond to a single elementary meaning which appears in the role of a separate word in the intermediary language.

Experimental translation rules (suitable so

far for the translation of merely one small text) have been formulated that make it possible to translate the text (from English and French) at first into the intermediary language and then to obtain various synonymic Russian translations for each of the sentences in the text being translated.

Continuing the work in this direction, it will be possible to raise the question of obtaining with the aid of machines sufficiently good word for word translations accompanied by several synonymic translations. The selection of the best translation can be done by either man or by the machine itself.

Such kinds of investigations are very important for studying the flexibility of language, i.e. various methods of transmitting one and the same meaning existing in one or another language. Thus, it has been established that in translating a single sentence, in principle several hundred Russian synonymic sentences can be selected that differ from one another by at least one word. It turns out to be very difficult for a person to envisage all the possible synonymic sentences in one language; therefore machines can be utilized for studying all the rich possibilities of a language. This problem is directly connected with the study of the language of fiction since the unconscious or conscious selection of the best possible methods of expression of one and the same thought is constantly made by every writer.

The formal study of poetic language can make use of the precise methods of statistics and the theory of probability for describing the rhythmic pattern of poetic texts. The fruitfulness of this approach was especially clearly shown in the works of Boris Tomashevsky and other Russian scholars of poetry. Quite recently the mathematical analysis of verse attracted the attention of eminent mathematicians. Academician Andrei Kolmogorov and his collaborators made a detailed investigation of Russian verse (in particular, iambic tetrameter). New ideas and methods of the mathematical theory of information were applied in these in-

vestigations. The results were reported to the first conference on the application of mathematical methods in the investigations of the language of belles lettres which took place in Gorky in September 1961.

The study of language by statistical methods, the theory of probability and the theory of information require that much statistical work be undertaken and that the frequency with which one comes across words, combination of words, of letters, etc. be investigated.

These statistical data can be compiled by using computers and analytical machines for processing the language of belles lettres, the colloquial tongue, and the language of the daily press and radio.

Machines can compile complete dictionaries of the language of an author, calculate the number of cases in which separate words are used, as well as their combinations, in which separate letters and their combinations are used, etc. Thanks to the possibility of such statistical processing of language texts, tests are now being started in decoding ancient inscriptions with the aid of computers.

Besides V. Ustinov, a staff worker of the Mathematical Institute of the Siberian Branch of the USSR Academy of Sciences, Yuri Knorozov, member of the Linguistics Section of the Cybernetics Scientific Council of the USSR Academy of Sciences and a specialist on inscriptions, took part in this work. Preliminary work was accomplished by making a catalogue of the Maya letters and recording the text to be fed to the machine. After that, work was launched at the Mathematics Institute on decoding these texts with the aid of machines. The initial results are quite close to the results of decoding conducted by Yuri Knorozov several years ago by hand.

The great resemblance of the results shows that in both the decoded texts the reading of the main mass of letters is correct. The results of these experiments revealed that the machine can perform a number of human operations in decoding the next.

Evidently it will be possible in the near future to transfer to the machine also several other types of work in the investigation of language besides purely statistical work and work on decoding. Thus, O. Kulagina reported at the

conference on the processing of information and machine translation, and at the Fourth USSR Mathematical Congress in Leningrad in 1961 on the experiments conducted at the Steklov Mathematics Institute in giving a syntactical description of the language by the machine itself. Sentences preliminarily analysed by man are fed into the machine; after that the machine gives, in accordance with definite rules, a description of the syntactical combinations of words that are encountered in these texts. Thus, linguistics (as partially mathematics, too) is on the way to partial automation, when the more laborious work in studying a language will be transferred to the machine while man will have only to make up rules for the machine's work and understand the results obtained by it.

Everyone speaking his native language follows the rules that have been mastered in early childhood and which have become involuntary. In cybernetic terminology this could be expressed by saying that with the recollections of childhood which are pushed back into the subconscious, a program is fed to man which afterwards acts in the course of his entire life. The instruction given in the schools on the grammar of the native tongue is aimed at helping the pupils to become aware of these involuntary rules. In this sense the teaching of the native language is an example of the conscious control of the unconscious, which is necessary not only in respect to language but also for the formation of an all-round developed individual.

The Soviet psychologist L. Vygotsky, in his work *Thinking and Speech*, has shown that the awareness of the rules of the native language is a turning point in the life of the child and marks his transition to the logical thinking of the adult. The role of the science of language in the development of mankind can be boldly compared with the role of the conscious study of the language in the development of a single man. Investigation of a language—the main means of communication and the instrument for understanding reality—by precise methods can have far-reaching consequences for the spiritual development of mankind.

VYACHESLAV IVANOV

*Cybernetics Scientific Council,
USSR Academy of Sciences*