THE ACADEMY OF SCIENCES, U.S.S.R.

D. Panov

CONCERNING THE PROBLEM OF MACHINE TRANSLATION OF LANGUAGES

1956

CONCERNING THE PROBLEM OF MACHINE TRANSLATION OF LANGUAGES

D. Panov

1. Introduction

Strictly speaking, my paper is not concerned with problems of the theory of information in its classical form, I mean the form it has adopted in the works of C.E. Shannon, Ph. M. Woodward, L.L. Davies and others, which problem have also been studied by the Soviet mathematicians Kolmogorov, Khinchin, Kotelnikov and Siforov.

The problem of machine translation of languages belongs, in my opinion, among certain specific problems, which until now have not been taken up by learned men working in the field of the theory of information, but which, possibly, deserve to be taken up.

Up to now, the theory of information has developed as a purely mathematical theory, applying the means of the calculus of probability to formalise the original material it has dealt with. Claude E. Shannon does not make any distinction between discrete sources of information and stochastic process (1, p. 00), thus ignoring all the individual qualities of the information given, preserving only its statistic characteristics. In speaking of telegraph messages, Shannon points out that these messages consist of letter sequences which are "not completely random" and which, "in general form sentences and have the statistical structure of, say, English". They go on to say that "this structure allows one to make a saving in time (or channel capacity) by properly encoding the message sequences into signal sequences" (l, p.00).

The successful application of statistical methods in cryptography naturally led some investigators to the idea that a similar approach could be as promising in solving problems of automatic translating. Thus Weaver thought it most tempting to say that "a book written in Chinese is simply a book written in English, which was coded into the "Chinese code" (2, p. 22), and to translate from one language into another one needs "to descend from

each language to the common base of human communication-the real but as yet undiscovered universal language- and then re-emerge by whatever particular route is convenient" to the other individual language (2, p. 23). This idea being developed, one very naturally comes to the necessity of analysing the logical structure of language as a means of solving the problem of automatic translation. Investigations in this field undertaken by O. Jespersen (4), Ch. C. Fries (3), and others seemed to be providing mathematicians with material highly promising. Should it be achieved, the problem of automatic translation would join as equal those profound problems united under the name of theory of information. Unfortunately, so it seems to me, we must refrain from this tempting road. The very nature of the problem of translation is such that individual features of the translated text cannot quite be ignored. This circumstance makes the direct application of the methods of the theory of information impossible, but it does afford certain other advantageous prospects. I believe here we are faced with a problem which, though statistical in character, requires special methods of analysis, similar to the experimental methods used in the study of natural phenomena. I shall note that it is just the problems, at the borderline of two branches of science, that often prove to be most helpful in scientific research work.

2. Some Peculiarities of Machine Translation.

I should like to give a few examples to illustrate the necessity of considering individual lexical peculiarities in translating. Investigation of linguistic structures can be of interest in many ways, but from the point of view of translation it aims, and here I am convinced, at secondary questions and, thus, leads away from the solution of our problem. The reason is that in taking to generalised structures, one can not but lose a number of peculiarities of the text which are of vital importance for the translation. Mathematically there does not exist a one-to-one correspondence between the sentence structures in different languages. Two English sentences of the same structure, when correctly translated into Russian may produce two Russian sentences of structures entirely different. (See Fig. 1). It is as easy to give sentences which are impossible to

translate correctly without knowing the context. In figure 2 a stanza from a poem by Swinburne is given. The pronoun "mine" in the last sentence of the stanza, can not be correctly translated into Russian without referring to the preceding sentence, while in the English text it presents no ambiguity. Some other examples are given by Charles Fries in his book, "The Structure of English". While analyzing possible meanings of the phrase, "The King of England' s empire" (3, p. 265) he comes to the conclusion that the structural grouping here is ambiguous. The meaning of the phrase can be determined only by some supplementary means (and Fries suggests intonation for one). Thus it becomes evident that in many cases, a structural analysis is useless, and other means must be utilized to determine the meaning. The authors of the structural methods in linguistics are well aware of the shortcomings of their system. Thus, Fries writes in the book just mentioned:

"In this book I have challenged the traditional uses of meaning as the tool of analysis in dealing with sentence structure".

" . . . I have, not, however, repudiated "all uses of meaning" in linguistic analysis, and I have deliberately insisted that all substitution procedures demand for their use the control of certain aspects of meaning". (3, pp. 293-294).

Our idea is that both lexical meaning and grammatical characteristics of the word can and should be considered in translating languages. It is highly unpractical to decline the information which can be thus obtained. We find in "Analytic Syntax" by Jespersen (4, p. 105) the following admission: "However much we may try to speak of pure syntax as apart from morphology (accidence) considerations of form will necessarily force themselves on us here and there". Precisely for this reason, in analysing the structure, recourse must be made to quite complex systems of supplementary elements introduced into the structural systems which finally lead to the classical morphological analysis.

The consideration of the lexical meaning of words as well as of their context may be useful in solving problems of coding. These problems are extremely important and I shall later return to them. Here I should only like to note that under certain conditions, great economy can be achieved in coding. There is a popular game for children in my country, which consists

in guessing a word, of which only three data are given: its first and its last letters, and the number of letters between these two. The guesser is allowed a certain number of attempts for suggesting any letter he likes; if any of them belong to the word, they are entered in their place. To win the game one must guess the whole of the word. I wonder whether children play this game in your country, but in any case it is quite instructive in solving code problems. It proves that under certain conditions, very few attempts are needed to guess a word. And with lexical meaning of the word taken into consideration the economy in coding still increases; take crossword puzzles as an illustration.
I should like to note that codes of this type can be easily used for a dictionary stored in the computer. As it seems, it would not take too long to get the computer accustomed to solving crossword puzzles.

3. Machine Translation from English into Russian at the USSR Academy of Sciences.

Having started work in machine translating, we very soon came to the conclusion that it should be organized on lines different from those described in the reports on American experiments. In our opinion, excessive contact between the translation programme and the dictionary ascription of the control codes directly to the words in the dictionary cannot but limit the possibilities of translation making the solution of the problem extremely complicated. Therefore, we made it our point to work out basic principles of machine translation before starting. Our five basic principles are the following:

1. Maximum separation of the dictionary from the translation programme. This enables us easily to enlarge the dictionary without changing the programme.
2. Division of the translation programme into two independent parts: analysis in the foreign language sentence, and synthesis of the corresponding Russian sentence. This enables us to utilize the same Russian synthesis programme in translation from any languages.
3. Storing all the words in the dictionary in their basic form. This enables us to make use of the standard Russian grammar in the synthesis of Russian words.
4. Storing in the dictionary a set of invariant grammatical characteristics of the word.

5. Determination of multiple meaning of the words from the context whereas their variant grammatical characteristics are defined by analyzing the grammatical structure of the sentence,

These principles have proved quite reliable in the practical test they were put to, and hence they must be considered as basic in the solution of the problem.

I shall describe here our dictionary and translation programme worked out for machine translating from English into Russian.

Our dictionary for machine translation consists of two sections.

The first section contains English words, coded into digit combinations. Here we have also dictionary information of the corresponding Russian words, together with their number in the Russian Section of the dictionary. Thus, for nouns the following information is given: gender, declension, soft or hard stem, presence or absence of sibilants in the stem, animate or inanimate object expressed, etc. ; verbs are given information concerning conjugation, aspect, etc.; in the case of adjectives the only information given is soft or hard stem. We call this section of the dictionary the English Section.

The other section does not contain anything but Russian words coded into digit combinations. The second section is called the Russian section of the dictionary. Every English word in the dictionary is provided with a certain ordinal number, that is with a special place- in- the-vocabulary indication. This compilation came as a result of linguistic analysis of mathematical texts, starting with Milne's "Numerical Solution of Differential Equations". For the practical experiments in MT carried out on BESM (the USSR Academy of Sciences high-speed electronic computer) a dictionary of 952 English and 1,073 Russian words was compiled.

For a number of English words (121 words, in our case), the place-in-the -vocabulary indication is replaced by special digit indication to show that these words have multiple meaning. The proper Russian word is chosen in this case by utilizing a special programme of automatic translation, which we call "the Polysemantic Dictionary"

If the spelling of the word in the text coincides exactly with that of a word in the dictionary (or rather their numerical codes coincide), it can easily be identified by the operation of matching. This is the principle used for

finding words in the dictionary.

In case the word in the text is inflected, i.e. has some grammar affix (say, "s", or, "ing", or, "ed"), a special programme for discarding these affixes is used. Then the search of the word with the discarded affix should be repeated.

To determine multiple meaning context analysis is performed which consists in analysing the surrounding words when both their meaning and grammatical characteristics are taken into consideration. The routine for determining the multiple meaning of a word has emerged from elaborate analysis of a great number of English texts.

It is the Polysemantic Dictionary that deals with idioms. A "key word" is found in every idiomatic expression and marked in the dictionary "polysemantic". This brings about checking of the words preceding and following it for participating in the idiom. If the answer is "yes", the translation is given to the idiomatic expression as a whole.

Take for instance, the English word "able" which is ordinarily translated into Russian as the adjective "способный" (capable), but when linked with any form of the verb "to be" (is, are, was, were, been, being) it can no longer be translated separately, instead it affects the translation of the verb, and the two of them acquire as their Russian equivalent the verb "мочь" (can). Therefore, before giving the translation "способный" for the word "able", the preceding word should be checked up for "is", "are", etc. If the answer is "yes", the correct translation for the whole word combination is "мочь" It should be noted that the system of the most simple and general criterions for determining the multiple meaning of a word (or a group of words) has required a lot of preliminary research work on the part of our linguists.

If a word in the English sentence is not found in the dictionary, it is stored unaltered in the memory of the machine. When the translated sentence is put out, such a word will be printed in Latin script.

A study for the design of an automatic dictionary presents special interest. In our group it has been carried out by L. N. Korolyov. Here are some of his results.

For the speed of the entire translating process the time spent in dictionary search is decisive. Therefore, much attention must be given to questions of speeding up this part of the programme. One of the quickest ways of finding words proved to be the following. The codes of the English words in the dictionary are arranged in the increasing order. Thus, by comparing the code of the word sought with that of a word taken at random from the dictionary, we can easily see whether to go on with our search in the section preceding or following our word. The method of search just described resembles very much that of finding roots of an equation by the "supposed conditions" method. It is remarkable that only 14 check- ups are required ($14 = \lg_2 10{,}000$) to find a word in a 10,000-word dictionary by this method, and in a 30,000-word dictionary, no more than 15 check-ups are necessary.

Another problem that has been studied by L. N. Korolyov is that of saving the memory space required for storing the dictionary. The "code compressing" method has been worked out for the purpose.

The operation of compressing the codes to value a makes any code $\leq\alpha$. Let the codes be represented by some combinations of binary digits i. e., of certain combinations of 0 and 1.

Here are a few examples of the operation $\nabla\alpha$, "compressing the code to value α".

1°. Discard the code positions, the ordinal number of which is greater than α.

2°. If the length of a code being compressed is greater than α, divide it into sections so that the lengths of each section be less than α.

This is followed by the binary arithmetic addition of the sections. Division into sections is correct when the code for the sum of these sections $\leq\alpha$ in length.

3°. Here too, long codes are divided into sections but instead of binary arithmetic addition positional addition is used (1+1=0, 0+1=1, 1+0=1, 0+0=0) in order to obtain the compressed code.

Codes for all the words in the dictionary are divided into several groups depending on the number of letters the word consists of. The operation of compressing $\nabla\alpha$ is performed for each of these codes. The words in the text undergo the same "code compressing" operation $\nabla\alpha$ before they are sought

in the dictionary. Thus matching is fulfilled for the "compressed code". Although the "code compressing" operation may result in having similar codes for two or more different words in the dictionary, the probability of the case has proved exceedingly small *).

Should it happen, however the codes of these words as, indeed, of all the words of that group will be compressed in a different manner, say with a shift by one position. The corresponding group of words in the text will be treated in a similar way before being matched in the dictionary.

The significance of this method lies not only in the fact that it provides considerable storage economy but in that it enables us to simplify the search routines as well as to speed up the whole translating procedure.

The word being found in the dictionary, all the information concerning this word, including its place-in the vocabulary indications both in the English and in the Russian sections, and grammatical information as well, is taken from the dictionary to form the numerical equivalent of the word. All subsequent operations are performed with these numerical equivalents. Every numerical equivalent is stored in two cells of the memory, the grammatical indications having constant positions. Thus necessary information of the word can be acquired automatically.

The value of the equivalents is not limited to the fact that they contain necessary information about words. The replacing of words by their numerical equivalents makes grammar part of the programme universal in application, since it no longer depends on concrete texts or dictionary.

The automatic translation programme concerned with grammar is divided into two main parts, these being Analysis and Synthesis.

In the first part the form of the English words, their place in the sentence, as well as their dictionary grammatical information are taken

---

*) For a 2^{10} - word dictionary, when the codes are compressed to 30 binary positions by means of the operation $\nabla\alpha$ as described in 3°., this probability is about 0.0005.

into consideration with a view to determination of both grammatical form of the corresponding Russian words and their place in the Russian sentence. The resulting information is recorded in the numerical equivalent of the word, thus permitting to pass on to the second part of the programme, that is "Synthesis of the Russian sentence". Here, Russian words, taken from the dictionary in their basic form, acquire grammatical form in accordance with the characteristics resulting from the analysis.

Both English and Russian Grammar is presented as a series of special schemes for the following parts of speech: verbs, nouns, adjectives, numerals. These are supplied with two more schemes, one dealing with syntactical analysis of the sentence, the other changing word-order in the Russian sentence if necessary. The working basis of each scheme is dichotomic analysis, i. e., a system of "checking up" for the presence or absence of a certain grammatical (morphological or syntactical) characteristics of the analysed word. These alternative check-ups may result in only two answers, either positive or negative, which mean either a final conclusion and development of a certain grammar indication in the numerical equivalent of the word, or passing on to the next check-up until the necessary grammatical characteristics of the word is completed.

Different parts of the programme are applied in a sequence which can ensure the development of the indications necessary to carry out further operations.

Figures 3 and 4 present part of the schemes for the Polysemantic Dictionary and for the English Noun. The following symbols are accepted in the schemes: A/B, C/ means passing on to No. B in the case of the positive answer, whereas negative answer will result in passing on to No. C. Obviously A/B, B/ means passing on to No. B in both cases (that is without special checking-up) A/0,0/ means that the final result is acquired and no further search is necessary. In the Polysemantic Dictionary scheme both figures and letters are used as symbols, whereas in the other schemes only figures are utilized.

S. N. Razumovsky has been studying the logical structure of the schemes and programmes for machine translation. He has developed a system of symbols to fix the contents of the schemes in a unifying manner. In Figure 5 the English Noun Analysis scheme is present in the symbolized form. The symbol system

of S. N.Razumovsky allows a unified recording of the formation processes, thus permitting an automatic compilation of the machine translation programme. Figure 6 presents part of the information used to automatically compile a programme of analysis for an English noun.

The Russian Synthesis programme is fully independent of the foreign language Analysis. The form of the Russian word is determined by the grammar characteristics found in the dictionary together with those developed in the English Analysis part of the programme. Hence, the Russian Synthesis part of the programme depends on Russian Section of the dictionary. Russian schemes may vary in form depending on the role relegated to the dictionary. It has been mentioned, that in our dictionary words are kept in their basic form, i. e., nominative singular for nouns, nominative singular masculine for adjectives, infinitive for verbs, etc. The primary concern of the Russian Synthesis programme is to modify the endings of the dictionary words, with necessary provision for the interchange of vowels and consonants, when required.

The main difference between the Russian Synthesis programme and the Analysis programme (English Analysis in this case) lies in the fact that separate words and their characteristics in the former programme are considered with no reference whatsoever to the neighbouring words.

Figure 7 presents part of the Russian Verb Synthesis scheme.

In Figure 8 principle steps in the machine translating of an English sentence are given. The sentence: "The cause of this phenomenon will be considered in the following articles", goes to the operative memory in coded form, as a 146 digit number:

21260800221620050800211400212612050024 26
08152881108152815001312272700060800222815 05
12300807083000121500212608001428272628
13121510001607211222270805

With every two digits coded in the binary system.

Then the sentence is broken up into words and the words are sought in the dictionary.

For only five words of this sentence, we can find in the dictionary the corresponding Russian words, with its grammar characteristics attached.

These words are:

1. phenomenon - явление (noun, neutral gender, first declension, soft stem)
2. consider - рассматривать (verb, first conjugation, imperfect aspect, takes the accusative case)
3. article - статья (noun, feminine gender, second declension, soft stem)
4. will - is not translated (modal verb; predicate)
5 be - быть (verb, first conjugation, perfect aspect)

Note that the machine could find the words "consider" and "article" only after it had discarded the endings "s" in "articles" and "ed" in "considered"

For the rest of the words the indication is given in the dictionary that they have multiple meaning, which means that a special analysis of the sentence is required in order to choose the correct meaning of the word.

The English word "in", most often translated as the preposition "в", may be translated in certain word combinations in a different way, say, preposition "к" interest in (интерес к) or "с", or "при " and so on. In our sentence context analysis shows that "in" should be translated as preposition "в".

All the dictionary information (the information taken from the Polysemantic Dictionary included) is recorded in the form of numerical equivalents of the words of the sentence. The Russian place-in-the-vocabulary indications of the equivalents furnish the following list of Russian words:

причина	этот	явление	быть	рассматривать
(cause)	(this)	(phenomenon)	(be)	(consider)
	в	следовать	статья	
	(in)	(follow)	(article)	

One can easily note that several words in the English sentence have not been translated into Russian (the, of, will), this being recorded as "omit"-indication in their equivalents.

The grammatical analysis programme supplements the originally received

from the dictionary equivalents by a series of characteristics required later for the Russian Synthesis programme. Thus, for the word "article", the English Analysis programme furnishes the following additional characteristics: plural (since the word in the text has the "s" ending), and prepositional case (since preposition "in" precedes the word). Together with characteristics obtained from the dictionary, programme characteristics furnish sufficient information for the Russian noun Synthesis programme to produce the correct ending of the word "статья" (article).

For the word "follow", the English Verb programme develops the following characteristics: present participle ("ing" having been discarded), plural, prepositional case, (the two latter characteristics being inferred from the noun following our participle).

Finally, the Russian Synthesis programme having been fulfilled, we get a Russian sentence, which is correct both semantically and grammatically. Here is the sentence:

"Причина этого явления будет рассмотрена в следующих статьях".

I should like to emphasize the fact that from the very input of the English sentence into machine, the entire translation process has been carried out automatically with no human intervention whatsoever. Enormous preliminary research work is required of philologists to make the machine translate in the manner just described. In our group the linguistic research work has been carried out by I. K. Belskaya, our Philologist-in-Chief, while the mathematical part of research has been done by the mathematicians I. S. Mukhin, L. N. Korolyov, S. N. Razumovsky, G. P. Zelenkevich, and partly by N. P. Trifonov.

In Figures 9 and 10 you can see sentences translated by machine from English into Russian. In Figures 11 and 12 a few illustrations are given of incorrect translations, with mistakes due to errors in the coding. The tapes shown in these photographs have actually been produced by the machine.

4. Further Studies in the Field of MT

Our opinion is that principally, the lines of which machine translation of languages should be organized, have been sufficiently developed by now and the time has come to consider the opportunities for practical work in this

field on a larger scale. With this view, we have started research work in automatic translation from German, Chinese, and Japanese into Russian. In Figure 13 part of the German noun scheme is shown, in Figure 14 you can see part of the Chinese numeral, and in figure 15 part of scheme analysing Japanese verbs. These are our recent achievements.

When taking up machine translation from Chinese and Japanese we had our doubts as to the problem of input in these languages. We have come to the conclusion by now, that Chinese telegraph code (Figs. 16 and 17) may perfectly solve the problem.

The group of linguists engaged in further studies in MT includes I.K. Belskaya, G. P. Zelenkevich, E. A. Khodzinskaya (German studies) A. A. Zvonov, V. A. Voronin (Chinese studies) and M. B. Yefimov (Japanese studies).

We intend soon to take up the problem of multilingual machine translation with foreign languages both at the input and output, while Russian serves as interlanguage into which the input is first translated. The choice of Russian interlanguage in our case proves highly practical. It is worth noticing that the method we have suggested is of great help in solving the problem of multilingual MT. The diagram in Fig. 18 illustrates the translation of an English sentence E into the corresponding Russian R. Letter V stands for Vocabulary, letter A stands for Analysis part of the programme; ER stand for "English-into-Russian". Letter S stands for Synthesis programme. VA symbolizes that both Vocabulary and Analysis part of work is over. It is evident that when so much as VA is accomplished, we are fully provided with Russian equivalents, with all their grammatical characteristics attached. Thus we can easily pass on to, say, a Russian-French dictionary, and get the corresponding French equivalents with their necessary grammatical characteristics. Thus using only the French Synthesis programme we can obtain a French sentence automatically translated from English.

We do not make any secret of our work, and last summer, when a group of American engineers were visiting our Institute, we mentioned our experiments in automatic translation carried out on BESM. We made a statement concerning our first achievements at the Conference on Computing Techniques in Moscow and at the two international conferences in London and Brussels. Certain

information on our first results was given in our first publications (5,6,7,8). Nevertheless, the French journal "TSF Phono-Ciné-Electricité" (1956, no. 730, p. 6) published not very long ago an article bearing the title "Pour percer le secret du cerveau electronique russe". I believe, it is clear from what was told above that we have no special "secret". We have found a method of solving the problem by combining mathematical approach with very concrete linguistic analysis, thus not going too far into formal methods of investigation. I shall conclude with a statement made by Jespersen, in his "Analytic Syntax" (4, pp. 13-14):

"The symbols here introduced to some extent resemble the wonderful system of symbols which during the last few centuries has contributed so much to make mathematics (and in some degree logic) exact and more easy to manage than was possible with the unwieldy word-descriptions used formerly. My system aims at providing linguists with some of the same advantages. But it cannot pretend to the same degree of universality as either the chemical or mathematical symbols. That is precluded simply because of the fact, which it is no use shirking, that language is everywhere socially conditioned and there is no getting away from that".

ТОЖДЕСТВЕННЫЕ РУССКИЕ СТРУКТУРЫ	РАЗЛИЧНЫЕ АНГЛИЙСКИЕ СТРУКТУРЫ
У нас есть черный кот.	We have a black cat.
У нас есть собственный дом.	We have a house of our own.

ТОЖДЕСТВЕННЫЕ АНГЛИЙСКИЕ СТРУКТУРЫ	РАЗЛИЧНЫЕ РУССКИЕ СТРУКТУРЫ
He should knock at the door before coming in.	Следовало бы ему постучать в дверь, прежде чем входить.
He would knock at the door before coming in.	Он бывало постучит в дверь, прежде чем войти.

Fig. 1

Can I forget? yea, that can I,
And that can all men; so will you,
Alive, or later, when you die,
Ah, but the love you plead was true?
Was mine not too?

Fig. 2

СУЩЕСТВИТЕЛЬНОЕ (Английская часть)

1(2,7)	Проверить данное слово на us.
2(3,5)	Проверить следующее слово на существительное.
3(0,0)	Выработать признак дательного падежа.
5(6,13)	Проверить предыдущее (непосредственно) слово на let.
6(0,0)	Выработать именительный падеж.
7(8,13)	Проверить данное слово на it.
8(13,10)	Проверить it на наличие признака какого-либо рода.
10(0,0)	Взять род от ближайшего предшествующего подлежащего.
13(14,15)	Проверка на наличие признака единственного или множественного числа.
14(0,21)	Проверка на наличие признака любого падежа
15(16,19)	Проверка на окончание - S.
16(17,17)	Выработка признака множественного числа.
17(18,14)	Проверка предыдущего слова на формулу без = .
18(0,0)	Выработка признака родительного падежа.
19(16,20)	Проверка предыдущего слова на much (*).
20(14,14)	Выработка признака единственного числа.
21(22,23)	Проверить предыдущее слово на let.
22(0,0)	Выработка признака именительного падежа и подлежащего.
23(24,28,	Проверить предыдущее слово на признак "союз однородный".
24(28,25)	Проверить непосредственно предшествующее и следующее (относительно союза однородного) слово на прилагательное.
25(26,27)	Проверить все слова на то же слово, что и данное.
26(0,0)	Взять падеж от найденного существительного.
27(0,0)	Взять падеж от ближайшего предшествующего существительного.
28(18,29)	Проверка на окончание 'S.

Fig. 3

61 (a,62)	**look**
a(в,c,)	Проверить следующее слово (или ближайший следующий предлог) на **for (+).**
в(0,0)	ИСКАТЬ (глагол, I спряжение, несовершенный вид + винительный падеж).
c (dd, e)	Проверить следующее слово (непосредственно) на **upon.**
d (0,0)	СМОТРЕТЬ (глагол, II спряжение, несовершенный вид).
e(0,0)	РАССМАТРИВАТЬ (глагол, I спряжение, несовершенный вид + винительный падеж). Вставить после него (непосредственно) КАК (частица).
62(a,63)	**many, much**
a(в,c)	Проверить предыдущее слово (непосредственно) на **how.**
в(0,0)	СКОЛЬКО (числительное, не склоняется).
c(d, e)	Проверить предыдущее слово (непосредственно) на **as.**
d(0, 0)	СТОЛЬКО ЖЕ (числительное склоняется).
e(g, i)	Проверить данное слово на **much.**
f(0, 0)	Не переводится (наречие).
g(f, k)	Проверить предыдущее слово (непосредственно) на **very.**
h(0, 0)	МНОГИЙ (прилагательное, твердая основа, с шипящим).
i(h,l)	Проверить предыдущее слово на предлог и следующее на существительное.
j(0, 0)	МНОГО (наречие).
k(l, j)	Проверить следующее слово на существительное.
l(0, 0)	МНОГО (числительное, склоняется).

Fig. 4

0* $[K = 0; 0]$

0 $\{\alpha[K+1](a_{1,K} = 1)\,1; (a_{2,K} = N_{точки})\,B; \alpha\}$

1 $\{(a_{2,K} = N_{US})\,2; 7\}$

2 $\{(a_{1,K+1} = 1)\,3; 5\}$

3 $[b_{1,K} = 3; 0]$

5 $\{(a_{1,K-1} = N_{let})\,6; 13\}$

6 $[b_{1,K} = 1; 0]$

7 $\{(a_{2,K} = N_{it})\,8; 12\}$

8 $\{(a_{6,K} = 0)\,10; 0\}$

10 $\{[i = K]\,\alpha[i-1]\{(a_{4\,i} = 1)\beta; \alpha\}\,\beta[a_{6,K} = a_{6,i}]\}$

Fig. 5

№40	011	63.15	33.0	19.0
1	212	0.0	46.0	0.0
2	101	34.0	0.0	37.0
3	011	63.14	35.0	19.0
4	121	44.0	0.1	36.0
5	113	39.0	16.1	16.0
6	121	41.0	0.1	40.0
7	211	0.1	46.0	39.0
8	131	16.2	16.0	41.0
9	211	0.1	40.0	49.0
10	121	46.0	63.14	38.0
11	011	0.1	37.0	39.0
12	131	16.2	16.0	41.0
13	212	10.13	40.0	0.2
14	121	46.0	0.7	38.0
15	011	0.2	37.0	39.0
16	131	16.1	17.0	41.0
17	212	0.1	40.0	0.3
18	121	46.0	0.5	38.0
19	011	0.3	33.0	6.1
20	321	16.0	0.3	46.0
21	210	0.0	34.0	0.5
22	111	3770	39.0	16.2
23	322	29.0	41.0	27.6
24	121	40.0	0.6	46.0
25	210	0.13	38.0	0.6
26	113	33.0	6.1	16.0
27	212	0.1	46.0	0.0
28	101	34.0	0.7	37.0
29	113	39.0	16.2	16.0
30	121	41.0	8.9	40.0
31	212	0.8	46.0	0.13
32	1	38.0		

Fig. 6

115(116,120)	Проверка на наличие признака "множественное число".
116(117,118)	Проверка на наличие признака "1 лицо".
117(0,0)	Добавить окончание -ЕМ.
118(0,0)	Добавить окончание -УТ.
119(115,115)	Взять данное слово без изменения из словаря, но перед ним поставить другое слово: БУД - с окончанием, найденным по схеме.
120(0,0)	Добавить (к БУД-) окончание -ЕТ.
122(123,124)	Проверка на слово ИСПРАВИТЬ.
123(0,0)	Отбросить три последние буквы и вместо них добавить к оставшейся части -ЬТЕ.
124(125,126)	Проверка на слово ПОКАЗЫВАТЬ.
125(0,0)	Отбросить последние шесть букв, затем к оставшейся части добавить окончание - ЖИТЕ.
126(0,0)	Добавить к оставшейся части слова окончание - ИТЕ.
128(129,130)	Проверка на слово ВЫЧИСЛЯТЬ.
129(157,157)	Перевод: ВЫЧИСЛИТЕЛЬНЫЙ. Выработка признака "твердая основа".
130(131,132)	Проверка на слово ИТТИ.
131(157,157)	Перевод: ПРОИСХОДЯЩИЙ. Выработка признаков "мягкая основа с шипящим".
132(133,134)	Проверка на слово ПРЕДШЕСТВОВАТЬ.
133(157,157)	Перевод: ПРЕДЫДУЩИЙ. Выработка признаков "мягкая основа с шипящим".
134(135,136)	Проверка на слово УДИВИТЬ.
135(157,157)	Перевод: УДИВИТЕЛЬНЫЙ. Выработка признака "твердая основа".
136(137,138)	Проверка на слово ОТЛИЧАТЬСЯ.
137(157,157)	Перевод: "различный". Выработка признака "твердая основа".

Fig. 7

«The cause of this phenomenon will be considered in the following articles».

БУКВЫ

2126080022162005080028140021261205002426
08152811081528150013122727000608002228150 5
1230080708300012150021260800 1428272728
131215100016072 11222270805

ЦИФРЫ

ВВОД

причина этот явление быть
рассматривать в следовать статья (ПРИЗНАКИ)

АНАЛИЗ

Причина этого явления будет рассмотрена в следующих статьях.

СИНТЕЗ

Fig. 8

Это было основано на дорогом эксперименте, проведенном мной и доктором R. H. Richens, от Кэмбриджского Университета, в котором мы разработали метод перевода малых отрывков выбранного текста на иностранные языки. Мы дали отчет о этом на конференции в Massachusetts в 1952, после которого I.B.M. компания в сотрудничестве с Джорджтаунским Университетом применили наши методы чтобы дать наглядную демонстрацию, которая была ограничена переводом нескольких предложений с русского на английский. Не имеется возможности в настоящее время перевода книги как произведения искусства.

This was based on an expensive experiment done by myself and Dr. R. H. Richens, of Cambridge University, in which we worked out a method of translating small sections of selected text in foreign languages. We gave an account of this at a conference in Massachusetts in 1952, after which the International Business Machines Company, in conjunction with Georgetown University, applied our methods to give a popular demonstration which was limited to translating a few sentences form Russian into English. There is no possibility at present of translating a book as a work of art.

Fig. 9

Equations involving more than one independent variable and the partial derivatives of the dependent variables with respect to the independent variables are called partial differential equations.

Suppose that both equations actually contain all the possible partial derivatives of second order.

It is necessary to find values between which the function f(x) is zero.

There are various numerical methods for this purpose.

Уравнения, содержащие более чем одну независимую переменную и частные производные зависимых переменных относительно независимых переменных называются дифференциальными уравнениями в частных производных.

Допустим, что оба уравнения действительно содержат все возможные частные производные второго порядка.

Необходимо найти значения, между которыми функция f(x) есть нуль.

Имеются разные численные методы для этой цели.

Fig. 10

THIS IS TRUE CERTAINLY OF THE VAST CATEGORY OF PROBLEMS ASSOCIATED WITH FORCE AND MOTION.

ЭТО ВЕРНО КОНЕЧНО ДЛЯ ОБШИРНОЙ КАТЕГОРИИ ЗАДАЧ СВЯЗАТЫХ С СИЛОЙ И ДВИЖЕНИЕМ.

Fig. 11

SO THAT WHETHER WE WANT TO KNOW THE FUTURE PATH OF JUPITER IN THE HEAVENS OR THE PATH OF AN ELECTRON IN AN ELECTRON MICROSCOPE WE RESORT TO DIFFERENTIAL EQUATIONS.

ТАК ЧТО ХОЧЕМ ЛИ МЫ ЗНАТЬ БУДУЩИЙ ПУТЬ ЮПИТЕРА В НЕБЕСАХ ИЛИ ПУТЬ ЭЛЕКТРОНА В ЭЛЕКТРОННОМ МИКРОСКОПЕ МЫ ПРИБЕГАЕМ К ДИФФЕРЕНЦИАЛЬНЫМ УРАВНЕНИЯМ.

Fig. 12

5(6,11)	Проверить предшествующее слово на неопределенный артикль.
6(0,0)	Выработать признак "единственное число".
7(38,19)	Проверить данное существительное на признак "местоименное".
11(12,13)	Проверить на наличие числительного перед данным существительным.
12(38,20)	Проверить найденное числительное на **zwei, drei, vier, beide.**
13(14,15)	Проверить на наличие перед данным словом определенного артикля.
14(6,18)	Проверить найденный артикль на **das, des, dem.**
15(16,17)	Проверить предшествующее существительное на признаки "местоименное" и склоняется, как прилагательное.
16(17,17)	Дальнейшая проверка одинаково распространяется на оба существительные: данное и найденное.
17(6,38)	Проверить на окончание **(e)m** или **(e)s** у данного (или найденного) существительного.
18(6,38)	Проверить наше существительное на окончание **e** и найденный артикль на **der.**
19(6,21)	Проверить на наличие перед данным существительным одного из слов: **ein, am, im, bein, vom, aufs, durchs, fürs.**
20(0,0)	Выработать признак "множественное число".

Fig. 13

КИТАЙСКИЙ ЯЗЫК	ЧИСЛИТЕЛЬНОЕ
23(24,25)	Проверить данное числительное на наличие перед ним форманта 第
24(0,0)	Переводится порядковым числительным, которое анализируется по схеме "прилагательное".
25(26,33)	Проверить данное числительное на наличие после него суффиксов единичности: 個(个), 本, 把(巴), 篇, 隻, 枝, 張, 座, 傢(家), 所, 根, 棵, 場, 次, 件, 陣, 方, 架, 枚, 名, 位, 盒, 台, 首, 頭, 番, 放, 家, 眼, 局, 部, 門, 箱, 股, 付, 匹,
26(27,29)	Проверить данное числительное на 一 (один).
27(28,30)	Проверить предыдущее слово на указательное местоимение: 這(这), 那. и определительное местоименение 每.
28(0,0)	Числительное 一 (один) не переводится?
29(0,0)	Переводится количественным числительным, которое анализируется по схеме "существительное".
30(28,31)	Проверить предыдущее слово на притяжательное местоимение 我, 我們, 我的, 我們的, 他, 他們, 他的, 他們的, 她, 她們, 她的, 她們的, 你, 你們, 你的, 你們的, 您, 您的, 咱, 咱們, 咱的, 咱們的 它, 它的, 其
31(28,32)	Проверить предыдущее слово на признак прилагательное.

Fig. 14

11(12,13) Проверить данный глагол на наличие окончаний
ばれる или はれる или ちれる или
がれる или られる или へられる или
びられる или されろ

12(24,24) Выработать признаки страдательного залога и изъявительного наклонения и перейти к определению времени по схеме.

13(14,15) Проверить данный глагол на наличие окончаний
ば(-) или は(-) или ち(-) или
が(-) или ら(-) или へら(-) или
びら(-) или さ(-).

14(25,25) Выработать признаки страдательного залога и произвести дальнейшую проверку.

15(16,17) Проверить данный глагол на наличие окончаний
めろ или けろ или ぜめろ или えろ

16(24,24) Вставить глагол МОЧЬ, выработать признаки неопределенной формы для нашего глагола и перейти к определению времени по схеме.

17(18,19) Проверить данный глагол на наличие окончаний
め(-) или け(-) или ぜめ(-) или え(-).

18(25,25) Вставить глагол МОЧЬ, выработать признаки инфинитива для нашего глагола и произвести дальнейшую проверку.

19(20,21) Проверить данный глагол на наличие окончаний
ぜろ или させろ .

20(24,24) Вставить глагол ДАТЬ, выработать признаки инфинитива для нашего глагола и перейти к определению времени по схеме.

21(22,23) Проверить данный глагол на наличие окончаний
ぜ(-) или させ(-) .

22(25,25) Вставить глагол ДАТЬ, выработать признаки инфинитива для нашего глагола и произвести дальнейшую проверку.

23(24,24) Выработать признаки действительного залога изъявительного наклонения и перейти к определению времени по схеме.

Fig.15

6600	6601	6602	6603	6604	6605	6606	6607	6608	6609	
迍	迎	近	迓	返	迠	迢	迤	迥	迦	
JTW	JTX	JTY	JTZ	JUA	JUB	JUC	JUD	JUE	JUF	
6610	6611	6612	6613	6614	6615	6616	6617	6618	6619	
迨	迪	迫	迩	迳	迷	迴	迷	迸	迹	
JUG	JUH	JUI	JUJ	JUK	JUL	JUM	JUN	JUO	JUP	
6620	6621	6622	6623	6624	6625	6626	6627	6628	6629	
追	迺	退	送	适	逃	迮	逆	逋	逸	
JUQ	JUR	JUS	JUT	JUU	JUV	JUW	JUX	JUY	JUZ	
6630	6631	6632	6633	6634	6635	6636	6637	6638	6639	
逍	透	逐	逑	途	逕	逖	逗	這	通	
JVA	JVB	JVC	JVD	JVE	JVF	JVG	JVH	JVI	JVJ	
6640	6641	6642	6643	6644	6645	6646	6647	6648	6649	
逛	逝	逞	速	造	逡	逢	連	逭	逮	
JVK	JVL	JVM	JVN	JVO	JVP	JVQ	JVR	JVS	JVT	
6650	6651	6652	6653	6654	6655	6656	6657	6658	6659	
週	進	逵	逶	逸	逿	逼	遇	遁	遂	
JVU	JVV	JVW	JVX	JVY	JVZ	JWA	JWB	JWC	JWD	
6660	6661	6662	6663	6664	6665	6666	6667	6668	6669	
遍	遈	遊	運	遍	過	遏	遐	遑	遒	
JWE	JWF	JWG	JWH	JWI	JWJ	JWK	JWL	JWM	JWN	
6670	6671	6672	6673	6674	6675	6676	6677	6678	6679	
道	達	違	遘	遙	遛	遜	遞	遠	遡	
JWO	JWP	JWQ	JWR	JWS	JWT	JWU	JWV	JWW	JWX	
6680	6681	6682	6683	6684	6685	6686	6687	6688	6689	
遣	遍	遝	遨	適	遭	遮	遯	遲	遴	
JWY	JWZ	JXA	JXB	JXC	JXD	JXE	JXF	JXG	JXH	
6690	6691	6692	6693	6694	6695	6696	6697	6698	6699	
遵	遶	遷	選	遹	遺	遬	遼	遽	避	
JXI	JXJ	JXK	JXL	JXM	JXN	JXO	JXP	JXQ	JXR	

Fig. 16

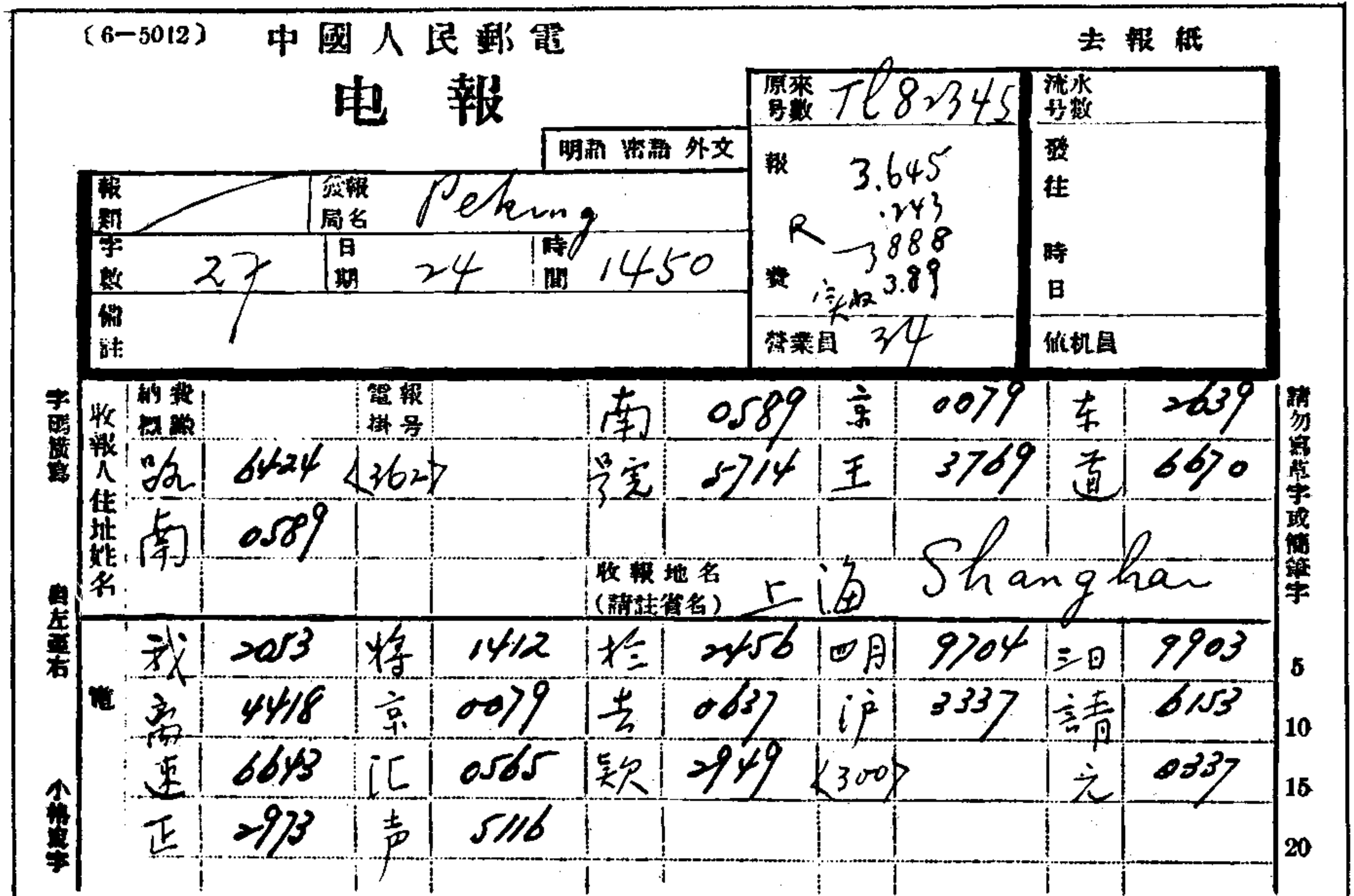
(6—5012)
中國人民郵電
电報
去報紙
明語 密語 外文
報類
發報局名 Peking
字數 27
日期 24
時間 1450
備註
原來号數 Tl 8·2345
報 3.645
R .243 3.888
費 3.89
營業員 34
流水号數
發往
時日
値机員
字碼譯寫
納費標識
電報掛号
收報人住址姓名
南 0589 京 0079 东 2639
路 6424 (3620) 號 5714 王 3769 道 6670
南 0589
收報地名（請註省名） 上海 Shanghai
由左至右
電
我 2053 将 1412 於 2456 四月 9704 三日 9903
离 4418 京 0079 去 0637 沪 3337 請 6153
速 6643 汇 0565 款 2949 (300) 元 0337
正 2973 声 5116
小佛寫字
請勿寫蕉字或簡筆字
5
10
15
20

Fig. 17

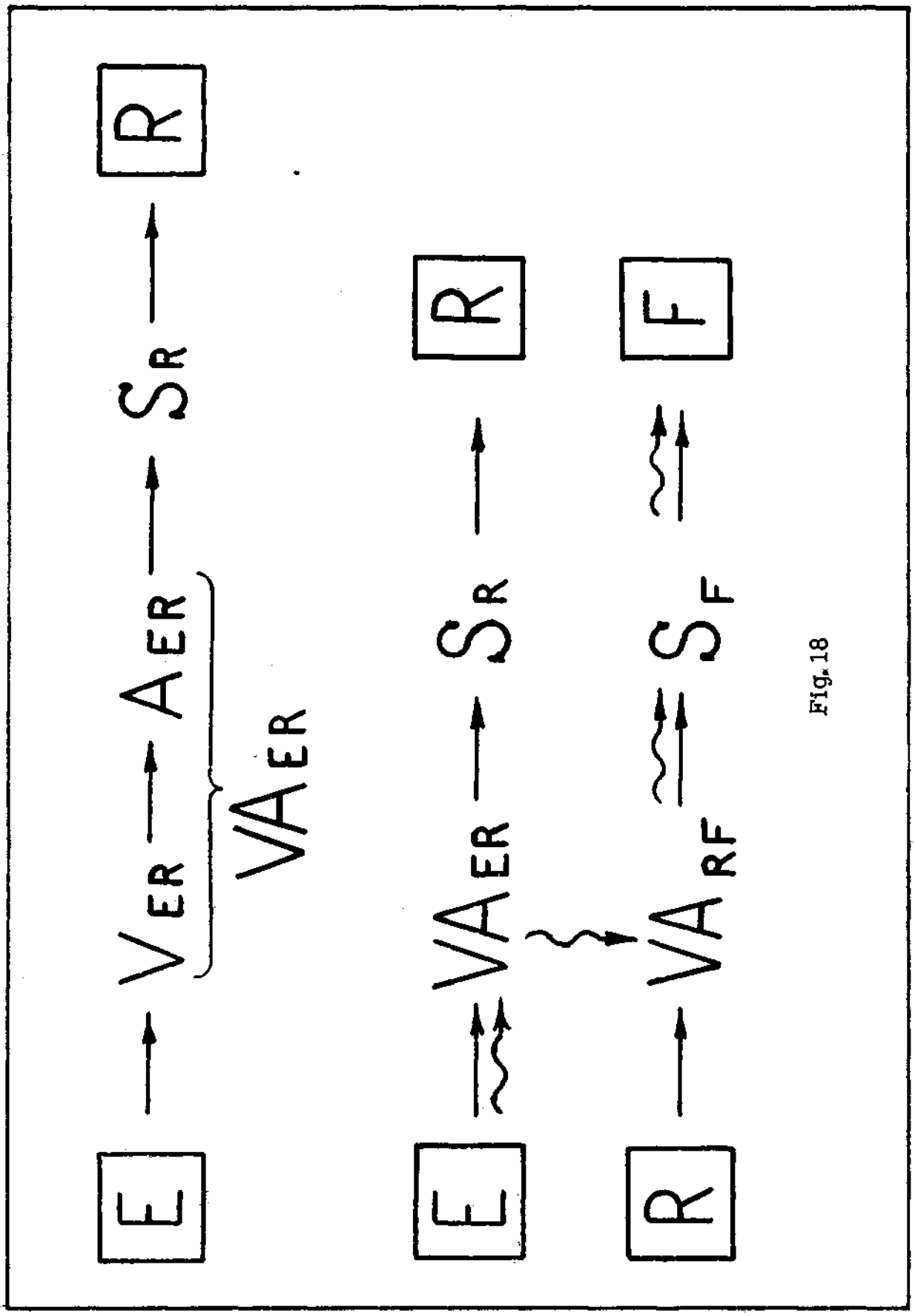

Fig.18

ЛИТЕРАТУРА

1. Shannon C. and Weaver W. Mathematical theory of Communication.

2. Machine Translation of Languages.
Edited by W. N. Locke & A. D. Booth. 1955.

3. Fries C.C. The Structure of English. N.Y. 1952.

4. Jespersen O. Analytic Syntax, Kopenhagen, 1937.

5. Д.Ю.Панов. Автоматический перевод. Москва, 1956.

6. Г.П.Зеленкевич, Л.Н.Королев, С.Н.Разумовский. Опыты автоматического перевода на электронной вычислительной машине БЭСМ, "Природа" No. 8, 1956.

7. L S. Mukhin. An Experiment of the Machine Translation of Languages Carried out on the BESM. Academy of Sciences, U.S.S.R.

8. И.К.Вельская, И.С.Мухин. Автоматический перевод с английского на русский с помощью БЭСМ - труды конференции "Пути развития советского математического машиностроения" /в печати/.