[From: *The Trend in Engineering at the University of Washington*, July 1956]

ENGINEERING PROGRESS IN MACHINE TRANSLATION

ROBERT E. WALL, JR.
Acting Instructor in Electrical Engineering

R E. Wall Jr

Research in the Electrical Engineering Department on the problem of the machine translation of language,* first reported by Dr. Thomas Stout in *The Trend in Engineering,*[1] and later by Mr. Geoffrey Douthwaite in the *Washington Engineer,* has proceeded at a steady pace. The phenomenal progress which has been made in computer components in the last few years, especially in memory devices, has brought machine translation closer to a possibility today than the contributors dreamed possible two or three years ago. The extreme scepticism of ten years ago has been steadily giving way to increasing optimism.

The specifications of the problem have remained substantially as listed by Dr. Stout in 1954. The machine-translation (MT) researcher is still considering seriously only the translation of scientific material, and the goal is still just understandability. The machine translator must be a special purpose machine, built specifically for MT and of no use for any other purpose.

The reader is referred to the papers by Stout and Douthwaite, and to a recent paper by the author[3] for consideration of some of the problems which will not be discussed in detail in this paper. However, sufficient background will be presented here so that the reader with no previous acquaintance with the MT problem will be able to follow the discussion.

The Translation Process

Four definite steps are required in the translation process: encoding, memory search, logical operations, and decoding. Let us consider these four steps in order.

Encoding. By "encoding" we mean the conversion of printed foreign language (source language) text material into machine code, perhaps teletype code. Electronic encoders have been announced in the literature,[4] but apparently their reliability is not yet up to that of the human transcriber. On the other hand, human transcribers would not be very efficient at this operation, since, for instance, an American typist transcribing Russian would be very slow and inaccurate. The development of a reliable electronic encoder should not be a very difficult project, and its importance is evident.

Dictionary Search. The permanent memory system of the machine translator may be considered to consist of two parts: the dictionary, in which the entire source language-target language dictionary is stored; and the stored programs for the necessary logical routines. The "source language" is defined as the language from which the translation is being made, and the "target language" as the language into which the translation is being made.

An incoming source-language text word must be stored in machine code in a rapid access device (perhaps a shift register) so that the code of the text word may be compared bit-by-bit with the source-language entries in the dictionary. As an example, suppose that the code that is being used is as shown in Tables I and II, and that the translation is from Russian to English. Suppose that the Russian word which is to be located is "эря." This would be coded "010001000011101." The method of entering material in the dictionary suggested by Hill[5] is used.

For example, эуб (tooth) and эря (in vain) would be stored in a section of the dictionary in binary code as follows :

. . . $I_1$010001001100010

$I_2$10100011110111110100010000 1,010001000011101

$I_2$010010111000000010110000010100101110 I_1 . . .

The machine would start comparing the code for the source-language text word ("эря" in this case) with the source-language word codes stored in the dictionary, one bit at a time starting with the left end of the word code. The symbol, I_1, means that the following word is in Russian, and the symbol, I_2, means that the following word is in English. The I_1 and I_2 symbols might be single space (five zeroes) and double space (ten zeroes), respectively, or, in some memory systems, some sort of punctuation symbols might be available. With this scheme, then,

* Engineering Experiment Station Project No. 157.

TABLE I

POSSIBLE RUSSIAN CODE*

SYMBOL	LOWER CASE CODE	CAPITAL CODE	SYMBOL	LOWER CASE CODE	CAPITAL CODE
а	00001	1111100001	п	01111	1111101111
б	00010	1111100010	р	10000	1111110000
в	00011	1111100011	с	10001	1111110001
г	00100	1111100100	т	10010	1111110010
д	00101	1111100101	у	10011	1111110011
е	00110	1111100110	ф	10100	1111110100
ё	1111011110		х	10101	1111110101
ж	00111	1111100111	ц	10110	1111110110
з	01000	1111101000	ч	10111	1111110111
и	01001	1111101001	ш	11000	1111111000
й	1111111010		щ	11001	1111111001
к	10010	1111101010	ъ	1111000010	
л	01011	1111101011	ы	11010	
м	01100	1111101100	ь	1111000011	
н	01101	1111101101	э	11011	1111111011
о	01110	1111101110	ю	11100	1111111100
			я	11101	1111111101
Space	00000		1	1111000100	
Comma	1111001110		2	1111000101	
Semicolon	1111010000		3	1111000110	
Period	1111010001		4	1111000111	
Etc.			5	1111001000	
			6	1111001001	
			7	1111001010	
			8	1111001011	
			9	1111001100	

*Where no capital code appears for a symbol, the Russian character never starts a word. Such characters also occur infrequently.

TABLE II

POSSIBLE ENGLISH CODE

SYMBOL	LOWER CASE CODE	CAPITAL CODE	SYMBOL	LOWER CASE CODE	CAPITAL CODE
a	00001	1111100001	n	01110	1111101110
b	00010	1111100010	o	01111	1111101111
c	00011	1111100011	p	10000	1111110000
d	00100	1111100100	q	10001	1111110001
e	00101	1111100101	r	10010	1111110010
f	00110	1111100110	s	10011	1111110011
g	00111	1111100111	t	10100	1111110100
h	01000	1111101000	u	10101	1111110101
i	01001	1111101001	v	10110	1111110110
j	01010	1111101010	w	10111	1111110111
k	01011	1111101011	x	11000	1111111000
1	01100	1111101100	y	11001	1111111001
m	01101	1111101101	z	11010	1111111010

when the search mechanism encounters an I_1 symbol, the machine starts to compare the source-language text word code with the following code sequence in the dictionary entry, one bit at a time. If any position shows lack of correspondence, the machine would discontinue the comparison and would start the comparison again, starting with the beginning of the incoming text word, when the next I_1 symbol is encountered. Thus in the above example, the machine would find correspondence between the code for "эря" and the dictionary entry until the ninth binary place is reached when the dictionary entry is found to be "1" and the code for "эря" has a "0" at this position. When the search mechanism encounters the next I_1 symbol, the comparison is started again, and this time complete correspondence is found to the I_2, signal, so the machine prints out the code sequence entered in the dictionary between this I_2 signal and the next I_1 signal. This sequence is the code for the English language equivalent for "эря," which is "in vain."

The Logical Operations. Word-for-word translation has proved surprisingly good, but it is not likely to be good enough for commercial translation. Some consideration of the context of words seems to be necessary to obtain a satisfactory output. Yngve suggests in an early article[6] the consideration of perhaps three or four words on each side of a particular word which is being processed. In a later article,[7] he suggests sentence-for-sentence translation. To the author's knowledge, no linguists are suggesting the consideration of words not located in the same sentence as the word being processed.

The logical operations revolve around three different problems. The first is the problem of determining whether the particular word being considered might be a compound of two smaller words, the second is that of determining a more accurate translation of a particular word or group of words by a consideration of the context, and the third is the problem of rearrangement of word order. These problems will be considered in detail later.

Decoding. In the decoding operation, which is not nearly so difficult as encoding, the machine code is translated into the target-language alphabet. At present there are several machines which would satisfactorily perform this operation. A high-speed teleprinter is one device that should be satisfactory for the first experimental models of the translator, and a recently announced electronic device[8] should be satisfactory for high-speed commercial translation.

The Problems

The problems may be classified, or lumped, into three classifications: reliability problems, memory and storage problems, and logical problems. Reliability is the most difficult problem in realizing an electronic encoder. Current encoders apparently have such a high probability of error that human transcribers are likely to be used for the first models of the translator. Although reliability is a very important factor, in this paper the reliability problem will not be considered. Only the memory and logical problems will be considered in detail.

The Low-Access Time Memory. Referring back to the descriptions of the way that material might be stored in the dictionary, we note that the entry between the I_1 and the next I_2 symbol is not necessarily just one source-language word, but is one source-language *semantic unit,* which often will be an idiomatic sequence of words. Thus there is a possibility of a storage of many words in this space. In order to compare an incoming source-language word sequence with such an entry it is necessary that sufficient low-access storage be provided to store at least this amount of incoming source language material at one time. The most common storage medium for such purposes is a "shift register."[9]

Shift Register

To the author's knowledge, no one has attempted to build a shift register which stores more than 200 bits at a time. The largest of these shift registers, which are designed for use in large-scale general purpose calculating machines, have capacities of about 70 bits. It is possible that the code for some idiomatic sequences might exceed 200 bits in length, although a cursory search by the author has not uncovered any such sequences in Russian, provided that the average code length per letter is not over six bits. A great many idiomatic sequences may be demonstrated which would exceed 70 bits when coded.

Schemes may be devised which will allow a register size somewhat smaller than the maximum size of word sequences which are to be handled as a unit, but all these schemes allow a certain amount of information to be lost. As an example of what might happen, consider the idiomatic sequence, "он мне подложил свинью." This sentence gives the literal translation "he slipped a pig under me," but in actual meaning it conveys the sense that "he played a trick on me." Actually the "он мне" part of the sentence is not idiomatic, since it merely states "he, to me . . .," but for the purpose of this example the entire sentence will be considered as an idiom in order to illustrate the principle. This idiomatic word sequence has 22 graphic symbols (including spaces), and, with the code of Table I, would require 115 bits in its coded form. One possible way to handle this sequence with a 70-bit register would be as follows. In such an entry, the register would be filled with as much of the text material as it would hold, or 70 bits. The entry would then be compared with the dictionary, and if, when comparing the contents of the register with a particular entry in the dictionary, the entire contents of the register correlated with the dictionary entry, but the I_2 signal had not yet been reached in the dictionary entry, then the contents of the register could be erased and the register filled with the next 70 bits of text material. The comparison of this new material with the dictionary would then proceed from the point where the last comparison had ended. If this second 70 bits of text correlated up to the I_2 signal in the dictionary, the complete correlation would be established, and the target-language equivalents would be read out of the dictionary. The difficulty with this scheme is that if the second 70-bit sequence did not correlate up to the I_2 signal, then the first 70-bit sequence, which had previously been erased, would not be available for any further comparison with the dictionary. For instance in the above example, the first 70 bits of the expression goes to the middle of "пложил." Thus if someone slipped anything other than a pig under a person, no translation could be made since correlation would be established up to the middle of "пложил," but then the dictionary entry of "свинью" (pig) would not correlate.

Another example which indicates that a reasonably large input storage should be provided is that of the translation of the various forms of negation in French. There are fifteen forms of negation in French, all starting with "ne" i.e., "ne . . . pas" (not), "ne . . . plus" (no more), etc. The words in the negation may be separated by several words in the sentence, for instance, "ne les ecoutez pas trop. . . ." It would seem likely that the best way to handle such sequences would be in the source language, i.e., when a "ne" is encountered in the incoming text, the following text would be searched for the other part of the negation and the English equivalent for the complete negation inserted in the target-language text in the "ne" location of the sentence. To do this it is necessary to store enough of the source-language text material so that the part of the sentence from "ne" to the "pas" in the above example is all available to the logical program of the machine. There are of course, exceptions in the use of the negation which will complicate the above routine somewhat.

Most of the logical processing will take place in the target language, or in the text material after the memory search. Thus a large low-access time buffer storage in the output is a necessity; at least one sentence at a time should be stored in the output in order to allow complete processing of the material.

The desirability of storing a substantial amount of input and output material, coupled with the difficulties encountered in attempting to construct shift registers of more than about 100 bits of capacity, encourages us to seek another method of storing the material as it is being processed. One medium that might be used is a magnetic core matrix. These devices are available commercially in units which store as much as 140,000 bits, and have access times of from six to eight microseconds. Some of the core matrices which are in the development stage have separate read and write circuits,[10] so it might be possible to construct a matrix which would allow constant write operations in order that the dictionary search could proceed continuously without delays while new source-language material was being introduced into the matrix. The core matrix is important enough so that it is worthwhile to explain the principle of operation of this device.

Core Matrix

The basic building block of a core matrix is a little doughnut-shaped piece of a ferromagnetic spinel material, commonly called "a ferrite." These little doughnuts (called cores) may be less than 1/16 in. in diameter. To understand the operation of these ferrite cores, consider the sketch in Fig. 1-A in which a wire carrying current into the paper threads one of the little cores, which is in the plane of the paper.

The magnetic flux density at any point in the field of the wire is proportional to the magnitude of current flowing in the wire, and, referring to Fig. 1-A, with the current flowing into the paper as shown, the flux will be clockwise about the wire. If the current were flowing out of the paper the flux would be in the opposite direction. Now with no current flowing in the wire, the ferrite doughnut will be a little permanent magnet with flux flowing in either the clockwise or counter-clockwise direction around the core. Suppose that the core is magnetized in such a way that the flux is flowing counterclockwise in the doughnut. The magnetization will be represented by point N in Fig. 1-B. If a current of magnitude I_1 is allowed to flow into the paper and this current is then removed, the state of the doughnut will move to point R, then fall back to point *N* in the hysteresis loop. If a current of 2 I_1 is allowed to flow, the magnetization of the doughnut goes to point Q on the hysteresis loop, so that now, when the current is removed, the state of magnetization of the core is represented by point P, or the magnetization in the core has "turned over" and flux is now circulating in the core in a clockwise direction.

Consider now Fig. 1-C, where a number of ferrite cores are threaded by three wires. Suppose that a current I_1 is passed through each of the wires, 2 and b, in the direction shown. In cores II and III the state will go to either point R or point S on the hysteresis loop of Fig. 1-B, depending upon whether

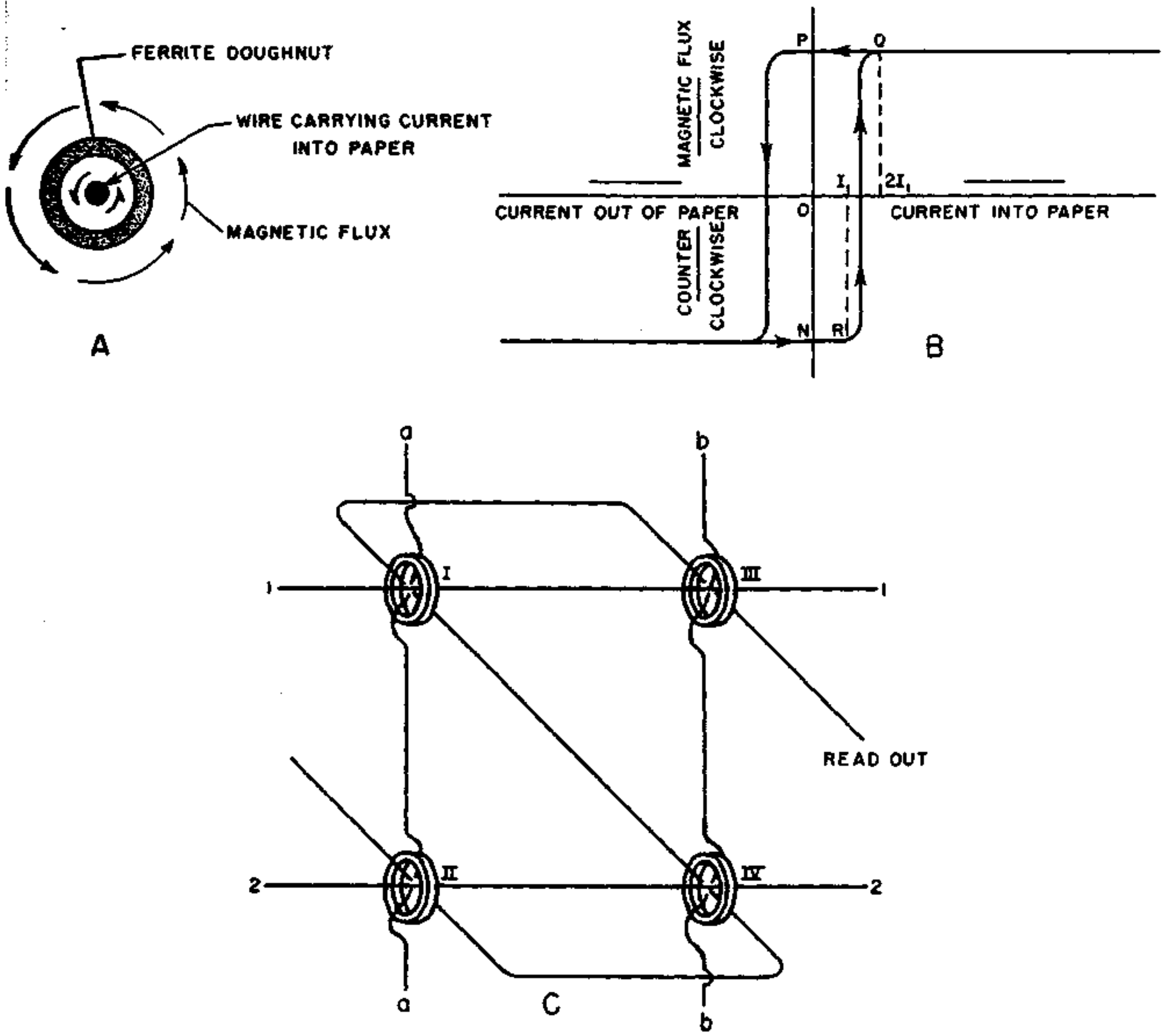

FIG. 1. OPERATION OF A CORE MATRIX

A. Ferrite doughnut in magnetic field of wire. B. Hysteresis loop of ferrite doughnut. C. Construction of core matrix.

the core had been at point P or X before the current flowed. Since current I_1 is insufficient to "turn over" the core, when the current is removed cores II and III will return to the same state that they had before current flowed. In the case of core IV, however, a current of 2 I_1 flows, so the state of the core will go to point Q regardless of the initial state, and when the current is removed, will return to point P. For the case of core IV, then, if the initial state was represented by point N, the core "turned over."

Now if a change in current through a wire can change the magnetic flux about the wire, then it is also true that a change in the flux about a wire will induce a current in the wire. Thus if core IV was initially at point P, essentially no change in flux took place in core IV as a result of the flow of the two currents I_1 in wires 2 and b. However, if the core had been at point N, a large change in flux (from $-\phi$ to ϕ) would have taken place. This change would induce a current pulse in the diagonally threaded wire which is labelled "read out." This pulse then tells the original state of the core IV. It is true that, in reading out the contents of core IV, the information stored on this core has been destroyed. This can be corrected by sending currents of I_1 through both wires 2 and b in the opposite directions from before, provided that a pulse appeared on the "read out" wire.

The rows and columns of the matrix are selected by reversible binary counters which operate diode or triode switches. Each column and each row has a

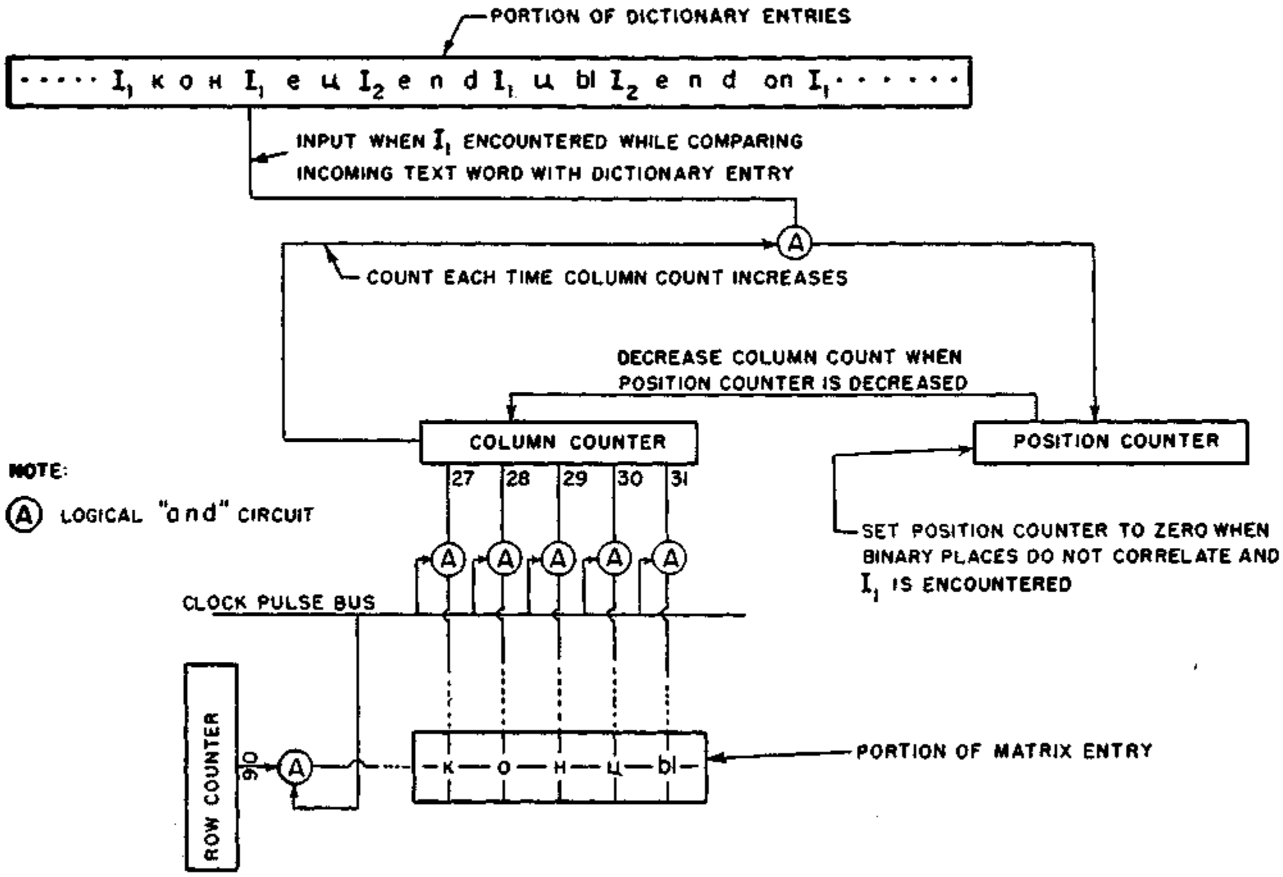
PORTION OF DICTIONARY ENTRIES
· · · · · I₁ к о н I₁ е ц I₂ e n d I₁ ц ы I₂ e n d on I₁ · · · · · ·
INPUT WHEN I₁ ENCOUNTERED WHILE COMPARING INCOMING TEXT WORD WITH DICTIONARY ENTRY
A
COUNT EACH TIME COLUMN COUNT INCREASES
DECREASE COLUMN COUNT WHEN POSITION COUNTER IS DECREASED
COLUMN COUNTER
POSITION COUNTER
NOTE:
A LOGICAL "and" CIRCUIT
27
28
29
30
31
SET POSITION COUNTER TO ZERO WHEN BINARY PLACES DO NOT CORRELATE AND I₁ IS ENCOUNTERED
CLOCK PULSE BUS
ROW COUNTER
90
к
о
н
ц
ы
PORTION OF MATRIX ENTRY

FIG. 2. IDEALIZED DICTIONARY AND MATRIX ENTRIES

unique number associated with it. One counter is provided for the rows of the matrix and another for the columns. When the number (address) associated with a particular row (or column) is entered in the counter, the counter "opens" the gate to that particular row (or column). All other rows (or columns) have their gates closed. When a pulse is introduced to the input of all the gates, only the row (or column) whose address is entered in the counter will actually be interrogated. Thus two counters which could store six bits each could control a matrix which stored $2^6 \times 2^6 = 4096$ bits.

The use of the matrix for an input storage is best illustrated by an example. Again the dictionary entries are assumed to be listed as suggested by Hill. In Fig. 2 a portion of the dictionary and the core matrix are shown in idealized form; i.e., the entries are shown in alphabetical form rather than in code in order to better illustrate the procedure. It is assumed that the text material in the matrix has been compared with the dictionary up to row 90 and column 28. It is also assumed that the dictionary has been searched for correlation with the incoming text entry, which starts with row 90 and column 28 in the matrix, as far as the dictionary entry which is shown in Fig. 2. When this point in the routine is reached, the machine, starting with the entry in the matrix in column 28 compares the dictionary entry with the first entry after I_1. These two entries correlate through the third entry, or past "кон." The next symbol in the dictionary, I_1 causes the column position counter to be activated, which means that each time the column counter increases its count by one, the position counter will increase its count by one. The machine will then resume the comparison, attempting to correlate the "ц" *in* the matrix with the "e" in the dictionary entry. No correlation is established at this point, so the machine does not attempt any more comparisons until the next I_1 symbol is reached in the dictionary. The machine then reduces the position counter to zero (in this case it is storing "1"), one count at a time. Each time the position counter is reduced by one count, the column counter is also reduced by one count. Thus when the position counter reaches zero count, the column counter is again at "30," and the machine again resumes the comparison. This time both "ц" and "ы" find correlation in the dictionary entry to the I_2 signal, so the target-language equivalent, "end on" is read out. In an actual machine, probably several auxiliary counters would be necessary, as may be readily seen If this procedure is attempted in the case of the negations of French. The principle would be the same, however.

Permanent Storage. A few years ago the great problem in MT was considered to be that of the large permanent memory for dictionary storage. Recent rapid progress made in computer components has brought to the horizon devices which promise to provide storage capacities with access times which will make this part of the problem much simpler than was dreamed possible. The most promising of the storage devices proposed at this time is one which would store information on a photographic plate and scan the plate optically.[3] As was pointed out in a previous paper, it should be possible to construct such a device which would store forty or fifty megabits of information with an access time of about one-tenth of a second. This rate would seem to qualify it for a commercially practical translator.

Logical Problems. The logical problems include multiple meaning problems, and problems in word order. The multiple meaning problem arises when a particular source-language word may have several possible meanings in the target language, and not

(Continued on page 27)

Engineering Progress in Machine Translation

(Continued from page 15)

necessarily closely related meanings at that. For instance the Russian word, "гитара," has the English equivalents "guitar" and "bracket." For reasonable understanding it is important to eliminate as many as possible of the multiple meanings which do not fit in with the context of the word. The multiple meaning problem for a particular word may be solved either by a consideration of the grammatical incident meaning of the word, or by considering the non-grammatical meaning of the word. As an example of the solution by considering the grammatical incident meaning of the word, Reifler[12] uses the German example "wegen dieser Schueler." The word "wegen" in isolation has the grammatical meanings of plural genitive or singular genitive. The word "dieser" has in isolation the grammatical meanings of singular masculine nominative, singular feminine genitive, singular feminine dative, or plural genitive. The word "Schueler" has in isolation the possible grammatical meanings of singular masculine nominative, plural genitive, singular masculine dative, singular masculine accusative, plural nominative, or plural accusative. Since only the plural genitive is common to all three, only the plural genitive meanings are selected, and the correct translation, "because of these pupils," is chosen.

The non-grammatical incident meanings of words present probably a more difficult problem, and a problem on which very little has been done. The example of the Russian "гитара" (guitar, bracket) listed above is one case where the non-grammatical incident meanings of the word must be determined from the word context. Another example of a problem in the non-grammatical incident meaning of a word is the case of the negation in French. The occurrence of "ne" has fifteen possible translations (not, only, etc.) depending on the occurrence of other words in the sentence.

The word-order problem has not been considered in great detail up to this time, mainly since it has been shown[13] that for scientific writing, in the great majority of cases foreign writers tend to use English word-order form even in those languages such as Russian which are actually word-order independent. There is no doubt that sometimes a rearrangement of word order would render the output more readable, so as the art of machine translation progresses, undoubtedly more consideration will be given to this problem.

The extensive use of compound words has led researchers to consider dissection schemes so that the storage of material in the dictionary can be reduced and so that extemporaneous compounds may be translated. The dissection scheme has been solved completely in the linguistic sense by Reifler.[13] The use of compound dissection schemes complicates the engineering problem since such a scheme must involve logical routines which are quite time consuming. This problem was considered in a previous paper.[3] Here we will note only that at least part of Reifler's dissection scheme will probably be included in even the first translation machine, but ultimately just what will be done about this problem cannot yet be determined.

Summary

The foremost question which an engineer must ask himself when evaluating a new idea is: "Will the results be useful enough to justify the cost?" or, in other words, "Will the device pay for itself?" To answer this question in the case of the machine translator, two other questions must first be answered: "Is there a large demand for translations of foreign language material?" and "Is the material available for translation?" The answer to the first of these questions is fairly well known. A tremendous amount of scientific work is being accomplished in foreign lands, particularly in Russia where scientific progress is going ahead at approximately the same rate as in the United States. A serious scientist will wish to avail himself of all published material in his field of interest, regardless of the source. The language barrier is something of a problem, since few scientists are able to handle more than two languages easily. Thus a machine translator should materially aid the dissemination of material. The question of availability of material may also be answered affirmatively. For instance, most of the Russian scientific journals are available for subscription to United States scientists. Russian books are also available in large numbers and at very reasonable prices. Thus there is an abundant supply of raw material for a translation machine, and a ready market for the machine output. The only question left to be answered is one of economics: "Can the material be processed economically?" At the present time there appears to be no reason why a machine translator cannot compete on a cost-per-word basis with well-trained polyglots.

In conclusion we note that interest in machine translation is by no means restricted to people in English speaking countries. For instance in the Russian newspaper, *Pravda,* on January *22,* 1956, an article appeared describing the recent translation of scientific material from English to Russian by means of a digital computer. A translation of this article appears below.

Fortunately for us, the problem of translation from Russian to English is much simpler than from English to Russian.

REFERENCES

1. STOUT, THOMAS M.. "Computing Machines for Language Translation." *The Trend in Engineering at the University of Washington,* Vol. 6, No. 3, (July, 1954).
2. DOUTHWAITE. GEOFFEY. "The UW Automatic Language Translator," *Washington Engineer,* Feb., 1956.
3. WALL, ROBERT E. Jr, "Some of the Engineering Aspects of the Machine Translation of Language," AIEE paper 56-693.
4. SHEPARD. DAVID H, and HEASLEY, CLYDE C., JR., "Photoelectric Reader Feeds Business Machines," *Electronics,* Vol. 28. No. 5 (May, 1955), p. 134.
5. HILL, W. RYLAND, "A Coding and Operational Program for Machine Translation," unpublished notes. University of Washington.
6. YNGVE, VICTOR H., "Syntax and the Problem of Multiple Meaning," *Machine Translation of Language,* ed. by William N. Locke and A. Donald Booth, John Wiley & Sons, New York, 1955.
7. -----------"Sentence for Sentence Translation," *Mech. Trans.,* Vol. 2, No. 2 (Nov., 1955), p. 29.
8. BLISS, W. H., and RUEDY, J. E., "An Electron Tube for High Speed Teleprinting," *RCA Review,* Vol. XVI, No. 1 (March, 1955), p. 5.
9. RICHARDS, R. K., *Arithmetic Operations in Digital Computers,* Van Nostrand & Co., New York, 1955, p. 144.
10. RAJEHMAN, J. A., and Lo, A. W., "The Transfluxor," *Proceed. IRE,* Vol. 44, No. 3 (March, 1956), p. 321.
11. REIFLER, ERWIN, "The Mechanical Determination of Meaning," *Machine Translation of Language,* ed. by William N. Locke and A. Donald Booth, John Wiley & Sons, New York, 1955, p. 159.
12. HARPER, KENNETH E., "The Mechanical Translation of Russian: A Preliminary Report," Progress Report, University of California at Los Angeles.

NOTE

The author, who is completing his doctorate in Electrical Engineering, chose Russian and French as the two foreign language requirements for his degree. He believes that the increasing quantity and importance of technical literature published exclusively in the Russian language should be recognized by engineers planning a research career. *Ed.*

A Machine Translates from One Language to Another

(From *Pravda,* January 22, 1956)

Translated from the Russian by Dr. Lew R. Micklesen, Assistant Professor of Far Eastern and Slavic Languages and Literature, University of Washington.

The first successful experiments in automatic translation from one language to another have been carried out in the Academy of Sciences of the USSR.* They were performed on the rapid-action electronic computing machine "BESM" described in *Pravda* Dec. 4 1,955.

At first glance it seems unlikely that a machine can automatically make a translation from one language to another. But if one ponders the question a little, then one can easily understand that there is nothing impossible in it. In reality, of course, a language represents a definite system in which the meanings of words and any of their modifications are reflected in lexical and grammatical devices.

The possibility should exist, therefore, of elaborating such dictionaries and such rules of translation which would consider all features of sentence construction and would allow the accurate and unambiguous determination of the meaning of the component words and their mutual relationship in the text. From this it also appears possible to make a translation by completely automatic means with the help of machines with programmed directions—for example, machines similar to those which automatically perform complex mathematical computations.

In order to realize such a translation, it is necessary first of all to encode the phrase to be translated into a special numerical code so that a certain binary number corresponds to each letter. By indicating, for example, as is done in the Baudot code, the letter "a" of the Latin alphabet by 16, the letter "n" by 15, the letter "d" by 30, we can substitute for the English word "and" (Russian "и"), the number 161530. The numbers, corresponding to words in this coding system, can be punched into a paper tape. A person who does not need to know English may do this on a simple apparatus like a typewriter having a keyboard with Latin letters. This tape is introduced into the electronic calculating machine, whose device contains a dictionary installed there beforehand. Each word

* The reader's attention is drawn to the fact that "January 7, 1954 marked the successful culmination of a joint project of the Institute of Language and Linguistics of the Georgetown University School of Foreign Service and IBM: one language (Russian) was successfully translated into another (English) by means of a high-speed electronic digital computer."—Tech. Newsletter 9, Applied Science Division, IBM, Jan., 1955. See also "The Georgetown-IBM Experiment," by L. E. Dostert in *Machine Translation of Languages,* Tech. Press, MIT, and John Wiley & Sons, Inc., New York, 1955. *Ed.*

of the dictionary, consisting of English and Russian parts, has also been replaced by a corresponding number; and the process of searching for a word in the dictionary amounts to a comparison of the number introduced into the machine and expressing a given word with all the numbers of the dictionary. Say that number corresponding to the word of the text that we want to determine is subtracted successively from all the numbers representing the words of the dictionary. When the remainder after subtraction is equal to zero, our search is over: the number of the dictionary which corresponds exactly to the number subtracted has been located and therefore the word corresponding to it has been found in the dictionary. This means that in the Russian part of the dictionary the word corresponding to the English word has been found. All this is done by the machine completely automatically and with great speed. For example, in the "BESM" machine, one matching operation takes about one ten thousandth part of a second, so that the machine can read a dictionary of 1000 words in a period of time measured in parts of a second.

The dictionary in the machine, however, differs from those dictionaries we customarily use. In it the rules which allow the automatic selection of the correct meaning of a word from the many possible ones must be foreseen. In order to determine the meaning of an ambiguous word, it is necessary to analyze the words surrounding it, to see which words stand before it and after it, what are their meanings and grammatical features.

Say we want to translate the English word "example," which has two meanings depending on whether the word "for" stands before it or not. In this case one must ascertain whether the preceding word coincides with the word "for"' or not. If it does, the translation will be "for example," if not, "example." Similar rules, but much more complex ones, can be elaborated for still other cases. They are formulated in the form of just such concrete questions about the other words of the sentence, but the machine must answer "yes" or "no" sometimes more than 20 times. Another system of rules is necessary in order to locate in the dictionary words which in English acquire endings. For example, we will not fnd the word "equations" (Russian упрaвнения) in the dictionary, since it has the ending -s denoting the plural. The singular is given in the dictionary. The machine removes the ending and then the word without its ending is again checked in the dictionary.

Finally the dictionary must contain the grammatical features of the words in it. However, in contradistinction to the usual English-Russian dictionaries, it contains the grammatical features not of English, but of the corresponding Russian words. This is necessary because without these features it is impossible to construct a good Russian sentence which would be the translation of the English. The English grammatical features are necessary only in so far as they aid in determining the Russian ones.

It is not necessary to record all the grammatical features in the dictionary. It is impossible, for example, to tell ahead of time that a given word will be a subject or a direct object. This situation can be clarified only by the juxtaposition of the given word with other words, by the analysis of the whole sentence. Therefore, after the dictionary work, the machine carries out a vast number of various checks of the same type with "yes" or "no" answers aiming at the determination of all the necessary grammatical features of the Russian words. When this is done, the English phrase being translated, for example, "problems associated with motion" acquires the following form: "задача" (substantive, feminine gender, second declension, plural, nominative case, soft stem ending in a hushing sound)—"связывать" (participle, past tense, feminine gender, plural, nominative case)—"с" (preposition, demands instrumental case)—"движение" (substantive, neuter gender, first declension, singular number, instrumental case, soft stem). Now, according to the rules of Russian grammar recorded in the form of such schemes, it is possible to obtain the required translation :

"задача, связанные движеннем"

The translation is then printed out on a teletype.

The following sentence taken from an English text may be quoted as an example of a translation performed by a machine.

"Elementary courses in differential equations present a long list of clever devices by means of which one is supposed to be able to solve differential equations."—

"элементарные курсы по дифференциальнм уравнениям дают длинный перечень искусных приемов, при помощи которых исследователь, как предполагается, может решать дифференциальные управнения."

Of course, it is difficult to postulate that in the near future the automatic translation of literary works will be successfully realized. This problem is much more complex than the translation of a scientific text.

Experimental work on automatic translation was carried out by a group of co-workers of the Academy of Sciences of the USSR on the universal calculating machine "BESM" not especially adapted for this. Besides the authors of this article, S. Razumovsky, L. Korolev, N. Trifonov, G. Zelenkevich and others were included in this group. The results obtained allow us to consider that on the principles worked out a specialized electronic machine for automatic translation can be built in the future.

D. Panov
I. Mukhin
I. Bel'skaya