[From: *Computers and Automation* 3 (2), February 1954]

LANGUAGE TRANSLATION BY MACHINE —

A REPORT OF THE FIRST SUCCESSFUL TRIAL

Neil Macdonald

On January 7, 1954, at a press conference in the office of International Business Machines Corporation in New York, the IBM 701 electronic data processing machine located there presented the first successful demonstration of meaningful translation from one language to another language by machine. Although this demonstration was simply a trial of machine translation, involving a vocabulary of only a few hundred Russian and English words, its success is full of significance and exciting promise for the future.

The success was due to a year and a half of cooperative effort by the Institute of Languages and Linguistics of Georgetown University, Washington, D.C., and the IBM Corporation . On the part of the Institute, those most concerned with the achievement were Dr. Leon Dostert, Director of the Institute, planner, organizer, and sparkplug of the project, and Dr. Paul Garvin, linguist in twelve languages, and main architect of the linguistic translation scheme. On the part of IBM, those most concerned were Dr. Cuthbert C. Hurd, Director of the Division of Applied Science, Mr. Peter Sheridan, mathematician and composer of the IBM 701 program which accomplished the translation, and Mr. Thomas J. Watson, Chairman of the Board of IBM, who authorized and encouraged the research project.

The Nature of the Trial

A total vocabulary consisting of 250 Russian words (in latinized spelling) relating to the fields of politics, law, mathematics, chemistry, metallurgy, communications, and military affairs was punched on punch cards. Associated with each Russian word and punched on the same card were one or two English equivalent words, and three codes designated as 1st, 2nd, and 3rd. These codes (linguistically they can be considered "diacritical marks") together with the program caused appropriate translation. For example, different meanings of words could be selected. The order of words in Russian could be left unchanged or could be altered in a specified way, as might be indicated. A word could be treated as a whole or could be divided into a root and a suffix. And so forth.

An extract from the dictionary is shown on page 10.

To set up the computer for the language translation trial, the punch cards were run into the machine and their information stored on the magnetic drums, taking up the space of 6000 machine words of 36 binary digits each.

Next, the program developed for purposes of translation was run into the machine. This program consisted of about 2400 program steps or instructions. The general scheme of the program is shown in Figure 1, Dictionary Syntax Flow Chart.

Finally, a number of Russian sentences staying within the vocabulary and the linguistic constructions planned for, were given to the machine. With about 5 to 8 seconds of machine computation for each one, the output printer of the IBM 701 proceeded to write out translations of the sentences. Examples follow:

KACHYESTVO UGLYA OPRYEDYELYAYETSYA KALORYIYNOSTJYU
The quality of coal is determined by calory content.

KRAXMAL VIRABATIVAYETSYA MYEXANYICHYESKYIM PUTYEM YIZ KARTOFYELYA
Starch is produced by mechanical methods from potatoes.

VYELYICHYINA UGLA OPRYEDYELYAYETSYA OTNOSHYENYIYEM DLYINI DUGI K RADYIUSU
Magnitude of angle is determined by the relation of length of arc to radius.

OBRABOTKA POVISHAYET KACHYESTVO NYEFTYI
Processing improves the quality of crude oil.

MI PYERYEDAYEM MISLYI POSRYEDSTVOM RYECHYI
We transmit thoughts by means of speech.

ZHIYELYEZO DOBIVAYETSYA YIZ RUDI XYIMYICHYESKYIM PROTSYESSOM
Iron is obtained from ore by chemical process.

VOYENNIY SUD PRYIGOVORYIL SYERZHANTA K LYISHYENYIYU GRAZHDANSKYIX PRAV
A military court sentenced a sergeant to deprival of civil rights.

VLADYIMYIR YAVLYAYETSYA NA RABOTU POZDNO UTROM
Vladimir appears for work late in the morning.

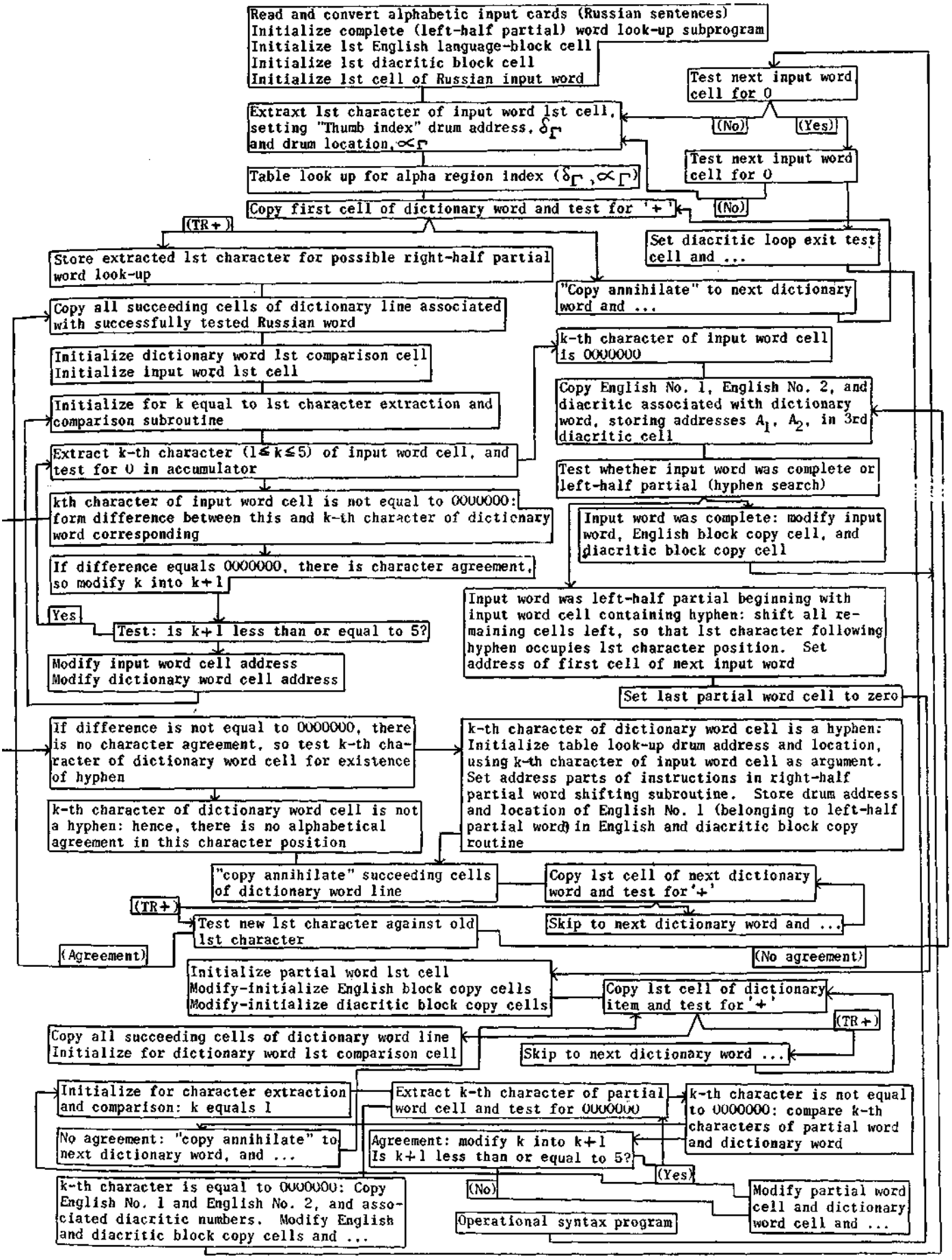

Figure 1 -- Dictionary Syntax Flow Chart

MYEZHDUNARODNOYE PONYIMANYIYE YAVLYAYETSYA VAZHNIM FAKTOROM V RYESHYENYIYI POLYITYICHYESKYIX VOPROSOV

International understanding constitutes an important factor in decision of political questions.

KOMANDYIR POLUCHAYET SVYEDYENYIYA PO TYELYEGRAFU

A commander gets information over a telegraph .

DOROGI STROYATSYA YIZ BYETONA

Roads are constructed from concrete.

DYINAMYIT PRYIGOTOVLYAYETSYA XYIMYICHYESKYIM PROTSYESSOM YIZ NYITROGLYITSYERYINA S PRYIMYESJYU YINYERTNIX SOYEDYINYENYIY

Dynamite is prepared by chemical process from nitroglycerine with admixture of inert compounds.

How Do the "Codes" Work?

Now it can be seen from the discussion so far that an important part of the success of this first trial run of. machine translation is the way in which the codes work. They cause words and word order to be selected and arranged according to six "rules of operational syntax". These six rules are stated in Figure 2, and under them is shown the way in which they work for a sample Russian sentence, which when translated is "magnitude of angle is determined by the relation of length of arc to radius".

According to Dr. Dostert, the rules of operational syntax that would be required for the translation of any Russian sentence into English might number a hundred; but for great numbers of sentences, and in particular for all the sentences of the type included in the experiment, the six rules were sufficient.

The effect of the codes and the rules is to fix the alternative meanings of the word, and enable the machine to determine the right meaning out of several possible meanings. The codes take into account the sentence structure, the word structure, the nature of the prefixes and suffixes, the context, etc.; they recognize a series of structural patterns or syntactic structures.

In fact, before the translation scheme was given to the IBM 701 programmers to convert into a series of instructions to the machine, Dr. Dostert and Dr. Garvin got hold of people who did not know Russian, gave them Russian sentences written in Roman characters, and a set of cards. The non-Russian speaker would look at a word and look it up in his cards. In his cards he found the word and instruction numbers. Then he would refer to the cards bearing the instructions. At the end of about five minutes, the non-Russian would come out with the correct translation of the Russian sentence in English. Then the Georgetown men knew they were on the right track, because they had succeeded in reducing the whole process to the capacity to read instructions and carry them out, which of course the machine could perform.

How Did the Translation Project Come About?

For about ten years a number of scholars in various institutions have been thinking as individuals about the possibility of formulating an adequate set of instructions so that an electronic computer would be able to transfer meaning from one language into another language. Much of the research was largely conjectural. Dr. Erwin Reifler of the University of Washington, Dr. Y. Bar Hillel then at Mass. Inst. of Technology, and others, formulated various theories and advanced various plans.

A conference on machine translation took place at MIT in June, 1952. This conference was held with the support of the Rockefeller Foundation and with the very active interest and support of Dr. Warren Weaver of the Foundation.

Dr. Dostert attended the conference; he went there rather skeptical about the whole idea, but came away convinced that the only way to put an end to many hypothetical disputes would be to try to make a simple, yet not too extensive, test of the feasibility of mechanical translation. After discussing it with some of his associates in Georgetown, he took the subject to IBM, and there met a sympathetic and helpful reception. Thus the project was launched, in terms of trial with a glossary of 250 Russian words.

What Will the Project Lead to?

Many exciting possible developments are indicated by the success of the trial, according to Dr. Hurd and Dr. Dostert.

Linguists will be able to study a language in the way that a physicist studies material in physics, with very few human prejudices and preconceptions, because the language has to be reduced to its operational characteristics in order to be handled electronically.

The technical literature of Germany, Russia, France, and the English-speaking countries will be made available to scientists of other countries as it emerges from the presses.

Technical know-how will be rapidly avail-

Rules of Operational Syntax

RULE 1: REARRANGEMENT

If first code is '110', is third code associated with preceding complete word equal to '21'? If so, reverse order of appearance of words in output (i.e., word carrying '21' should follow that carrying '110')-otherwise, retain order.

In both cases English equivalent I associated with '110' is adopted.

RULE 2: CHOICE-FOLLOWING TEXT

If first code is '121', is second code of the following complete, subdivided or partial (root or ending) word equal to '221' or '222'? If it is '221', adopt English equivalent I of word carrying '121'; if it is '222', adopt English equivalent II.

In both cases, retain order of appearance of output words.

RULE 3: CHOICE-REARRANGEMENT

If first code is '131', is third code of preceding complete word or either portion (root or ending) of preceding subdivided word equal to '23'? If so, adopt English equivalent II of word carrying '131', and retain order of appearance of words in output -if not, adopt English equivalent I and reverse order of appearance of words in output.

RULE 4: CHOICE-PREVIOUS TEXT

If first rode is '141', is second code of preceding complete word or either portion (root or ending) of preceding subdivided word equal to '241' or '242'? If it is '241', adopt English equivalent I of word carrying '141'-if it is '242' adopt English equivalent II.

In both cases, retain order of appearance of words in output.

RULE 5: CHOICE-OMISSION

If first code is '151', is third code of following complete word, or either portion (root or ending) of following subdivided word equal to '25'? If so, adopt English equivalent II of word carrying '151' -if not, adopt English equivalent I.

In both cases, retain order of appearance of words in output.

RULE 6: SUBDIVISION

If first code associated with a Russian dictionary word is '*', then adopt English equivalent I of alternative English language equivalents, retaining order of appearance of output with respect to previous word.**

SOURCE SENTENCE: **vyelyichyina ugla opryedyelyayetsya otnoshyenyiyem dlyini dugi k radyiusu.**

ANALYSIS:

RUSSIAN WORD	ENGLISH EQUIVALENTS		1st	2nd	3rd	RULE
	I	II	CODE	CODE	CODE	NO
vyelyichyina	magnitude	---	***	***	**	6
ugl-	coal	angle	121	***	25	2
-a	of	---	131	222	25	3
opryedyelyayestsya	is determined	---	***	***	**	6
otnoshyenyi-	relation	the relation	151	***	**	5
-yem	by	---	131	***	**	3
dlyin-	length	---	***	***	**	6
-i	of	---	131	***	25	3
dug-	arc	---	***	***	**	6
-i	of	---	131	***	25	3
k	to	for	121	***	23	2
radyius-	radius	---	***	221	**	6
-u	to	---	131	***	**	3

TARGET SENTENCE: **magnitude of angle is determined by the relation of length of arc to radius.**

Figure 2 — Rules of Operational Syntax, and a Sample Sentence Translated

able to the under-developed areas of the world, such as Pakistan, Indonesia, Yugoslavia, the Arab world, in their own languages. Divisions of the U. S. Government may well be interested in picking up and carrying forward this development.

A problem in an entirely new field of the social sciences has been solved.

Information from this experiment will be of considerable use in the design of information-handling machinery particularly adaptable to language translation.

But of course it must be emphasized that a vast amount of work is still needed, to render mechanically translatable more languages and wider areas of a language. For 250 words and 6 syntactical structures are simply a "Kitty Hawk" flight.

EXTRACT FROM DICTIONARY

	English Equivalents:		1st	2nd	3rd
Russian Word	I	II	Code	Code	Code
k	to	for	121	***	23
kyislorodn-	oxygen	***	***	***	**
lyishyenyi-	deprival	***	***	222	**
matyeryial-	material	***	***	***	**
mi	we	***	***	***	23
mislyi	thoughts	***	***	***	**
mnog-	many	***	***	***	**
myedj	copper	***	***	***	21
myest-	place	site	151	***	23
myexanyichyesk-	mechanical	***	***	242	**
myezhdunarodn-	international	***	***	***	**
na	on	for	121	***	23
napadyenyi-	attack	attacks	121	***	**
nauka	a science	***	***	242	**
obrabotka	processing	***	***	***	**
obwyekt-	objective	objectives	121	***	**
ofyitsyer-	an officer	the officer	***	***	**
-ogo	of	***	131	***	23
-on	by	***	131	***	**
opryedyelyayet	determines	***	***	***	**
opryedyelyayetsya	is determined	***	***	***	**
optyichyesk-	optical	***	***	***	**
orudyiye	gun	***	***	241	**
otdyel-	section	***	***	***	**
otdyelyenyiye	division	squad	121	242	**
otnoshyenyi-	relation	the relation	151	***	**