Mechanical Translation: Needs and Possibilities

The invention of electronic computers has revolutionized not only the performance of certain mathematical operations but also our whole attitude toward "routine" operations. The solution of, say, a system of certain differential equations was regarded until recently as a performance of which only a highly gifted and thoroughly trained human brain is capable. There are machines which solve such equations in a matter of split seconds. Since this is so, could not also other operations considered so far as the privilege of intelligent minds, be carried out by computer-like machines?

One such task challenged the thinkers' imagination as early as 1945, i.e., immediately after the recently invented computers and their operators were freed from the necessity of working for military aims exclusively. In that year, Dr. Warren Weaver, now Director of the Science Division of the Rockefeller Foundation, started thinking and talking about the possibility of using computer-like machinery for carrying out part or all of the operations involved in translating from one language to another. I shall not tell you here the interesting history of the further development of this problem. Let me say only that it was taken up by at least five independent groups, and according to my information, by not more than that. There is the English group, consisting of Dr. Booth, a computer engineer, and Dr. Richens, a geneticist who is in charge of an abstraction and translation service in his field and got thereby interested in our problem, two Los Angeles groups, one comprised of Dr. Huskey, a computer engineer, Prof. Oswald, and Prof. Bull, both linguists who became fascinated by the linguistic intricacies of the task, another group within RAND Corporation, who had to deal with this problem from an angle different from that adopted by all other groups: whereas the others were primarily interested in high-accuracy translation of mainly scientific material, the RAND group investigated the possibilities of high-speed mass translation of propaganda output of potential or actual enemies and related material. The fourth group consists of just one man, Prof. Reifler from the University of Washington, a linguist who in search of the common semantic foundations of all languages was induced to look for mechanical ways of transforming one language into another. The fifth consists of myself, who, starting from an extremely intriguing problem belonging to the Philosophy of Language, namely to what degree linguistic studies can be carried out without taking meanings into account, became involved in the problem of translation by machines which are unable, at least so far, to understand the meanings of the expressions upon which they operate. Working on our problem only by myself, I was able to benefit from discussions with people at MIT who in addition to being highly interested in this topic, were also actively engaged in research in closely related fields.

The present Conference, of which this meeting is the first session, the only one open to the general public, is the first opportunity for the members of these groups to come together and to exchange experiences, successes and failures, between themselves and with some of the creative thinkers in closely related fields, with the aim of coordinating their future research to a much higher degree.

The aim of this session, especially, is to give to the interested scientific public an account of some of the less technical aspects of mechanical translation. Due to the fact that this open session precedes our conference instead of winding it up, it may well turn out that the views presented by the various participants in the opening symposium or by the panel members will be much divergent. I hope that at the end of our meeting these divergencies will be reduced to some degree but I am sure that they will not be completely eliminated. More specifically, I would say that the aim of this session is threefold: to present some arguments why mechanical translation should be attempted, second to show that this task is not superhuman, or should I say

supermechanical, third to prove that it is by no means an easy and trivial task, as many enthusiasts of the forthcoming electronic age are led sometimes to believe and leading others to believe.

The other participants of this symposium will elaborate upon one or the other of these aims, either explicitly or by implication. Let me therefore make only the following introductory remarks. The translation load of just one agency, the Translation Service of the United Nations, is about 175,000 pages per year and is in constant increase. Though the use of some international auxiliary language might considerably reduce this load, such a solution is apparently still in the far future. A mechanical translator that would do most or even only some of this work should be most welcome. The output in Chinese, Russian, and Korean, in print and telecast, connected with the war in Korea, is of the order of a million words daily. A translation machine that could present an even rough-and-ready English version of this stuff would do an important job.

But it is a difficult, even a very difficult job. To get a good look on it, imagine yourself on the task of translating from one language you do not understand into another you do not understand either, and this without the benefit of dictionaries from the source language into English and from English into the target language. This looks quite formidable but would still be manageable to a certain degree for an intelligent translator, provided he were given - what? What would you require to be given before undertaking such a task? Remember the restriction put upon you: you will be forever denied the understanding of the text you are going to translate. I presume that you would ask for at least the following prerequisites: first, either a complete set of instructions, in English, how to split each word in the source language into stem and affixes or a list of all words of the source language together with their splitting up into these two parts, second a bilingual stem-dictionary giving for each stem in the source language one or more correspondents in the target language, third a complete set of instructions, in English telling how the syntactic functions corresponding to the affixes in the source language operate upon the stems in the target language as well as specifying the order into which the transformed stems have to be reassembled in the target language. This is only a very rough description of your minimal needs, but it gives some idea of the great difficulties in providing for these prerequisites - and as a matter of fact, I do not think that anything like the first and last prerequisites have ever been prepared for any natural language whatsoever. You will also have noticed that you will usually end up not with one equivalent to a sentence of the source-language but with many, since nothing will tell you which of the stems corresponding to a stem in the source-language has to be chosen. Unfortunately no single method is in view that might enable the intelligent translator to come out with a single answer. It seems that he will have to tell the man who ordered him to do the translation what the newly hired secretary told her employer who asked her to perform a certain addition: "Well, sir," she said after a few minutes, "I checked the sum six tines, and here you have the six results."

Let me stress immediately, for the linguists among you, that I would not say that a unique outcome is impossible, in principle. Given a dictionary of great elaborateness, amounting in effect, to a phrase-dictionary instead of the customary stem-dictionary, the translation would be unique, of course only if the original sentence was unambiguous. But since a dictionary of three-word-phrases will not always be sufficiently long to enforce uniqueness -, unique translation under the mentioned conditions is practically out of the question. (Let me remark in parentheses that if the field to which the source-text belongs is sufficiently narrowed down, almost-uniqueness might perhaps be achieved under practically realizable conditions. But I do not think that there is anything like unanimity in this respect among the experts.)   It seems therefore that in general even a human translator would be unable to arrive at a unique translation under the mentioned conditions. Since a machine translator will differ from a human translator mainly in that its understanding of the language, in which the instructions are given and the syntactical structure of the source and target languages explained, is highly restricted, amounting to something like being able to match, copy, shift, and select its next operation according to whether something matches something else or not, and to understand orders in which these

operations are required to be carried out in a certain *s*equence, it is hardly conceivable that the machine will be able to do better than its human competitor, in spite of the fact that it is able to perform those questions at much higher speeds. We arrive therefore at the result that completely automatic and autonomous mechanical translation with unique correlates to the original text is, in general, practically excluded, even *w*ith respect to scientific texts. (Incidentally, let it be perfectly clear that none of the participants in this conference intends to translate poetry by machines.)

This being so, machine translation means no more than <u>mechanical</u> <u>aids</u> <u>to</u> <u>translation.</u> Only some kind of <u>brain-machine</u> <u>partnership</u> is envisaged. As soon as this is recognized, a whole gamut of possible solutions to our problem is now in view. Returning to the situation envisaged before, it is quite possible that in spite of the fact that the machine output will be multiple, perhaps a few million of correlates to an average sentence of the source language, the reader of this output who understands the target language may well be able to pick up a unique and adequate correlate, if he is sufficiently familiar with the topic. Since, however, not every prospective reader will have the required training, it is, of course, preferable to have a special post editor select the unique correlate together with some stylistic smoothening. Notice that the post-editor must understand the target-language and be an expert in the respective topic to be able to fulfill his task adequately, but - and here is the decisive point - he need not understand the source-language! That under suitable conditions an intelligent post-editor will be able to pick out quickly a unique correlate can be shown experimentally. A machine-post-editor partnership seems feasible, on principle. There are however many obstacles to overcome, both on the linguistic and the hardware side, which I have no time to specify.

Other partnerships that can be shown to work, on principle, are of the <u>pre-editor</u>-machine type, pre-editor-machine-post-editor type, and finally of a type in which a bilingual human translator participates, but such that the time-span during which he is engaged in translations is only a small fraction of the time he might have had to spend in a complete autonomous translation. All these partnerships are of high practical importance, since in each of them the major bottleneck that exists today, namely that of a scarcity of bilingual experts, is effectively overcome. It is clear that, in general, the less the machine is required to do, the simpler its construction will be. But here we reach another problem, which is not so much scientific as economical. Is it worthwhile to develop machine translators? Will it not be much more efficient and cheaper to train additional bilinguals?

I am by no means in a position to give a categorical answer to this question. Let me therefore answer in a conditional way. Even if it should turn out that none of the possible machine-brain partnerships would be more effective than a human translator, in the sense that they will be neither quicker nor cheaper nor more exact than the human translator, under conditions existing today or in the near future, I would strongly advocate a continuation of this research. Electronic machines will doubtless become cheaper, human brains probably more expensive. A partnership that could not stand free competition today may well outbid its human competitors in some not too remote future. Let me finish with the following remark. Mechanical translation, if and when put into action, will have many sociological feedbacks. One might be that certain people will write their reports and books in such a way as to facilitate mechanical translation. Another might be that training in foreign languages, insofar as its sole aim was to enable the student to understand foreign texts will become superfluous. I do not know how to evaluate in dollars and cents the saving in millions of student hours per week which would result. This factor should certainly be brought into consideration when the final account of the financial worthwhileness of mechanical translation is made. I hope that it is perfectly clear that not teaching of foreign languages in general might be abolished but only such teaching as has understanding of foreign scientific texts as its <u>sole</u> aim. But here I am far beyond the limits of scientific sobriety and deep in speculation - so let me better come to an end.