

THE MACHINE AND THE MAN*

Victor H. Yngve

Research Laboratory of Electronics
Massachusetts Institute of Technology
Cambridge, Massachusetts

WHEN extensive mechanical translation becomes a reality, many new jobs will be created. Some of these jobs will be closely related to existing occupations. In this category are those occupations connected with the construction of the machines — electronic design and construction, machine shop work, and the like. Then there will be others involved in the daily running of the machine — typists, operators, office workers, and administrative personnel. In addition to these rather obvious occupations, there are some that may be less obvious. In the following article we shall discuss several of the less obvious roles that humans may play in relation to a translating machine.

Man in the role of creator of the machine, the designer of the system by which it translates, was one of the earliest concepts to be found in MT literature. This idea is implicit in practically all of the work that has been done on mechanical translation. The machines that have been considered are slave machines, built by man and tirelessly carrying out to the letter the instructions originally given them. The burden that this throws upon man is the task of designing the machine and instructing it in detail in the routine it is to use to translate everything fed into it.

Perhaps Y. Bar-Hillel has given the most detailed statement of the tacit assumption that underlies the thought of many others when he writes of the necessity for man to provide "an operational syntax" for the machine. By this he means a program that the machine can carry out in sequence, at each point being given the exact criteria for determining what to do next. This program is to be capable of translating all possible sentences from the input language to the output language. Furthermore Dr. Bar-Hillel has outlined the things that he considers necessary for man to do before the machine can get to work. He envisions the compilation of a complete word index giving the stem-ending analysis; a complete dictionary giving for each word the various meanings and all the other information that will be needed for the grammatical analysis; and an operational syntax "giving a complete sequential program for the analysis of every sentence." The construction of this program constitutes a great

challenge to the linguist, since it requires him to consider language as it actually is and to specify exactly and completely all the operations necessary for translation. As has been pointed out, the machine will be in the position of a person trying to translate from language A to language B, using a set of rules expressed in a third language and never knowing the meaning of what is being translated. The challenge to the linguist and to man as the creator and designer of the machine is to provide this set of rules.

Another widely held assumption is that a machine may never be able to produce a perfect translation. For this reason, a good deal of thought has gone into the possibility of man-machine combinations. One of the great difficulties that man as the creator of the machine will have to face is the fact that the input language does not have sufficient semantic explicitness in many cases to provide a machine with enough information to solve the many problems in grammar, syntax, and multiple meanings. Prof. Erwin Reifler pursued this problem and suggested a number of ways in which a human pre-editor could make the input text more explicit. The job description of the pre-editor is to be found in Reifler's first paper, abstracted in the last issue. "Whatever the native reader has to do by way of interpretation in the case of non-distinctive features of the FL (foreign or input language) text, can at least at the present stage of computer development, not be mechanized. Therefore, all that an FL text leaves to the FL reader to determine concerning lexical meaning, connotations, grammatical meaning, and word order, has to be added to the FL text before it is fed into the computer. And it has to be added in a form that the computer can 'digest'."

Perhaps his most far-reaching suggestion, as far as its possible impact on man, was his universal MT orthography. He proposed that the pre-editor capitalize the first letter of nouns, as in German, the second letter of verbs, the third letter of attributive adjectives, and so on. Reifler further proposed that this orthography could become universal and be applied to all languages that are written in scripts that allow capitalization. Thus the machine would have at the input a specification of the grammatical categories of the words to assist it in making a proper translation. This orthography would be taught in the schools. Here we have MT changing our conventional script, and thus affecting nearly everyone by

* This work was supported in part by the Signal Corps; the Office of Scientific Research, Air Research and Development Command; and the Office of Naval Research

requiring a change in the conventional method of writing. This concept of changing the input language to fit the needs of the machine is carried to the extreme by Stuart Dodd, who proposed that English and other languages be regularized along the lines of his proposed "Model English." Writers of material to be translated would be required to write according to the rules of Model English. The output of the machine could also be in a "modelized" language.

It seems to be a fair statement, however, that the idea of the pre-editor, and all other tampering with the input text or language, is nearly dead. Most workers now seem to consider that probably all of the tasks formerly assigned to the pre-editor can be mechanized. Perhaps the greatest stimulus to this thinking came from the work of Oswald and Fletcher, who proposed routines by which a machine could recognize blocks of words of a German text, and by which "the fluid German word order is resolved into a rigid English sequence." This suggestion, together with the suggestion of Booth and of Oswald and Lawson of strictly limiting the dictionary of the machine to those words and meanings required to translate in a particular field, brain surgery, for example, was supposed to eliminate the pre-editor for all but a very few routine problems, such as the splitting of long German compounds into their component parts. Even this problem seems amenable to solution by methods suggested by Reifler.

With or without the pre-editor, the output of the translating machine may still be no literary masterpiece. But it may be satisfactory for some purposes. For example, it might be adequate for the use of the scientist in keeping up with the foreign literature in his field. Much of the problem of keeping up with the literature is concerned with looking over articles in a rather cursory manner and deciding which ones merit more careful attention. For every important article, there are usually many that are unimportant for that particular person. If the scientist or engineer can scan and discard 100 documents by seeing only a rough translation made by a machine, and can select the one in which he is particularly interested, this one can be translated for him carefully by an expert human translator. If imperfect mechanical translations are given a fairly wide circulation to people who are interested in following the literature in a given field, the demand for translations of good quality, made by standard methods, will increase greatly. Thus the wide use of imperfect but useful mechanical translations may actually increase the demand for human translators.

The output of the machine itself, of course, could be made the basis for the more careful job of translation. This leads us to the concept of the post-editor, which has also been discussed in detail in the MT literature, particularly by Reifler.

A post-editor is a person skilled in the out-

put language but who may be entirely ignorant of the input language. His task is to take the imperfect output from the machine and edit it into a polished or at least easily comprehensible document. This puts man in the role of partner with the machine. Or, as some would have it, the machine helps him produce the output text by doing much of the routine work that he would otherwise have to do to produce an acceptable translation. Although man has been reduced to a link in the chain, he does not have to solve the large number of routine problems, but can concentrate on the real difficulties. It has been shown that the post-editor is better able to do his job if he also knows the input language; thus we have the bilingual post-editor. It has also been shown that the post-editor is better able to do his job if he is an expert in the particular field of knowledge. If a mathematics text is being translated, the post-editor should be an expert in mathematics. Various authors have specified different ideal qualifications for the post-editor. It seems obvious that the amount of work done by a post-editor depends upon the ultimate purpose for which the translation is being made. If the purpose is to provide a translation in a literary style that could be published in a journal, possibly with large circulation, the post-editor might have a big job. If the purpose is to provide a rough copy that can be used by experts to determine whether or not the material is of interest to them, the post-editor would have a smaller job, or might not be needed at all. His utility depends upon how perfect a translation the machine makes and how perfect a translation is desired.

If the output of translating machines is imperfect, but adequate for screening purposes, the ultimate user or reader of the translation can be regarded as his own post-editor. He may be strongly motivated to acquire the skills necessary to do his own post-editing as he is now motivated to learn several languages so that he can keep up with the literature in his field. Thus there may be a considerable change in language teaching in the schools, with more emphasis on the skills of post-editing and less emphasis on reading ability of foreign scientific material.

Let us at this point dispose of the post-editor by saying that a machine can probably be constructed which will give a translation that is sufficiently accurate for any purpose that we happen to have in mind, if we don't have in mind a translation which reflects accurately the literary quality of the original. We now inquire what is the relation of man to the machine under these circumstances.

We still have man as designer and creator of the machine; but let us not be so demanding as to say that he must create the machine and the translation system in its final form before the switch is thrown and the machine starts carrying out its built-in destiny. Let us

suppose that man as the creator does not do as good a job as this, but first designs and builds a machine that can translate some things, but not all things. To be specific: the machine may have only a limited vocabulary; it may be able to handle only a limited number of grammatical or syntactic problems. Man in this new role, which we might call monitor and program adjuster, watches the machine translate, checks the output, notes its shortcomings, and alters the design or the program or the contents of the memory of the machine in such a way that the machine gradually builds a larger vocabulary, gradually becomes more proficient. Such a man may actually post-edit, but if the output is already satisfactory he will not have to do this. His duty is to instruct the machine, taking his cues from the machine's shortcomings as revealed by its output. We might say that the man is providing feedback of the type required for learning and that he is altering the machine in such a way that it

behaves as if it were learning by its mistakes.

There are some who believe that this learning loop can be closed inside the machine, that the machine can be programmed to learn by its own mistakes with no human intervention other than the original design and construction of the machine. Perhaps experiments with the more deterministic type of machine will help show how to realize a learning type of translating machine at some time in the more distant future.

We have briefly discussed some of the ways in which man and machine may be related in the future mechanical translation industry. Beside these more or less obvious connections, the easy availability of mechanical translations of the most important foreign scientific and cultural writings is bound to have a great effect upon international communication and understanding; on our own culture, science and technology; and thus on nearly all of the occupations of man.