MACHINE TRANSLATION

The field of machine translation (MT) has had a long and turbulent history. Indeed, it was the first nonnumerical application suggested in the 1940s for the then nascent field of computer science. Within a few years, more than two dozen MT projects had been organized in the United States and in Europe. Research and development continued until the early 1960s, when it became clear that a high-quality, fully automated translation system could not be developed within the then current state of the art. This realization led to the cut-off of research support by the United States government and signified the end of the early MT paradigm. In the late 1970s, research and development in MT intensified significantly, especially in Europe and Japan. From then, until the present, the number of MT projects has continued to grow. Among the contributing factors to this renascence are the enormous improvement in the quality of computer hardware and software, realistic expectations, and progress in such essential for MT areas as theoretical and computational linguistics and artificial intelligence (AI). At present MT is a vibrant research and development topic, actively pursued in the United States, Japan, and Europe. A number of good surveys of the history of machine translation are available, notably, Zarechnak [1] and especially Hutchins [2], who discusses the topic in considerable detail. For additional descriptions of current issues in the field, see, e.g., Nirenburg [2a].

HISTORY OF MT RESEARCH

In the late 1940s MT seemed a very attractive and feasible application of computer technology. This opinion was bolstered by the following considerations. First, in the era of information explosion translation became a very important business. As in every other business, automation is supposed to enhance efficiency. Second, translation is a common task regularly performed by humans. Therefore, the specification of the task is relatively straightforward: the conceptual design of a potential MT system can be modeled after the organization of the translation process performed by humans. Third, the dictionary lookup, which may account for a very significant part of the time spent on translation, can be reduced to an insignificant level when on-line dictionaries are used. Finally, spectacular successes of cryptography during World War II called for applying its methods to other fields. Translation could indeed be understood as a code-breaking task (“When I look at an article in Russian, I say: ‘This is really written in English, but it has been coded in some strange symbols. I will now proceed to decode,’” [3], and this was yet another impetus to the development of MT. Feasibility considerations tended to be influenced by the perception of translation as a common everyday task, performed with relative ease by humans. Such was the rationale behind the exciting development of MT research from the late 1940s until the early 1960s.

It is customary to consider the so-called Weaver memorandum as the starting point of research in MT. In 1949, Warren Weaver, then a Vice President of the Rockefeller Foundation, distributed 200 copies of a letter in which he suggested the concept of MT to some of
the people who could have an interest in developing it. Even though the memorandum was predominantly a strategic document, several important theoretical and methodological issues were discussed in it, including the problem of multiple meanings of linguistic units, the logical basis of language, the influence of cryptography, and the necessity to analyze language universals. Not all of the scientific ideas in the memorandum were appropriate and useful (notably, the entire cryptographic angle soon proved to be inapplicable), but it aroused significant scientific and public interest in the concept of MT. In 1948, the University of London team led by Andrew Booth and Richard Richens was the only one to carry out research and experiments in MT. In the first two years after the Weaver memorandum, work on MT started in earnest at a number of scientific research institutions in the United States, including the Massachusetts Institute of Technology, the University of Washington, the University of California at Los Angeles, the RAND corporation, the National Bureau of Standards, Harvard University, and Georgetown University.

The major concepts, topics, and processes of MT—such as morphological and syntactic analysis, pre- and postediting, homograph resolution, interlingual representation of meaning, work in restricted vocabularies, automating dictionary look-up, etc.—were first defined and debated at that time. The first scientific conference on MT was held in 1952 at MIT, and the first public demonstration of a translation program took place in 1954 at Georgetown University.

The Georgetown experiment involved translating 49 Russian sentences, selected from texts on chemistry, into English. The dictionary included about 250 words, and the Russian grammar consisted of only six rules. No pre-editing of the source language sentences was required, and the output was of adequate quality. This experiment was perceived by the general public and sponsors of scientific research as strong evidence for the feasibility of MT. The wide publicity and resonance of this experiment has also led to the establishment of MT projects outside the United States, notably, in the Soviet Union.

Through the 1950s and into the following decade, research in MT continued and grew. The requirements of this application gave an impetus to significant theoretical developments in linguistics and what would later become known as the discipline of artificial intelligence. The quality of actual translations, however, still remained largely below an acceptable level and required extensive postediting, as can be seen from the following example, an excerpt from a 1962 demonstration of the Georgetown GAT system, translated from the Russian original:

> By by one from the first practical applications of logical capabilities of machines was their utilization for the translation of texts from an one tongue on other. Linguistic differences represent the serious hindrance on a way for the development of cultural, social, political, and scientific connections between nations. Automation of the process of a translation, the application of machines, with a help which possible to effect a translation without a knowledge of a corresponding foreign tongue, would be by an important step forward in the decision of this problem.

Still, researchers in MT remained largely optimistic about the prospects of the field. “The translation machine. . .”—wrote Emile Delavenay in 1960—“is now on our doorstep. In order to set it to work, it remains to complete the exploration of linguistic data.” When Yehoshua Bar Hillel published his critique of contemporary MT research [4,5], his was a minority opinion. Bar Hillel’s central claim was that fully automatic high-quality machine translation was unattainable because of the inability of building computer programs for lexical disambiguation. His now famous example was the following paragraph:
Little John was looking for his toy box. Finally, he found it. *The box was in the pen.* John was very happy.

The word “pen” in the emphasized sentence above has at least two meanings; a writing pen and a playpen. Bar Hillel’s conclusion was that “no existing or imaginable program will enable an electronic computer to determine that the word *pen* in the given sentence within the given context has the second of the above meanings.”

Since Bar Hillel was one of the early champions of MT and had intimate knowledge of the research in the field, his critique has had a wide resonance in the public attitudes toward MT as well as among its sponsors in U.S. government and industry. Coupled with the increased difficulty of problems facing MT research after the initial successes, and notwithstanding the fact that many of the then current projects (notably at Georgetown and IBM) pursued exactly the type of MT research recommended by Bar Hillel, namely, a combination of machine translation with human postediting, this criticism started the process of reassessment of attitudes toward the field. The reassessment culminated in the publication in 1966 of a report by the influential Automatic Language Processing Advisory Committee organized in 1964 by the National Academy of Sciences. The ALPAC [6] report, as it came to be known, was critical of the state of the art in MT and recommended drastic reductions in the level of support for MT research. The ALPAC report was sharply (and appropriately) criticized as biased and inaccurate. Many of its assessments, however, were correct, especially those dealing with the evaluation of practicality of MT research. In retrospect, the strongest negative effect of ALPAC was not so much the reduction in funding as the damage to the status of MT as a scientific endeavor in the United States.

The early MT projects, indeed, failed to reach their stated goal of building systems of good quality, fully automated translation in broad domains. The principal mistake of the early MT workers was, however, one of judgment: the complexity of the conceptual problem of natural language understanding was underestimated. The variety and the sheer amount of knowledge that must be used in any solution of this problem proved to be enormous, so that the success of MT as an application became dependent on the solution of this problem. It would take more than fifteen years for MT to start a scientific comeback.

While the ALPAC report effectively brought MT programs in the United States to a halt, research and development continued in several scientific groups in the Soviet Union, Canada, Germany, France, and Italy, as well as in a small number of commercial institutions in the United States. Notable achievements of MT in the 15 years after the ALPAC report included the development and everyday use of the first unquestionably successful MT system, TAUM-Meteo, developed at the University of Montreal, Canada, and used routinely to translate weather reports from English into French. The MT program Systran has been used during the *Soyuz–Apollo* space experiment in 1975, and in the following year was officially adopted as a translation tool of the European Economic Community (EEC).

The beginning of the revival of MT as a scientific discipline and an application of linguistic and computer technology must, however, be traced to the establishment of the Eurotra project and the MT effort in Japan. Begun in 1978, Eurotra is an ambitious, well-supported project aimed at providing MT capability among all official languages of the EEC (Danish, Dutch, English, French, German, Greek, Italian, Portuguese, and Spanish). At present, Eurotra employs about 160 researchers in a number of national groups and at the project headquarters in Luxembourg. The current generation of Japanese MT efforts started around 1980, supported both by the government (notably, within the Fifth Generation Computing project) and by industry.
The above developments gradually led to a revival of MT research in the United States. In 1983, the National Science Foundation approved a research grant, albeit a small one, for MT work, the first such action in two decades. MT activities at various scientific meetings have significantly intensified, and several conferences devoted specifically to MT have been organized recently. New research groups have been set up, notably, the Center for Machine Translation at Carnegie-Mellon University, with a staff of about 30. The importance of MT has been underscored by the resonance of MT Summit and MT Summit II, a congress, held in Japan in 1987, and in West Germany in 1989, respectively, which attracted researchers, users and sponsors of MT from all over the world, members of academia, governmental institutions, industrial entities, and multinational bodies, such as EEC. The general mood of the conference was one of optimism and the realization that the need for MT in modern world is even more pressing than it was 40 years ago.

The new optimism of MT researchers and sponsors is based on spectacular advances in computer technology (drastic improvements in processing speed and memory capacity, advances in computer architecture, emergence of database technology, development of high-level programming languages, and interactive programming environments, etc.) and computational linguistics (in particular, techniques for morphological and syntactic analysis and synthesis of natural language texts). Advances in automatic processing of meaning and techniques of human-computer interaction are also an important component of the current MT paradigms. With the knowledge of the past difficulties, and therefore, with a realistic assessment of the possibilities of MT technology application, the current MT projects are well equipped to produce a new wave of scientifically nontrivial and practically adequate machine translation systems both to compete and to cooperate with humans in translating a wide variety of scientific, industrial, official, journalistic, and other texts.

MACHINE TRANSLATION TECHNIQUES

The task of MT can be defined very simply:

1. Obtain a text in one language (SL, for source language) and understand its meaning
2. Produce a text in another language (TL, for target language), so that the meaning of the TL text is the same as the meaning of the SL text

The successful completion of the former task is a prerequisite for the latter. We will therefore concentrate at first on the understanding facet of MT work. A number of important questions can be raised at this point:

1. How does one set out to extract the meaning of a text?
2. How does one represent the meaning of a text?
3. What is this meaning?
4. Is it absolutely necessary to extract it (or at least all of it) in order to translate?

Question 3 is a basic problem in linguistics and philosophy of language. We cannot even circumscribe all of its facets here. We will have to take a more operational approach to discussing the problem of the meaning of MT. Question 2 relates to the problem of knowledge representation, either in the style it is done in artificial intelligence or in a more traditional sense, as in lexicography. Question 1 highlights the computational problems of
such an enterprise as MT. All the above are difficult problems, and no definitive solutions have been suggested at this stage of the development of the field of computational linguistics and AI. It is in this light that one must interpret Question 4. Is there a possibility that success in a particular application area, such as MT, is not contingent on producing workable solutions for the above problems? The rest of this section is devoted to a discussion of the depth of meaning analysis necessary for determining the translation of a text.

Human translators use dictionaries as sources of information about SL and TL. The type of dictionary that is most often used by humans is the bilingual dictionary, which incorporates correspondence units of SL and TL. By design, this SL–TL mapping seeks to preserve meaning. And, since meaning is the invariant between the SL and TL texts in translation, such dictionaries must serve the purpose adequately. An important point to remember, however, is that bilingual dictionaries, as we know them, are designed for human use. People possess a great ability to “make sense” of language units. This makes the task of the lexicographer simpler in that not all aspects of meaning have to be absolutely laid out; people will be able to understand even a flawed explanation. The situation is quite different when the dictionary is used by a computer program. Let us illustrate the types of dictionaries and processing modules that will become necessary in this case.

BUILDING AN MT SYSTEM

Suppose the system obtains the German text (1) as input.

(1) Das Buch liegt auf dem Tisch.

The English translation of (1) is (2).

(2) The book is on the table.

The dictionary necessary to perform this translation is as follows:

<table>
<thead>
<tr>
<th>German</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>auf</td>
<td>on</td>
</tr>
<tr>
<td>Buch</td>
<td>book</td>
</tr>
<tr>
<td>das</td>
<td>the</td>
</tr>
<tr>
<td>dem</td>
<td>the</td>
</tr>
<tr>
<td>liegt</td>
<td>is</td>
</tr>
<tr>
<td>Tisch</td>
<td>table</td>
</tr>
</tbody>
</table>

The translation program supplied is one that substitutes the English words for their German counterparts, one by one. Do we have an MT system? Yes, a system of machine translation of the German sentence (1) into English. Indeed, it is only under certain constraints that “liegen” should be translated as “be.” Can we use the same system to go from English into German? No, because we have a one-to-many relationship from “the” to “das” and “dem.” Our knowledge, as recorded in the dictionary, is insufficient to resolve this ambiguity, and we have no additional knowledge to help us make the proper choice. This is the first time we observe that MT research can be viewed as a process of accumulating knowledge that facilitates making correct choices of output. Now, to translate (2) into Russian, we will need the dictionary as follows:
English  Russian
a         #
book      kniga
is        #
on        na
table     stole
the       #

Sentence (2) will be, therefore, translated into Russian as (3).

(3) Kniga na stole.

Interestingly enough, (3) is also the translation of (4):

(4) The book is on a table.

The above suggests, somewhat unexpectedly, that the articles do not have meaning in English. This, of course, is not true. We will return to the question of how the articles influence the translation later. Note also that the Russian word “stole” is in fact one of 10 different words (corresponding to different case and number values) that will in real texts correspond to the English “table” in its meaning of a piece of furniture. We will not, however, discuss morphological analysis here. Morphological analyzers have been built for many languages, including such morphologically rich ones as Hebrew or Finnish.

Once again, we can see that it is impossible to use the same dictionary for a back translation of (3) into English. There is no indication where to put (if at all) the words that correspond to zero strings in Russian. This example shows that more knowledge has to be introduced into the system. For instance, the fact that class nouns, when used in their singular forms in English sentences, must be preceded by an article (a, the), a demonstrative (this, that . . .), a possessive (e.g., my, their), a question-word (e.g., what, which, whose), or the quantifier one. This is a part of the knowledge about the syntax of English.

The knowledge and difficulties described above were well known to the early MT researchers. One way of coping with these problems for them was to build the so-called “direct” translation systems that were specifically designed with a particular SL and a particular TL, and in which analysis of TL actually depended on what the TL was. Most of the early MT systems were direct, and the large production-level systems of the 1970s, such as Systran or Spanam, were also built on this principle. The rules on which they operated can be exemplified by the following excerpt from the set of rules used in the Ramo-Woolridge Russian-English MT project (quoted from Hutchins, [2] and slightly edited to facilitate understanding):

If in the SL text the word DO follows within eight words of OT in any sentence, translate the Russian OT as the English FROM and the Russian DO as the English TO.

If in the SL text ZA is followed by a symbol, translate it into English as PER, otherwise, as FOR or DURING.

To translate the Russian NET do the following: if OF was inserted into the TL before a potential genitive at the beginning of a word block preceding or following NET, suppress the English equivalent for NET and substitute THERE IS NO for OF.

In general, the processing in direct translation systems is organized as illustrated in Figure 1.
In an advanced direct translation system the processing is typically divided into many sequential steps, most of them are language pair-specific. The following is the set of processes used by Systran for translating a SL sentence:

1. Morphological analysis of SL words
2. Resolution of homographs based on grammatical categories of words in immediate context
3. Detection of compound nouns (e.g., error message)
4. Identification of the boundaries of noun and verb groups and other phrases
5. Identification of dependency relations (agreement, government, parataxis) inside phrases
6. Identification of conjunction (including punctuation) and conjoined phrases (needed, e.g., to determine that in a phrase such as “old men and women” the conjoined elements are the two nouns, while the adjective modifies the entire conjunction, not just the noun “men”)
7. Identification of the functional structure of the sentence, primarily, of its subject and predicate
8. Analysis of prepositional phrases
9. Translation of idiomatic expressions
10. Translation of prepositions
11. Resolution of remaining translation ambiguities, using the information in the main bilingual dictionary
12. Translation of remaining lexical units
13. Word order rearrangement

The above process uses a type of syntactic knowledge that was not tapped in our earlier examples, specifically, the knowledge of lexical categories of words. However, unlike in earlier MT systems, we can now derive the syntactic structure of SL texts within a theoretically motivated framework of generative grammar. The knowledge of the syntactic structure of (5) helps us to decide whether “coach” is a noun or a verb; “lost,” a verb or an adjective; “set,” a noun, a verb, or an adjective. The knowledge of English syntax is sufficient to eliminate this 12-way ambiguity and choose the correct reading. This type of knowledge is recorded in a “grammar” of a SL in a modern MT system. A special processing unit (a syntactic parser) applies this knowledge to the input text and produces its syntactic structure. The dictionary now can have a separate entry for every distinct syntactic
reading of an SL word (so that, e.g., “set” will appear as the head of three entries, one each for this word’s syntactic meaning of noun, adjective, and verb).

(5) The coach lost a set.

With the development of generative grammar methods for representing the syntactic structure of natural language sentences, it became clear for MT researchers that such analysis is best done based on the SL information alone, without a reference to a particular TL. The architecture of MT systems underwent a change, and the so-called transfer approach supplanted the direct translation paradigm. The architecture of a typical transfer MT system is illustrated in Figure 2. The process of translation in the transfer paradigm consists of three stages:

- Analysis of SL text
- SL—TL transfer
- Synthesis of TL text.

At the first stage, the SL text is analyzed, most often, morphologically and syntactically, without any reference to a particular TL language. A grammar of SL is used as the main source of knowledge. The result of this stage is typically a syntactic structure of the sentences in the input text, usually represented as a tree. Figure 3 shows some of the tree structures that can be used for representing the results of syntactic analysis of (5).

The results of analysis disambiguate the SL text to the point when it is easier for the transfer stage mechanisms to substitute lexical units and syntactic structures of SL by their counterparts in TL. Transfer processes rely on the information stored in bilingual SL—TL dictionaries and special transfer grammars that contain data about correspondences of syntactic structures in a particular SL—TL pair. Finally, the TL synthesis stage is responsible essentially for linearizing the TL trees obtained after transfer, according to the rules of a TL grammar.

Unfortunately, in some cases, syntactic knowledge is not sufficient for disambiguation. Thus, on purely syntactic grounds it is impossible to determine whether in (6a) the conjunction connects two nouns or a noun and a noun modified by an adjective (i.e.,
conjunction connects two nouns or a noun and a noun modified by an adjective (i.e., whether the chairs are also white). Note that one cannot neglect to extract this type of knowledge, because the form of a potential translation may depend on the intended meaning. Thus (6a) will be translated into Hebrew as either (6b) or (6c).

(6a) White tables and chairs
(b) Shulhanot levanim vekisaot: (white tables) and (chairs)
(c) Shulhanot vekisaot levanim: White (tables) and (chairs)

Further types of ambiguities that syntactic knowledge fails to take care of include prepositional phrase attachment and decomposition of noun–noun compounds in English.

Even more profound evidence of the insufficiency of syntactic analysis for MT is presented by the commonplace lexical–semantic ambiguity of natural language. Thus, in a standard English–Russian dictionary the words “coach,” “lose,” and “set” from (5), in their correct syntactic meanings, detected by the syntactic analysis, have 6, 10, and 34 readings, respectively. This is a 2040-way ambiguity. (Incidentally, even though syntactic disambiguation leaves us with this multiple ambiguity, syntactic analysis is still very useful. After all, if the correct parts of speech are not detected, the number of readings for the three words is 11, 15, 96, respectively, producing a 15,840-way ambiguity.)

It certainly says something about the disambiguating powers of humans that we can effortlessly assign this sentence a single meaning, with the only source of uncertainty in whether the game played was tennis, squash, or racquetball.
Coach lost set total
ways ambiguous, syntactically and semantically
11 15 96 15840
ways ambiguous, syntactic ambiguity eliminated
6 10 34 2040

As another example, consider the Russian sentence (7).

(7) Novaja partija byla luchshe vo vsekh otnoshenijax

Taken out of context, this sentence does not contain any clue as to the appropriate meaning of the highly polysemous Russian word *partija*. The texts in (8) illustrate the correct translations when the context is provided.

(8a) [The old Liberal Center was too doctrinaire for his taste.] The new party was better in all respects.
(b) [The previous consignment contained substandard supplies.] The new batch was better in all respects.
(c) [In the previous game he did not notice a fork that cost him a rook, and lost.] The new game was better in all respects.

The choice of the appropriate TL correlate was facilitated for humans by the context. A Russian–English dictionary that would be able to distinguish the alternatives will have to have special context identification markers, as in (9). These context markers are semantic in nature. They suffice for human translators. MT systems should, however, possess special means of identifying the semantic context and acting upon its recognition. This task has been a central concern of the subfield of natural language processing within AI, and was found to require much more information to be stored in the computer than simply the semantic markers such as those in (9) or those used in the dictionaries of such MT systems as Metal. In addition, the set of semantic markers should be made consistent and expressive enough to cover all the necessary shades of meaning in a particular domain of discourse, since even the descriptions of the same objects will require different granularity of distinctions for translating texts in different domains of discourse. Thus, describing the concept of land in a political context will be different from the description of the same concept in the domain of, say, agronomy.

(9) *Partija*, noun, feminine. 1. (political) party; 2. (commerce) batch; 3. (chess) game.

In order to automate semantic analysis one must, first of all, devise a principled way of representing the meaning of the input text with the help of a complete system of semantic markers and then provide rules of using such a representation to extract the necessary knowledge about the context. The large number of semantic markers necessary to describe a reasonably rich subworld, the fact that they stand in well-defined relationships to other markers, and the total absence of any natural language material in the representation are characteristics of AI-oriented approaches to meaning extraction. The notations used in AI systems are rich enough to be called semantic interpretation languages. In AI they are called knowledge representation languages; in MT they are traditionally called *interlingua*. If the analysis module of an MT system can produce a representation of an SL text in terms of a natural language-independent *interlingua* (IL), and the synthesis module of such a
system can obtain this IL text as input and produce a TL text, then the transfer stage of the MT process can be avoided. Moreover, bilingual dictionaries will be rendered superfluous, since SL and TL will not be in any direct contact. And this means that once an IL text is produced from a certain input it can then be used to generate texts in any number of TLs, thus facilitating multilingual MT. A significant economy of effort ensues, due to the fact that there will be no need to compile bilingual dictionaries and transfer grammars for each SL/TL pair, but instead only two dictionaries for each natural language: one for the SL-IL translation and the other for the IL-TL translation. This economy of effort is illustrated in Figure 4.

In general, the interlingua approach can be considered a logical extension of the transfer approach, where the "length" of the transfer stage has shrunk to become a point (cf. Fig. 5).

MT researchers understood the advantages of the interlingua approach to MT very early. However, the task of semantic interpretation of SL texts has proved too difficult for the early MT efforts. One of the first attempts to use semantic analysis, that is, to use a knowledge representation language in MT, was made within the Yale AI school. The Conceptual Dependency knowledge representation language [e.g., 7] was used to represent the meaning of the input sentence. The experimental MT systems that were built in this school used background knowledge about the world to infer information not explicitly mentioned in the input sentence, in order to be able to disambiguate it. This background knowledge usually described typical complex event sequences, called "scripts," that are common in certain subworlds. Thus, the knowledge that the event described in (5) belongs to the $\text{STENNIS}$ script, helps disambiguate this sentence completely by suggesting the

---

**FIGURE 4.** (a) Transfer approach: $N(N-1)$ processes. (b) Interlingua approach: $2N$ processes. For $N = 72$, transfer = 5,112, interlingua = 144 processes.
appropriate readings for the words “coach” and “set.” Once the meanings of the input sentence have been represented in the interlingua, with the help of scripts and using conceptual dependency, a set of discrimination nets, with choice points marked by particular units of semantic knowledge, are used at the generation stage of the system to connect interlingua meanings and TL words and phrases.

Schematically, the process of translation in such systems can be illustrated as follows [we use the example from Ref. 8]. Suppose, for instance, that the sentence (9) has been supplied as the input to the translation program. The sentence will be analyzed (translated into the conceptual dependency language) as (10). The dictionary that the analyzer will use will connect the English verb “hit” with a “frame” in which there will be a slot for “action,” occupied by the marker PROPEL (which is not to be interpreted as an English word!) and a slot for “force” which will be filled by the marker ABOVE-AVERAGE. The slots for “agent,” “object,” and “instrument” will be listed in the dictionary without fillers. It is the responsibility of the analyzer to create an interlingua structure for the new event, using the representation of “hit,” taken from the dictionary, as its nucleus, and to find fillers for the slots that are unoccupied. It is predominantly for the purpose of identifying these slot fillers that the system uses scripts and other background knowledge.

(9) Mary hit John.

(10) (event EV001
    (action PROPEL)
    (agent MARY)
    (object JOHN)
(instrument *UNKNOWN*)
(force *ABOVE-AVERAGE*)
(intentionality *POSITIVE*)

In order to translate (10) into Spanish, a discrimination net such as that in Figure 6 must be used. The aim is to choose the appropriate verb to render the IL action PROPEL in Spanish.

In the case of generating (10) *pegar* will be chosen.

The kind of analysis performed by this type of system is, however, far from sufficient. A number of choices still remain unresolved even at this level of semantic processing. It appears that the script information alone is insufficient to resolve all the text ambiguities. Additional types of choices remain unaccounted for. Let us briefly illustrate them.

If (11) is uttered by a boss in a conversation with a subordinate, it should be translated into Russian as (12); if it is uttered in a conversation between a homeowner and a reluctant housepainter, it should rather be translated as (13). This example highlights the influence of the speech act character of the utterance (an order in the first case; a plea in the second) on the representation of its meaning and, therefore, its eventual translation.
(11) Will you please start working on the project?
(12) Bud’te dobry, nachnite rabotu nad proektom.
(13) Ne mogli by vy nachat’ rabotu nad proektom?

In a standard conceptual analysis system, all six Russian sentences in (14) will be assigned the same meaning. The word order permutations are, however, significant in that they contribute to the meaning of the sentence. In (15) the English translations of the nonemphatic readings of the sentences in (14) are listed. The sentences in (15) differ in what is considered (by the speaker) already known and what is considered new information in these sentences. Establishing these distinctions is known as thematic analysis of text. Note that, while in Russian these distinctions are marked by word order, in English word order is accompanied by the choice of indefinite and definite articles (indefinite articles typically introduce noun phrases that are new). Thus, we find that English articles do, after all, carry a meaning. Note that in (14a) the new information can be either the prepositional phrase v komnatu or the entire verb phrase voshel v komnatu. In the former case the indefinite article must be used; in the latter, the definite. The remaining ambiguity in (14f) is of a different origin: the new information in it is that the man came into a/the room and not into some other place.

(14) 
(a) Chelovek voshel v komnatu
(b) V komnatu voshel chelovek
(c) Voshel chelovek v komnatu
(d) V komnatu voshel chelovek
(e) V komnatu chelovek voshel
(f) Chelovek v komnatu voshel

(15) 
(a) The man came into a/the room
(b) Into the room came a man
(c) Came the man into the room
(d) A man came into the room
(e) It was a man that came into the room
(f) It was into a/the room that the man came

Sometimes it is difficult or impossible, while processing a text, to evoke a standard script or even a more general memory organization packet (MOP) that relates texts with typical abstract settings and events remembered from past experience. But it may help enormously if just the subworld to which the text belongs could be determined. Thus, (16) [9] will be translated in two clearly distinct manners depending on whether the text belongs to the subworld of jail or prizefighting. Of course, it is not an easy task to detect the subworld automatically.

(16) Rocky slowly got up from the mat, planning his escape. He hesitated a moment and thought. Things were not going well. What bothered him most was being held, especially since the charge against him had been weak. He considered his present situation. The lock that held him was strong, but he thought he could break it.

One must also have means of understanding elliptic construction. Ellipsis is an ordinary and necessary feature of all input texts, not an aberration. Thus, one has to be able to understand (17) if it appears in a text after (18). Special types of knowledge and processing arrangements are necessary for this task.
Six, to be precise.
There are several flights from Atlanta to Pittsburgh on Tuesday.

The problem of anaphoric reference also involves a number of knowledge-based choices. Knowledge is needed for the computer to be able to find the referents for "there" and "they" in (19), as well as the beginning of the list of problems referred to there. Note that one cannot, in the general case, translate (19) without understanding, that is, simply using corresponding pronouns in TL. This is because in some languages additional choices have to be explicitly made. For example, translating (19) into Hebrew, one will have to make a choice between the masculine and the feminine gender form of "they." In order to do this, one has to determine the anaphoric referent for the pronoun. Anaphoric phenomena cover not only pronouns but also definite noun phrases. In order to translate (20) [10, p. 56], one has to understand that all the italicized noun phrases in this text refer to one concept token. Indeed, a straightforward translation of this passage into, say, Russian would be difficult. A human translator would not use ispanec (Russian for Spaniard) to translate the noun phrase in the last sentence. It will be similarly impossible to translate the indefinite noun phrase a man as such into Russian if complete understanding of coreference relation in this text is not achieved.

There they found many additional problems.

Priest is charged with Pope attack.
(Lisbon, May 14) A dissident Spanish priest was charged here today with attempting to murder the Pope. Juan Fernandez Krohn, aged 32, was arrested after a man armed with a bayonet approached the Pope while he was saying prayers at Fatima on Wednesday night. According to the police, Fernandez told the investigating magistrates today he trained for the past six months for the assault. He was alleged to have claimed the Pope 'looked furious' on hearing the priest's criticism of his handling of the church's affairs. If found guilty, the Spaniard faces a prison sentence of 15-20 years.

expert system approach may prove interesting, because it will be a way of extracting this additional knowledge and learning to use it in a computer system. Naturally, it may prove impossible to extract that knowledge efficiently. It seems, however, that such knowledge exists: we intuitively believe that experienced translators know how to do translations better than novices. This may have to do with their ability to express their understanding of the SL text better. Indeed, it is the case that the translatorial agencies consider it more important for a (human) translator to know TL better than SL.

THE CHOICES

The deeper the desired level of analysis of a text, the more difficult it is to achieve with the help of computers, and, therefore, given the state of the art in the field, the less feasible it is at present to build an MT system that benefits from that level of analysis. Indeed, even syntactic parsers and grammars of sufficient generality cannot be taken off the shelf and used without major modifications. Semantic analyzers are scarce and provide at best a partial coverage of the world of concepts necessary for translation. Modules capable of analyzing the rhetorical content of the utterance or its discourse structure are even more remote. A significant amount of research is currently under way on the problems of semantic and pragmatic analysis of natural language. But this research is predominantly theoretical.

Is there a possibility to simplify or avoid complete analysis and still achieve tangible results in MT? The answer is a qualified yes. There are two major avenues of circumventing the problem of completely automatic disambiguation. First, one can restrict the grammar and the vocabulary of the input text in such a way that most of the ambiguity is thus eliminated. This is the sublanguage or subworld approach to MT. Second, one can drop the requirement of complete automation and allow humans to get involved in the translation process. As we will see, there are a number of ways in which this process of machine-aided translation can be organized. The difference between these approaches is not only in the tactics of interspersing the automated and manual steps in the process of translation, but also in the nature of the subtasks for which humans are responsible.

Those who contemplate building an MT system must weigh the particular requirements in terms of quality, allowed development time and breadth of coverage of their projects before deciding what level of automation is the most appropriate for them. A simplified rule of thumb is: the less time allowed for research and development, the shallower the analysis module and, therefore, the deeper the involvement of humans in the translation process.

Restricting the Ambiguity of SL Text

The best example of the sublanguage approach is the operational MT system TAUM-Meteo, developed at the University of Montreal and delivered to the Canadian Weather Service for everyday routine translations of weather reports from English into French. The system operates very successfully, practically without human intervention. Its vocabulary consists of about 1500 items, about half of which are place names. The syntactic constructions that occur in the variant of English used as SL in TAUM-Meteo constitute a relatively small subset of English syntactic constructions. There is very little semantic ambiguity in the system, since it is always expected that ambiguous words are used in the context of their meanings which belongs to the subworld of weather phenomena (for instance, the word front in TAUM-Meteo will be understood unequivocally as a weather front). Finding well-
delineated, self-sufficient and useful small sublanguages is a very difficult task, and this is one reason why the success of TAUM-Meteo was not repeated by any other operational system.

**Partial Automation of Translation**

The demand for MT is high and growing, most of it in subject areas whose corresponding sublanguages are much richer, and, consequently, less feasible, than that of weather forecasts (as the Montreal group quickly learned when they tried, with quite limited success, to extend their system to the subject area of aircraft maintenance manuals). To deal with the demand for automation of translation today one has to think of ways of using the knowledge and knowhow already available in the field to provide a certain degree of automation of the translation process. Practically all operational and experimental MT systems feature some human involvement, and it is safe to say that in the immediate future one can expect this involvement to decrease but not to disappear completely.

**There are two major classification dimensions for MT systems featuring partial automation. They may be classified according to:**

1. The actual share of work performed by humans rather than the computer (the degree of automation)
2. The strategy of human involvement: whether the humans work on the text before, during or after the computer deals with it (there is, of course the possibility of combining some or all of these three strategies)

In accordance with the degree of automation involved, MT systems range from relatively simple interactive editing and dictionary look-up tools for the use of human translators (the type of activity supported by such systems is known as machine-aided human translation, MAHT) to quite sophisticated systems, most of them still experimental, that involve syntactic and sometimes even semantic analysis and provide a much higher level of automation of the translation process (these systems perform human-aided machine translation, HAMT). The achievement of fully automated good quality translation remains a distant but, for some projects, persistent goal.

The philosophy of the MAHT approach is best expounded by Alan Melby of Brigham Young University. He recognizes three levels of human-computer interaction in developing what he calls a translator workstation. Level One Workstation, the least sophisticated one, essentially presupposes a complete and convenient word processing environment with convenient means of accessing on-line dictionaries and encyclopedias. Level Two Workstation adds the spelling checks, concordances and text dictionaries and presupposes that the text to be translated is in machine-readable form. Level Three Workstation involves a degree of automatic processing, including, possibly, some analysis. Melby is not specific about what particular means are available at this level, but it is clear that such a workstation is somewhere in between MAHT and HAMT.

With respect to the strategy of human involvement, the three basic possibilities have come to be known as preediting, postediting, and interactive editing. A human preeditor reads the input text and modifies it in such a way that the MT system is able to process it automatically. Difficult and overly ambiguous words and phrases are replaced with those that the editor knows the program will handle. A human posteditor obtains the output from an MT system and eliminates all inaccuracies and errors in it. An interactive editor engages in a dialogue with the MT system, in which the human resolves ambiguities that the machine is not capable of resolving itself. It is, of course, necessary to build a special interface to
maintain the dialogue. The types of questions asked can also vary: the computer may ask the human to provide a TL correlate for an ambiguous SL unit or may ask to be provided with the meaning of an SL unit, in the language in which the meanings are represented in the system.

Interactive editing has become a viable alternative with the development, over the past decade, of sophisticated interactive programming environments. A disambiguation-oriented dialogue can include interactions as follows:

The word ‘pen’ means:
1) a writing pen
2) a play pen

To resolve referential ambiguity, a system can ask a question in the following manner:

The word ‘she’ refers to
1) ‘Cathy’
2) ‘my mother’
3) ‘the sailboat’

The interactive word sense disambiguation can indeed be accomplished relatively easily. The referential ambiguity, though, will present a problem, because the program will have to be able to find the candidate (pronominal and other) referents, which is a nontrivial task in itself. The design of the interactive component to perform syntactic disambiguation may be difficult and the component itself, cost ineffective simply because it is not an everyday task for a human to compare syntactic structure trees. Semantic analysis, however, has a stronger disambiguating power and, therefore, syntactic disambiguation can be rendered unnecessary in an interediting system that relies on human intervention to choose the appropriate word senses.

Until very recently human intervention in the process of MT predominantly took the form of postediting, whereby the product of an MT system is submitted for editing by a human editor, exactly like human translations are in better translatorial agencies. (Preediting the SL text before submitting it to an MT program has also been discussed and used.) The important feature of this approach is that the posteditor is not required to know SL. In practice, however, many of the outputs of such systems come in such a garbled way that it becomes a major problem to edit them without dipping into the SL text.

In the systems that espouse the HAMT strategy with postediting (and a majority of current experimental systems belong to this group) feasibility and cost effectiveness become the major criteria for success. The postediting approach is based on the premise that MT can (and should) be performed without a complete understanding of SL texts by the computer. This belief is justified in terms of feasibility. What this approach means is that an MT system is essentially an aid in human-controlled translation.

The quality and depth of the disambiguation process, as determined by the quality of underlying conceptual models of language and world used in an MT system is an alternative criterion on the basis of which MT systems can be judged. If one accepts the position that
the nature of human involvement should be not in correcting the (errorful) texts produced by
the system, but rather provide the system with additional disambiguation knowledge that
was not recorded in its knowledge base, one becomes able to judge an MT systems in terms
of the latter criterion. The methodological basis for this approach is in the interactive
editing approach with the dialogue aimed at gaining disambiguation knowledge, not the TL
correlates that will be eventually obtained based on this knowledge. This approach is
compatible with the research strategy of gradual movement toward fully automated transla-
tion and, therefore, its success depends on significant advances in basic research; the former
approach is more of the engineering variety in that it aims at partial automation within the
realm of what is feasible today. However, the postediting approach already produces re-
sults that are considered adequate for a number of purposes, while no large-scale system
based on deep understanding has yet been put into operation. The tension between these
two positions enlivens the MT research and development scene of the 1990s.

REFERENCES

   (B. Henisz-Dostert, R. R. McDonald, and M. Zarechnak, eds.), Mouton, The
2a. S. Nirenburg (ed.), Machine Translation: Theoretical and Methodological Issues,
3. W. Weaver, Translation, reprinted in Machine Translation of Languages
4. Y. Bar Hillel, “Report on the State of Machine Translation in the United States and
5. Y. Bar Hillel, “The Present Status of Automatic Translation of Languages,”
8. J. G. Carbonell, R. E. Cullingford, and A. V. Gershman, “Steps Toward Knowledge-
   Based Machine Translation,” IEEE Transactions on Pattern Analysis and Machine
   Intelligence, 1981.
9. R. Anderson, R. Reynolds, D. Schallert, and E. Goetz, “Frameworks for Compre-

SERGEI NIRENBURG