

Limitations of Machine Translation



Gayane Shmavonyan

Under the rapid technical and technological advances throughout the world, **Machine Translation**, as well as **Computer Assisted Translation** is becoming widespread nowadays. However, there has been some disappointment regarding this issue and even Chomsky washed his hands of the matter stating that “machine translation seems to be pointless and probably quite hopeless” (Chomsky 1975:40). Indeed, correct, precise and high quality machine translation will hardly be achievable at least in the nearest future.

Even if the computer were suddenly able to communicate meaning, it would still fall short of what humans do with language in a number of ways. This is because linguists have long been aware that communication of meaning is only one among functions of language. These difficulties would tend to place machine translation outside the limits of solvable problems.

For instance, if we compare Text A meaning exactly the same thing as text B, translation would be no problem at all, and no professional translator would be needed. Absolutely anyone able to read would be able to translate any text between these two languages by looking up the number for the words in one language and then substituting the words with the same number in the other language. It would not be necessary to know both languages and, consequently, computer translation in such a case would be easy.¹

But the real truth of the machine translation efforts is that word 60, for example, in language A does not mean exactly what word 60 in language B means. A good experiment of how imperfect computer translation is was carried out by French and English translators, who translated various brief passages. The final results of such a process bore almost no resemblance to the original, however hard technical experts called translators tried.

There is at present no evidence suggesting that a computer can turn out High Quality text at a rate faster than a human. Indeed humans may in some cases be faster than the computer. A human translator, for instance, can only handle 200 to 500 words per hour, which is often true, whereas computers can do far better.

Many translators report that their non translating colleagues believe it should be perfectly possible for a translator to simply look at a document in language A and just type it out in language B as quickly as though it were the first language. If human beings could do this then there might be some hope for computers to do so as well. But they cannot make the correct choice of the terms in the majority of cases.

Computer translation developers prefer to ignore many of the limitations of computer assisted translation. They are convinced that computers with limitless power will be able to perform a number of operations today’s computers are not capable of. They are supposed to provide terminologists with a complete list of possible terms to use. But they

may not reliably make the correct choice between the terms in most cases. Even in the 1940th prominent linguist L. Bloomfield acknowledged that “there must be some way around all the difficulties connected with computer translation. The only hope for a thorough solution seems to lie with technology” (Bloomfield 1944).

Distinction should be made between the two types of translation systems: **Machine Translation (MT)** and **Computer Assisted Translation (CAT)**. Machine translation performs the use of computer software to translate a text or a speech from one natural language to another. This technique is used for more complex translations such as translation of idioms; it is also effective in places where formal language is used, such as translation of official and legal documents. MT has proven useful as a tool to assist human translators; it produces a rough text for a human translator to revise. In a very limited number of cases it can be used in e.g. weather reports. The translator must interpret and analyze all the features of the text: grammar, semantics, syntax, idioms of the source language.

The problem how to program a computer that will “understand” a text as a person does and that will “create” a new text in the target language that “sounds” as if it has been written by a person can be solved in terms of two approaches to Machine translation: **Interactive and Interlingua Approaches**.

An Interactive system prints and displays the text created with no help from the translator, it only consulted with the translator on various words.

The so called Interlingua Approach encodes the information as a universal stage enabling translation back and forth between languages. Interlingua machine translation can use a method based on linguistic rules, which means that words will be translated in a linguistic way, the most suitable (orally speaking) words of the target language will replace the ones in the source language.

The Interlingua approach to machine translation has advantages and disadvantages. The advantage of multilingual machine translation is that no transfer component has to be created for each language pair. The obvious disadvantage is that the definition of an Interlingua is difficult and maybe even impossible for a wider sphere. The ideal context for Interlingua machine translation is thus multilingual machine translation in a very specific sphere. The machine translation system provides the target text segments that are paired with the source text segments so that the end product is a quality translation.

Computer Assisted Translation (CAT) is a form of translation where a human translator translates texts using computer software designed to support and facilitate the translation process. Computer Assisted translation is sometimes called Machine aided translation, where the software is designed to help translators do their work. It is a broad term converting a range of tools, from simple to complicate: spell checkers, grammar checkers, word processing software, database created in a program.

Having discussed the types of translation technology, it is important to turn to translation memory. It should be noted that **Translation Memory** is a multilingual text archive containing multilingual texts, allowing storage and retrieval of aligned multilingual text segments against various search conditions. In other words, ASE (Assisted Stock Exchange) stores source and target language pairs of text segments that can be

retrieved for use with present texts and texts to be translated in the future (Webb 1992).

A source text sentence or sentence like unit (heading, title) may be segmented into larger units such as paragraphs or small ones such as clauses. As the translator works through a document, the software displays each source segment in turn and provides a previous translation for reuse if the program finds a matching source in its database. If it does not, the program allows the translator to enter a translation for a new segment. After the translation for a segment is completed, the program stores the new translation and moves into the next segment.

Translation memory is a technology, too. It has proven to be useful, especially in technical fields where electronic documents are constantly being updated and revised. Individual translators, translation agencies, clients, and companies with in-house translation divisions can all benefit from this technology. Translation memory saves the user time and money under the appropriate circumstances. The key for success is known when, where and how to use it.

Translation memory should be considered one of the many tools of the translation trade. A painter has a brush and palette; a pianist has sheet music. How painters and pianists use their brushes, palettes and sheet music is what counts. The same can be said of translators and translation memory.

Unlike the brush, palette and sheet music, however, not everyone recognizes that translation memory is an invaluable tool for the translator and that the translator, like the artist, is the one who must fit all the pieces together to form a work of perfection. The more that is written about this technology, the more translation professionals and their clients will be made aware of its important role in the translation process.

How does TM (Translation Memory) differ from MT (Machine Translation)? MT creates automated translations and requires an advanced terminology database that includes all grammatical elements of a language. The MT system uses comprehensive dictionaries to translate the source text while at the same time applying the grammatical rules, or rule sets, from the database in order to produce the resulting grammatically correct target sentences. The technology sounds like an excellent solution. However, there is a catch: the source and resulting target text segments are not stored away in a database for future use. If a similar text (such as an automobile user's manual for the same model but different year) needs to be translated, the MT system would have to start from scratch. On the other hand, a TM system is used as a translator's aid, storing a human translator's text in a database for future use. TM can be used in a few different ways. One way would be to have a translator or a machine translation system translate the original text, using translation memory to store the paired source and target segments. The translator could then reuse the stored texts to create the revised or updated version of the text. Only the segments of the new text that do not match the old one would have to be translated. The alternative would be to use an MT system or a different TM system to translate the original. The new TM system could then be used by a translator to translate the revision or update by aligning the texts produced by the MT system or other TM system and storing them in the TM database for present and future work. The translator could then proceed to translate only the segments of the new text, using TM as described above.

What distinguishes TM from other Computer-Assisted Translation (CAT) tools? There are many CAT tools available to assist the translator, such as bilingual and multi-lingual dictionaries, grammar and spell checkers and terminology software but TM goes one step further by making use of these other CAT tools while at the same time matching up the original source document stored in its database with the updated or revised document through exact matching. Normally, the basic unit of the text in a TM database is a sentence; and the TM user can define what the unit will be. The basic unit might even be a sentence fragment or a paragraph. The translator does not have to retranslate the work he or she has already completed.

It is necessary to turn to language as a social phenomenon. How can the computer, a nonhuman, possibly assist the language learning process?

Language is very much a part of being human. Despite our best efforts over 50 years, we still can't teach the computer how to function in human language like a human being does. Despite amazing advances in artificial intelligence and computational linguistics, and promises of products that we see on the market, we will probably always be reliant on humans to negotiate language. This is good news for those of us involved in the language translating business – our job as language translators is certainly secure. No one who knows what he or she is talking about suggests that the computer can take the translator's place.

Perhaps the computer would be a failure as a language translator. So would a translator himself. But both the translator and the computer have valuable contributions to make to the language translation process.

According to Jeff Allen, a research linguist and translation lab supervisor at the Center for Machine Translation at Carnegie Mellon University "it is true that many large companies are looking into automated translation solutions; however, this does not necessarily mean that the companies are trying to replace their valuable and experienced translation personnel with computers, but rather that the companies are trying to reconsider if they are using the time of their translation specialists most effectively and how they can improve it" (Allen 2005).

Translators, agencies and corporations will have to adapt to the new technologies and solutions in order to remain competitive over the long term.

Notes:

1. Human translators are often too inconsistent or too slow in choosing the right equivalents of the given words and expressions.

References:

1. Allen, J. (2005) *Machine Translation*. Pennsylvania Pittsburgh: Carnegie Mellon University Press.

2. Bloomfield, L. (1984) *Language*. Chicago: Chicago University Press.
3. Chomsky, N. (1975) *The Logical Structure of Linguistic Theory*. Chicago: Chicago University Press.
4. Gross, A. (1992) *Limitations of Computers as Translation Tool*. London: London University Press.
5. Garner, D. H. (1996) *FAQs and Fictions about Computers and Language Teaching*. Cambridge: Cambridge University Press.
6. Hackett, Ch. F. (1970) *A Leonard Bloomfield Anthology*. Bloomington: Bloomington University Press.
7. Webb, L.E. (1992) *Advantages and Disadvantages of Translation Memory*. San Francisco: San Francisco State University.

Համակարգչային թարգմանության սահմանափակումները

Ամբողջ աշխարհում տեխնոլոգիաների զարգացմանը զուգընթաց աճում է համակարգչային թարգմանության պահանջարկը: Տիրապետելով բազմաթիվ ու բազմազան թարգմանական հնարավորությունների՝ համակարգչային թարգմանությունը այնուամենայնիվ ունի իր թերություններն ու սահմանափակումները: Տեխնոլոգիայի զարգացման ներկայիս փուլում թարգմանչի կողմից կատարված թարգմանությունը ավելի որակյալ, արդյունավետ ու արժեքավոր կլինի, եթե այն համատեղվի համակարգչային թարգմանության հետ: