COMPUTATIONAL LINGUISTICS, HUMAN AND ARTIFICIAL INTELLIGENCE IN LEARNING A FOREIGN LANGUAGE

Camenev Zinaida, assoc. prof. PhD, FIUM

Articolul este dedicat lingvisticii inginereşti, o metodică experimentală care exprimă legătura dintre limbă, limbaj, funcţionarea mecanismului interior al limbii naturale cu calculatorul. Problema constă în pregătirea materialelor pentru a asigura realizarea traducerii automate, deoarece lingvistica contemporană manifestă un interes deosebit pentru problemele de cercetare sistematică a limbii la nivel gramatical cît şi lexical.

La baza soluţionării problemei stă procesarea automată care include segmentarea automată, schematizarea automată şi traducerea automată precedare de filtrarea sensurilor atribuite unităţii lexicale din dicţionarele generale şi de specialitate înfăptuite de lingvist, traducerea umană în contrast cu traducerea artificială efectuată de calculator. Aici sunt incluse procesarea şi producerea concordanţei la traducerea unităţilor lingvistice selectate.

Cuvinte-cheie: lingvistică, inteligenţă artificială, limbă străină, computer

Computational Linguistics is the study of computing systems for understanding and generating natural language. Three classes of applications are supposed in this field: machine translation (MT), information retrieval and human-machine interfaces. Also included in this section is language data processing, specifically the production of concordances and word indexes for literary and linguistics research. It is a simple application of the computer in computational linguistics [Ahmad et al., 1985, p. 39]. Hutchins [1986, p. 15] defines MT as "the application of computers to the translation of texts from one natural language to another".

Communication across language boundaries has never been more important and necessary than today. In the 21st century, emerging digital communication technologies and the World Wide Web (WWW) facilitated an exponential growth in the volume of information in

different languages and blurred geographical boundaries once hard-to-find data easily accessible online.

Rapid growth in international commerce, in political cooperation and in technical cooperation led to a "translation explosion", while at the same time the cost of computers fell. Before long an increasing number of translation users found machine translation could be worthwhile, provided it was treated sensibly. MT now is a reality, and it is used for a variety of purposes in translation services in Eastern and Western Europe, North America and the Far East. It is fast, its quality has improved, and the need for it is much greater, since human translation often cannot supply the quantity at the right speed or price.

The growing number of users from non-Anglophone countries created a great demand for cross-cultural interaction and understanding of online content in the era of globalization speeding up every aspect of modern social interactions, including consumption of information and communication. Under these circumstances translation plays an extremely important role.

The term MT refers to computerized systems responsible for the production of translations with or without assistance. The boundaries between machine-aided human translation (MAHT) and human-aided machine translation (HAMT) are often uncertain and the term computer-aided translation (CAT) can cover both, but the central core of MT itself is the automation of the full translation process.

While comparing the human translation with the machine one we see that the human translation is the best. The accuracy of MT is much lower than competent human translation, but it can be improved in some way by introducing elements of postediting of the MT.

MT is connected with human intelligence but this one is connected with artificial intelligence. Many people consider intelligence connected uniquely with the human quality. Michael Levy in his book "Context and Conceptualization in Language Learning" spoke about Artificial Intelligence (AI).

Computer programs for learning employ AI techniques – intelligent tutoring systems (ITS) or intelligent computer assisted instruction (ICAI). Though the process of transferring the teacher's pedagogical knowledge into a program can be complex, the essential structure and content of the program is prepared in advance in the way of a vocabulary. ITS can provide adaptive, individualized instruction.

Although the ideal goal of MT systems may be to produce high-quality translation, in practice the output is usually revised (post-edited).

The translation quality of MT systems may be improved either by developing more sophisticated methods or by imposing certain restrictions on the input. The system may be designed, f.e., to deal with texts limited to the sublanguage (vocabulary and grammar) of a particular subject field (f.e. economy) and/or document type (patents). Alternatively, input texts may be written in a controlled language, which restricts the range of vocabulary and avoids homonymy and polysemy and complex sentence structures. A third option is to require input texts to be pre-edited with indicators of prefixes, suffixes, phrase and clause boundaries, or of different grammatical categories (homonymy of the parts of speech). Finally, the system itself may refer problems of ambiguity and selection of equivalents to human operators (translators) in an interactive mode [Charmak E.]. In other words, SL analysis is oriented specifically to one particular TL. Typically, systems consist of a large bilingual dictionary and a single monolithic program for analyzing and generating texts; such "direct" translation systems are necessarily bilingual and unidirectional.

In the overall system design, there have been three basic types. The first type is generally referred to as the 'direct translation' approach, which is designed for one particular pair of languages, f.e. Russian as the SL and English as the TL.

The second basic design strategy is the interlingua approach which assumes that it is possible to convert SL texts into representations common to more than one language. Translation

is made in two stages: from the SL to the interlingua (IL) and from the IL to the TL. Procedures for SL analysis are intended to be SL-specific and not oriented to any particular TL; likewise programs for TL synthesis are TL-specific and not designed for input from particular SLs. A common argument for the interlingua approach is economy of effort in a multilingual environment.

The third basic strategy is the less ambitious transfer approach. It generates the final target texts. Transfer sustem consists of three types of dictionaries (SL dictionary containing detailed morphological, grammatical and semantic information; similar TL dictionary and a bilingual dictionary relating base SL forms and base TL forms) and various grammars (for SL analysis, TL synthesis and for transfer of SL structures into TL forms).

The direct translation approach was typical of the "first generation" of MT systems. The indirect approach of interlingua and transfer-based systems is characterizing the "second generation" of MT system. Both are essentially on the specification of rules. Corpus-based methods have changed the traditional picture.

During the last years there emerged a "third generation" of hybrid systems combining the rule-based approaches of the earlier types and the more recent corpus-based methods. These recent developments underline that MT research and MT systems adopt a variety of methodologies in order to solve the full range of language phenomena, complexities of terminology and structure, misspellings, 'ungrammatical' sentences, neologisms, etc. The development of an operational MT system is necessarily a long-term 'engineering' task applying techniques which are we-known, reliable and well tested.

The translation offered by the computer may not be perfect or even acceptable. The program makes according to the American "Business Week" "occasional errors", they cannot conceal their mistakes. In an English text in the European Commission, the Italian delegate "appealed to the Commission … to achieve rapid harmonization". "Appealed" was machine-translated into French as 'a fait plasir' instead of "urged". That was the version of the systran system.

Human assistance may be needed before, during or after a text is translated by computer, or not at ll. There are therefore 4 modes of machine translation>

with restricted input,

with interactive editing,

with post-editing,

or raw.

For some applications different modes are combined. If raw MT is used, of course, it must be labelled clearly: warning: machine translation.

MT, like human translation, is very varied.

In 1989 Veronica Lawson wrote: "…humans do the actual translating" […, p. 192], usually human assistance is required. The job of postediting (revising the computer's draft translations) is not the only new career avenue. Translators can also build up the MT dictionaries and identify the improvements needed in its software – both highly skilled and satisfying occupations to which translators are uniquely suited.

In conclusion we can say that it is not necessary to build up a word-for-word translation dictionary but a dictionary of word combinations because their translations include grammar and lexical cohesion of the words which is very important in MT. our dictionary includes 2500 such terminological word combinations in the domain of shipbuilding.

Bibliography:

1. Ahmad, K., Corbett, G., Rogers, M., and Sussex, R. (1985). Computers, language learning and language teaching. Cambridge, Great Britain: Cambridge University Press.
2. Charniak, E. and McDermont, D. An Introduction to Artificial Intelligence. Massachusetts, 1985.
3. Hutchins, W.Y. Machine Translation: Past, Present, Future. N.Y., 1986.

4. Lawson. Practical Machine Translation: An Introduction. In: Fremdsprachen, 1988, v. 32, N 1-4, ISSN0016-0970.
5. Levy M. CALL. Context and Conceptualization in Language Learning. Clarendon. Paperbacks.