The Development Prospect of English Translation Software Based on Artificial Intelligence Technology

Zhenhua Wei*
Xi’an Siyuan University, No.28 Shui’an Road, Xi’an, 710038, P.R. China

*Corresponding author e-mail: james@xasyu.info

Abstract. At present, information technology has developed by leaps and bounds. In particular, some advanced information technologies such as artificial intelligence technology and computer technology have been widely used in all aspects of social production and life. How to effectively use artificial intelligence technology to solve problems in translation is the ultimate purpose of this paper. This paper first introduces the development of computer translation. Secondly, the paper mainly introduces the basic principle and system composition of the case based machine translation. On the basis of the above research, a simple english-chinese machine translation system is explored to realize the help of computer-aided translation to domestic teaching. The basic process is to extract the key words from the English sentences or sentence fragments input by the user or the system, and search the sentences containing the key words by the key words, so as to obtain the possible translation of examples. Finally, the realization, experimental results and evaluation of the translation system are introduced. The development prospect of English translation software is analyzed from the experimental results.

Keywords: Machine Translation, Bilingual Alignment Tree, Statement Similarity, Artificial Intelligence Technology, Software Prospect

1. Introduction

Computer-aided translation, or CAT, has been around for more than a decade. When using computer-aided translation, translators use Trados, MemoQ and other translation software to add various dictionaries in the software, and in the translation process, translators constantly add translation results to the translation software to build corpus to achieve the purpose of auxiliary translation. When processing the same or similar translated text, translators can extract the translation results from the database to improve the translation efficiency and quality. The CAT translation model is widely used by many translation teams. It is not difficult to find that computer-aided translation is different from artificial intelligence translation in that it takes human translation as the leading role and USES computer data as the auxiliary tool. The computer data here is the corpus built by translators in their own translation work. The core technology of computer-aided translation is the extraction of translation memory, and the quality of the translation is controlled by the translator.

The quality and quantity of articles in each encyclopedia language vary greatly. Translating from
another wikipedia is a natural way to add more content, but the translation process is not properly supported in the software used by wikipedia. Computer-aided translation tools built for wikipedia in the past are not commonly used. Laxstrom creates a tool that ADAPTS to the specific needs of the open community and wikipedia content. Qualitative and quantitative data show that new tools help users translate articles more easily and faster [1]. EBMPracticeNet is a national electronic medical point information platform in Belgium. The project included, among other tasks, the translation of 940 EBM guides from English to Dutch and French from the ten fingers medical publications. In view of the scale of the translation process, it was decided to adopt computer-assisted translation, with certified translators with limited expertise in medical translation. Our consortium USES a hybrid approach that includes a human translation supported by translation memory (using SDL Trados Studio), term recognition from the medical terminology database (using the SDL multi-term terminology database), and support from online machine translation. This produces a proven translation memory that is now used to translate new and updated guidelines. The purpose of Stijn Van De Velde is to evaluate the performance of a hybrid manual and computer-assisted translation approach versus one that does not support translation memory and term recognition. The translation efficiency was compared with that of medical translation experts [2]. This paper discusses how to make better use of lexical and statistical translation evidence for cross-language information retrieval (CLIR). Word translation evidence comes from wikipedia and large machine-readable dictionaries, statistical translation evidence comes from parallel corpus, and co-occurrence evidence in document language provides the basis for limiting the adverse effects of translation ambiguity. Coverage statistics from the NII test bed and information access research community (NTCIR) queries confirm the complementary advantages of these resources. Experiments on translation evidence from small parallel corpora have shown that even fairly rough estimates of translation probabilities can yield greater improvements than a powerful technique of translation weighting based on jansen-shannon divergence as a measure of relevance to terms [3].

Since the great development of information technology and artificial intelligence technology, there has been a worry in the society that "whether machine translation will one day replace human translation", which obviously puts "machine translation" and "human translation" in the opposite position. From the perspective of the development of machine translation and computer-aided translation, the author of this paper will discuss the relationship between machine translation and human translation by means of empirical analysis.

2. Proposed Method

2.1 Artificial Intelligence

Artificial intelligence is an interdisciplinary subject. Although it sounds like a course, it is actually a comprehensive subject including psychology, philosophy, neurophysiology and computer technology. Although it contains the word intelligence, it has no flesh and blood. It is mainly an artificial intelligence that simulates a series of complex activities related to human intelligence such as perception, learning, reasoning and communication through computers. It differs from human natural intelligence in that it produces organic activities such as metabolism and reproduction that can be largely eliminated. Its creation is like the realization of human intelligence on a machine. A major goal of ai research is to enable machines to perform complex tasks that would normally require human intelligence. The development and use of intelligent machines can simulate and extend the functions of human organs, replacing many simple, repetitive and even complex tasks that are difficult to do in daily life, resulting in a huge change in the demand for workers [4].

1) Characteristics
2) Deep learning

The primary technical feature of ai should be deep machine learning. By using the hierarchical deep learning architecture, we can calculate the semantic expression of hierarchical objects at different knowledge levels through deep learning, so that we can effectively help further analyze and solve
more complex and abstract artificial intelligence problems and make their complexity more precise. At present, deep machine learning has made a series of significant advances in the application of speech, image and natural language. In the application fields of image and natural language processing, it is widely used in speech machine translation and semantic data mining. It has also made many achievements in the field of speech and image processing [5].

2) Cross-border integration

The effective fusion of artificial intelligence and various fields can achieve the effect of "1+1>2". In the field of education, artificial intelligence assistants and real-time monitoring devices can liberate teachers from a large number of repetitive work, better carry out humanistic care for students and teach students according to their aptitude, and help students to develop education and become better themselves. These are the manifestations of cross-border integration [6].

3) Man-machine collaboration

In the era of intelligence, people and intelligent machines are required to work together to complete the dull, repetitive work done by intelligent machines and innovative work and humanistic care work is up to us. In the future, as AI advances, the circle in the lower left corner will expand to the right. Therefore, in the era of artificial intelligence, the machine is just our assistant and serves us. In our future life, we will deal with the machine, which is the era of human-machine collaborative interaction [7, 8].

2.2 English Translation Software

The so-called computer translation, is the use of computer assisted manual translation, to create a human-computer interaction of the translation process. CAT, in its broad sense, includes the integration of various computer operating systems and application software, including the input of paper text, editing, typesetting, storage, printing, backup, transmission, network resources, and the use of relevant electronic software, etc., which can be summarized as the electronization of human translation. In the narrow sense, machine-aided translation is an application software system that utilizes the principle of translation memory in computer software, which is also the technical core of machine-aided translation, supplemented by other auxiliary translation means, such as electronic dictionary, term management, alignment tools, quality check, word count statistics, text analysis, project management, etc. [9]. In actual translation activities, sentences in different documents are more or less duplicated. By reusing these translation resources, the goal of avoiding duplication of effort can be achieved. To take advantage of these resources, the machine will first need to existing corpus segmentation idiom, alignment, and then to translation unit that is bilingual sentence plus other form of information stored in the database, the system will automatically by the sentence similarity algorithm to find whether there is with the current in the library for the same or similar sentences, if any, are given the corresponding translation as a reference; If not, the translator can also cross-query to find the translation of the words in the sentence. The management of translation project is also one of the contents of CAT technology translation. For the translation with heavy workload and tight time, it requires several translators to divide the work and assist. Inconsistent terminology is bound to cause difficulty and confusion. For example, "specification", "specification", "specification", etc., when the project manager determines a unified term, there will be no confusion in English corresponding to several translations [10].

3. Experiments

3.1 Experimental Background

At present, the main way of teaching computer-aided translation in China is to offer computer-aided translation courses in traditional translation courses and add computer-aided translation technology modules. The number of colleges and universities that offer this course is not large. Based on the analysis of relevant data, the following reasons can be obtained: first, due to the limitation of the economic strength of the university, many colleges and universities are unable to set up the computer
aided translation laboratory and purchase the hardware and software equipment that can respond to the demand. Second, the computer-aided translation course has strict requirements on the computer knowledge of teachers, and many schools do not have such teachers. In analysis, therefore, if a computer aided translation technology can get open in ordinary colleges and universities in China must change ideas, really reduce the cost of computer aided translation and related technical requirements, make of computer aided translation is neither dependent on expensive lab conditions, also don't have to have the professional knowledge of computers. We can add the computer aided translation technology module to the undergraduate translation teaching instead of setting up special computer aided translation courses.

3.2 Experimental Design

(1) Basic operations: including creating a translated project, merging and splitting the original text, exporting bilingual reference files, saving project files, etc.

(2) Terms: including the extraction of translation terms, the selection of translation terms, the translation of translation terms, the production of translation terminology library, the management of translation terms, etc.

(3) Translation memory: including translation text recognition, bilingual alignment, revision of the translation memory bank, retrieval of the translation memory bank, etc.

(4) Revision of translation: including revision of MT translation, revision of terms, revision of student translation, etc.

(5) Project management: including the overall process of the translation project, the overall cooperation of the translation project, the monitoring of the translation quality, the summary of the translation project, etc.

After 2 weeks of experimental data collection, the application of computational translation assistance software in teaching translation is shown in Table 1.

|          | Completion time | Complete the quality | Student acceptance | Degree of problem excavation |
|----------|-----------------|----------------------|--------------------|-----------------------------|
| The first translation | Eight days | 60                   | 20                 | 60                          |
| The first computer translation | Seven days | 80                   | 80                 | 80                          |
| The second translation | Eight days | 70                   | 20                 | 40                          |
| The second computer translation | Four days | 90                   | 80                 | 80                          |

4. Discussion

4.1 Application Analysis of English Translation Software Based on Artificial Intelligence Technology

According to the application of computer aided translation technology module in the above experimental data, computer aided translation technology and translation teaching are an organic whole. According to test students in contact with CAT aided translation translation intention after investigation, from the completion time, quality, students’ acceptance, translation problems mining degree to compare the four dimensions, CAT technology is introduced into translation teaching can
shorten the teaching time and completion time, improve the teaching efficiency and the efficiency of the translation by each grade students mutual and self-reported questionnaire, the result is shown in Figure 1, the results prove that improve the quality of student learning, students generally think that acceptance is higher than the traditional teaching method, students can more directly reflects the translation problems in the translation. The introduction of CAT technology into translation teaching can shorten the teaching time, improve the teaching efficiency and improve the learning quality of students. The degree of acceptance is higher than that of traditional teaching methods, and it can also more directly reflect the translation problems of students in translation.

![Score results](image)

**Figure 1.** Students' ratings of computer-aided translation and manual translation

The introduction of the "translation proofreading" module in the computer-aided translation technology improves the efficiency of students in the mutual evaluation of the translation, and at the same time fully embodies the perfect combination of translation technology and translation teaching. For another example, the second step of the comparative translation teaching method is for students to review each other's translations. After the introduction of the "translation proofreading" module of CAT, the computer-aided technology improves the efficiency of mutual evaluation of translations, which is a perfect combination of the comparative translation teaching method and translation technology. In the introduction of computer aided translation technology, the project management module is very suitable for the translation of specific projects and project management. The combination of corpus technology and computer-aided translation technology has significantly improved the efficiency of translation teaching. It can be seen that English translation software based on artificial intelligence technology well meets the needs of teaching resources and is more popular among teachers and students because of its fast and accurate characteristics.

5. Conclusions
In the era of big data, artificial intelligence is developing rapidly, black technologies are emerging in an endless series, and it is also turning to real product output, such as travel translation treasure and translation pen, which are launched to meet the needs of business scenes and the growing number of outbound trips. The emergence of artificial intelligence translation has drawn wide attention. Based on the development of English translation software under artificial intelligence technology, this paper has verified the high efficiency and high accuracy of computer assisted translation technology through experiments. Through experiments, it can be concluded that the English translation software combined
with artificial intelligence has greatly met people's needs for translation at the present stage and has a broad development prospect.

References
[1] Laxström, Niklas, Giner, Pau, Thottingal, Santhosh. Content Translation: Computer-assisted translation tool for Wikipedia articles[J]. Computer Science, 2015, 2008(2):75-79.
[2] Velde S V D, Macken L, Vanneste K, et al. Technology for Large-Scale Translation of Clinical Practice Guidelines: A Pilot Study of the Performance of a Hybrid Human and Computer-Assisted Approach[J]. Jmir Medical Informatics, 2015, 3(4):e33.
[3] Kim S, Ko Y, Oard D W. Combining lexical and statistical translation evidence for cross-language information retrieval[J]. Journal of the Association for Information Science and Technology, 2015,32(2):66.
[4] Jang, Chung-Hyok, Kim, Kwang-Hyok. The Improvement of Negative Sentences Translation in English-to-Korean Machine Translation[J]. Computer Science, 2015,21(1):655.
[5] Pust, Michael, Hermjakob, Ulf, Knight, Kevin. Using Syntax-Based Machine Translation to Parse English into Abstract Meaning Representation[J]. Computer Science, 2015,13(4):482-489.
[6] Qi Ruihua, Huo Yuehong, Guo Xu. Authorship Identification in English Translations of Chinese Classics[J]. New Technology of Library & Information Service, 2015,26(11):73-77.
[7] Robert Krajewski, Henryk Rybinski, Marek Kozlowski. A novel method for dictionary translation[J]. Journal of Intelligent Information Systems, 2016, 47(3):491-514.
[8] Man Ho Ivy Wong, Helen Zhao, Brian MacWhinney. A Cognitive Linguistics Application for Second Language Pedagogy: The English Preposition Tutor[J]. Language Learning, 2018, 68(2):92.
[9] Luong, Minh-Thang, Le, Quoc V, Sutskever, Ilya. Multi-task Sequence to Sequence Learning[J]. Computer Science, 2015,4(2):511-514.
[10] Zhang, Jiajun, Liu, Shujie, Li, Mu. Beyond Word-based Language Model in Statistical Machine Translation[J]. Computer Science, 2015,31(2):145.