# Automatic Machine Translation System with Artificial Intelligence Theory

## Lanqin Wang

Heilongjiang Polytechnic, Harbin, Heilongjiang, China, 150000

**Abstract:** At present, with the continuous development of computer technology, automatic machine translation system has been adopted by many universities and research institutes, and it has entered the stage of practical and commercialization. In recent years, with the continuous development of artificial intelligence technology, a lot of research has been conducted on it. Based on this current situation, this paper conducts a research and analysis on the automatic machine translation system of artificial intelligence theory.

Keywords: Artificial Intelligence, Automatic Machine Translation, Translation Systems

## 1. Introduction

Machine translation is a new language system supported by big data and with computer as the core, which can continuously optimize its internal structure, and thus improve its quality while ensuring accuracy and real-time. A translation is only a refinement rather than a complete language service. It can be used to handle both the initial translation of a language service, and to deal with aspects such as content management and project management <sup>[1]</sup>.

## 2. Research Process of Artificial Intelligence and Machine Translation

Generally speaking, computer translation is machine translation and computer-assisted translation. Machine translation reflects by computer, while computer-assisted translation focuses more on the language conversion that people can achieve by computer. However, so far there is no machine translation system that can completely replace the work of human translators. This is due to the fact that in addition to its own institutional characteristics, language also has a cultural characteristic. With the continuous development of arti"Icia' Intelligence technology, machine translation has become better able to solve most general translation problems. However, it is still difficult for machine translation to replace human beings when facing professional, diverse and detailed communication problems that contain human emotions. In the current development of artificial intelligence, there is a development trend of mutual as mirror, mutual as embedding and mutual as information [2].

The development of translation skills is also a constantly changing and interacting process. Translation is not a linear process dominated purely by progressive input, and its machine translation often encounters significant challenges in understanding and processing metaphors between languages. The perception of metaphors also varies greatly across language systems. This requires artificial intelligence with high-speed computing ability to quickly identify, store as well as compare these metaphorical information, and to be able to perform high-speed data processing on these massive linguistic information, establish a grammatical metaphorical pattern set and a lexical metaphorical pattern set, and calculate the amount of metaphorical information by traditional or instance, and also by inference and statistics. Through this method, translators can effectively identify and understand the metaphors between various languages and improve the translation quality of translators.

## 3. The Promotion of Artificial Intelligence for Automatic Machine Translation System

In practical applications, machine translation faces many complex problems, especially the complexity of language knowledge and cultural diversity. To solve these difficulties, we must resort to modern technology. At the technical level, since machine translation involves two or more languages, it is necessary to have the ability of linguistic parsing of the source language first. The syntactic-based

and semantic-based intermediate structures, which play a great role in the process of understanding and analyzing the source language. Linguistic analysis has a significant effect on machine translation, and at the same time, the level of linguistic analysis will be directly related to the quality of machine translation. In terms of language parsing, automatic machine translation with artificial Intelligence undoubtedly has the absolute upper hand. There are usually significant differences between different languages, especially in grammatical rules and expressions, and different linguistic symbol systems show asymmetry in linguistic information, such as vocabulary, semantics and context, which poses many challenges for automatic machine translation work, it also requires the use of continuously developing artificial intelligence technology, as a way to deal with some special problems that arise in practical applications. To a certain extent, a translation machine with artificial intelligence, which is a continuous cognition of the operation rules between different language systems, achieves fast and accurate decoding and encoding through deep recognition of multiple contextual information, such as texts, paragraphs and sentence groups. In addition, language collocation is a linguistic phenomenon that manifests itself in a certain pattern in vocabulary or fixed phrases. In a natural language processing such as machine translation, word collocation plays a very critical role. The goodness of machine translation is closely related to the quality and usage of the corpus. A large and effective matching resource can greatly improve the efficiency of autonomous language understanding. The development of translation skills is also a constantly changing and interacting process. Translation is not a linear process dominated purely by progressive input, and its machine translation often encounters great challenges in understanding and processing interlingual metaphors. The perception of metaphors also varies greatly across language systems. This requires artificial intelligence to have high-speed computing ability, which can quickly identify and compare these metaphorical information, and it can perform high-speed data processing of these massive linguistic information, or use inference and statistical methods to perform the calculation of metaphorical information data. Through this method, metaphors between various languages can be effectively identified and understood, and the efficiency of translation can be improved<sup>[3]</sup>.

The development of translation skills Is also a constantly changing and interacting process. Translation is not a linear process dominated purely by progressive input, and its machine translation often encounters significant challenges in understanding and processing metaphors between languages. The perception of metaphors also varies greatly across language systems. This requires artificial intelligence with high-speed computing ability to quickly identify, store as well as compare these metaphorical information, and to be able to perform high-speed data processing on these massive linguistic information, establish a grammatical metaphorical pattern set and a lexical metaphorical pattern set, and calculate the amount of metaphorical information by traditional or instance, and also by inference and statistics. Through this method, translators can effectively identify and understand the metaphors between various languages and improve the translation quality of translators.

#### 4. Machine Translation Based on Artificial Intelligence Technology

With the continuous development of artificial intelligence technology, the ability of machine translation is also improving, and this trend is well reflected in the field of professional translation. In essence, computer network information is a huge data and information base, which contains a large number of sub-databases that provide a large amount of data materials. Some translation corpus, due to its wide range of fields involved and its high technical and professional nature, coupled with the large amount of information it contains, often results in a large number of repetitions in the translated text.

From the relevant statistics, some technical documents that require machine translation can have a reproduction rate of more than 40% of the relevant content. To solve such problems, AI can automatically translate the presented corpus on the basis of translation memory. In order to improve the efficiency of machine translation and ensure the consistency of the translated content, AI often builds a translation memory and word management components when using machine translation. The so-called "translation memory" is an equivalent corpus corresponding to the original text and the translated content established by AI in machine translation.

With the development of artificial intelligence technology, the performance of machine translation system has been greatly improved, and it can automatically set the minimum matching degree between the original text and the translated text according to the actual situation. After that, the fuzzy matching procedure is used to retrieve the corpus in the translation memory, and even for the sentence types that cannot complete a perfect match, AI is able to use machine translation to complete similar content confirmation. Currently, this way of working not only ensures high-level translation, but also enables

machine translation to improve the quality and efficiency of translation with the help of artificial intelligence.

In the future, having stronger learning capabilities and more data processing power will help improve the efficiency of machine translation. The translation memory that is already available can also be used to automatically identify and store new translations using the self-learning mechanism of AI. As the scale of translation memory becomes larger and larger, AI can greatly improve the efficiency of machine translation with the support of big data technology. On this basis, using network sharing, AI can make accurate algorithmic analysis of a large amount of network translation data and improve its self-learning ability, making it close to the level of human in terms of quality.

#### 5. System Hardware Design

#### 5.1. Client Architecture Design

The system architecture allows users to upload photos to their computers, thus facilitating access to the information to be translated, saving time in user input and increasing its efficiency and accuracy. Users can use their cell phones to take pictures, capture the text to be translated, and upload the converted content to the cloud server as a photo. Users can also select a pre-taken text image from the local gallery and upload the photo to the cloud server, and finally set the focus and flash of the camera in the setting options of the translation system. The settings of the language that needs to be recognized according to the user's requirements. The selected photos are uploaded to the cloud server, it can convert through HTTP protocol, and then through the OCR system installed on the cloud server, it can convert the images into text, and then edit them. The translation of the source language text is completed using Google Translate technology, and the source language text content is converted into the target language text content is fed back to the client. Users can edit the received source language text content and target language text content appropriately, or they can go online to search for text content of interest <sup>[4]</sup>.

#### 5.2. Artificial Intelligence Processor Design

Among the many business units, AI processing is a processor with high computational power, which is also the bottleneck that limits the operational efficiency of the whole business unit. This requires multiple AI processors to process different users simultaneously, while the number of AI processors is based on the number of user requests, and the number of AI processors determines how fast their translation speed is. Artificial intelligence techniques are applied to digital images containing textual information to be recognized. After pre-processing the digital image, the text information to be recognized is extracted by using algorithms for locating, segmenting and extracting text information, and then pattern recognition algorithms are used. Based on this, the standard code of the extracted text was determined by analyzing the morphological features of the extracted text, and then the data processing was performed. The main function of the processor is built on Tesseract- OCR2.3, which is a laboratory-developed artificial intelligence engine, which has been improved continuously by Google, and it has developed into an open source engine with high accuracy rate in AI field, and it can support Chinese language with command line mode. Based on the analysis of the customer's needs, the developer has designed the business process of the customer and the architecture of the system on this basis. Using artificial intelligence technology, the hardware design was carried out <sup>[5]</sup>.

#### 6. System Software Design

#### 6.1. Calculate sentence similarity

The sentence similarity algorithm firstly performs a coarse selection of similar sentences based on lexical attributes, and then performs a refinement filter to finally obtain the similarity between identical utterances. Although this algorithm fully takes into account the frequency properties of each word, which assigns different weights to each word, the lack of word adhesion makes the similarity between words with longer or lower word frequencies appear more different, which affects the fast translation of the system. In the calculation of sentence similarity, the inverted index file is first used to obtain the number of sentences to be derived, and then the content of the sentences to be calculated is determined based on the number of sentences. Following the process of utterance similarity calculation, the n

selected utterances are subjected to similarity analysis, and they are input to the utterance similarity analysis module. The utterance similarity analysis can improve the standard of utterance translation <sup>[6]</sup>.

#### 6.2. Sentence Ambiguity Removal

The elimination of ambiguity in syntax can effectively solve the problem, and thus achieve the purpose of automatic machine translation. On the one hand, because of the ambiguity caused by word classes, the same word can have different word classes, thus causing ambiguity in the translation. On the other hand, because the same word in different situations can have different meanings. To eliminate the ambiguity of sentences caused by word classes, we must first distinguish the word classes of words clearly, use the calculated similarity to mark the word classes, and use the marked word classes as the basis to judge the true meaning of the word in the sentence, thus eliminating the ambiguity and finally realizing the translation of the whole sentence. In order to solve the problem of sentence ambiguity due to contextual differences, this requires the search and search of the statement to be translated, treating the word as a specific word meaning item and defining it as a word with a specific meaning, so as to eliminate the ambiguity of the word, and thus realizing automatic machine translation. On this basis, the control of the user side Is realized through the design of the architecture and artificial Intelligence processor of the user side.

# 7. Development Trend of Automatic Machine Translation System with Artificial Intelligence Theory

In the context of economic globalization and informationization, people's demand for translation is growing, which provides new opportunities and new challenges for the development of international language service industry. In the process of dealing with multilingualism, how to quickly and effectively discover the intrinsic operation mechanism among multiple languages, which is a great challenge faced by automatic machine translation at present. Artificial intelligence can help machine translation to efficiently uncover the inner operation relationship of language systems, so as to achieve the purpose of Improving the efficiency of machine translation. It does so by mining a wide range of data, including raw data of information in multiple languages such as sentence clusters, chapters and genres, and then encoding or decoding this information under various contextual conditions set up. With the rapid development of the Internet and artificial intelligence, the application of information technology in the field of translation is becoming more and more widespread. The statistical model of reconstructed word alignment technique contains a large number of high-quality scalable semantic units, which can provide strong support for the further development of automatic machine translation. If this technique can be applied to machine translation, its response speed will be greatly improved. In the context of big data, how to collect and use the huge amount of Internet data is an urgent problem. The next step will be to make automatic machine translation more in line with the language habits and characteristics of users through big data technology<sup>[7]</sup>.

### 8. Conclusion

In summary, although the quality of machine automatic translation system translations is far inferior to that of human translations, they are now commonly used in many fields, and machine automatic translation systems are a way to supplement and correct human translations. In some cases, it can enhance the translator's work efficiency and improve the accuracy of the translation, so it has become a great helper tool for translators.

## References

[1] Li Rong, Zhou Meili. Design of automatic machine translation system based on artificial intelligence processor design[J].Modern Electronics Technology,2022,45(02):183-186.
[2] Guo Mingyang, Zhang Xiaoling, Tang Huiling,Meng Qingduan,Ren Longbo. Research on the Application of Artificial Intelligence in Machine Translation[J]. Journal of Henan University of Science and Technology (Natural Science Edition),2021,42(03):97-104+8.
[3] Ma Wanzhong,Liu Junhua,Zhu Xiaojie. Standardized research on the evaluation of artificial intelligence machine translation ability level[J]. Information Technology and Standardization, 2020(Z1):21-26.

[4] Feng Zhiwei. Parallel Development of Machine Translation and Artificial Intelligence[J]. Foreign Language(Journal of Shanghai Foreign Language University), 2018, 41(06): 35-48.

[5] Wei Z. The Development Prospect of English Translation Software Based on Artificial Intelligence Technology[J]. Journal of Physics: Conference Series, 2020(3):1533.

[6] Zheng S, Zhu S. A Study of Computer Aided Translation Based on Artificial Intelligence Technology[J]. Journal of Physics: Conference Series, 2020(1): 1646.

[7] Jiang K, Lu X. The Influence of Speech Translation Technology on Interpreter's Career Prospects in the Era of Artificial Intelligence[J]. Journal of Physics: Conference Series, 2021(4):1802.