Algorithmic Translation Correction Mechanisms: An End-to-end Algorithmic Implementation of English-Chinese Machine Translation

Lei Shi[1,*]

[1] School of Foreign Studies, Shandong University of Finance and Economics, Jinan 250014, Shandong, China

Abstract

INTRODUCTION: Machine translation is a modern natural language processing research field with important scientific and practical significance. In practice, the variation of languages, the limitation of semantic knowledge, and the lack of parallel language resources limit the development of machine translation.

OBJECTIVES: This paper aims to avoid duplicating neural networks during the learning process and improve the ability to generalize complex neural network machine translation models with limited resources.

METHODS: Textual material in the source language was studied, and a suitable textual material representation model was used to express complex, high-level, and abstract semantic information. Then, a more efficient neural network machine translation integration model was developed based on the control of written data and algorithms.

RESULTS: Data mining must be applied to complex neural network machine translation systems based on transfer learning to standardize finite neural network models.

CONCLUSION: Neural network-based embedded machine translation systems based on migration training require a small number of labelled samples to improve the system's permeability. However, this adaptive migration learning region approach can easily lead to over-learning problems in neural network machine translation models, thus avoiding excessive correspondences during the learning process and improving the generalization ability of the translation model with limited neural network resources.

*Corresponding Author. Email: shilei@sdufe.edu.cn

1. Introduction

People use Natural language in their daily lives, and it is an essential tool for learning and living. Knowledge recorded and transmitted through language accounts for more than 80% of all knowledge and contributes to the development of human civilization. Natural language processing refers to using computers to process human linguistic information, including language manipulation and processing. Machine translation is a particular application of natural language processing[1-3]. People use machines to translate one language into another, a particular information exchange between humans and machines. Realizing the accelerated integration of the

global economy, the growth of international trade, and the rapid growth of online knowledge in areas such as cross-border e-commerce, tourism, foreign trade, cross-border finance, social media, and national security and intelligence services, the military and other units are challenged to process multilingual data in their business processes rapidly. Machine translation technologies can accurately and efficiently translate natural language and play an essential role in facilitating political, economic, and cultural exchanges.

Due to the high cost and slow speed of traditional human translation, machine translation came into being. Machine translation technology has developed with information technology, linguistics, and information technology[4-5]. It is the crown jewel of natural language processing and a significant breakthrough and milestone in artificial intelligence. Compared with human translation, automatic translation can quickly process linguistic data. Studies have shown that professional and experienced translators need help to meet the growing demand for translations, as each translator translates about 2,000 words in 8 hours. However, machine translation systems can translate at a rate and speed thousands of times faster than humans. In practice, machine translation can reduce delivery times and significantly improve efficiency. They improved translation quality. Translation quality in the translation industry is very high. Some more minor languages and dialects need the right people. In machine translation, the quality of the translation meets the basic tasks and compensates for the strengths and weaknesses of the translation. They reduced transfer costs. For some translators, the difference between the cost of interpreting and machine translation is insignificant. If the number of interpreters increases, the cost of interpreting will be much higher than that of machine translation. Developing small language skills with specialized skills requires much time and labor. Therefore, the development of machine translation technology is expected to fulfil people's communication and content delivery needs, whether in government, the military, business, or general users, and is no longer limited to language skills. However, language formation has gradually evolved into a long-term archive of people, regions, and cultures containing many complex semantic data points[6-7]. Compared to machine translation, human translators are more likely to integrate context, understand semantic information, and select the most appropriate words for translation. The capabilities of translation systems also include the expression of bilingual knowledge, cultural knowledge, and physiological and psychological factors. Machine translation has yet to understand semantic knowledge, and computers must have knowledge and understanding.

The development of machine translation technology can be divided into three stages: rule-based approach. Turing was one of the earliest founders of natural language thinking. In the 1950s and 1970s, researchers believed that natural language processing was a cognitive process of language, so natural language processing mainly used a rule-based approach. The rule-based approach was implemented by programmers and linguists manually writing translation rules, but it had apparent drawbacks[8]. First, it is only possible to create rules for some proposals. Only the correct translation can satisfy one's sentence rules because the previous sentence rules are untranslatable. Second, handwritten translation rules are costly, limited in scope, and constantly changing because developers need to master not only computers but also languages. If they did not meet the requirements of the rules, they could not be used.

The first generation of conventional machine translation technology is far from being able to meet the practical needs of a rapidly developing society. Statistical machine translation techniques the researcher revealed the statistical framework of English through Markovian thinking and showed that these modelling models of the framework (by analyzing the relative probabilities of letter and word combinations) could generate actual language based on statistical rules. Later, he showed through a series of text tests that he could estimate the probability of a letter or word appearing in each text. Moreover, the results became more apparent as statistical models became more complex. With the rapid growth of the Internet, the abundance of libraries, and the rapid improvement of hardware performance, statistical methods are becoming more and more dominant. Statistical machine translation is achieved by searching for sentences that may change from the source to the target language. However, statistical machine translation requires step-by-step processing of text groups, grammar, syntactic analysis, semantic analysis, and human educational activities. The processing result of each module affects the performance of the final machine translation system. With the development of deep learning technology, deep learning has gradually been adopted to train multilevel neural networks to perform specified tasks[9-10]. Progress has been made in processing natural language, such as machine translation, question-and-answer systems, reading, and comprehension. One of the key technologies is to support statistical machine translation frameworks. Still, deep learning techniques have improved some intermediate modules, such as translation models, language models, sorting models, and other crucial components. Secondly, statistical machine translation is removed as a base (without basic preprocessing, such as basic lexical and manual design skills), an integrated machine translation system based on the Oxford Framework 2013 neural network.

2. Background of the study

Machine translation is the application of artificial intelligence to machine translation. Machine translation automatically translates computer input source code into text in the target language. Since the 21st century, the popularization of the Internet, the significant increase in international travel, and the development of the national "One Belt, One Road" strategy have led to an increase in

the exchange of information between multilingual populations, which has increased the demand for high-quality translation services. However, interpreting is expensive, and very few interpreters are available to provide translation services[11]. Fortunately, many researchers are involved in machine translation technology, and machine translation quality has improved significantly over the past two decades. The goal of machine translation is "faithfulness, expressiveness, and elegance". Loyalty, clarity, and elegance accordingly mean loyalty to the translated content, fluency in words means expression, and elegance in literature means elegance. Traditional statistical machine translation methods have achieved the goal of "validation". Nowadays, the fast-developing neural machine translation techniques pursue "fidelity". These methods are far from "elegant" and require continuous machine translation. The main goal of this paper is to improve the quality of machine translation to a practical level by proposing new translation methods and approaches.

Since the introduction of machine translation by American mathematicians in 1949, machine translation has evolved since the 1970s. The continuous reduction of hardware costs and the use of graphic cards have created conditions for improving machine translation efficiency. Combined with these conditions, machine translation theory and techniques have developed rapidly. The machine translation process uses rule-based, pragmatic, statistical, and deep neural network methods. Current rule-based machine translation methods require the manual organization of several translation rules[12]. However, due to the limitations of manual rule correction, rule-based machine translation tends to be of poor quality. It plays a limited role in certain areas that fulfil the minimum requirements of users of machine translation systems. Unlike machine translation based on manual translation rules, the example-based approach does not require much manual organization of translation rules. The sample library searches for sentences similar to the sentence simulates the corresponding translation of the sentence, and ultimately creates the translation. The disadvantage of this approach is that the translation system requires limited text coverage and many texts. A statistical model for automatic translation based on the translation instance model is proposed. The model is based on statistical data from a large bilingual parallel dataset that searches for translations in various research areas, significantly improving translation efficiency. Unlike the previous three machine translation methods, the deep neural network-based machine translation method continuously models the machine translation problem and improves the translation quality. Machine translation neural models are commonly used in machine translation and online translation evaluation systems[13-15]. Simulation of translation processes for statistical machine translation and neural machine translation has become one of the core topics in natural language

processing and artificial intelligence. Machine translation systems are frequently used for online and machine translation[16]. Machine translation technology has attracted the attention of the scientific community and industry because of its high efficiency, short development cycle, and high reliability.

The basic idea of machine translation is to assign a certain probability to each translation and ultimately choose the most likely translation. Translation can be summarized as a formal mathematical model: sentences are in the source language X, and there are many ways to translate them into the target language Y. Port X, call the p(Y-X) translate: Y method to port the input set X. The specific results are shown below:

$$y* = \arg\max pr(y \mid x) \tag{1}$$

However, translation between different languages is so complex that it is unrealistic to list all sentences in the target language in the final translation: y-hy*. In addition, solution P is a problem that needs to be solved for machine translation.

3. Research methodology

3.1 Neural Network Translation

Significant progress has been made in recent years since the introduction of complete neural machine translation models. The quality of large-scale translation tasks will be significantly improved compared to statistical machine translation methods. The main idea of complex neural translation is to perform mechanical translation between natural languages through deep neural networks. Thus, machine translation neural models typically use encoder and decoder frameworks to convert sequences from the source to the target language[17-19]. When translating a sentence from the source language, the NMT model first creates a vector representation for each word in the source language and then creates a vector representation of the entire sentence from left to right through a recurrent neural network called an encoder. Meanwhile, the target language uses a second neural network loop to extract the first sentence into the final translation of the actual vector. The original NMT model faced a severe problem: regardless of the sentence length in the original language, it was encoded as a standard vector. The proposed due diligence mechanism will effectively solve this problem. The basic principle of the attention mechanism is that the final translation process of the NMT decoding model not only uses the original language to create fixed sentence vectors but also considers the current hidden state of each term in the original language. During this operation, the decoder dynamically searches for words related to the source language and adds contextual information to the decoder function. Thus, the attention mechanism changes the data transmission method by dynamically computing the source language context, which is crucial for decoding

the current lexicon, effectively solving the transmission problem for remote data processing, and significantly improving NMT translation[20]. Therefore, attention-based conference code modelling has become a significant approach to machine translation and has been widely used.

The model differs from the original machine translation neural system based on repetitive neural networks and uses only the attention mechanism to build the whole machine translation model. The model achieved good translation results and significantly improved the parallelism and decoding speed of translation models based on neural network learning. Despite the great success of NMT, machine translation researchers have high expectations for the attention mechanism and have proposed for the first time a sentence-based NMT model and a neural machine translation model for the attention-deficit mechanism, both of which have achieved good translation results. On the decoder side, the neural machine translation model uses a sleep-wake network to simulate sentences focused on the target language. Inaccurate translation, the order of words and the target language varies depending on the sentence structure. The method introduces the native sequence level into the source language sentences to fulfil the requirement that the text sequence translated by the sleep-wake network match the source language sentence sequence.

3.2 Correction Algorithm

During actual translation, there are often rotational and shear deviations between the link head's actual position and the hinge joint's theoretical position. This can lead to severe problems, such as damage to the connecting joints and work platforms, hinge deviation, and hinge deviation. So, after breaking down the connection paths of the nodes, it is necessary to compare the actual positions of the connection nodes with the positions of the theoretical connection nodes, calculate the differences between them based on the actual connections, and fix the connection paths based on the differences in the position matrix. The connection wall repair method is based on local loop characteristics. The loop data of the connector nodes is used as calibration data for the next iteration of the algorithmic neighbor (ICP), which is significantly more effective when combined with the loop properties of the connector nodes.

Since the function-based calibration method takes a long time to separate the characteristic points from the characteristics of the two-point clouds, the properties of the two-point clouds are not very similar, and the calibration could be more effective. The ICP point cloud calibration algorithm has good reliability. It does not need to extract the characteristic points, which is widely used in point cloud reconstruction and calibration. The algorithm data trajectory is shown in Figure 1.
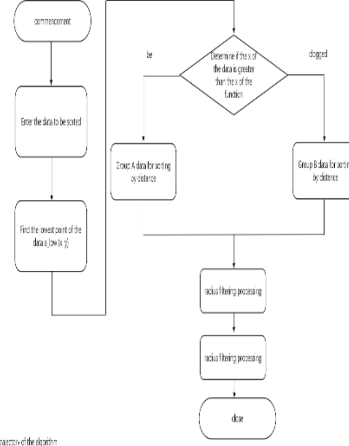
Figure 1: Algorithmic data traces

The basic principle of the ICP algorithm is to find the following corresponding points (Pi, Qi) in the two-point clouds of Q and P. The rotation matrix R is computed, and the matrix T is transferred to the two-point cloud to minimize the error of the function E. The ICP algorithm is based on the following principles:

$$E(R,T) = \frac{1}{n}\sum_{i=1}^{n} \| q_i - (R_{pi}+t) \|^2 \tag{2}$$

The n in the equation is the number of connection points; the absolute value and the squared difference are taken better to explore the sum of errors of the function.

To further calculate the distance from Pi to Qi, the following equation is used:

$$d = \frac{1}{n}\sum_{i=1}^{n} \| p'_i - q_i \|^2 \tag{3}$$

If D is below the specified threshold or the number of iterations exceeds the specified number, return the optimal shear and rotation matrix at the end of the iteration. Otherwise, return to step 2 and continue the iteration. Numerous studies and experiments have shown that the accuracy of the ICP algorithm record depends on the initial positions of the two-point clouds, is susceptible to noise compatibility, and can lead to optimal solutions to localized problems. This paper separates the point cloud contours from the calibration data to eliminate the noise, simplify the point cloud data, and improve the calibration accuracy. To find a match between the corresponding point cloud and the target cloud, the KD tree searches for the nearest neighboring matches, which speeds up the search for matches. The total time required to retrieve a profile is reduced by simply retrieving data points from the local profile and pasting them into the complete profile of the target cloud. This approach helps maintain registry accuracy and speed.

Many researchers have studied and improved the ICP calibration algorithm, a traditional calibration algorithm, to increase its accuracy and speed. The improvement method focuses on improving the objective function of the ICP writing algorithm. The original ICP recording algorithm mainly stores the stored and target data points.

It receives the sum of the squared differences of the corresponding points, which requires a lot of calculations. Therefore, some researchers have proposed a new method to improve the lens operation, which calculates the distance from a point to a specific cloud storage point and from a point to the logger. However, experimental results show that the above improvement is specific to calibration. The calibration accuracy of the SAC-IA algorithm depends on how well the two-point cloud attribute histogram (FPFH) is extracted, which is better if there are more similar points in the histogram of the separation function. Before determining the SAC-IA, the histograms of the two point clouds need to be computed, and then the rotation matrix and shear matrix are computed to obtain the thickness of the two point clouds. The registration process for the SAC-IA is as follows:

(1) To distinguish between n FPFH data points, it is necessary to compute the distance between n data points randomly. The second operation can only be performed if the distance between them is large enough to exceed the defined minimum distance limit, D. Otherwise, the random selection will take a finite amount of time.

(2) N data points with similar FPFH properties are found in the cloud of object P. If there are multiple data points with similar properties, one of the similar points is randomly selected as the corresponding point. The point cloud W computes the rotations and intersections between the corresponding points of the rotation and shear matrices and transforms the rotations and intersections of the N data points into W to compute the values of the Huber penalty function between them:

$$f = \sum_{i=1}^{n} H(l_i) \tag{4}$$

H functions in Eq. (3) as follows:

$$H(l_i) = \begin{cases} \dfrac{1}{2} l_i^2, \| l_i \| < m_i \\ \dfrac{1}{2} m_i (2 \| l_i \| - m_i), \| l_i \| \ge m_i \end{cases} \tag{5}$$

The H function is for better calculation of the distance difference; the f function discriminates the booking value. The finite iteration uses a sequence of transformations starting with the value of the minimum error function as the final output. The SAC-IA algorithm is calibrated to be well-positioned between the two clouds and is accurately calibrated using the ICP algorithm to improve the calibration efficiency. However, the combination of these two algorithms requires the computation of the FPFH characteristics of the two-point clouds, which takes a lot of time and affects the application of the algorithm. The flow of the algorithm, as shown in Figure 2.
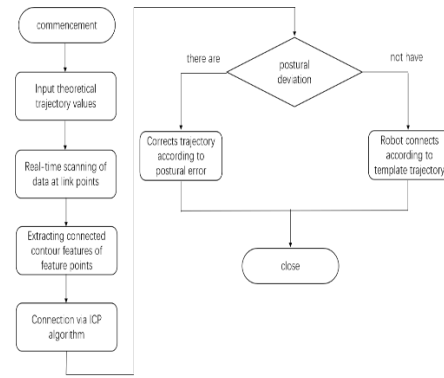
Figure. 2: Flowchart of the Algorithm

Based on the characteristics of the local loop, the connected path correction algorithm is mainly used to pull local points from the connected node loop function as calibration data. This helps to simplify the data, get rid of noise points, and improve the accuracy of writing. Meanwhile, the real-time local loop recording theory is utilized to decompose the complete loop of the combined node, which significantly shortens the extraction time and improves the calibration speed. When selecting ICP hit points, the KD tree search method speeds up the recovery of hit points.

4. Results and discussion

4.1 Results

Human translation, machine translation, and post-editing are three standard stages in the field of translation; each has its characteristics, advantages, and disadvantages in various situations.

4.1.1 Human translation and machine translation

A statistical analysis of the output evaluation of human and machine translation reveals both advantages and disadvantages. The advantages of human translation are:

(1) More accurate: human translators can understand the context and semantics and accurately translate the source language into the target language.

(2) More natural: Human translators can express culture and emotions, making the translation more natural and fluent.

(3) More flexible: human translators are more flexible than machine translators in terms of the quality of translation. (4) More flexible: human translators can judge and process the text according to the actual situation to ensure the quality of the translation.

The advantages of machine translation are:

(1) Faster: machine translation can complete a large number of translation tasks in a short time and improve translation efficiency.

(2) Cheaper: Machine translation only requires a one-time investment cost.

(3) Wider field: Machine translation can be adapted to a broader range of translation fields, such as medicine,

technology, etc. However, machine translation has the following drawbacks: Due to the limitations of computer algorithms and corpus quality, the quality of translation could be more stable, resulting in consistent content.

4.1.2 Machine Translation and Post-Editing (MTPE)

Machine translation inevitably has problems, such as unstable translation quality, incorrect grammar, easy confusion, and contextual errors. Therefore, post-editing has become an essential step in machine translation. Post-editing can improve the accuracy and readability of machine translation and translate results more in line with readers' expectations. However, it should be noted that post-editing is relatively time-consuming and requires specialized personnel, which may increase translation costs and time.

4.1.3 Human translation and post-editing

The combination of human translation and post-editing is necessary to ensure the accuracy and fluency of the translated text. Regarding work distribution, some projects may require full-time translators. At the same time, others may require translators to perform post-editing functions to ensure the quality of the final translation. During the translation process, if there are questions or uncertainties, the translator may discuss them with the post-editor to determine the best translation solution. This iterative approach to revision and discussion ensures accuracy and consistency in the final translation result.

Machine translation, human translation, and post-editing are the three main approaches in the field of translation, and their relationship can be described as complementary. First of all, machine translation and human translation are complementary. Human translation can accurately convey the original text with high translation quality. Therefore, in translation tasks, machine translation can be used for preliminary translation, and then the translator will proofread and touch up the result to improve the translation quality. Secondly, post-editing is a standard complement to machine translation and human translation. Whether machine or human, the results may inevitably contain errors and inaccuracies. Therefore, post-editors need to proofread, revise, and embellish the translated text to improve the translation's quality and readability. Finally, despite the rapid development of machine translation and post-editing, human translation remains the most critical translation method. Machine translation and pre-editing can improve translation efficiency and quality, but the accuracy and naturalness of the translated content still require the specialized knowledge and skills of human translators. A comparison of the development and testing of algorithmic translation is shown in Figure 3.
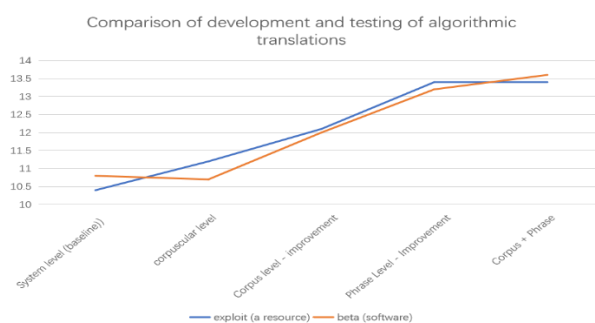

Figure. 3: Comparison of the development and testing of algorithmic translation of

4.2 Empirical analysis

From the above analysis in Chapter 4, there are three main problems in machine translation in addition to formatting errors, and corresponding solutions are proposed based on these problems.

4.2.1 Vocabulary Problems and Solutions

Lexical mistranslation includes problems such as incorrect application of words and misinterpretation of words. For example, English and Chinese are similar in that they both have multiple meanings for words. However, English is a morphological language, while Chinese is a meaningful language that needs more formalization in vocabulary, leading to the problem of lexical mistranslation in the machine translation process.

Machine translation is primarily word-for-word based on the order of the original words, and machine translation needs the logical reasoning ability to analyze and reorganize the original sentence based on intellectual and cultural context, making it challenging to convey the original sentence.

's meaning adequately.

The scientific and technical English vocabulary usually consists of specialized vocabulary, subspecialized or semispecialized vocabulary, and idiomatic vocabulary. The use of specialized vocabulary is limited to specific professions. Semi-specialized vocabulary is adapted from standard vocabulary and has various specialized meanings in different disciplines in addition to the original. In generalized vocabulary, scientific and technical English differs from ordinary English. Therefore, inconsistency in translating the same specialized vocabulary in a text is an important issue when translating specialized vocabulary.

In conclusion, although these problems are likely to be solved slowly, a specialized corpus can be designed to collect fixed word collocations and help analyze discourse modules to remove ambiguities. A text-specific terminology base can be created to maintain terminological consistency and harmonize translation styles. Secondly, since a corpus is a database of information that can improve the speed of the system and the accuracy of interpretation, a larger corpus can be developed to analyze sentence features, extract raw

sentence structure, and feature information, and use it to solve some problems related to phrase grammar and fixed collocations. The corpus should be organized to match the restriction domains, and the raw sentences should be automatically segmented to distinguish between each word in the sentence's lexical and contextual meaning.

4.2.2 Grammatical problems and solutions

Long sentences and passive voice are frequently used, and nominalized and non-finite verbs are widely used. According to the classification of sentence functions, narrative is frequently used in English for electrical engineering, which directly elaborates relevant theoretical knowledge. Various utterance types reflect the logic and conceptual clarity of electrical engineering English. In technical English translation, the appropriate language should be chosen according to the degree of formality of the text. According to stylistic formality, language can be categorized into five types: serious style, formal style, deliberate style, informal style, and intimate style. Even though this classification isn't apparent and needs clear criteria, it has been set up from bottom to top, from informal to formal style reference indexes. Authors carefully look at changes in the environment and the level of formality of the terms in the style of technical English to choose the correct language for translation.

In post-translation editing, the translator needs to adjust the machine translation order according to the expression habits of the target language to form a text structure and sentence expression that are more in line with the reading habits of the target readers. Therefore, there are two solutions. First, a large corpus is collected through data-capturing software such as Python; a corpus is built and used as a model training knowledge base; and the machine translation software is trained to translate long and complex sentences by adjusting the structure and form of utterances to make the English translation of scientific and technological texts closer to human translation. The English-Chinese translation is compared with the Chinese-English translation, as shown in Figure 4.
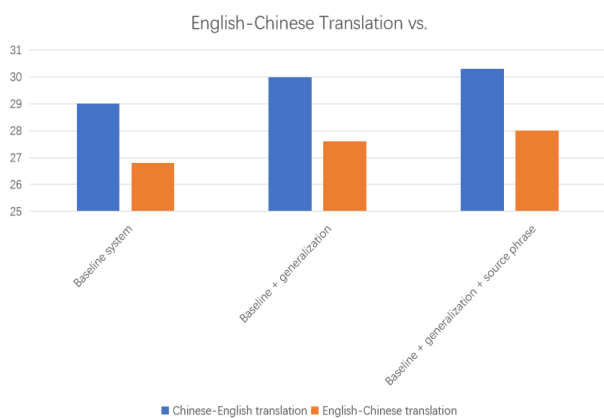

Figure. 4: English-Chinese Translation vs. Chinese-English Translation

4.2.3 Discourse problems and solutions

When dealing with translations of various discourse types, machine translation systems run into several issues, such as producing better translation results for medical and scientific texts with less semantic complexity, which causes a lack of integrity and coherence when dealing with texts made up of interrelated and constrained sentences; these issues require post-translation editing to fix.

As for post-translation editing, the Author believes that it can be used in combination with computer-aided translation (CAT). Desktop or cloud translation software such as MemoQ, MateCat, TradosStudio, Memsource, etc., provides integrated translation environments with machine translation, translation memory, and post-translation editing functions. For example, MemoQ supports machine-translated pre-translations and light or full post-translation editing of the text, depending on the quality of the translation.

By comparing the translation quality of four different machine translation software applications and manual translations, the following results were obtained:

First, comparing the translation quality of four machine translation programs (Ringo Cloud Translator, Yodo Translator, Sogou Translator, and Baidu Translator) shows that machine translation technology still has a specific error rate. Still, Yodo Translator is the best choice for significant texts. A comparison of the translation quality of the machine translation programs and the human translation shows that the machine translation is better than the human translation. Machine translation often requires editing and customization to optimize the results, requiring additional time and staff workload. Based on the above analysis, while machine translation technology advances, manual translation remains preferred when prioritizing translation quality, accuracy, and readability. Manual translation provides the quality, accuracy, and expressiveness of Joyoung, which improves readability and comprehension. Therefore, the quality of machine translation software is at most that of manual translation. However, when dealing with many translation tasks under time constraints, the machine translation post-editing (MTPE) method can be chosen to improve translation efficiency and quality.

This translation experiment report points out several limitations. As mentioned in Chapter 1, the selected background materials are two English texts of electrical equipment manuals, which are characterized by certain lexical, grammatical, and textual aspects and, therefore, cannot fully reflect the problems of machine translation in other types of texts. Regarding the evaluation metrics, the QA self-test option of X-Bench is the main limitation. In addition, due to text confidentiality, the experimental part of human translation includes only the Author's translation for comparison, and the lack of comparative data from different authors for the exact text in human-translated texts limits the evaluation of the quality of human translation. The lexical incremental analysis is shown in Figure 5.
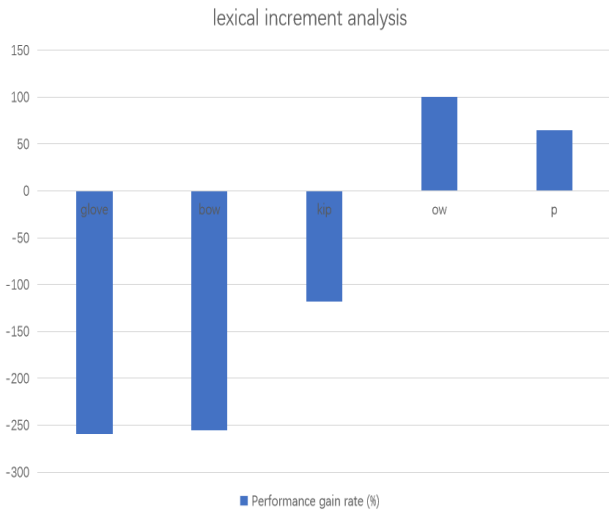
Figure. 5: Lexical Increment Analysis

4.3 Empirical Discussion

Neural Machine Translation (NMT) works well in resource-solved languages but not in resource-solved Chinese and English. How to fully utilize different language skills to fill the resource gap is challenging for Chinese and English NMT. Integrating multiple levels of language skills is crucial to improving the efficiency of machine translation in Chinese and English. This paper examines the lack of bilingual resources in Chinese and English, examines different levels of language proficiency, and finally designs and implements a machine translation system based on multilevel functional knowledge.

4.3.1 A parallel Chinese-English repository was created

Using crawler network technology, Chinese-English bilingual information can be found on the Internet by hijacking paths. Depending on the case, the source code of the XPath path is obtained, the data is removed from the final index, and the 144K Chinese-English bilingual data is organized. Parallel 2000 is randomly selected from a test series of Parallel 000 and a study series of 140K Chinese-English bilinguals.

4.3.2 The Chinese-English neuro-translation method is based on different perceptions of language characteristics

In this paper, the characteristics of the English language are analyzed and divided into four different entry levels. A series of knowledge particles of speech properties are formed by adding deep collapsible and detachable neural networks and moving them into the model at different levels. Local sequence data is explored to analyze the impact of different levels of functional knowledge on machine translation. It includes different levels of functional information in Chapter 3.

4.3.3 A Chinese-English neuron translation method based on multilevel functional knowledge is proposed

Three different Chinese characters, words, and sentences are proposed to be combined to address the problem of using different language levels in neural machine

translation to fill the resource gap. First, a bidirectional LSTM and an attention mechanism combine symbols in words with level semantic information and represent optional words based on the dynamic combination of symbol components. Second, a sentence tree encoder is built into the standard sequence sensor, and the sentence information is integrated into the sentence through the NMT sequence transformation process. Experiments show that this method effectively integrates attribute data at different levels and somewhat improves NMT performance.

In recent years, neural machine translation has been a research center and a future research direction. In the field of neural machine translation, resource-rich languages have been extensively studied with good results. However, neural machine translation is relatively limited and resource-poor, and many studies still need to be in the early stages. However, with the trend of global connectivity, machine translation plays a vital role in international trade, and the UK has a very close relationship with South China. The performance comparison of different algorithms in machine translation is shown in Figure 6.
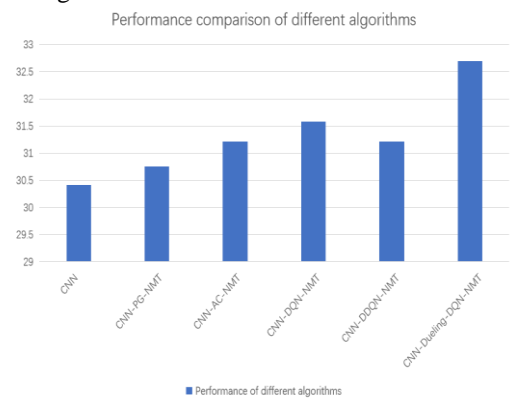
Comparison of different algorithms for quota performance in machine translation

Figure. 6: Performance comparison of different algorithms in machine translation

5. Conclusion

In this paper, the Southern language's grammatical variations and different grammatical structures, words, and sentences are considered the multilevel representation of sequences of linguistic symbols representing different linguistic features in Chinese-English NMT. This is based on understanding the linguistic features of different language levels. This improves the performance of NMT to some extent. Second, more research needs to be done: the collection of Chinese-English bilingual libraries is relatively small, while the collection of Chinese-English parallel libraries is small. Therefore, the next step is to study how to increase the number of Chinese-English bilingual groups while ensuring the quality of the groups. The Chinese NMT method selects four levels of separation of linguistic auxiliaries by analyzing ESL language features based on understanding different

language features. Based on the validation mechanism, a depth separation coil is used to improve the neural machine levelling model. Depth can tell the difference between the entangled sequences that are used to entangle particles with different sizes, study the local sequence data and pull out more attribute data, reduce the number of parameters needed for entanglement operations, speed up the theoretical computation of the model, and make the model easier to see. However, the scale of the studied linguistic features is relatively small because English grammatical structures have many morphological variations and different grammatical structures that entirely use the grammatical knowledge of the sentence. In Chinese-English Neural Machine Translation (NMT-RRB), the knowledge of linguistic attributes is combined on several levels. First, the semantic attributes on the surface of the words are put together, and then, based on the sequential encoder, a sentence tree encoder is made. The sentence information contained in the sentence is integrated into the NMT sequence transformation process. However, the method has drawbacks because the input text must be analyzed using a sentence structure tree. Therefore, errors in the sentence structure analysis will inevitably affect the model's overall performance. In the future, the development of a method for integrating different language skills into the NMT model will continue to be explored.

References

[1]   Shafei B. The Treatment of Some Deviation Cases in Non-Standard Data Using Artificial Intelligence and Natural Language Processing Applied to Arabic Language. 6. 2022;5(4):12–9.

[2]   Sauerland MBU, Sauerland U, Uli Sauerland Uli SauerlandLeibniz-Centre General Linguistics ZAS G by US Schützenstrasse, Berlin, and, Gould MBKA, Gould KA, et al. Risks and Benefits of Large Language Models for the Environment. Environmental Science And Technology. 2023;57(9):3464–6.

[3]   Liu Y, Rui X, Li Y. Long-term development archive of the Yellow River since the Neogene in the central Jinshaan Gorge, China. Palaeogeography, Palaeoclimatology, Palaeoecology: An International Journal for the Geo-Sciences. 2022;5(591-):591.

[4]   Allen HW, Paesani K. Genre instruction, textual borrowing, and foreign language writing: Graduate teaching assistant perspectives and practices: Language Teaching Research. 2022;26(4):755–76.

[5]   Liu C, Wang K, Wu A. Management and Monitoring of Multi-Behavior Recommendation Systems Using Graph Convolutional Neural Networks. International Journal of Foundations of Computer Science. 2023;8(27):120-124..

[6]   Zhang J, Wang F. A Better Medical Interpreting Service: Interpreter's Roles and Strategies Under Goffman's Participation Framework. International Journal of Translation Interpretation and Applied Linguistics. 2021;3(1):1–14.

[7]   Alhassan A, Sabtan YMN, Omar L. Using Parallel Corpora in the Translation Classroom: Moving towards a Corpus-driven Pedagogy for Omani Translation Major Students. Arab World English Journal. 2021;12(1):40–58.

[8]   Wang J, Bai L. Unveiling the Scoring Validity of Two Chinese Automated Writing Evaluation Systems: A Quantitative Study. International Journal of English Linguistics. 2021;2(2):56–60.

[9]   Qiu G, Su F, Gao L, Chen CC. Research on Translation Style in Machine Learning Based on Linguistic Quantitative Characteristics Perception. Sensors and materials: An International Journal on Sensor Technology. 2021;5(6 Pt.2):33.

[10] Bundgaard-Nielsen RL, O'Shannessy C. When more is more: The mixed language Light Warlpiri amalgamates source language phonologies to form a near-maximal inventory. Journal of Phonetics. 2021;85(4):101037.

[11] Xiang H Aijun Su. Fault detection of wind turbine based on SCADA data analysis using CNN and LSTM with an attention mechanism. Measurement. 2021;175(1):34–40.

[12] Novitskiy N, Maggu AR, Lai CM, Chan PHY, Wong P. Early Development of Neural Speech Encoding Depends on Age but not Native Language Status: Evidence from Lexical Tone. Neurobiology of Language. 2021;12(8):1–59.

[13] Gulcu S. Training of the feed-forward artificial neural networks using dragonfly algorithm. Applied Soft Computing. 2022;12(12):124.

[14] Zhang H, Liu Y, Yang L, Hou G, Shao B, Xia J, et al. Numerical modelling of porphyroblasts rotation in the brittle-viscous transition zone. Journal of Structural Geology. 2021;24(148-Jul.):1–13.

[15] Fantinuoli C, Prandi B. Towards evaluating simultaneous speech translation from a communicative perspective. 12. 2021;18(10):23–30.

[16] Fu Y, Nederhof MJ. Automatic Classification of Human Translation and Machine Translation: A Study from the Perspective of Lexical Diversity. 10. 2021;6(6):16–20.

[17] Kan W, Dechao L. Are translated Chinese Wuxia fiction and Western heroic literature similar? A stylometric analysis based on stylistic panoramas. Digital Scholarship in the Humanities. 2022;3(4):4.

[18] Fabunan AG. Machine Translation's Role: in Literary Translations. Multilingual. 2023;12(1):34–46.

[19] Wang Y, Daghigh AJ. Cognitive effort in human and machine translation post-editing processes: A holistic and phased view. FORUM. 2023;21(1):139–62.

[20] Zhang Y, Li L, Wu Y, Su Q, Sun X. Alleviating the Knowledge-Language Inconsistency: A Study for Deep Commonsense Knowledge. IEEE/ACM transactions on audio, speech, and language processing. 2022;20(30-):30.