Journal of Combinatorial Mathematics
and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Exploration of Language Translation Problems under the Background of Artificial Intelligence Technology

Jinming Cui[1], Lin He[2,✉]

[1] *School of Chinese studies/Institute of Humanities and Social Sciences Xi'an International Studies University, Xi'an, Shaanxi, 710061, China*

[2] *Faculty of Education, Shaanxi Normal University, Xi'an, Shaanxi, 710062, China*

ABSTRACT

With the deepening of globalized cooperation among countries around the world, the human society's demand for machine translation has increased rapidly, and the advancement of artificial intelligence technology has also put forward new requirements for the quality of machine translation. In this paper, we propose a neural machine translation model with adversarial training algorithm, using the Transformer model as the baseline model, by incorporating prior knowledge in order to strengthen the correlation between the intermediate hidden layers. In addition, besides introducing lexical knowledge into the neural machine translation approach, a neural machine translation model PhraseNet with phrase memory capable of storing bilingual phrases in the form of symbols is also incorporated. Finally, the effectiveness of the language translation model is tested using the NIST corpus. The results show that compared with the baseline model, the BLEU values of this paper's method in the NIST 2010, 2012, and 2016 test sets are improved by 1.81, 2.63, and 2.48, respectively, which significantly outperforms the baseline translation system. It shows that the method is able to bring a large improvement for both language translation accuracy and training execution cycle. Meanwhile, it also has a good guiding significance for the research and development of language translation in the context of the whole artificial intelligence technology.

Keywords: machine translation, Transformer model, prior knowledge, PhraseNet

1. Introduction

The emergence of artificial intelligence language services highlights the characteristics of language information in the context of globalization and big data and the advent of the era of intelligent technology platform for translation [1-2], and the future is about to be the era of the "Belt and Road" of specialized language services and the surge in demand for intelligent translation [3]. At present, the

✉ Corresponding author.
E-mail address: helinxisu444@163.com (L. He).
Received 20 February 2024; Revised 25 April 2024; Accepted 20 November 2024; Published Online 16 April 2025.
DOI: 10.61091/jcmcc127b-219

development trend and status quo of the domestic language service industry are: imbalance between supply and demand in the language service talent market, imperfection of the language service talent training system, insufficient use of translation technology and software [4-5], in-depth cooperation between language service enterprises and universities in innovation and collaboration, and the need for new technologies and practices in the development of the network translation platform and the application of research [6-7].

At present, AI technology is being widely and deeply used in various fields, and combining AI technology with translation services is the trend of the times [8]. In this context, language service enterprises should make full use of AI technology as an assistant to collect and mine data information, understand customer needs, actively carry out technological innovation, improve efficiency, optimize the customer relationship management system, create tailor-made accurate and convenient services, provide mutually beneficial solutions, and achieve the healthy and sustainable development of the enterprise [9-11]. The application of artificial intelligence translation technology plays an increasingly important role in the language service industry. The traditional simple and mechanical human translation mode can no longer adapt to the rapidly developing market demand, and will be gradually replaced by machine translation, speech recognition technology and synchronized multilingual subtitle translation technology, neural network machine translation technology in the future [12-14].

The level of AI translation is progressing rapidly, and the speed of its progress may greatly exceed people's expectations, and AI translation allows people to free up more energy and time to engage in activities that are creative and meaningful [15-17]. The current level of AI translation still has a lot of gaps from people's expectations, there are problems such as mechanical translation, stiffness, poor overall matching, etc. The machine can only replace the part of the translation in which the syntax and structure are relatively simple, and people do not have high requirements for the translation results [18-19]. However, it is believed that in the near future, AI translation will be completely comparable to human translation. In terms of market application prospects, the AI translation market has great potential [20].

In this paper, we propose a neural machine translation method based on fusing prior knowledge. The method encodes the syntactic tree in a recursive form bottom-up at the encoder side to obtain the vector representation. During the decoding process, joint learning aligns the words and phrases at the source end to realize the generation of target words. On this basis, a tree encoder is appended to the conventional sequence encoder. And the phrase memory is scanned by PhraseNet to identify candidate phrase pairs to optimize the recurrent neural network word-by-word generation. Finally, the effectiveness of the model is verified by comparing the BLEU values and perplexity of different models in the NIST dataset.

2. Artificial intelligence-based machine language translation technology

2.1. Neural Machine Language Translation

Throughout the development trend of machine translation, it is essentially reducing the dominant role of human in the translation process. Rule-based methods rely entirely on human experts to compile translation rules, and statistical-based methods require human design of features and hidden structures to automatically learn translation knowledge from data. Deep learning-based approaches, on the other hand, can directly utilize end-to-end ideas to describe the complete translation process. In fact, it is a process that requires ever-increasing computer modeling capabilities.

2.1.1. Encoder-Decoder Framework. Unlike general deep learning tasks, natural language processing tasks are generally sequence learning. The mainstream neural machine translation model is "Sequence to Sequence" learning (Seq2Seq), which is the most mainstream "encoder-decoder" model in machine translation. The machine model consists of two neural networks, the first of which

is called the encoder, which encodes the source language sequence to be translated and outputs it as a fixed-length vector representation. The second neural network, called decoder, encodes the vector representation after encoding by the encoder and decodes it into the target language sequence by the decoder. Therefore, this model is also called "encoder-decoder" model.

The output sequence of the encoder is called the "context", and C is the vector representation of the context, which is a vector summarizing the sequence of inputs $X=\left(x_1,x_2,\ldots...,x_{T_x}\right)$ at the encoder side, both as outputs of the encoder and as inputs of the decoder. The current output y_t of the decoder is in the form of a probabilistic output, which is a prediction based on the output $\left(y_1,y_2,\ldots..,y_{t-1}\right)$ of the previous time step and the vector representation C of the context, i.e.:

$$P(Y)=\prod_{t=1}^{T_y}P\left(y_t\mid\left\{y_1,y_2,\ldots...y_{t-1}\right\},C\right) \tag{1}$$

$$P\left(y_t\mid\left\{y_1,y_2,\ldots..y_{t-1}\right\},C\right)=g\left(y_{t-1},h_t,C\right) \tag{2}$$

where $Y=\left(y_1,y_2,\ldots..y_{T_y}\right)$, T_y are the lengths of the output sequence Y.

The "encoder-decoder" structure is currently used as the general framework for neural machine translation, and almost all machine translation research work has been carried out on this basis.

2.1.2. Self-attention mechanisms. The attention function can be described as mapping a query and a set of key-value pairs to an output, where query, key, value and output are all vectors. The output is computed as a weighted sum of values, where the weight assigned to each value is computed by the compatibility function of the query with the corresponding key.

Dot product attention with scaling and multiple attention are shown in Fig. 1, the left figure demonstrates the attention method for dot product similarity calculation, K, V correspond to key-value respectively. The inner product is solved with each element in Q and then the softmax method is executed to get the similarity between the elements in Q and V. The new vector is obtained after weighted summation to get the final dot product attention result. The figure on the right then shows the method for computing the multi-head attention result, where the three states are linearly transformed and input to the scaled dot product. One point to note, however, is that the process needs to be performed h time, and its allows the model to jointly attend to different subspace information from different locations.

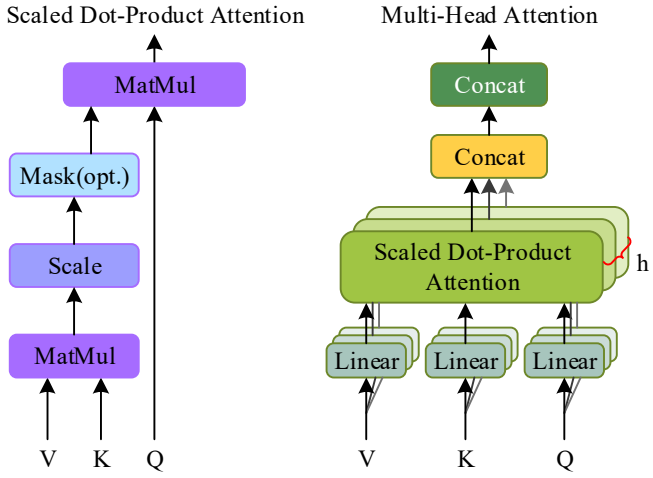


Fig. 1. Point-of-focus with attention and multi-head attention

Figure 2 shows the main working process of the decoder part of the machine translation system,

which mainly adopts the sequential sequential generation of words, and generates translated words according to the currently obtained state information in each time step. And the information of the current state is passed to the next time step to participate in the generation process of the next word. After executing the above method sequentially for many times, the final translated sentence is generated.

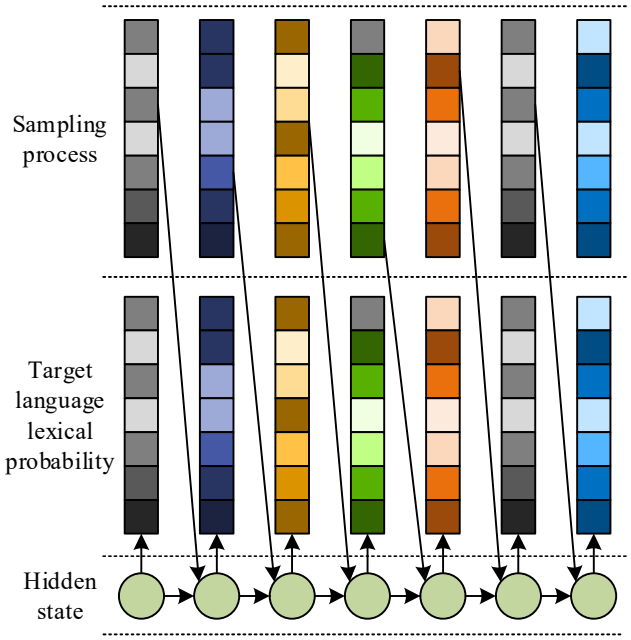


Fig. 2. Decoder structure of machine translation system based on neural network

2.2. Neural machine translation with the introduction of prior knowledge

In natural language processing, there are two main forms of syntactic analysis, phrase structure analysis and dependency analysis. Regardless of which analysis method, sentences are regarded as a recursive tree structure, so the analysis results will correspondingly generate a syntax tree containing all syntactic information. Therefore, in order to overcome the problem that the end-to-end neural machine translation approach oversimplifies the bilingual conversion process and lacks knowledge support, this paper incorporates a priori knowledge into the neural network translation model.

2.2.1. Introduction of syntactic information. The key steps in the method of introducing syntactic information are firstly phrase structure syntactic analysis of the source language sentences. Secondly, the syntactic tree is encoded at the encoder side in a recursive form from the bottom up to obtain the vector representation. Finally, the decoding process joint learning aligns the words and phrases at the source end and generates the target words one by one. Fig. 3 shows the schematic diagram of the tree-to-sequence based attentional neural machine translation model.

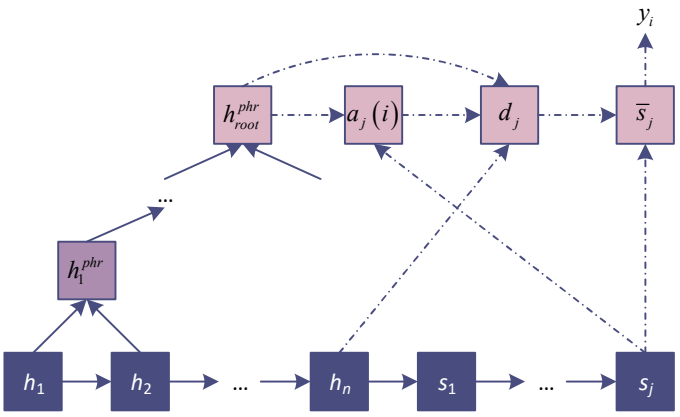


Fig. 3. The attention neural machine translation based on tree to sequence

In a core word-driven phrase structure grammar (HPSG), sentences usually consist of several phrases and their data structure can be abstracted as a binary tree as a whole. Based on this, a tree encoder is appended to the traditional sequence encoder. In this case, the hidden state h_k^{phr} of the parent node of the k st phrase can be calculated by the hidden states h_k^l and h_k^r of its left and right children nodes based on the following equation:

$$h_k^{phr} = f_{tree}\left(h_k^l, h_k^r\right) \tag{3}$$

Where f_{tree} is a nonlinear activation function. At this point, the parameters of the hidden state h_k^{phr} and the memory cell c_k^{phr} of the k nd parent node are updated as follows:

$$i_k = \sigma\left(U_l^{(i)} k_k^l + U_r^{(i)} h_k^r + b^{(i)}\right) \tag{4}$$

$$f_k^l = \sigma\left(U_l^{(fl)} k_k^l + U_r^{(fl)} h_k^r + b^{(fl)}\right) \tag{5}$$

$$f_k^r = \sigma\left(U_l^{(fr)} k_k^l + U_r^{(fr)} h_k^r + b^{(fr)}\right) \tag{6}$$

$$o_k = \sigma\left(U_l^{(o)} k_k^l + U_r^{(o)} h_k^r + b^{(o)}\right) \tag{7}$$

$$\overline{c}_k = \tanh\left(U_l^{\overline{c}} h_k^l + U_r^{\overline{c}} h_k^r + b^{(\overline{c})}\right) \tag{8}$$

$$c_k^{phr} = i_k \square\ \overline{c}_k + f_k^l \square\ c_k^l + f_k^r \square\ c_k^r \tag{9}$$

$$h_k^{phr} = o_k \square\ \tanh\left(c_k^{phr}\right) \tag{10}$$

Where $i_k, f_k^l, f_k^r, o_k, \overline{c}_k$ is the input gate, the left and right child nodes of the forget gate, the output gate and the memory cell of the LSTM, respectively. c_k^l and c_k^r are the left and right child nodes of the memory unit, while $U(\cdot)$ and $b(\cdot)$ are the weight matrix and bias vector, respectively. Two sentence vectors h_n and h_{root}^{phr} can be obtained at the end of the encoding process, which are generated by the sequence encoder and the tree-based encoder, respectively. In the decoding process, the decoder is first initialized using the hidden state s_1. The computation of s_1 is as follows:

$$s_1 = g_{tree}\left(h_n, h_{root}^{phr}\right) \tag{11}$$

where g_{tree} and f_{tree} are the activation functions.The attention model implemented by Eriguchi et al. is optimized accordingly by dynamically jointly attending to the sequence and phrase hidden states one at a time to reflect the weight of which words or phrases in the original text are contributing to the

currently generated target word. The j th context vector d_j is obtained by weighted summation of sequence and phrase vectors:

$$d_j = \sum_{i=0}^{n} \partial_j(i) h_i + \sum_{i=n+1}^{2n-1} \partial_j(i) h_i^{phr} \tag{12}$$

The formula for the current target hidden state s_j is as follows:

$$s_j = f_{dec}\left(y_{j-1}; s_{j-1}\right) \tag{13}$$

By performing both sequence coding and phrase coding, at the encoder side. And by jointly paying dynamic attention to sequences and phrases in an attention model, the performance level of machine translation methods can be improved.

2.2.2. Introduction of external knowledge. Although Neural Machine Translation has shown the best translation performance on both bilingual translation tasks such as English and French, the end-to-end model architecture still faces the challenge of difficulty in incorporating external knowledge such as dictionaries. In this paper, we propose a framework that can connect the neural machine translation model with the bilingual dictionary, and apply the bilingual dictionary, which contains information such as low-frequency words and unregistered words from the training data, to neural machine translation, thus significantly improving the translation performance.

Two models are used for the above framework:

(1) Hybrid word/character model, where words with frequencies above a certain threshold in the corpus are retained, while low-frequency words and out-of-set words are replaced using their character sequences.

(2) Pseudo-sentence pair synthesis modeling, i.e., constructing J sets of sentence pairs containing (D_{xi}, D_{yi}) for low-frequency word pairs or unlogged word pairs (D_{xi}, D_{yi}) to extend the training corpus.

In addition to introducing lexical knowledge into the neural machine translation approach, this paper proposes PhraseNet, a neural machine translation model with phrase memory.PhraseNet has the advantage of being able to store bilingual phrases in the form of symbols, which can be either mined from a corpus or specified by an expert.

The translation principle of phaseNet is illustrated in Fig. 4. For any given source language sentence, phaseNet scans the phrase memory to identify candidate phrase pairs and accordingly integrates the annotation information in the representation of the sentence. The decoder utilizes both lexical and phrase generation components in a specially designed binary classification strategy, thus integrating the generation of vector sequences of single words or multiple words at a time. PhraseNet is not only able to incorporate external knowledge into neural machine translation, but also extends the word-by-word generation approach of recurrent neural networks.

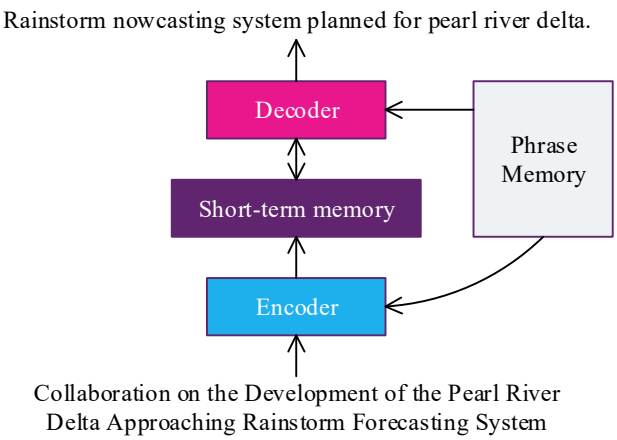


Fig. 4. PhraseNet translation principle

2.3. Evaluation Criteria for Machine Translation - BLEU

The evaluation of the effectiveness of machine translation can be done manually by judging the quality of the translation, but in order to save human resources, the quality of the translation is evaluated by comparing the degree of similarity between the output translation of the machine translation model and the translation given in the original parallel corpus.

The effectiveness of machine translation is usually evaluated with the Bilingual Evaluation Understudy (BLEU) score, which is one of the most widely used automatic evaluation metrics for assessing the difference between the target language generated by a trained model and the actual target language. Its basic idea is to consider that if the result of machine translation is close to the result of human translation, then the translation effect of the model is good. Theoretically, some candidate translations are selected from the results of machine translation and compared with the translations translated by human beings, and the closer they are, the better the translation effect of the model is proved to be. In the specific operation, the model is evaluated by counting the number of identical n -grams in the machine-translated and human-translated translations, where n is usually set to 5. Finally, the number of counted 1-grams is divided by the total number of n -grams in the machine-translated translation candidates to get the result of effect evaluation.

3. Analysis of the results of the language translation experiment

The hardware environment for the AI-based machine language translation experiment is Intel (R) Core (TM) i7-10200, 3.2GHz, 32GB RAM, 512SSD, Windows 10 Professional operating system, the model is trained and tested using Google's open-source AI deep learning framework Tensorflow, and the computing resource Four-card RTX 3060TL experimental data analysis software environment and test environment for Pycharm 2021 Professional Edition, the length of training is about 48 hours.

3.1. Experimental data set and parameterization

3.1.1. Experimental data set. The training corpus consists of three major sources, namely, the 2021 AI Challenger English-Chinese Bilingual Parallel Corpus, which includes 3M Chinese-English bilingual parallel corpus, 5M Chinese monolingual corpus, and 2M English monolingual corpus, where M denotes millions of pairs of sentences.The bilingual parallel corpus of the 2021 CWMT for training Chinese-English machine translation consists of 21M, with 8M Chinese monolingual corpus and 5M English monolingual corpus. The bilingual parallel corpus of CWMT 2021 is 5 M, the Chinese monolingual corpus and the English monolingual corpus are 1 M each, and these monolingual sentences do not have any relationship. In other words, a sentence in Chinese cannot find a corresponding translation in English. The translations corresponding to these monolingual data are used as reference translations for comparison with the generated translations.

The NIST 2008 dataset was selected for the validation set, and the NIST 2010, NIST 2012 dataset, and NIST 2016 dataset were selected for the test set. Among them, in each test set, it corresponds to 3 reference translations. For Chinese-English comparison, the first English sentence of the 3 reference translations is used as the source language sentence, and the Chinese sentence is used as an independent reference translation. In order to limit the size of the word list, the first 50,000 high-frequency words in the training corpus are used to construct the word list and the rest of the low-frequency words are labeled as <UNK>. In the training corpus, the dimension of the word vector is set to 256 and the length of the sentences is limited to 30.

3.1.2. Parameter setting. The Transformer model was used as the baseline model for translation training. In the experiment, the relevant important parameters of the Transformer model are set as

shown in Table 1. Where m denotes the number of heads in the Transformer model, specifically 10. d denotes the dimension size of scaling, specifically 128. $r_{dropout}$ denotes the $dropout$ proportion in the network, i.e., the proportion of randomly discarded neurons. α and β denote the scaling factor and translation factor in layer normalization, respectively, and λ denotes the learning rate of gradient update.

Table 1. The main parameters settings of Transformer model

Model	Chinese-English translation	English-Chinese translation
m	10	10
d	128	128
$r_{dropout}$	0.2	0.5
α	0.05	0.05
β	0.95	0.95
λ	0.0001	0.0005

In order to have a clearer comparative description of the experiment, the models and methods of different types of corpus resources are compared in the experiment. The three types of corpus resources are categorized into three types, which are all bilingual parallel corpus (Method 1), only monolingual corpus resources (Method 2), and a mixture of bilingual parallel corpus and monolingual corpus resources (Method 3). Specifically, the baseline methods for the corpus of Method 1 include ConvS2S and Transformer model, while the baseline methods for the corpus of Method 2 include "RNN + Reverse Translation" and "Unsupervised PBSMT". The corpus methods of Method 3 include the baseline method "Pairwise Machine Translation", and the neural machine translation method "Transformer + PK", which introduces prior knowledge in this paper. The effectiveness of this paper's method can be clearly verified through the comparison of machine translation methods based on three different types of corpus, as shown in Table 2. It can be concluded that in the two tasks of Chinese-to-English and English-to-Chinese translation, this paper's method (Transformer + PK) obtains BLEU value evaluation results above 30, which is relatively good, and these experimental results verify the effectiveness of this paper's method. That is, the performance of machine translation models can also be improved by utilizing a priori knowledge for monolingual corpus collocation. Compared with using only bilingual data, this paper's method reduces the amount of bilingual parallel data required and the workload of corpus collection, while at the same time obtaining significant results. The method in this paper ensures the quality of the generated corpus, which reduces the training burden of the model, and is highly cost-effective as compared to this unsupervised machine translation using monolingual data exclusively, with lower training cost and better performance than the baseline model.

Table 2. Models performance results on different test datasets in the experiment

Material type	Model	BLEU values on different data sets					
		Chinese-English translation			English-Chinese translation		
		NIST 10	NIST 12	NIST 16	NIST 10	NIST 12	NIST 16
Method 1	ConvS2S	24.35	23.82	24.65	29.08	28.92	29.81
	Transformer	27.41	25.95	27.71	32.55	33.84	34.11
Method 2	RNN +Reverse	26.30	25.89	26.17	29.76	31.40	33.29
	PBSMT	27.16	27.37	28.01	35.62	34.82	36.64
Method 3	Dual machine	27.48	27.98	28.67	31.85	36.45	37.27
	Transformer + PK	30.72	30.44	30.81	38.99	38.15	39.20

3.2. Language Translation Performance and Perplexity Analysis

3.2.1. Analysis of language translation performance results. The results of language translation performance after introducing lexical, named entity, shallow semantic, and dependent syntactic knowledge in the Chinese-English and English-Chinese translation tasks are shown in Table 3. The translation performance of the baseline system as well as the a priori knowledge-based neural machine translation system proposed in this section can be seen on the Chinese-English translation task on different test data. Where the bolded fonts are the highest BLEU values and * indicates significantly better results than the baseline system in the case of $p < 0.05$. In the Chinese-English and English-Chinese translation tasks, the maximum BLEU values obtained on the test dataset are 36.89 and 28.66, respectively.Therefore, embedding the additional four linguistic knowledge in the a priori knowledge-based language translation method proposed in this section can effectively improve the translation performance of the translation system.

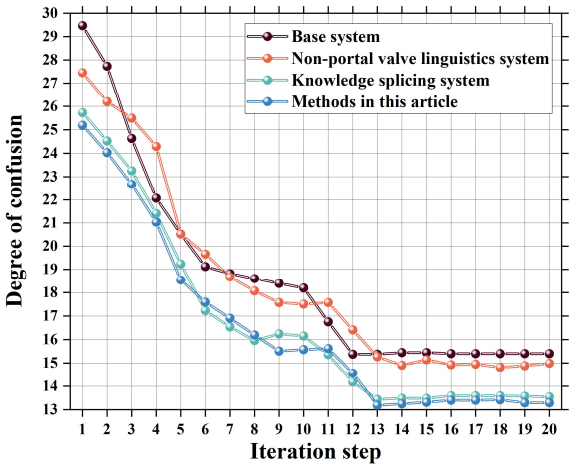
Table 3. Language translation performance results

System	Chinese-English translation		
	NIST 2010	NIST 2012	NIST 2016
Base system	35.08	27.14	23.49
+Lexical system	36.13*	29.01*	24.58*
+Named entity system	36.25*	28.54*	24.73*
+ Shallow semantic system	35.97*	28.63*	25.45*
+ Shallow semantic system	36.84*	28.93*	24.19*
+ All systems	36.89*	29.77*	25.97*
	English- Chinese translation		
	NIST 2010	NIST 2012	NIST 2016
Base system	26.33	27.15	24.16
+Lexical system	27.56*	27.45	24.55
+Named entity system	27.43*	27.56*	24.83*
+ Shallow semantic system	27.75*	27.93*	24.69*
+ Shallow semantic system	28.66*	28.61*	25.06*

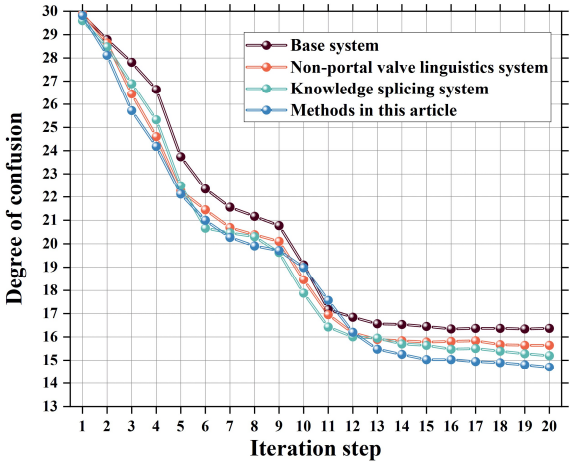
In the Chinese-to-English translation task, the embedded lexical-only condition (+Lexical system) improves the linguistic translation performance on the test set by 1.05, 1.87, and 1.09 BLEU values, respectively, compared to the baseline system. The "+Named entity system" improves the BLEU values by 1.17, 1.4, and 1.24, respectively, and finally, the simultaneous use of all four linguistic knowledge improves the BLEU values by 1.81, 2.63, and 2.48, respectively, i.e., the translation system that uses all four linguistic knowledge simultaneously has a higher BLEU value compared to the translation system that uses any of them alone. Linguistic knowledge alone, the translation results are better and significantly outperform the baseline translation system.

3.2.2. Confusion Analysis. In information theory, perplexity is a measure of how well the predictions of a probability distribution or probability model fit the sample. The lower the perplexity, the more accurate the fit is, indicating a better model, and this quality can be used to compare the advantages and disadvantages between different models. Fig. 5 shows a graph comparing the perplexity of the baseline system, the gate-less linguistic system, the knowledge splicing system, and the a priori knowledge-based neural machine translation method proposed in this section on the calibration set for both the Chinese-English and English-Chinese translation tasks. Where Figure 5(a) shows the Chinese-English translation task and Figure 5(b) shows the English-Chinese translation task. It can be seen that the perplexity of the methods proposed in this section on the check set is always lower than

that of the baseline system, the gate-less linguistics system, and the knowledge splicing system. The lowest perplexity for the Chinese-English translation task is stabilized at around 13.5, and the lowest perplexity for the English-Chinese translation task is stabilized at around 15, and the perplexity performance is consistent with the BLEU values in the previous section, which verifies the high efficiency of the method in this paper again.



(a) Chinese-English translation



(b) English-Chinese translation

Fig. 5. Perplexity as a function of epochs on validation dataset

4. Conclusion

As a bright pearl in the upper-level application of artificial intelligence, the development of machine translation theory and technology is a product of the joint development of linguistics, cognitive science, information theory and computer science. However, since neural machine translation only utilizes a single nonlinear neural network to realize the conversion between two languages, it brings a series of advantages but also has shortcomings:

1) Neural machine translation is more sensitive to long sentences with complex sentence structure than statistical machine translation.

2) The end-to-end implementation process does not explicitly utilize linguistic knowledge to further enhance translation performance.

Aiming at the language translation problems existing in the context of the above two AI technologies, this paper proposes a neural machine translation method that incorporates prior knowledge. The

empirical results show that the maximum BLEU values obtained by this method in the Chinese-English and English-Chinese translation tasks are 36.89 and 28.66, respectively, which are significantly higher than those of other models. This method obtains a certain degree of improvement over the baseline system.

References

[1] Wenge, M. . (2021). Artificial intelligence-based real-time communication and ai-multimedia services in higher education. Journal of multiple-valued logic and soft computing(1/3), 36.

[2] Bi, S. . (2020). Intelligent system for english translation using automated knowledge base. Journal of Intelligent and Fuzzy Systems, 39(5), 1-10.

[3] Zhao, X. , & Jiang, Y. . (2022). Synchronously improving multi-user english translation ability by using ai. International Journal on Artificial Intelligence Tools.

[4] Du, M. . (2022). Application of digital technology-based tpack in english translation. Mobile information systems(Pt.2), 2022.

[5] Wu, H. . (2021). Multimedia interaction-based computer-aided translation technology in applied english teaching. Mobile Information Systems.

[6] Zongli, J. , & Wei, W. . (2017). Translation recommendation system with information retrieval technology. Journal of Harbin Engineering University.

[7] Li, Y. . (2017). The application of parallel corpora based computer-aided translation technology in chemical english translation. C e Ca, 42(6), 2433-2437.

[8] Liu, X. . (2017). Construction of foreign language translation and training platform based on computer technology. Revista de la Facultad de Ingenieria, 32(15), 382-387.

[9] Jing, C. , & Liu, G. . (2022). Acquisition of english corpus machine translation based on speech recognition technology. Scientific Programming.

[10] Wang, X. . (2017). A translation aesthetics in the translation of english for science and technology. Boletin Tecnico/Technical Bulletin, 55(13), 412-417.

[11] Pan, W. . (2021). English machine translation model based on an improved self-attention technology. Scientific programming(Pt.14), 2021.

[12] Yao, S. . (2017). Application of computer-aided translation in english teaching. International Journal of Emerging Technologies in Learning, 12(8), 105-117.

[13] Yan, Y. M. , Zhang, B. , & Guo, J. . (2018). A state-space approach for service-oriented sla translation. Journal of Internet Technology, 19(1), 197-206.

[14] Guo, J. . (2017). Translation processing algorithm of complex long sentence based on multiple strategy analysis. Boletin Tecnico/Technical Bulletin, 55(18), 705-710.

[15] Ban, H. , & Ning, J. . (2021). Design of english automatic translation system based on machine intelligent translation and secure internet of things. Mobile information systems.

[16] Helen, E. . (2024). Transcribing in the digital age: qualitative research practice utilizing intelligent speech recognition technology. European Journal of Cardiovascular Nursing.

[17] Wang, R. . (2021). Research on intelligent english translation method based on the improved attention mechanism model. Scientific Programming.

[18] Song, X. . (2020). Intelligent english translation system based on evolutionary multi-objective optimization algorithm. Journal of Intelligent and Fuzzy Systems(10), 1-11.

[19] Bei, L. . (2020). Study on the intelligent selection model of fuzzy semantic optimal solution in the process of translation using english corpus. Wireless Communications and Mobile Computing, 2020(5), 1-7.

[20] Qiwei, Y. , Yu, D. , & Guangming, L. . (2023). Exploration of english speech translation recognition based on the lstm rnn algorithm. Neural computing & applications(36), 35.