Real-Time Automatic Translation Algorithm for Chinese Subtitles in Media Playback Using Knowledge Base

Li Yan

Foreign Languages Department, Xinjiang Teacher’s College, Urumqi 830000, China

Correspondence should be addressed to Li Yan; 2021201003@stu.sicau.edu.cn

Received 14 April 2022; Revised 11 May 2022; Accepted 27 May 2022; Published 18 June 2022

Academic Editor: Muhammad Babar

1. Introduction

Presently, machine learning (ML) is undergoing a revolution, with deep learning (DL) serving as the primary driving factor. Deep neural networks are incredibly strong ML systems that can perform admirably on difficult issues like speech processing [1] and visual object identification [2]. DNNs, in particular, have enabled significant advances in natural language processing because of their ability to capture complex language patterns [3]. Machine translation is an effective representation of NLP activities, and its major job is to use computer program to convert words or phrases from one language to the other. With the more and more frequent exchanges among ethnic groups in the world, there is an urgent need for machine translation to solve the problem of language barriers among ethnic groups. Machine translation is the process of utilizing a computer to automatically convert one natural language into another that has the exact same meaning as the original language being translated. Machine translation is closely related to syntactic analysis and semantic understanding. Its research goal is to solve a series of linguistic problems [4–6]. It is widely assumed that machine translation has evolved in three stages: rule-based machine translation [7], statistical machine translation [8], and NMT [9]. During the last two decades, SMT has been the dominant driving factor. This technique, unfortunately, may ignore the extended dependence beyond the duration of phrases, resulting in discrepancies in conversion outcomes, for example, erroneous gender agreements. Separate components including word aligners, translation of rule extractors, and other characteristic extractors suffer as well. In comparison to the SMT approach, NMT has a simpler design and can capture extensive dependencies in sentences, indicating that it has the potential to become a new trend in translating.

Subsequent research has concentrated on huge pre-trained systems, which largely depend on the model of transformer, nonetheless with substantially more ability [10]. This method employs a mix of pretraining and supervision well. The capacity of transformer language models
has grown dramatically, from hundred million variables [11] to 1.6 billion variables [12], and eventually 18 billion variables [13]. Even though each increment improved downstream NLP effectiveness, exercise of such frameworks needs huge scale specialized computer gear, such as TPUs of Google [14]. These computer groups are often out of reach for small- and medium-sized businesses. On the other hand, these systems are too intricate and have too much potential for humans to comprehend. That is, we understand the models work well, yet we do not distinguish the reasons. The most common method for developing neural network models is to gather a collection of training instances displaying the right action for the target job, train a system to reproduce these behaviors, and then assess its performance on unrelated held-out cases. This method, which is generally acknowledged, may create skilled models that behave as domain experts. Many previous studies have demonstrated that employing a domain-specific bilingual vocabulary can improve prediction accuracy [15]. Depending on this discovery, the authors developed a technique for discovering and retrieving naturally existing Chinese-English simultaneous sections. Their assessment findings show that the automatically generated simultaneous data improves translation quality significantly. The author of [16] created the UM-Corpus as a multidomain and balancing equivalent corpus. It is a 2 million-word English-Chinese aligned corpus drawn from 8 distinct text domains, comprising education, media, science, speech, and subtitles. Though employing a domain-specific corpus for training the model has given encouraging results, we feel that a multilingual corpus from subject matter experts is lacking. As a result, choosing how to harness domain knowledge and generate a novel corpus is crucial.

As previously said, machine translation has a challenging development path. People have worked hard on it and made significant progress. Unfortunately, based on the existing scenario, machine translation has been unable to provide an acceptable translation. In the course of investigating machine translation, researchers increasingly learn that to produce high-quality translation outcomes, it is important to effectively evaluate and comprehend natural language meanings. The architecture of a language and its meaning are intricately intertwined. A language structure can have numerous meanings, and various structures can represent multiple semantics. As a result, the application of semantic knowledge in machine translation has taken center stage. The highest-level academic conference in the field of natural language processing and computational linguistics was held in Beijing in 2015. Here a session on semantic-based machine translation was expressly organized for the conference. Scholars generally believe that the semantic era has come, and the development of machine translation should move towards semantics [17]. Dong Zhendong’s HowNet is the most comprehensive of the existing semantic knowledge system. In 2013, the HowNet English-Chinese machine translation system came out. The semantic knowledge of HowNet plays an important role in the system, which is also a typical application of semantics in machine translation [18]. Because of the aforementioned concerns, this paper thoroughly explores HowNet, a machine translation system, as well as the theoretical framework. It also discusses the translation principle and procedure, along with the system knowledge base, which includes the HowNet knowledge base, the axiomatic rule base, and the translation rule base. In addition, it evaluates the HowNet machine translation system by analyzing and summarizing the advantages and problems of the system. On the base of these, it further puts forward an effective solution to the problem of unregistered words and translation word selection in domain-oriented translation testing, which is the expansion of domain-oriented system knowledge base. The scheme solves the problem of unlisted words by expanding the HowNet knowledge base or concept knowledge base of the system and solves the problem of translation word selection by expanding the translation rule base of the system. The main work of this paper is summarized as follows:

(1) In this study, the HowNet machine translation system is built based on HowNet. The theoretical basis of the whole system is the knowledge base system, including the HowNet knowledge base, axiomatic rule base, and translation rule base. Through the in-depth analysis of the HowNet knowledge base, this paper has a deeper understanding of HowNet, which is a network knowledge system. This network relationship is the soul of HowNet and runs through the whole translation system. Axiomatic rules act on the inference engine to lay the foundation for meaning community disambiguation, and translation rules control the translation process from semantic analysis and transformation.

(2) To have a deeper understanding of HowNet’s machine translation system, the system was evaluated and analyzed in the patent title, aviation corps, and people’s daily. In the extracted 260 sentences, the problems of the system were classified and summarized according to the two parts of logical semantic analysis and translation conversion generation, and 15 error categories were obtained. Among them, unregistered words and selected words accounted for the highest proportion of all problems, and the problems of unregistered words and selected words in the aviation corps were more serious. Because of this, this paper proposes a domain-oriented knowledge base extension of the HowNet machine translation system, which is respectively HowNet knowledge base extension and translation rule base extension.

The remainder of our research work is divided into four sections: Sections 2–5, of which Section 2 is based on the current state of machine translation research, Section 3 explains our knowledge base and machine translation, Section 4 describes the experimental results and their simulations, and Section 5 concludes this work.

2. Research Status of Machine Translation

This section introduces the stages of machine translation. The development of machine translation has experienced three different periods: upsurge, low tide, and development,
from rule-based translation methods to statistical translation methods, NMT, and so on. The three basic machine-based translation approaches, rule-based, statistics basis, and natural translations are discussed in depth here.

2.1. Rule-Based Translation Method. The rule-based machine translation method relies on the manual compilation of bilingual dictionaries and various forms of translation rules summarized by experts. The dictionaries and translation rules are decoded by computer, and the source language sentences are translated into target language sentences. The rule-based translation system can deal with the ordering problem in translation well, and the system resources occupied by the runtime are also small. The early mainstream machine translation methods are rule-based. At present, most of the machine translation products sold in the market are rule-based. The machine translation system developed in the early stage mainly serves national and international government organizations and the military. A typical example of this system was jointly developed by the University of Montreal and the Translation Bureau of the federal government of Canada, which began to provide weather forecasting services in 1976 [19].

2.2. Statistical-Based Translation Method. In the early 1990s, Brown and others proposed IBM series models, which marked the emergence of statistical machine translation and the research of machine translation entered a new era. The main content of this method is the alignment of bilingual sentence pairs. The probability of mapping a word of one language to a word of another language is calculated by the possibility of word cooccurrence. Statistical machine translation has many advantages, such as unsupervised learning ability and good mathematical model, more integration of syntactic structure and semantic grammar information, and easy large-scale training. Its biggest advantage is that the training of a large number of datasets does not require too much manual participation [20].

At this point, there are three types of statistical machine translation models that have been developed: word-based translation models, phrase-based translation models, and syntax-based translation models [21]. The phrase-based translation model is the most developed model, while the syntax-based translation model is the focus of current research. Figure 1 depicts the pyramidal structure of the statistical machine translation model.

According to the above figure, the closer to the top of the pyramid, the more thorough the language analysis, the stronger the disambiguation ability, and the higher the translation quality. The deeper the analysis is, the more problems may arise. The quality of the translation may not be the result expected by mankind [22]. The development of statistical machine translation has also presented many problems, resulting in low translation quality. For example, the training time is too long, and the resource consumption is too large. The training effect will be better if the corpus has high requirements for parallel corpus, a large amount of data, and good quality, which requires more accuracy in the process of corpus selection and processing; lack of language knowledge, and almost no linguistic knowledge is applied in the model based on words and phrases [23].

2.3. Neural Machine Translation. In several fields of natural language processing, neural network models are used. The authors developed a probabilistic language model based on a multilayer feedforward neural network in their paper [24]. In contrast to traditional language models, words in the neural language model are automatically translated to a low-dimensional space to represent the language based on word similarity. Similarly, the expert of [25] presented language modeling of neural network models reproduction to forecast current words with infinite-length information. Another well-known use in [26] is the processing of a multilayer convolutional neural network for word annotation, named entity recognition and semantic role annotation. In this connection, the work of [27] proposed a neural network machine translation, with a simple structure composed of a deep neural network encoder and decoder. The encoder receives the input sentence, and the decoder produces a translation of the output. Bahdanau proposed an attention-based neural network machine translation model in 2015, in which attention is based on the source language words related to the target word. In addition, there are many other works of MT, such as taking the attention-based NMT model as features in the phrase model, decoding together with other features, and using minimum risk (MRT) to tune neural networks. The translation quality of these works is greatly improved than the traditional phrase model.

3. The Knowledge Base and Neural Machine Translation

3.1. The Knowledge Base. HowNet machine translation system is designed based on HowNet. This chapter first introduces the HowNet knowledge base, analyzes the concept and usage of HowNet and the relationship between concept attributes and attributes, then introduces and
analyzes HowNet disambiguation tools showing how HowNet uses common and individual rules to realize concept-related connection and ambiguity elimination, and finally introduces how the translator invokes translation rules. The translation process of the HowNet machine translation system can be seen in Figure 2. Among them, the knowledge base, axiom rule base, and translation rule base in the above picture are collectively referred to as the knowledge base of the machine translation system of the knowledge network. As the theoretical basis of the machine translation system of the knowledge net, the three resource bases play an important role in the machine translation of the knowledge network. It will be described in detail below.

3.1. HowNet Knowledge Base. HowNet knowledge base serves a critical role in translation as the linguistic resource for the HowNet machine translation system. Based on HowNet, the knowledge base is built. HowNet is a description of world knowledge. It uses “semantic origin” and “semantic relationship” to describe a “concept,” including all changes in everything. Dong Zhendong proposed that HowNet’s grasp of the world is that all things move and change in a specific time and space, changing their attributes, which are reflected in the corresponding attribute values. Adding “components” to these foundations is the whole world, which is described by the formula as follows:

\[
\text{World} = \text{everything} + \text{time} + \text{space} + \text{attribute} + \text{attribute value} + \text{Event} + \text{part}. \quad (1)
\]

3.2. Neural Machine Translation. The goal of machine translation is to use the computer to realize automatic translation of natural language. Its basic idea is that in the case of a given source language input sentence X, the system gives a target language sentence Y to maximize the conditional probability \( p(Y|X) \). As a machine translation system with the best performance at present, NMT mainly uses a neural network to realize automatic mapping of natural language. Its basic idea is to give a given source language input sentence X and parameter \( \mu \), and the system generates the optimal target language sequence \( Y \), which maximizes the conditional probability \( p(Y|X,\mu) \). Parameter \( \mu \) is a parameter in the neural translation model, which is mainly obtained by model training and automatic adjustment. When a sentence in one language is converted into a sequence in another language using a NMT model, the encoder and decoder work together to create semantic context vectors are then used to build the target language sentence sequence. RNNs are employed in the encoder and decoder to encode source language utterances and create target language sentences [28].

An encoder-decoder architecture of cyclic neural networks is used in this paper, with the addition of a soft attention mechanism. The system’s architecture can be depicted in Figure 3.

In the training stage, dropout is added to the network output layer to prevent the overfitting of the model. In order to obtain the optimal translation sequence, the beam search algorithm is used to select the optimal target sequence result during decoding.

3.2.1. Cyclic Neural Network Model. An artificial neural network known as a cyclic neural network is one such example. Its neurons build a directed graph as they interact over time. As a result, it can change its behavior over time. When compared to a feedforward neural network, an RNN can process input sequences using its internal neuron state (memory). Intelligent speech, language models, text production, and other fields rely on the RN model [29].

The hidden layer state \( h_t \) of each time step \( t \) of the cyclic neural network is obtained by a nonlinear transformation between the hidden layer state \( h_{t-1} \) at time \( t - 1 \) and the current input \( x_t \), as shown in

\[
h_t = f(h_{t-1}, x_t). \quad (2)
\]

Among them, \( f \) is a nonlinear function, and the commonly used nonlinear functions include sigmoid, ReLU, and Tanh. According to the output characteristics of the RNN network, sentences of any length can be semantically encoded by using RNN, and the output results of each time can be saved to obtain the output sequence of the whole input sentence. This paper mainly uses RNN to construct the encoder-decoder network model of NMT.

Long-Short-Term Memory (LSTM) is a type of cyclic neural network that is used for short-term memory. To each neuron in the hidden layer, memory units have been introduced based on regular RNN neurons. As a result, automatic control of the memory information on the time series has been achieved. Compared with the simple cyclic neural network architecture, the LSTM network solves the problems of long-short distance dependence and gradient disappearance and can better encode sentence information. It has achieved the most advanced results in testing artificial datasets that rely on learning ability for a long time and challenging sequence processing tasks [30]. The network architecture of LSTM is shown in Figure 4. Based on the external RNN model, LSTM adds an internal circulating “LSTM cell unit,” which is similar to the ordinary circulating neural network. Each cell unit has the same input and output but adds a gate control unit. The model structure is relatively complex, and the training iteration cycle is long.

The forgetting gate mainly determines what useless information should be discarded from the cell state. The gate will output a value between 0 and 1. 1 represents “complete retention,” and 0 represents “complete abandonment.” Its calculation formula is shown in

\[
f_t = \sigma(W_f x_t + b_f), \quad (3)
\]

where \( x_t \) is the input vector at the current time, \( h_{t-1} \) is the hidden layer state vector at the previous time, and \( b_f \) and \( W_f \) are the offset and cyclic weight of the forgetting gate, respectively.
The update gate is primarily responsible for determining, which is divided into two sections: first, the sigmoid layer is referred to as the “input gate layer,” and it is this layer that chooses which information in the model will be updated, and second, the sigmoid layer is referred to as the “output gate layer.” Further, a Tanh layer is required to create a new candidate value vector, and its calculation formula is shown in

\[
C_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c),
\]

(5)

where \(x_t\) is the input vector at the current moment, \(h_{t-1}\) is the hidden layer state vector at the previous moment, \(b_c, W_c\) are the bias and loop weights when updating the cell state, respectively.

For the update of cell state values, the model multiplies the old cell state with the forgetting gate, discarding useless information, plus the information to be updated to the cell state, and the calculation formula is shown in

\[
C_t = f_t \cdot C_{t-1} + i_t \cdot C_t.
\]

(6)

Finally, the model determines which part of the useful information will be output the next time through an output gate. The output gate and the final output result are calculated as shown in

\[
o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o),
\]

(7)

\[
h_t = o_t \cdot \tanh(C_t),
\]

(8)

where \(x_t\) is the input vector at the current time, \(h_{t-1}\) is the state vector of the hidden layer at the previous time, and \(b_o, W_o\) are divided into bias and cyclic weight of different output gates.

Threshold recurrent unit (GRU) is another variant of the recurrent neural network, and its model complexity is lower than that of LSTM. Different from LSTM, GRU uses a single threshold unit to control the forgetting factor and update the status unit at the same time. The update operation is shown in

\[
u_i^{(t)} = u_i^{(t-1)} h_i^{(t-1)} + \left(1 - u_i^{(t-1)} \right) \sigma \left( b_f + \sum_j U_{i,j} x_j^{(t)} + \sum_j W_{i,j} f_j^{(t-1)} h_j^{(t-1)} \right),
\]

(9)

where \(u\) stands for “update door” and \(r\) stands for “reset door.” Their calculation methods are shown in

\[
u_i^{(t)} = \sigma \left( b_f + \sum_j U_{i,j} x_j^{(t)} + \sum_j W_{i,j} f_j^{(t-1)} h_j^{(t-1)} \right),
\]

(10)

\[
r_i^{(t)} = \sigma \left( b_f + \sum_j U_{i,j} x_j^{(t)} + \sum_j W_{i,j} f_j^{(t-1)} h_j^{(t-1)} \right).
\]

(11)
Figure 3: Encoder-decoder model.

Figure 4: LSTM network architecture.
3.2.2. Semantic Knowledge Pretraining Model. This paper uses the TransE model to pretrain the semantic knowledge triples to obtain the word vector and semantic relationship vector. The dataset information used in its pretraining is shown in Table 1.

For the purpose of minimization of an objective function, the TransE model uses the maximum interval method during the model training phase. This paper uses the open-source TransE model tool by the NLP Laboratory of Tsinghua University for pretraining. The parameters use the default configuration. The word vector and relationship vector obtained from the pretraining will be used in the subsequent NMT model.

3.2.3. Knowledge Graph Coding Model. The purpose of the knowledge map encoder is to better understand the input text sentence. It can enhance the semantic information of words by encoding the corresponding knowledge map of words. The structure of the knowledge map encoder is shown in Figure 5. The knowledge map encoder uses each word in the input sentence as the key value to retrieve the related triple \( g_i = \{T_1, T_2, \ldots \} \) in the local semantic knowledge base. Each retrieved triplet contains two words and the relationship information between words. For common words (such as "of") that may not be able to retrieve the corresponding triples in the knowledge base, a special symbol "non" is used. Then, the knowledge graph encoder uses the graph attention mechanism to calculate the graph vector \( g_i \). By splicing the word vector \( w(x_i) \) and the knowledge map vector \( g_i \), the spliced vector \( c(x_i) = [w(x_i); g_i] \) will be sent to the encoder of NMT for semantic coding.

3.2.4. Graph Attention Model. The graph attention model is used to generate the vector representation of the retrieved knowledge graph and dynamically encode the semantic knowledge related to word semantics into the system. The graph attention model used in this paper encodes more semantic information, considering not only all nodes in the graph but also the semantic relationship between nodes. The graph attention model generates a semantic vector representation for each knowledge graph to enhance the semantic information of words in the input sentence.

Formally, the graph attention model takes the knowledge vector \( K(g_i) = \{k_1, k_2, \ldots \} \) in the retrieved knowledge graph \( g_i \), as the input to generate the graph vector \( C \). The calculation process is shown in

\[
 g_i = \sum_{n=1}^{N_g} \alpha_n^s [h_n, t_n], \\
 \alpha_n^s = \frac{\exp(\beta_n^s)}{\sum_{j=1}^{N_g} \exp(\beta_j^s)}, \\
 \beta_n^s = (W_r r_n)^T \tanh(W_h h_n + W_t t_n + W_s s_{i-1}),
\]

where \( (h_n, r_n, t_n) = k_n, W_h, W_r, W_t, W_s \) are the weight matrices, \( s_{i-1} \) is the state of the hidden layer at the previous time, and the attention weight measures the degree of correlation among \( r_n \), the headword \( h_n \), tail word \( t_n \), and the historical information.

Graph vector \( g_i \) is the weighted sum of the head and tail word vectors of triples contained in the knowledge graph, which comprehensively considers the coding information of words and the relationship between words.

4. Experiment and Simulation Results

This section describes the experiment of incorporating a semantic knowledge base to enhance the translation performance of uncertain terms in Mt. WordNet is used to extract semantic concept information. This paper uses a wn18 semantic knowledge base, which contains 18 semantic relationships and 140000 sets of semantic knowledge triples. This paper makes a comparative experiment on the bilingual data of Britain, China, and Germany and makes a detailed comparative analysis of the experimental results.

4.1. Experimental Data. This experiment is also conducted on two different language datasets, one is the English-German news corpus selected from WMT2014, including 4.5 million double sentence pairs, and the other is the English-Chinese conference news corpus selected from the United Nations Corpus, which contains 5 million bilingual sentence pairs. The Anglo-German experiment uses WMT2013 provided by WMT as the verification set of the evaluation task, WMT2012, and WMT2014 as the test set of 3000 sentences each. The Chinese-English experiment uses an English-Chinese public verification set and a test set of 4000 sentences, each provided by the United Nations Corpus. English and German word segmentation uses the word segmentation tool provided by the open-source system Moses, and Chinese word segmentation uses Jieba Chinese word segmentation system. Conduct lexical analysis on the training set (TRS), verification set (VES), and test set (TES) of the two datasets, respectively. The detailed data information after word segmentation is shown in Tables 2 and 3.

According to Table 2, the TRS dataset has 8400000-sentence logarithm in both English and German with corpus words 568266 and 1623598, respectively. The VES dataset has 5000-sentence logarithm in both English and German with corpus words 36589 and 6589, respectively. Similarly, the TES dataset has 3000-sentence logarithm in both English and German with corpus words 56236 and 8456, respectively.

According to Table 3, the TES dataset has 8000000-sentence logarithm in both Chines and English with corpus words 256589 and 562888, respectively. The VES dataset has 3000 sentence logarithm in both Chines and English with corpus words 8802. Similarly, the TES dataset has 3000 sentence logarithm in both Chines and English with corpus words 36589 and 6589, respectively.
4.2. Experimental Setup. The following super parameters were established in the experimental section: the dimension of hidden layer neurons in the encoder and decoder is 512, and the vector dimension of source language words and target language words is also 512 in the experimental section. To optimize network parameters, the model makes use of the AdaDelta method. The model is trained in batches of 128 sentences to increase the speed with which it converges to its final state. The maximum sentence length for the English-German experiment in this research is set to 50 words, and the maximum sentence length for the Chinese-English experiment is set to 80 words. Dropout is utilized at the output layer of the decoding end to prevent parameter overfitting. The dropout rate is adjusted to 0.5 to prevent the overfitting of parameters. For this test, the decoder employs the beam search algorithm, with the beam size set to 10.

4.3. Experimental Evaluation Method. In this paper, the ambiguous word experiment uses the BLEU value and NST value as the automatic evaluation criteria of translation results. The higher the BLEU value is, the more accurate the translation result is, and the closer it is to the result of professional manual translation. BLEU-4 is used as the final evaluation result in the experiment. The proposed approach uses accuracy, recall, and $F_1$ value as evaluation indexes. The higher the $F_1$ value is, the more effective the method proposed in this paper is in dealing with ambiguous words.

4.4. Experimental Results. The comparative experiment in this section includes five experimental systems that can be explained in Figure 6.

4.4.1. RNN Search. It is a baseline NMT system with an improved attention mechanism. Unk in the generated translation is restored using the attention mechanism.

4.4.2. NBOW-Concat. The system is proposed by Liu et al. By encoding the context semantic information of ambiguous words into NMT, the NBOW method is used to obtain the word context semantics, and the context semantic information is integrated into MT through splicing operation.

4.4.3. BiLSTM-Concat. The system is proposed by Gonzales et al. It encodes the semantic vector information of ambiguous words into NMT and integrates the semantic vector information into NMT through splicing operation.

4.4.4. Our Model. The system proposed in this paper uses the NMT model based on the graph attention mechanism to
combine the semantic knowledge encoded with the word vector.

4.4.5. Our Model and BiLSTM-Concat. Our model is combined with BiLSTM-Concat to add the semantic knowledge retrieved from words and the semantic vector information of words into the MT model at the same time so as to improve the semantic discrimination ability of ambiguous words.

In order to further analyze and demonstrate the method proposed in this paper, based on the original dataset, this paper selects the subdataset with high N18 coverage and re verifies the effect. The details of the extracted British-German and English-Chinese datasets are shown in Tables 4 and 5.

The subdatasets extracted in this paper are used for English-German translation tasks. The proposed method (our model) improves the BLEU value of the verification set and test set by 1.25 on average compared with the baseline system RNN Search and 0.89 and 0.35 BLEU values compared with the other two traditional word sense disambiguation methods (NBOW-Concat and BiLSTM-Concat), respectively. The method proposed in this paper is combined with the traditional method, and the contextual semantic information and the relevant semantic information of words are integrated at the same time. This method has better performance than the system integrating single information and improves the BLEU value by 1.52 compared with the baseline system.

As shown in Tables 6 and 7, for the English-Chinese task, our model has an average of 1.14 BLEU values higher than the baseline system RNN Search in the verification set and test set and 0.82 and 0.22 BLEU values higher than the other two traditional ambiguous word processing methods (NBOW-concat and BiLSTM-Concat), respectively. Combining the method proposed in this paper with the traditional method and integrating the context information and word-related semantic information at the same time improves the BLEU value by 1.43 compared with the baseline system.

Figure 7 shows the comparison of English and German BLEU scores (%) in different systems. This figure compares five systems such as RNN Search, NBOW-Concat, BiLSTM-Concat, our model, and our model & BiLSTM. According to this data, the average score of our model and our model & BiLSTM is 33.56% and 45.26%, respectively, demonstrating that the efficiency of these models is better than the others.

Figure 8 shows the comparison of English and Chinese BLEU scores (%) in different systems. This figure compares five systems such as RNN Search, NBOW-Concat, BiLSTM-Concat, our model, and our model & BiLSTM. According to this figure, the average score of our model and our model & BiLSTM is 37.69% and 45.29%, respectively, demonstrating that the efficiency of these models is better than the others.
In conclusion, from the results of English-German and English-Chinese translation tasks in Figures 7 and 8, it can be seen that the effect of our model on the standard TRS and test set is poor, and the translation performance is lower than that of traditional word sense disambiguation methods. Through the statistical analysis, it is found that the wn18 semantic knowledge base covers only 30% of the words in the TRS and 15% of the words in the test set.

5. Conclusion

Machine translation came into being to overcome language barriers with the emergence of computers. Machine translation based on rules and statistics has developed one after the other, bringing a new era for machine translation. Facing the problems of lack of language knowledge and ambiguity in machine translation, more and more scholars place their hopes on the expression of semantic information. At present, semantic machine translation is mostly in the research stage. Researchers mostly add semantic information to the existing machine translation methods to improve the translation quality. In 2013, the English-Chinese machine translation system based on HowNet was completed, bringing new ideas for the application of semantic knowledge in machine translation. Therefore, this paper thoroughly explores HowNet as well as the theoretical framework. It also discusses the translation principle and procedure along with the system knowledge base, which includes the HowNet knowledge base, the axiomatic rule base, and the translation rule base. In addition, it evaluates the HowNet machine translation system by analyzing and summarizing the advantages and problems of the system. The scheme solves the problem of unlisted words by expanding the HowNet knowledge base or concept knowledge base of the system and solves the problem of translation word selection by expanding the translation rule base of the system.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that he has no conflicts of interest.

References


