International Journal for Housing Science and Its Applications
Publish August 3, 2025. Volume 46, Issue 3 Pages 920-932

https://doi.org/10.70517/ijhsa46369

A Computational Model of Cross-Cultural Semantic Association for German Literary Translation and Its Application to Goethe's Works

Yue Yu[1,*]
[1] Jiangsu University of Technology, Changzhou, Jiangsu, 213001, China
Corresponding authors: (e-mail: yuyuerita2008@163.com).

Abstract This study proposes a cross-cultural semantic association computational model for German literary translation, aiming to address the shortcomings of traditional machine translation in terms of cultural differences and semantic complexity. By fusing RNN and Self-Attention Network SAN encoder, the semantic associations of German words, sentences and paragraphs are quantified layer by layer, and a cross-language semantic ontology structure model is constructed to realize accurate semantic mapping between German and Chinese. The hybrid translation model is further designed to combine four pre-trained encoders with Marian decoder to optimize the translation generation in literary context. The experiments are based on the German-Chinese parallel corpus of WMT2018 and WMT2022.In the German-Chinese translation task, the correlation misalignment rate of this paper's model is only 4.82%, which is 49.05% lower than that of the baseline model, QE-BERT.The BLEU value is up to 26.21, which is significantly better than that of the comparison model, e.g., 20.05 for Gen-Det.In addition, the model performs stably in the multi-language task. The highest BLEU value of 31.86 in German-English language translation and content word ratio experiments show its robustness to complex semantic elements. The study shows that cross-cultural semantic association computation and hybrid model design can effectively improve the accuracy and cultural appropriateness of literary translation.

Index Terms German literary translation, cross-cultural semantics, semantic association computation, Goethe's works, machine translation

I. Introduction

Under the impetus of globalization, cultural exchanges among countries have become more and more frequent and in-depth [1]. As one of the important forms of cultural exchange, theater is a key bridge for different countries to display their cultural charms and export their cultural values [2]. Through the introduction of cross-cultural works, it can not only promote mutual understanding and integration between different cultural backgrounds and artistic concepts, but also draw on the advanced technology and artistic essence of Western theater, further promote Chinese theater culture to the international stage, enhance its artistic level and international influence, and contribute to the inheritance and development of Chinese culture [3], [4].

Johann Wolfgang von Goethe was an outstanding German writer, dramatist, thinker and poet, not only one of the greatest German writers, but also one of the giants in the world literary world [5]. Among them, the Faust poetic drama usually refers to the 12,111 lines long Faust Part I and Part II, which is the literary masterpiece of Goethe's dedication, and the most classic of all Goethe's works [6]. Along with The Divine Comedy, the Homeric epics of Ancient Greece, and Hamlet, it is one of the four great European classical masterpieces [7]. Goethe, together with Dante, Shakespeare, and Homer, ranks among the great men of Western literature, establishing Goethe's high status in the field of world literature [8].

Goethe's works have made an important contribution to the intermingling of Chinese and world cultures [9]. Its literary thought and artistic structure, in the process of Chinese literary creation, inspired the creative inspiration of Chinese literati and artists such as Zong Baihua, Feng Zhi, Guo Moruo, etc., promoted the national thinking related to self-worth, the meaning of life, and the liberation of the nation, and exerted a vital influence on the ideological emancipation and enlightenment movement in China [10]-[12]. Its social background breaks the stereotyped thinking about Germany and finds a focus of resonance in the local environment and German society [13]. Its translations, mediations, commentaries, and dramatic creations have promoted the mutual exchanges of Chinese and Western culture and art, and have taken on a new significance in contemporary society, both in terms of the study of Goethe's works themselves and the process of adapting his dramatic forms [14].

In German literary translation, the precise capture and transmission of cross-cultural semantic associations is the central challenge to ensure the fidelity and literariness of the translated text. In this study, we propose a cross-cultural semantic association computation method for German literary translation, aiming at constructing an interpretable translation model adapted to complex literary contexts through multi-level semantic association computation and ontology structure modeling. Specifically, based on RNN and Self-Attention Network SAN encoder, the semantic associations are parsed layer by layer from words, sentences to paragraphs to quantify the deep logic inside German texts. And by constructing a cross-language semantic ontology structure model, we realize semantic mapping and knowledge sharing between German and Chinese bilinguals, and solve the problem of disambiguating culture-specific expressions in translation. Finally, a hybrid model of machine translation incorporating multiple pre-trained encoders is designed to optimize the translation generation process by combining the results of semantic association computation and cross-cultural ontology knowledge. In the design of the hybrid model, the semantic relevance weights extracted by RNN and SAN encoders are embedded into the multi-head self-attention mechanism of the Transformer architecture to ensure that the translation generation process fully takes into account the cross-cultural semantic logic of the source text. At the same time, the semantic modification relations and grammatical analysis scheme defined in the cross-language semantic ontology model further constrain the generation path of the decoder.

II. Computational methods and modeling of cross-cultural semantic associations

II. A. Calculation of semantic relatedness of German words

The above-designed fusion RNN and SAN encoder is used to encode the information of German passages, and on the basis of which, the semantic relevance of German words is computed, which provides support for the construction of the subsequent sequencing model of German passage translation information.

According to the requirements of German paragraph machine translation, RelArtNet algorithm is selected to calculate the semantic relatedness of German words, and the flow of calculating the semantic relatedness of German words is shown in Fig. 1.

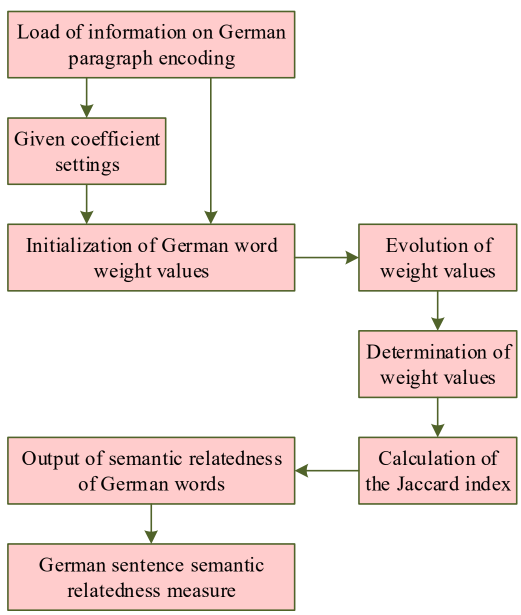

Figure 1: Calculation process of semantic correlation degree of German words

As shown in Fig. 1, weight values, Jaccard index calculation are most important in the process of German word semantic relatedness calculation. Among them, the weight value calculation is mainly determined by the German word importance. German word importance is expressed as

$$tf_{i,L} = \frac{\eta_{i,L}}{\sum_{k} \eta_{k,L}} \tag{1}$$

In equation (1), $tf_{i,L}$ denotes the importance of the German word t_i in the German passage L; $\eta_{i,L}$ denotes the number of times the German word t_i occurs in the German passage L; and $\sum_k \eta_{k,L}$ denotes the number of sentences k containing the German word t_i the total number of sentences k.

The Jaccard index is called Jaccard correlation coefficient, which takes the value of $[0,1]$. The closer the Jaccard index is to 1, the stronger the semantic correlation between the two; on the contrary, the closer the Jaccard index is to 0, the weaker the semantic correlation between the two. Based on the calculation result of formula (1), the semantic relatedness between any two words is calculated as

$$J(t_i,t_j)=\frac{\alpha^*\left|tf_{i,L}\cap tf_{j,L}\right|}{\beta^*\left|tf_{i,L}\cup tf_{j,L}\right|} \tag{2}$$

In Eq. (2), $J(t_i,t_j)$ represents the semantic relatedness between German words t_i and t_j; α^* and β^* represent the auxiliary parameters for the calculation of Jaccard's index, with the value range of $[1,10]$, which should be set according to the actual situation.

The components of a German paragraph are paragraphs, sentences and words, so it is necessary to measure the semantic relevance of German sentences scientifically in order to guarantee the accuracy of German paragraph translation.

Based on the results of German word semantic relatedness calculation, the value of German sentence semantic relatedness is measured, and the formula is

$$J(p_i,p_j)=1-\frac{Levenshtein(p_i,p_j)}{\max(|p_i|,|p_j|)} \tag{3}$$

In Eq. (3), $J(p_i,p_j)$ denotes the semantic relatedness of German sentences p_i and p_j; $Levenshtein(p_i,p_j)$ denotes the editing distance between German sentences p_i and p_j; and $|\cdot|$ denotes the number of words in a German sentence.

It should be noted that the process of measuring the semantic relatedness of German sentences mainly measures the semantic relatedness of neighboring German sentences, which is convenient for the translation of subsequent German passages.

The calculation of German word and sentence semantic relatedness is completed by the above process, which provides support for the realization of information ordering for the translation of subsequent German passages.

II. B.Modeling the Semantic Ontology Structure of Associative German

The calculation of semantic relatedness of German words lays a fine-grained foundation for cross-cultural semantic mapping; however, the complexity of literary translation requires higher-level semantic structure modeling. For this reason, this section further constructs the associative German semantic ontology structure modeling to extend the semantic associations of words and sentences to the cross-linguistic knowledge system in order to solve the problem of semantic disconnection caused by cultural differences.

II. B. 1) Semantic Mapping of German Translations for Cross-Language Information Retrieval

In order to realize the most relevant German semantic translation selection in cross-language information retrieval, it is necessary to firstly construct a model of the most relevant German semantic ontology structure in cross-language information retrieval, and carry out machine learning and training of German semantic translation for cross-language information retrieval based on semantic similarity calculation method.

Classification of semantic mapping relations for cross-language information retrieval is carried out based on semantic relatedness calculation and semantic selection structure of German words.

Definition 1: German semantic mapping: the formal definition of an ontological semantic mapping model for cross-lingual information retrieval of German translations for the grammatical analysis scheme A_i is a quintuple $O=\left\{C,H^C,R,I,A\right\}$, where:

C : the set of semantic modification concepts. The statement C_s in C has m different syntactic classifications, and semantically, the semantic similarity analysis yields a cross-linguistic database containing multiple clauses. In the ontology, A_i is a postpositive determiner, which satisfies the semantic structure of the utterance of the basic unit.

I : the set of instances. I is a unique individual obtained by semantic mapping for each instance of a simple clause selected one at a time. In the ontology, instances are ontology-mapped representations of the semantic structure of an utterance, and hence are also called semantic modification targets.

H^C : a collection of categorical relations on the semantic relevance of utterances. This class of relations allows to derive feature mapping relations for non-statement backbones, and is represented by the function $match(W_i, W_c)$ to denote the different syntactic analysis schemes between parent and child concepts in the representation ontology.

R : the categorized set of elements within the ontology of the current Cross-Language Information Retrieval Linked German database. The relations contained in R can be divided into two main categories: semantic targeting information indexing behavior relations and concept affiliation relations.

A : a collection of semantic modification targets. Each semantic modifier target in A represents the number of real words in the semantics of a German translation, which can be used to describe the mapping relation between concepts and instances for cross-language information retrieval by the semantic correlation between the semantic modifier targets W_{c_i} , or to describe the constraint relation for the mapping of a semantic feature under the i th syntactic analysis scheme.

II. B. 2) Model for calculating the relevance of semantic ontologies

On the basis of the above definition of semantic mapping of German translation that has been carried out for cross-language information retrieval, based on the multiple syntactic analysis schemes of German utterances in cross-language information retrieval, the syntactic analysis party scheme of the most relevant German language translation is constructed, and the syntactic analysis scheme of the most relevant German language translation is shown in Figure 2.

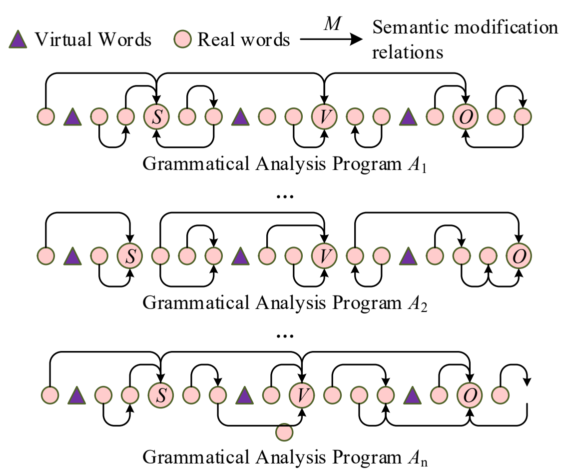

Figure 2: The most relevant grammatical analysis scheme for German translation

Assume that the mathematical model of the set of ontology-mapping three-tier integrated distributional concepts for associative German in cross-language information retrieval databases is:

$$x_j^i = f_j^i\left(x_j^i, u_i, u_j\right) \tag{4}$$

where x_j^i in j denotes the number of m grammatical analysis schemes for the utterance, and i denotes the semantic relevance; the weight coefficient of utterance stem $x_j^i \in R^n$. There are m grammatical analysis schemes for German clauses, and the most semantically logical grammatical analysis scheme A_i satisfies all lexical translation relevance constraints as:

$$A_i = \arg\max\left(f_{A_i}\right) \tag{5}$$

Based on cross-language information retrieval of word knowledge, semantic mapping relations are characterized by differences. Through the grammatical analysis of German translation of different ontologies, we decompose the matching degree of conceptual contextual association, consider the equivalent semantic mapping in the conceptual node Mountain, and obtain the optimal scheme of grammatical analysis after the calculation of semantic ontological relevance.

Definition 2: Semantic ontology modeling. Cross-language information retrieval associates the German translation ontology $O=\langle C,I,H^{C},R,A\rangle$ between $O'=\langle C',I',H^{C'},R',A'\rangle$ and the Semantic modification of mutual information features is represented by the mapping function M, $M:C*C'\to rel$. Where C is the ontology semantic relevance and rel is the set of real words of German resource information in the cross-language information retrieval database, called the German clause attribution relation.

Adopting the semantic directional information indexing method for contextual semantic mapping of German translations enables the expected use of vocabulary to effectively reflect German semantics in clause range selection. Based on the simple semantic units to establish the semantic modification relations of German translation, the semantic modification relations of German translation are shown in Figure 3.

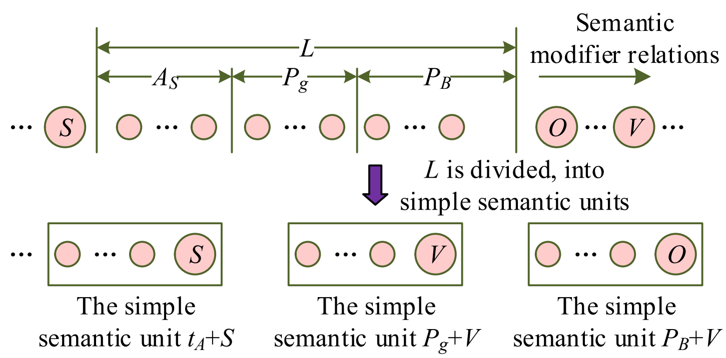

Figure 3: Semantic modification relations in German translation

Due to the heterogeneity of establishing inter-ontology German translation semantic modification relations, the semantic structure is $L\to A_A P_D A_G$, and the conceptual nodes (C_s and C_s') judge whether semantically related clauses in German translations belong to the main or subordinate clauses, and the judgment criterion includes the following: the internal between C_s and C_s' grammatical structure mapping relation $\left(\overset{\subset}{\to}\right)$, simple semantic unit semantic feature mapping relation $\left(\overset{\supset}{\to}\right)$, and self-organizing mapping generalization relation $\left(\overset{\sim}{\to}\right)$, the clause weight relation $\left(\overset{L}{\to}\right)$ and the mapping relation $\left(\overset{\perp}{\to}\right)$ for each syntactic analysis scheme. Calculate each simple semantic unit semantic ontology relevance, get the value of semantic relevance between heterogeneous ontologies, such as focusing on the concentration of pheromone of German translation vocabulary of semantic block n to realize the knowledge sharing and reuse, get the semantic relevance of semantic pointing information indexing in the collection of German translation words and phrases rel in cross-language information retrieval denoted as:

$$match\left(W_i,W_{C_s}\right)=\alpha*sim\left(W_i,W_{C_s}\right)+\beta*rel\left(W_i,W_{C_s}\right) \tag{6}$$

where $sim\left(W_i,W_{C_s}\right)$ denotes the similarity of the associated German semantic translation in the utterance C_s, and $rel\left(W_i,W_{C_s}\right)$ denotes the cross-language semantic relevance of the two sets of ontology fragments. Information retrieval semantic relatedness, and $\alpha+\beta=1$.

II. C.Machine Translation Hybrid Model Design

Based on the construction of cross-linguistic semantic ontology model, the semantic associations between German and Chinese bilinguals have been realized to be expressed in a structured way. However, how to effectively integrate this knowledge system into the generation process of machine translation is still the key to improve the quality of translations. This section proposes a hybrid model design for machine translation, which achieves accurate mapping from source language to target language by integrating multiple pre-trained encoders with cross-cultural semantic association knowledge, with particular focus on translation optimization of long-distance dependencies and cultural metaphors in literary contexts.

Machine translation is a sequence-to-sequence task, modeling the mapping relationship from source language $x=(x_1,x_2,\cdots,x_m)$ to target language $y=(y_1,y_2,\cdots,y_n)$. The current mainstream neural machine translation model architecture is shown in Figure 4.

The source language words are first converted to hidden representation Z by the encoder, Eq:

$$Z = Encoder(x) = (z_1, z_2, \cdots, z_n) \tag{7}$$

Next, the decoder generates the t th target word representation h_t using Z and the previously generated target word sequence $y_{<t} = (y_1, y_2, \cdots, y_{t-1})$ as inputs with Eq:

$$h_t = Decoder(Z, y_{<t}) \tag{8}$$

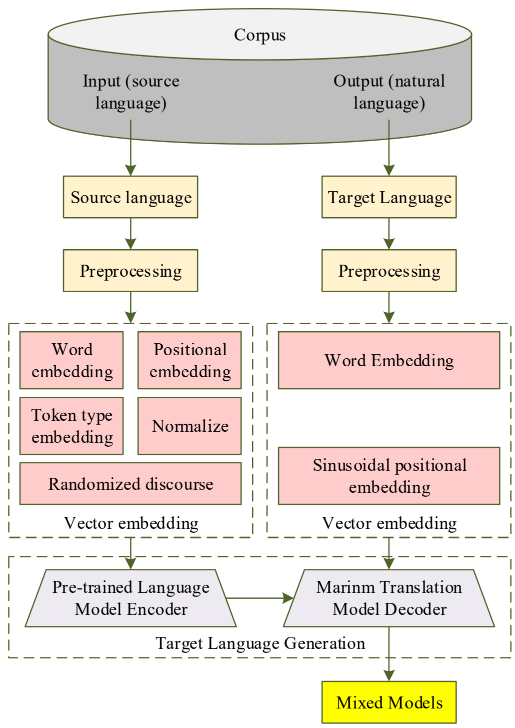

Figure 4: Machine translation hybrid model architecture

Subsequently, a linear projection is used to map h_t to the dimension $|V|$ of the target vocabulary list, and the Softmax function is used to compute the probability distribution of the generating word y_t at the moment t, which is given by:

$$p_\varphi(y_t \mid Z, y_{<t}; q) = soft\max(Linear(h_t)) \tag{9}$$

Ultimately, the word with the largest probability value is taken as the target word y_t predicted at moment t. The probability of the whole sentence is then expressed as the joint probability distribution of the individual words of the target language with the formula:

$$p_\varphi(y \mid x) = \sum_{t=1}^{m} p_\varphi(y_t \mid Z, y_{<t}; q) \tag{10}$$

In order to construct this hybrid model for machine translation, four different pre-trained models were chosen: luke-base, distilroberta-base, distilbert-base-uncased, and electra-small-discriminator. the encoders for these models are all based on the The encoder-decoder models of these models are all based on the transform architecture, which can effectively handle long-distance dependencies in sentences, and have excellent performance especially in longer sentences. Four different Seq2Seq hybrid models were generated by combining the encoder of each of the four models with the decoder of the Marian model.

The principle of these hybrid models is based on the Transformer architecture, in which each input lexical element is converted into a word vector using a word embedding algorithm, and the word embedding process occurs only in the lowest level encoder. After the input sequence is word embedded, each word element in the source language flows through the multi-head self-attention layer and feed-forward network layer in the encoder. The multi-head self-attention sublayer in the Transformer uses a “self-attention” mechanism, which adequately represents the closeness of the lexical element-to-lexical element connections. The self-attention mechanism allows the model to dynamically assign weights based on the intrinsic relevance of the input sequences, which is calculated as follows:

$$Attention\left(Q,K,V\right)=soft\max\left(\frac{QK^{T}}{\sqrt{d_{k}}}\right)V \tag{11}$$

where: Q denotes the query vector; K denotes the chain vector; V denotes the value vector; and d denotes the dimension of Transformer, where Q , K , and V are calculated as:

$$Q_h=W_h^Q\cdot X \quad K_h=W_h^K\cdot X \quad V_h=W_h^V\cdot X \tag{12}$$

where: W_h^Q , W_h^K and W_h^V are the weight matrices for each attention head; X is the input vector.

The hybrid model injects sequential information by adding position coding, by combining different frequencies of sine and cosine waves with each position of the input vector. The frequency of these waveforms increases with position, providing a unique encoding for each position. The position encoding is computed as:

$$P(pos)=\sin\left(\frac{pos}{10000^{\frac{2}{d}}}\right)+\cos\left(\frac{pos}{10000^{\frac{2}{d}}}\right) \tag{13}$$

where: pos is the position in the sequence.

Through the construction of such hybrid models, the four hybrid models proposed in this paper are able to utilize the rich linguistic knowledge learned from the pre-trained models and are able to combine the advantages of different pre-trained encoders to further enhance the performance of machine translation. The design concept of this hybrid model is to utilize the pre-trained encoder's ability in language understanding, combined with the Marian decoder's advantage in translation generation, so as to achieve higher quality translation results.

III. Experimental validation of cross-cultural semantic association model and multilingual translation performance analysis

Through the cross-cultural semantic association computational model and hybrid translation architecture constructed in Chapter 2, this chapter further verifies its effectiveness in multilingual translation tasks through experiments and analyzes in-depth the role of semantic association mechanism in improving the quality of translations.

Table 1: Experimental data scale

Task release time	Type	Dataset	Number
WMT2018	Parallel corpus	Training set	3.6M
	Evaluation data	Training set	14823
		Validation set	1123
		Test set	1028
WMT2022	Evaluation data	Training set	25374
		Validation set	1638
		Test set	1304

III. A. Experimental design and baseline modeling

III. A. 1) Data sets

The German-Chinese translation quality assessment data and parallel corpus for the experiments in this paper come from the WMT translation quality assessment task, where the quality assessment dataset contains two parts, WMT2018 and WMT2022, and the parallel corpus is the German-Chinese parallel corpus on Goethe's works provided by the WMT2022 quality assessment task. Table 1 lists the size of each corpus.

III. A. 2) Experimental setup

The optimizer used in this paper is AdamW, where the parameters of the optimizer are $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\varepsilon = 1e\text{-}8$, and the weight decay ratio is 1e-5. In the pre-training phase, the total number of training steps is 1 million, and the learning rate is adjusted according to the linear warm-up method, which reaches a maximum value of 1e-4 after 10,000 steps, and decays linearly in the subsequent training. The batch size was 15 sentences, and the gradient was accumulated as a factor of 2. In the fine-tuning stage of translation quality assessment, by performing hyperparameter search on the validation set, this paper uses the hyperparameter values that perform best on the validation set. Where the learning rate is 3e-5 and the batch size is 8. The graphics card used for training is GTX 1080ti.

III. A. 3) Baseline model

In order to verify the effectiveness of the three pre-training methods proposed in this paper, the models proposed in international conferences in recent years for the task of translation quality assessment are selected as the comparison systems, which are firstly briefly introduced.

QE-BERT: a pre-training method for the task of translation quality assessment, using a bilingual corpus, while inserting [GAP] representation intervals at the target language end, the pre-training task is to randomly mask the parallel corpus and restore it. The encoder used by QE-BERT is m-BERT, which is different from the transformer based on this paper.

NMT-style: uses NMT-style seq2seq model, and the pre-training task is to predict subwords of the target language using parallel corpus. The encoder is a bidirectional RNN or transformer, when the encoder is an RNN only simple attentional mechanisms are used to fuse the bilingual information, when the encoder is a transformer, no semantic association enhancement module or task is designed for the QE task.

NULL-Ins: Since there is a leakage problem in real translations, null characters are inserted into the parallel corpus, and the pre-training task is to randomly mask and restore on the parallel corpus where null characters are inserted, and the model needs to judge whether the masked part is a null character or a normally masked subword.

Gen-Det: A “generator-detector” model, whose pre-training task consists of mask prediction by the generator and translation quality detection by the detector, where the detector's pre-training data comes from the generator. The “generator-detector” is the highest performance model at this stage, but due to the weak generating ability of the generator, it is unable to generate diversified translation quality assessment data.

Table 2: Fuzzy decision attributes in German-Chinese translation

	R_1	R_2	R_3	R_4	R_5
X_1	1	0	1	1	1
X_2	1	0	1	0	1
X_3	1	0	1	0	0
X_4	0	1	01	1	1
X_5	1	0	1	1	0
X_6	1	1	1	01	1
X_7	0	0	1	1	0
X_8	1	1	0	1	1
X_9	01	0	0	1	1
X_{10}	1	1	1	0	1

III. B. German-Chinese Translation Fuzzy Decision Fusion and Performance Verification

On the basis of completing the construction of the dataset and the selection of the baseline model, this section further quantifies the semantic alignment accuracy of German-Chinese translations through the fuzzy decision fusion method, and verifies the robustness of this paper's model in complex semantic scenarios.

The performance testing experiment of German-Chinese translation model is established in C++ simulation environment, the order of AHP-Gray correlation model is set to 5, and the AHP-Gray correlation model is used to calculate the associated semantic defuzzification parsing factor to realize the fuzzy decision fusion of German-Chinese translations, and the obtained fuzzy decision attributes of German-Chinese translations are shown in Table 2.

Based on the results in Table 2, the association semantic fuzzification parsing factor is calculated, and the solution space adaptive optimization method is used to realize the accurate semantic alignment in German-Chinese translation and the optimal decision-making of German-Chinese translation. Fig. 5 Comparison results of association misalignment rate of German-Chinese translation using different methods.

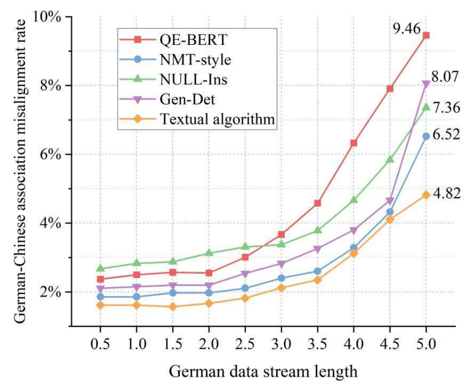

Figure 5: Comparison of correlation misalignment rates in German-Chinese translation

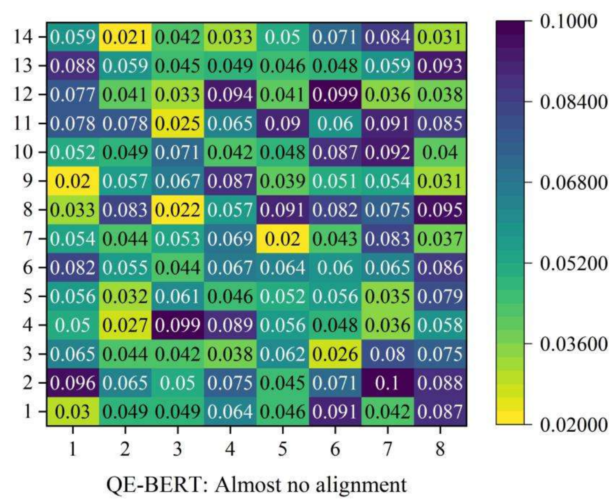

Figure 6: Heat map of attention head weight distribution in QE-BERT model

The analysis reveals that the German-Chinese translation using this paper's method has a low misalignment rate, when the length of the German data stream is 5×10^5, the misalignment rates of QE-BERT based on the encoder as m-BERT, NMT-style using seq2seq with NMT style, NULL-Ins, and the "Generator-Detector "Gen-Det model have German-Chinese association misalignment rates of 9.46%, 6.52%, 7.36% and 8.07% respectively, while the association misalignment rate under this paper's algorithm is only 4.82%, which is 49.05% lower than that of QE-BERT and close to half of that of the traditional model, indicating that the associative German semantic ontology structure model constructed in this paper performs German-Chinese translation with a high degree of association and better translation accuracy.

III. C. Word alignment performance of cross-language pretrained models

A very intuitive way to assess the semantic relevance ability of a model is to observe the weight distribution of the model's self-attention module, and a good attention weight distribution should have strong interpretability. Fig. 4 shows the heat map of this paper's visualization of attention, which contains the four models mentioned above, as well as this paper's hybrid model based on transformer machine translation.The heat maps of the attention head weight distribution of QE-BERT, NMT-style, NULL-Ins, Gen-Det and this paper's model are shown in Fig. 6, Fig. 7, Fig. 8, Fig. 9, and Fig. 10, respectively.

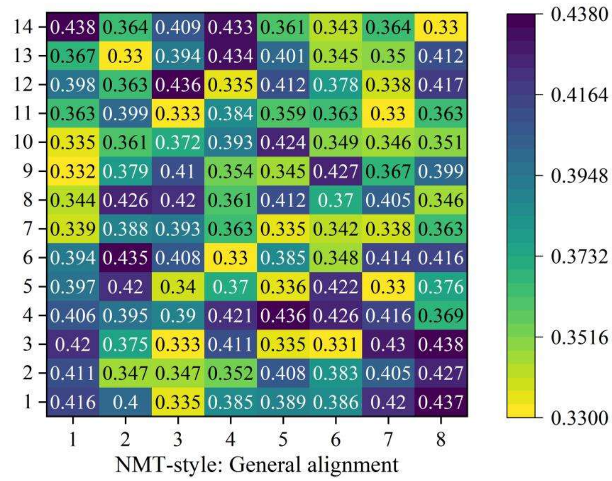

Figure 7: Heat map of attention head weight distribution in NMT-style model

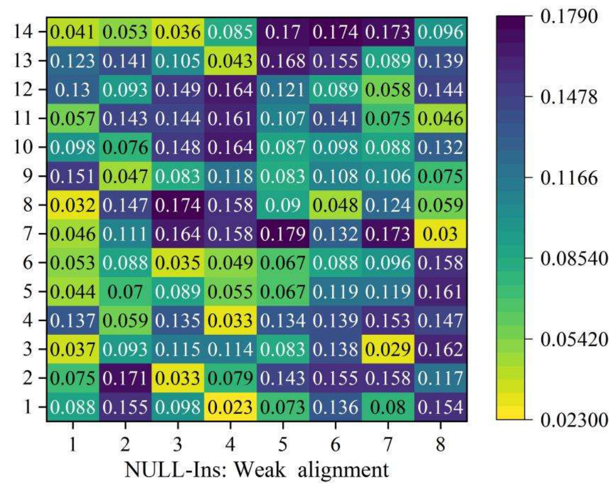

Figure 8: Heat map of attention head weight distribution in NULL-Ins model

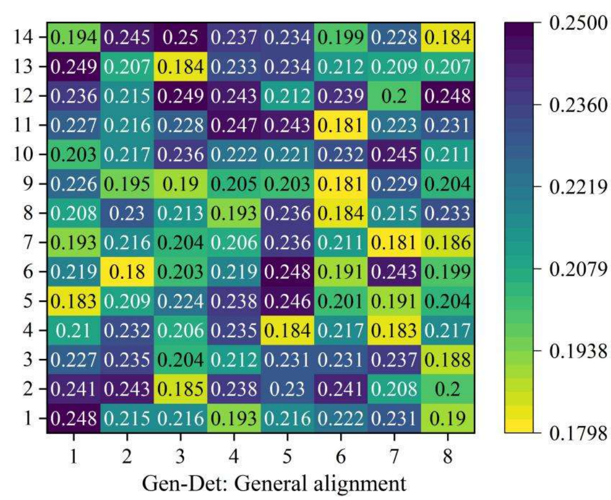

Figure 9: Heat map of attention head weight distribution in Gen-Det model

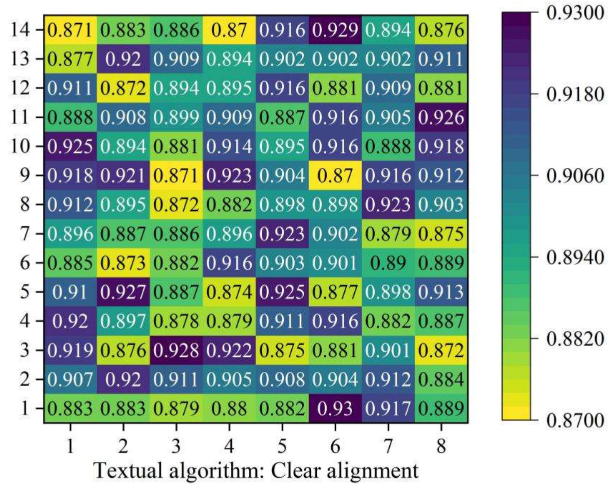

Figure 10: Heat map of attention head weight distribution in textual model

The heat map shows that the QE-BERT based on the encoder as m-BERT has almost no alignment, with an attention head weight distribution between 0.02-0.10; the NULL-Ins model has only weak alignment, with an attention head weight distribution between 0.023-0.179; the "generator- Detector" Gen-Det model has an average alignment phenomenon, with an attention head weight distribution between 0.18-0.25; the seq2seq model NMT-style, which uses NMT style, and like the models in this paper, the encoder uses a bidirectional RNN or transformer, has a better alignment phenomenon than the other three compared models , with attention head weight distribution between 0.33-0.438; while this paper's embedded the semantic relevance weights extracted from RNN and SAN encoders into the Transformer architecture of the multi-head self-attention mechanism model, obtaining a very clear bi-directional alignment phenomenon, with attention head weight distribution between 0.87-0.93.

III. D. *Analysis of the ratio of multilingual translation tasks to content words*

Based on the word alignment performance study of the cross-language pretrained model, this section further extends it to the multilingual translation task to explore the generalization ability of this paper's model in different language pairs and cultural contexts, and reveals its advantage in capturing key semantic elements through content word ratio experiments.

III. D. 1) Multi-translation task performance

The BLEU values of this paper's method and the baseline system on different datasets are shown in Table 3.

Table 3: Performance on several translation tasks with the BLEU metric

	German-English	German-Chinese	German-Spanish	German-French
QE-BERT	18.81	13.16	16.12	10.39
NMT-style	28.02	22.36	26.64	17.06
NULL-Ins	21.98	17.19	20.28	12.86
Gen-Det	24.96	20.05	23.79	14.39
Ours	31.86	26.21	29.14	20.92

Table 3 shows the comparison of the BLEU values of the different models in the four language groups for German-English, German-Chinese, German-Spanish and German-French. The data show that: this paper's model leads the way across the board: in the German-English task the BLEU value reaches 31.86, which is significantly higher than the baseline model, e.g., 28.02 for NMT-style; in the German-Chinese task, this paper's model far outperforms the other models, e.g., 20.05 for Gen-Det, with a BLEU value of 26.21. NMT-style performs the next best in the German-English task, with a value of 28.02, but its performance plummets 17.06 in the German-French task, indicating its lack of adaptability to distant language pairs; QE-BERT performs the worst among all tasks, with only 10.39 for German-French, highlighting the limitations of its cross-cultural semantic modeling.

The BLEU values of the German-Chinese task are generally lower than those of the German-English, reflecting the challenge of cultural differences between China and Germany for translation, but this paper's model still maintains a significant advantage through the mechanism of cross-language ontological knowledge sharing. This paper's model performs well in all the multilingual tasks, especially the advantage is more prominent in the language pairs of German-Chinese with greater cultural distance, which verifies the effectiveness of cross-cultural semantic association computation and hybrid model design.

III. D. 2) Proportion of content words

The selection of content words is obtained by a semantic association algorithm. It is obvious and easy to see that the performance of the model may be affected by the number of content words. Therefore, the effect of the number of content words on the model performance is verified by varying the proportion of content words present in the source language sentences. The German-English and German-Chinese datasets are selected for the study, and the experimental results of the performance of different proportions of content words in the German-English dataset and the German-Chinese dataset are shown in Fig. 11 and Fig. 12, respectively.

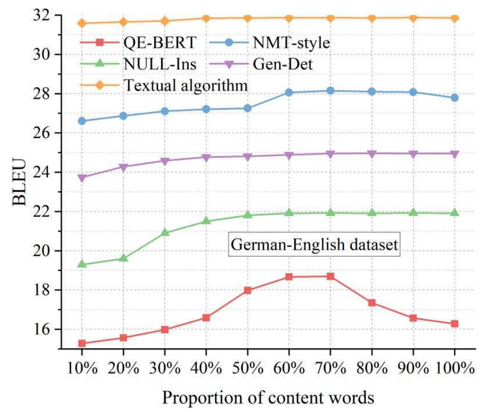

Figure 11: Performance of different proportions of content words on German-English

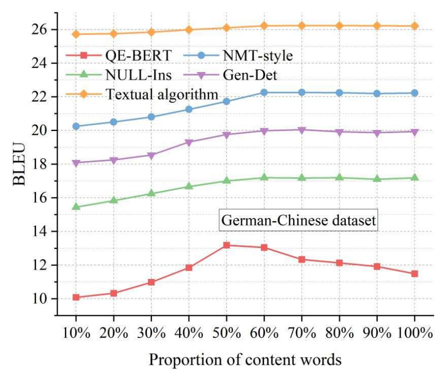

Figure 12: Performance of different proportions of content words on German-Chinese

Fig. 11 and Fig. 12 show the translation performance BLEU values of different models under the variation of the content word ratio of the source language sentences. The content word ratio reflects the density of key semantic information in a sentence; the higher the ratio, the more complex semantic elements such as cultural metaphors and logical associations in the sentence that need to be accurately translated.

In the German-English dataset, the method maintains the highest BLEU value in all content word ratios, and the performance fluctuates only slightly with the increase of content word ratio, and the BLEU value always stays stable at around 31.86, which shows its robustness to complex semantics.The performance of NMT-style, NULL-Ins, and QE-BERT reaches its maximum at 60% of content word ratio, and then either stabilizes or starts to fluctuate. After that, they either level off or start to decline, e.g., the QE-BERT model starts to decline when it reaches a BLEU value of 18.81, and the BLEU value is 16.28 when the content ratio is 100%, reflecting that it lacks cross-language semantic correlation modeling and is difficult to cope with the complexity of literary translation.

The model performance tends to be similar in the German-Chinese dataset, but the model performance is far less high than that in the German-English dataset, and the BLEU value of this paper's model stabilizes at 26.21 in the German-Chinese dataset.The German-Chinese task shows a more drastic decay of the baseline model performance due to the greater cultural distance, while this paper's model bucks the trend by the cross-language knowledge sharing mechanism.

IV. Conclusion

This study significantly improves the quality of German literature translation by constructing a cross-cultural semantic relevance computation model and a hybrid translation architecture.

The semantic association computation and cross-language ontology structure model based on RNN and SAN encoders reduces the association misalignment rate of German-Chinese translations to 4.82%, which is nearly half of the 9.46% of the traditional model, e.g., QE-BERT, and verifies the mitigating effect of fine-grained semantic analysis on the problem of cultural disconnection.

The hybrid model incorporating multiple pre-trained encoders has a BLEU value of 26.21 in the German-Chinese task, which is more than 30% improvement over the baseline model. With a German-English BLEU value of 31.86 in the multilingual task, its long-distance dependency processing ability and cross-cultural knowledge sharing mechanism are outstanding.

The content word ratio experiment shows that the BLEU value of this paper's model fluctuates only ±0.5 when the subsemantic density of the source utterance increases, and the German-English task stabilizes at 31.86, which is significantly better than the baseline model, confirming its ability to accurately capture literary metaphors and logical associations.

References

[1] Raikhan, S., Moldakhmet, M., Ryskeldy, M., & Alua, M. (2014). The interaction of globalization and culture in the modern world. Procedia-Social and Behavioral Sciences, 122, 8-12.

[2] Bukunmi Deborah, A. J. I. T. O. N. I. (2024). CROSS-CULTURAL EXCHANGE IN COMMUNITY THEATRE: GLOBAL PERSPECTIVES. Journal of Development Communication And Applied Theatre, 2(1).

[3] He, C. (2011). World drama and intercultural performance: Western plays on the contemporary Chinese stage. Neohelicon, 38(2), 397-409.

[4] Haiping, Y. (2015). Theater and Society: An Introduction to Contemporary Chinese Drama. In Theatre and Society: Anthology of Contemporary Chinese Drama (pp. ix-xlvi). Routledge.

[5] Goethe, J. W. V. (2013). Goethe, Johann Wolfgang von. Goethe's letters and letters to Goethe, 2, 1786-1805.

[6] Schweiger, C. (2022). „Da steh ich nun, ich armer Tor!“: Futures Literacy am Beispiel der Non-Futures Literacy von Goethes Faust. # schuleverantworten, 2(2), 160-168.

[7] Ibrahimov, T. (2015). A Comparative Study Based on the Author's Collection of Classical Music Masterpieces and Other Similar Classical Collections. Dirasat: Human & Social Sciences, 42(3).

[8] Pizer, J. (2022). Johann Wolfgang Goethe: Origins and Relevance of Weltliteratur. In The Routledge companion to world literature (pp. 3-10). Routledge.

[9] Purdy, D. (2014). Goethe, Rémusat, and the Chinese Novel: Translation and the Circulation of World Literature. German Literature as World Literature, 43-60.

[10] Chen, Y. (2015). Who is the Other? Goethe’s encounter with “China” in his Concept of Weltliteratur. In Major versus Minor?–Languages and Literatures in a Globalized World (pp. 241-252). John Benjamins Publishing Company.

[11] Jianguang, W. (2023). The Goethean Poetics in the Perspective of Seasonal and Phenological Culture: A Study of Goethe's Poem Chinese-German Book of Hours and Seasons. Comparative Literature in China, (3), 180.

[12] Uhlig, S. H., & Zhang, C. (2018, May). Introduction: Goethe, Worlds, and Literatures. In Seminar: A Journal of Germanic Studies (Vol. 54, No. 2, pp. 121-132). University of Toronto Press.

[13] Murnane, B. (2022). The Worldliness of Weltliteratur: Goethe's" Handelsverkehr" between China and Weimar. Goethe Yearbook, 29(1), 73-93.

[14] Zhang, C. (2017). Goethe’s Chinesisch-deutsche Jahres-und Tageszeiten: Vernacular Universal, Erotica Sinica, and the Temporality of Nachträglichkeit. In Reading the Past Across Space and Time: Receptions and World Literature (pp. 245-264). New York: Palgrave Macmillan US.