Hindawi
Mobile Information Systems
Volume 2022, Article ID 4863623, 10 pages
https://doi.org/10.1155/2022/4863623

Research Article

Research on Machine Translation Method of English-Chinese Long Sentences Based on Fuzzy Semantic Optimization

Zhaofeng Dong

School of Foreign Languages, Henan University of Urban Construction, Pingdingshan, Henan 467036, China

Correspondence should be addressed to Zhaofeng Dong; 30140506@hncj.edu.cn

Received 18 July 2022; Revised 8 September 2022; Accepted 3 October 2022; Published 11 October 2022

Academic Editor: Santosh Tirunagari

The machine translation of English-Chinese long sentences is out of order under the constraint of the mixed corpus. In order to improve the accuracy of machine translation of English-Chinese long sentences, a machine translation method based on fuzzy semantic optimization is proposed. By adopting the method of systematic semantic configuration structure fusion, the ontology mapping model of the English-Chinese long sentence subject list under the constraint of the mixed corpus is constructed. In the distribution structure space of English-Chinese long sentences under the constraint of the mixed corpus, the method of extracting the best semantic relevance and analyzing the concept tree of context semantic mapping is adopted to mine the dynamic features of the subject list of English-Chinese long sentence translation under the constraint of the mixed corpus. In the intrinsic distribution sequence of semantic mapping between concepts, the fuzzy semantic combination features of English-Chinese long sentences under the constraint of the mixed corpus are extracted. The distribution results of word knowledge priority selection sequence, the automatic translation of English-Chinese long sentences under the constraint of the mixed corpus is realized by scheduling a grammatical error correction task list. According to the simulation results, machine translation of long English-Chinese sentences under the restriction of mixed corpora is capable of good semantic configuration. The degree of linguistic structural similarity between English and Chinese is great, which enhances the automatic translation capacity of English-Chinese long sentences under the limitation of the mixed corpus.

1. Introduction

The syntactic causal structure of English-Chinese lengthy sentences is constrained by the mixed corpus [1], and the causative action and causative consequence are reflected by various lexical forms. The causative meaning of this kind of causative structure is not determined by causative verbs but by the result of the interaction between verbs in English-Chinese long sentences and other components in the structure under the constraint of the mixed corpus. As a result, verbs in English-Chinese long sentences under the constraint of the mixed corpus can be pure causative verbs, causative verbs with specific lexical meanings, and noncausative verbs, but they all have causative meanings after entering the corresponding sentence patterns, expressing the causative process in transitivity system. Events can express different process types in a transitivity system, and the process is caused by a causative external force and the combination of the two forms causative action process, semantic relationship process of English-Chinese translation of long sentences under the constraint of the causative mixed corpus, causative psychological process, etc.

According to the feature distribution and phase space feature distribution of English-Chinese long sentences under the constraint of the mixed corpus, a fusion model of English-Chinese long sentences under the constraint of the mixed corpus is established. Combined with the information feature extraction results of English-Chinese long sentences under the constraint of the mixed corpus, the automatic translation of English-Chinese long sentences under the constraint of the mixed corpus is realized. It is of great significance to study the translation methods of English-Chinese long sentences under the constraint of the mixed corpus to improve the accuracy of machine translation of English-Chinese long sentences.

Generative grammar overcomes the fact that generative semantics stays in the research of the meaning of causative verbs. The primary focus of the causative study is on the development and modification of causal structures with the intention of examining the degree to which form and language are universal. The explanatory capacity of its generalised one-sentence rule has several limits, and it does not pay enough attention to the pragmatic and semantic qualities of causal structures. Generally speaking, generative linguistics has carried out some research on the causative structure, but the description of syntactic and semantic features of the structure is not thorough and systematic. Automatic translation of English-Chinese long sentences under the constraint of the mixed corpus is based on optimized clustering and feature fusion analysis of data. Among the traditional methods, machine translation methods of English-Chinese long sentences [2, 3] under the constraint of the mixed corpus mainly include machine translation method of English-Chinese long sentences under the constraint of the mixed corpus based on correlation degree feature analysis, resource dynamic translation method of PCA principal component analysis [4], and machine translation method of English-Chinese long sentences under the constraint of mixed corpus of K-means fusion clustering [5]. Statistical feature extraction and autocorrelation feature detection are used to realize machine translation of English-Chinese long sentences under the constraint of the mixed corpus. However, the traditional method of machine translation of English-Chinese long sentences under the constraint of the mixed corpus has poor adaptability and weak feature recognition ability.

To solve the above problems, a machine translation method of English-Chinese long sentences under the constraint of fuzzy semantic optimization mixed corpus is proposed. Ontology is a discipline of philosophy that deals with the study of many things that exist in the world. Its primary objective is to categorise all of the things that exist. From the perspective of computer science, ontology is a specification of a conception. In this sense, ontologies aim to characterise concepts already present in a domain and relate them based on those concepts' traits [6]. Firstly, the ontology mapping model of English-Chinese long sentences under the constraint of the mixed corpus is constructed by the method of systematic semantic configuration and structure fusion, and the semantic mapping between concepts is in the intrinsic distribution sequence. Then, the fuzzy semantic combination feature quantity of English-Chinese long sentences under the constraint of the mixed corpus is extracted. According to the distribution result of the word knowledge priority selection sequence, the automatic translation of English-Chinese long sentences under the constraint of the mixed corpus is realized by scheduling the grammatical error correction task list. Finally, the simulation analysis shows the superior performance of this method in improving the machine translation ability of English-Chinese long sentences under the constraint of the mixed corpus.

2. Concept Tree Model of Upper and Lower Semantic Mapping of Machine Translation

2.1. Machine Thesaurus Construction. In order to optimize the machine translation of English-Chinese long sentences under the constraint of the mixed corpus, the natural semantic processing method [7] is adopted to construct the subject word list of English-Chinese long sentence machine translation under the constraint of the mixed corpus. By combining structural information with ontology mapping, the graph model parameters for English-Chinese long phrase machine translation under the restriction of mixed corpus are established. Figure 1 illustrates how semantic editing is used to create the topic word list for English-Chinese long sentence machine translation under the restriction of mixed corpus.

According to the keyword list distribution of English-Chinese long sentence machine translation restricted by the mixed corpus shown in Figure 1, using the data enhancement method of error correction model back translation, the training samples of machine translation are as follows:

$$D = \left\{ S_{i,j}(t), T_{i,j}(t), U_{i,j}(t) \right\}, \tag{1}$$

where in, $S_{i,j}(t)$ represents the correlation dimension [8] of the word context of English-Chinese long sentence machine translation under the constraint of the mixed corpus, $T_{i,j}(t)$ represents the similarity propagation graph model of English-Chinese long sentence machine translation under the constraint of the mixed corpus, and $U_{i,j}(t)$ represents the fuzzy similarity feature quantity (correlation) of English-Chinese long sentence machine translation under the constraint of the mixed corpus [8]. $O = \{C, H^C, R, I, A\}$ is defined as a five-tuple ontology structure model of English-Chinese long sentence machine translation under the constraint of the mixed corpus. The data enhancement model parameters of English-Chinese long sentence machine translation under the constraint of the mixed corpus are constructed by using an ontology mapping method, and the output of neurogrammatical error correction is as follows:

$$S_{i,j}(t) = \frac{p_{i,j}(t) - sp_{i,j}(t)}{p_{i,j}(t)}. \tag{2}$$

For the basic unit of knowledge storage of semantic ontology in machine translation, the following correction vector set representing semantic ambiguity in the process of machine translation of English-Chinese long sentences under the constraint of the mixed corpus is calculated and expressed as follows:

$$T_{i,j}(t) = \frac{\left| p_{i,j}(t) - \Delta p(t) \right|}{p_{i,j}(t)}, \tag{3}$$

where in, $U_{i,j}(t)$ is used to define the frequent item sets [9, 10] of semantic autocorrelation in machine translation of English-Chinese long sentences under the constraint of the mixed corpus, and ontology integration is used to exchange knowledge and search keyword information in the

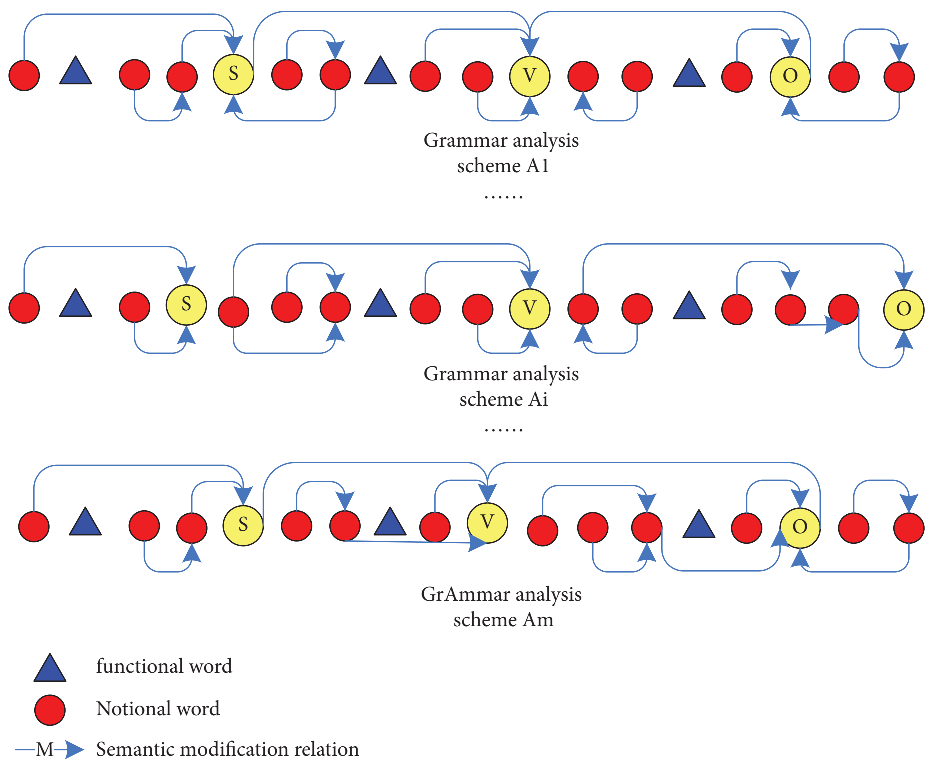

Figure 1: List of keywords in machine translation of English-Chinese long sentences under the constraint of the mixed corpus.

translation process so as to realize the correlation analysis measure of machine translation of English-Chinese long sentences under the constraint of the mixed corpus, and its calculation formula is as follows:

$$U_{i,j}(t) = \exp\left[-b\left[z_i(t) - z_j(t)\right]^2\right], \quad (4)$$

where in, $p_{i,j}(t)$ is the semantic similarity of machine translation of English-Chinese long sentences under the constraint of mixed corpus; $sp_{i,j}(t)$ is the multistrategy similarity in machine translation of English-Chinese long sentences under the constraint of the mixed corpus; $\Delta p(t)$ combines the reference values for the Bayesian network model [11] of English-Chinese long sentence machine translation under the constraint of the mixed corpus; $z_i(t)$ and $z_j(t)$ respectively represent ontology mapping of machine translation thesaurus. According to the above analysis, the overall structure model of machine translation of English-Chinese long sentences under the constraint of the mixed corpus is shown in Figure 2.

2.2. Concept Tree of Context Semantic Mapping of Machine Translation. The establishment of semantic mapping between machine translation ontologies of English-Chinese long sentences under the constraint of the mixed corpus needs to consider the context relationship of machine translation. The context mapping method is used to construct the concept tree model of machine translation. The multi-fuzzy semantic automatic judgment is performed via model. The big data fusion cluster analysis method is used to realize dynamic resource reorganisation and feature screening control [12]. The phase space distribution structure model of English-Chinese long sentences under the constraint of the mixed corpus is constructed. By using the method of fuzziness detection, the characteristic scalar time series of English-Chinese long sentences under the constraint of the mixed corpus is $x(t), t = 0, 1, \cdots, n-1$. Given the numerical attribute and classification attribute feature quantity of English-Chinese long sentences under the constraint of the mixed corpus, the distribution sequence of fuzzy association rules [13] of English-Chinese long sentences under the constraint of the mixed corpus is $x_1, x_2, \cdots x_n \in C^m$ (m-dimensional complex space). In the fuzzy clustering center, the sparsity feature points of English-Chinese long sentences under the constraint of the mixed corpus are $P_i = (p_{i1}, p_{i2}, \cdots p_{iD})$, where

$$j \in N_i(k), N_i(k) = \left\{\left\|x_j(k) - x_i(k)\right\| < r_d(k)\right\}. \quad (5)$$

In the distributed subspace of English-Chinese long sentences under the constraint of the mixed corpus, considering the relevance of resource attributes, the fuzzy set [14] of English-Chinese long sentences under the constraint of the mixed corpus is obtained by the method of fuzzy association rules scheduling, and the semantic correlation function of grammar analysis with the largest semantic correlation of English-Chinese long sentences under the constraint of the mixed corpus is obtained as follows:

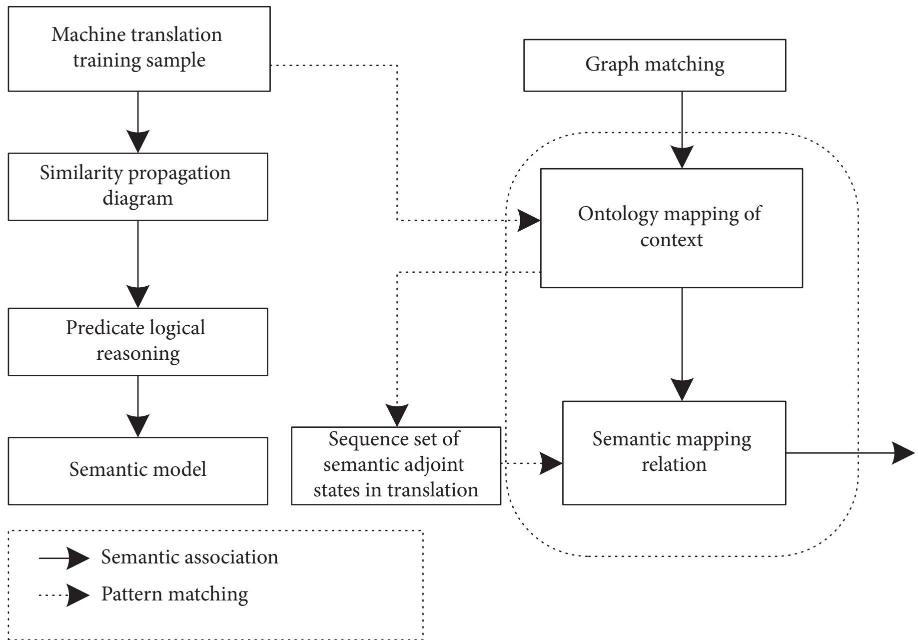

FIGURE 2: Structure model of machine translation of English-Chinese long sentences under the constraint of the mixed corpus.

$$f(k) = \begin{cases} f(k-1) - \frac{1}{n}, & 1 \leq k < n, \\ 1, & k = n. \end{cases} \tag{6}$$

The decision tree Wi of residual data segmentation is constructed under the mixed corpus constraint by weighting the semantic modification target between two words of English-Chinese long phrases. Under the I-th grammar analysis scheme, the concept set [15] of context semantic mapping of English-Chinese long sentence allocation semantic tree vocabulary semantic library under the constraint of the mixed corpus can be recorded as follows:

$$D''_{i,j}(t_{n+1}) = \frac{D'_{i,j}(t_{n+1}) + f(n)D'_{i,j}(t_n)}{2}. \tag{7}$$

Under the constraint of the mixed corpus, the main features of prepositional function words in English-Chinese translation are extracted as follows:

$$f_{Ai} = K_{SVO} * (match(S,V) + match(O,V)) + \sum_{i=1}^{n} match(W_i, W_{Gi}). \tag{8}$$

The fuzzy clustering [16, 17] of English-Chinese long sentences under the mixed corpus of the i-th classification attribute is taken, the quantitative regression analysis results of English-Chinese long sentences under the mixed corpus is obtained, and the semantic similarity at time $T$and time T+1 is analyzed, finds all clauses, and calculates the maximum semantic relevance of English-Chinese long sentences under the mixed corpus as follows:

$$I_{i,j}(t) = \frac{\sum D''_{i,k}(t)D''_{k,j}(t)}{\sum D''_{i,k}(t)}. \tag{9}$$

In the context of semantic mapping of English-Chinese long sentence translation under the constraint of the mixed corpus, the simple semantic unit of English-Chinese long sentence attribute under the constraint of the premixed corpus is expressed as follows:

$$\widehat{S}_w = \sum_{i=1}^{c} p_i 1/n_i \sum_{k=1}^{n_i} \left[\left(\overrightarrow{X}_k^{(i)} - \overrightarrow{m}_i \right) \left(\overrightarrow{X}_k^{(i)} - \overrightarrow{m}_i \right)^T \right]. \tag{10}$$

The binary structure feature distribution set of English-Chinese long sentences constrained by the mixed corpus is established in the virtual database, and the context semantic mapping features of simple semantic units of English-Chinese long sentences constrained by the mixed corpus are arranged in descending order as follows:

$$S\left(\overrightarrow{X}_1\right) \geq S\left(\overrightarrow{X}_2\right) \geq \cdots \geq s\left(\overrightarrow{X}_l\right). \tag{11}$$

The conceptual tree structure model of context semantic mapping for machine translation of English-Chinese long sentences under the constraint of the mixed corpus is obtained by different boundary division schemes, as shown in Figure 3.

In Figure 3, the target attribute values of the semantic modification of English-Chinese long sentences are mapped into the semantic mapping idea tree, and the fuzziness of sentences is automatically assessed. In the experimental process, the dynamic feature mining of the subject word list in English-Chinese long sentence translation under the constraint of the mixed corpus is realized by the methods of extracting the best semantic relevance

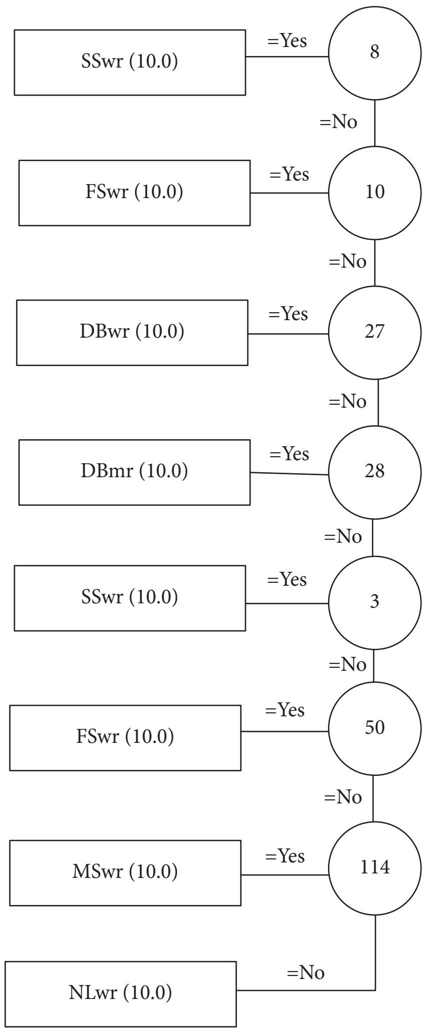

Figure 3: Conceptual tree structure of context semantic mapping of English-Chinese long sentences under the constraint of the mixed corpus.

and analyzing the concept tree of contextual semantic mapping. In the intrinsic distribution sequence of semantic mapping between concepts, the high-order spectrum of English-Chinese long sentences constrained by the mixed corpus is extracted, and it is transformed into the progressive semantic feature analysis of simple semantic units and the adaptive adjoint tracking recognition [18, 19] of machine translation. According to a microlevel analysis of the semantic components of English and Chinese causative resultative constructions, the semantic properties of English and Chinese causative semantic elements are relatively similar, but the syntactic units that reflect the semantic components are different. The sorts of transitivity processes, semantic configuration structure, and syntactic functions of English and Chinese causative resultants are found to differ significantly on a macro level when the functional semantic syntax of these resultants is examined. English only expresses the causative process but lacks the causative action process and the causative psychological process, which is related to the nature of syntactic units that embody English and Chinese causative results. After semantic analysis and machine translation modification in the C4.5 decision tree, the schematic diagram of the process of converting long sentences into short sentences is shown in Figure 4.

3. Realization of Machine Translation of Long Sentences

3.1. Rule Reduction of Machine Translation of Long Sentences from English to Chinese. In the distribution structure space of English-Chinese long sentences under the constraint of the mixed corpus, the method of extracting the best semantic relevance and analyzing the concept tree of context semantic mapping is adopted to mine the dynamic features of the keyword list in English-Chinese long sentence translation under the constraint of the mixed corpus. In the intrinsic distribution sequence of semantic mapping between concepts, a fuzzy semantic feature extraction model of text information is established. The reduction rule function of machine translation of English-Chinese long sentences under the constraint of the mixed corpus is described as follows:

$$\begin{aligned} x(2k+1) &= (x(2k)+x(2k+2))*a+x(2k+1)x(2k) \\ &= (x(2k-1)+x(2k+1))*b+x(2k)x(2k+1) \\ &= (x(2k)+x(2k+2))*c+x(2k+1)x(2k) \\ &= (x(2k-1)+x(2k+1))*d+x(2k). \end{aligned} \tag{12}$$

The English-Chinese long phrase machine translation feature decomposition formula is expressed as an S, V, and O decomposition operation after numerous iterations of target clause creation. The maximum spanning tree matrix A of each simple semantic is then calculated, which is a real matrix of SD. English-Chinese double object structure includes double-name structure and dative structure, which is set as the target clause, with M-order orthogonal semantic reduction rule matrix U and N-order orthogonal matrix V. After SVM decomposition, the weight of English-Chinese long sentence machine translation under the constraint of the mixed corpus is determined.

$$A = USV' = U\begin{pmatrix} \sum & 0 \\ 0 & 0 \end{pmatrix}V', U*U' = I, V*V' = I \tag{13}$$
$$\sum = \operatorname{diag}(\sigma_1, \sigma_2, \sigma_r)\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0,$$

where, A is $m \times n$ matrix, $A*A'$ and $A'*A$ are the square root of information fuzzy semantic feature extraction in the lexical-semantic calculation of English-Chinese long sentences under the constraint of the mixed corpus. The rule information set of English-Chinese long sentence translation under the constraint of the mixed corpus is described as follows:

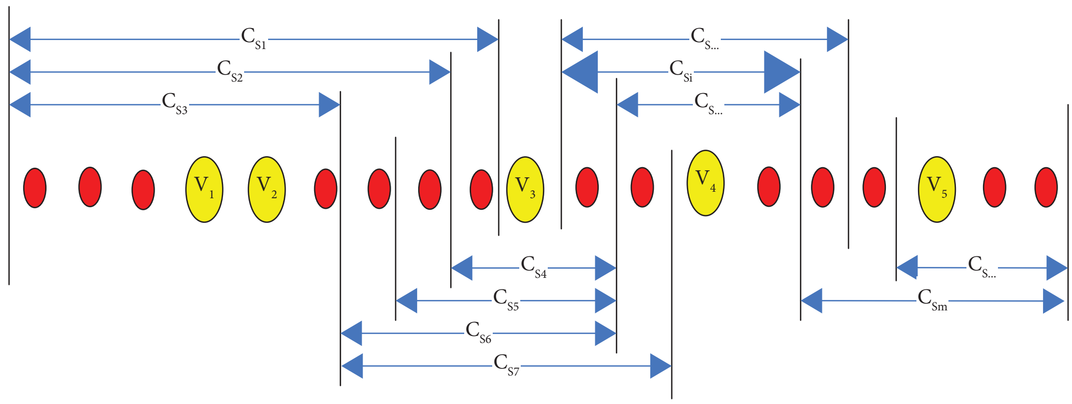

Figure 4: Schematic diagram of the process of converting long sentences into short sentences.

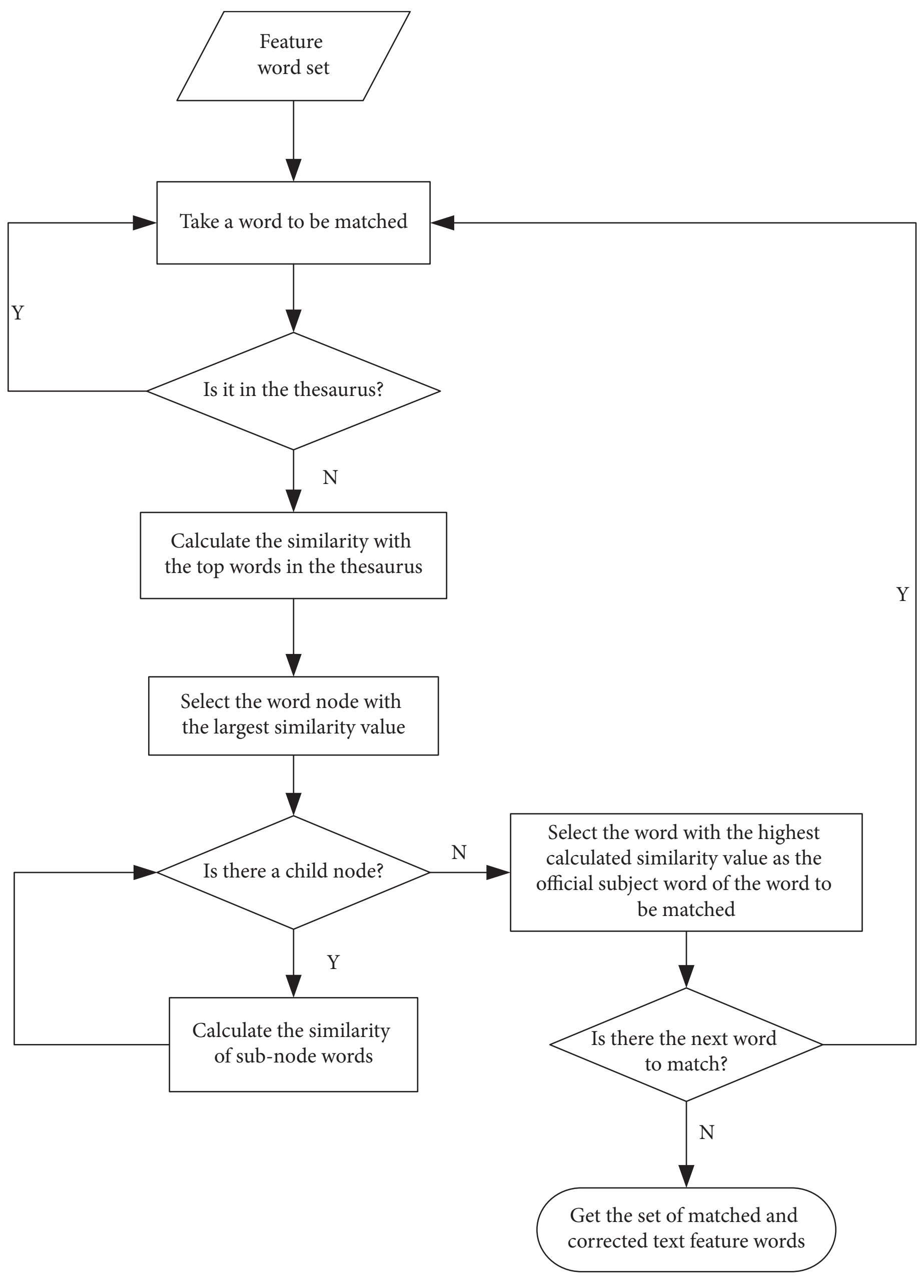

Figure 5: The process of machine translation of English-Chinese long sentences under the constraint of the mixed corpus.

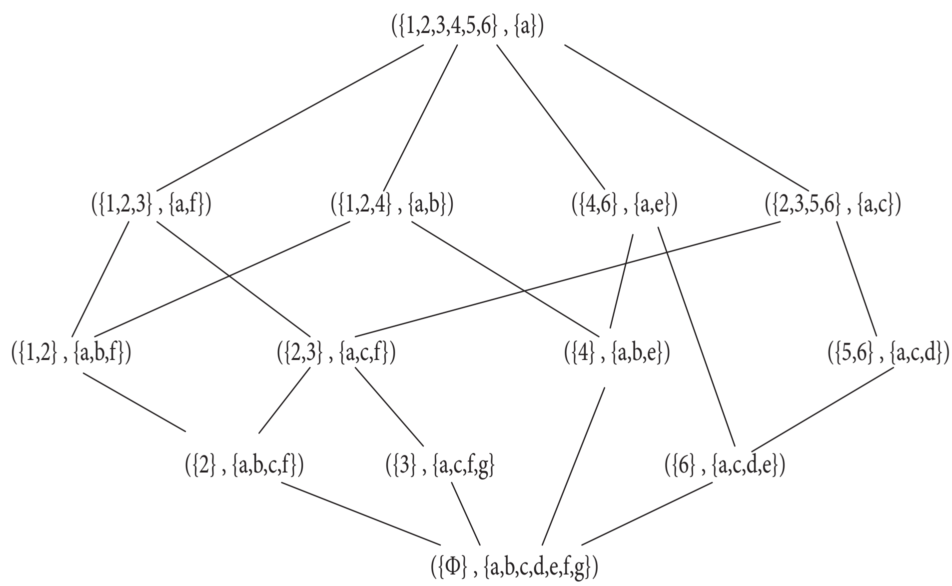

Figure 6: Parameter setting of concept tree of context semantic mapping.

$$u(t) = A(t)\exp[j\theta(t)] = A(t)\exp\left[-j2\pi K\ln\left(1-\frac{t}{t_0}\right)\right], \quad (14)$$

wherein, $A(t)$ is the envelope amplitude of vocabulary meaning of English-Chinese long sentences under the constraint of the mixed corpus. According to the multi-transformation method, the maximum utility fitness function of English-Chinese long sentences under the constraint of the mixed corpus is calculated. According to unsupervised machine learning, the information fusion scheduling model of English-Chinese long sentences under the constraint of the mixed corpus is s_h^w, and the global optimal solution $E(T_k^w - \xi_k^w(\omega) \mid T_k^w \geq \xi_k^w(\omega))$ of dynamic attribute translation of English-Chinese long sentences under the constraint of each mixed corpus can be expressed as follows:

$$f_i(t) = \frac{1}{2\pi}\frac{\mathrm{d}\theta}{\mathrm{d}t} = \frac{K}{t_0 - t}. \quad (15)$$

The attributive clauses of machine translation are summed up, the best grammar analysis is carried out, and the reduction rule function of English-Chinese long sentences under the constraint of the mixed corpus is constructed as follows:

$$u_a(t) = \frac{1}{\sqrt{a}}A\left(\frac{t}{a}\right)\exp\left[-j2\pi K\ln\left(1-\frac{t}{at_0}\right)\right]. \quad (16)$$

The fuzzy semantic combination features of English-Chinese long sentences are extracted under the constraint of the mixed corpus, and the rule reduction output of machine translation of English-Chinese long sentences is as follows:

$$f_{i,a}(t) = \frac{K}{at_0 - t} = \frac{K}{t_0 - [t + (1-a)t_0]}. \quad (17)$$

Binary structure combination control is adopted, $\tau^* = (1-a)t_0$ is adopted, and the reduction output of English-Chinese long sentence machine translation is restricted by the mixed corpus.

$$f_{i,a}(t) = f_i(t + \tau^*). \quad (18)$$

In the intrinsic distribution sequence of semantic mapping between concepts, the fuzzy semantic combination feature quantity of English-Chinese long sentences constrained by the mixed corpus is extracted. According to the distribution result of word knowledge priority selection sequence, multifuzzy semantic automatic judgment is made to provide accurate information input basis.

3.2. Step Description of English-Chinese Translation of Long Sentences under the Constraint of the Mixed Corpus. The fuzzy semantic feature [20] decomposition of English-Chinese long sentences under the constraint of the mixed corpus is described as follows:

(1) According to the thesaurus of multi-fuzzy semantic object set O of English-Chinese long sentences under the constraint of the mixed corpus, the concept subset clauses of English-Chinese long sentence translation under the constraint of the mixed corpus are obtained.
(2) Select clauses of dynamic matching words in English-Chinese long sentences under the constraint of the mixed corpus for S, V, and O decomposition and determine reasonable weights to obtain several simple main sentence units of multi-fuzzy semantics in English-Chinese long sentence translation under the constraint of the mixed corpus.
(3) Calculate the semantic correlation values of common attributes of fuzzy semantic feature units of English-Chinese long sentences under the constraint of the

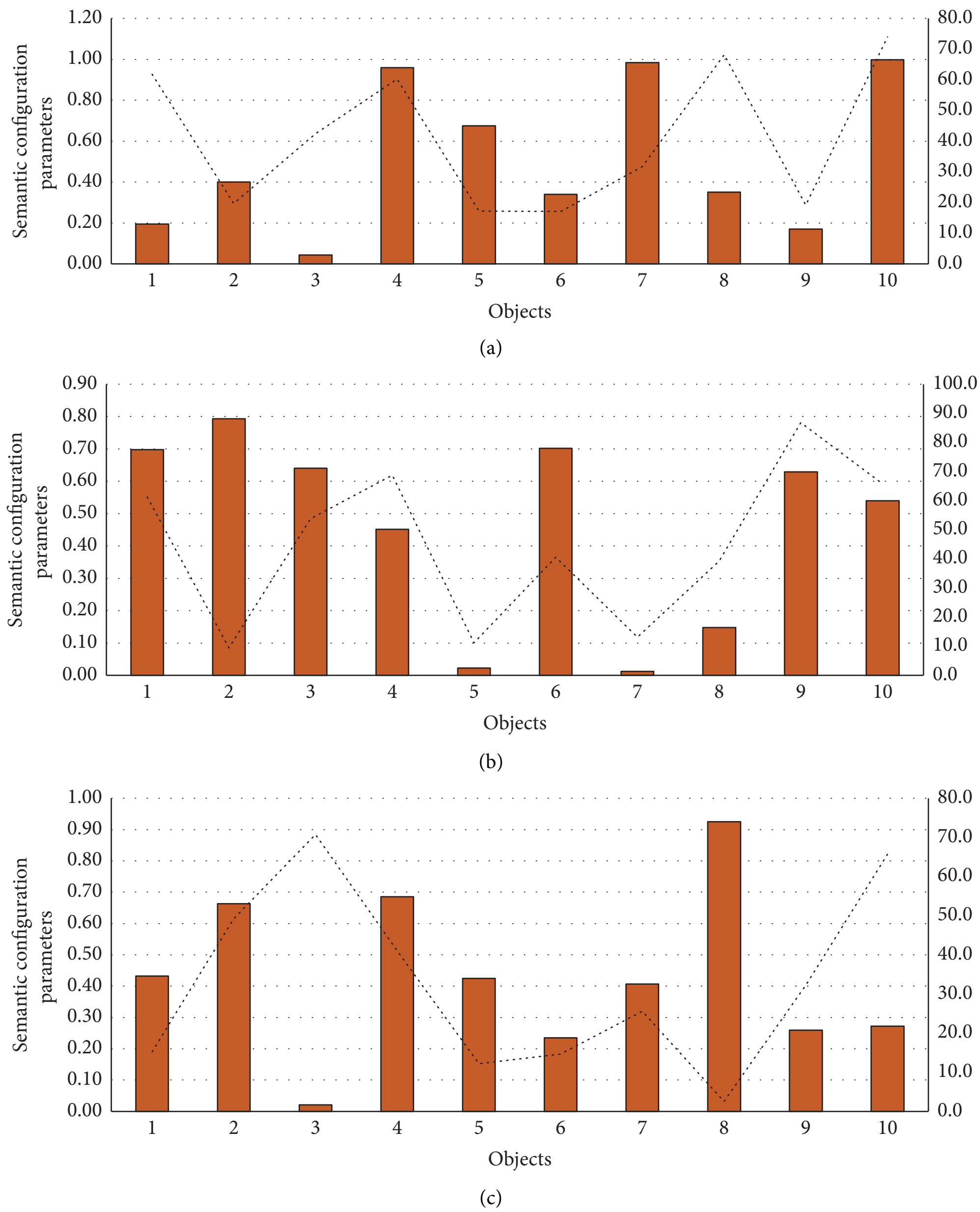

Figure 7: The results of semantic linear fitting of English-Chinese long sentences under the constraint of the mixed corpus. (a) Corpus A. (b) Corpus B. (c) Corpus C.

mixed corpus, and define the fuzzy semantics of English-Chinese long sentence translation under the constraint of the mixed corpus as g(I): = {*o* O| A I, *o* R A}.

(4) Automatically judging and searching according to the topmost node in the subject word list of English-Chinese long sentences under the constraint of the mixed corpus and calculating the characteristic word value of the object set O of the concept tree of English-Chinese long sentences under the constraint of the mixed corpus.

(5) Reduce the matched and corrected text to a vocabulary. For the concept sets *o* O, A A of English-Chinese long sentences restricted by the mixed corpus, if the convergence conditions are met, the machine translation results obtained are fuzzy, so go back to step 2 and redefine the basic units of the thesaurus; otherwise, go to the next step.

(6) According to the matching algorithm of the subject word list of English-Chinese long sentences under the constraint of the mixed corpus, the multi-fuzzy semantic automatic judgment of machine translation [21] of English-Chinese long sentences under the constraint of the mixed corpus is realized, and the corresponding best grammar analysis result is obtained.

(7) Adjust the headings (predicates) of English-Chinese long sentences under the constraint of the mixed

corpus, automatically register and judge the fuzzy semantics of English-Chinese long sentences under the constraint of the mixed corpus, and get the weight coefficient KS of English-Chinese long sentences under the constraint of the mixed corpus and make an experimental comparative analysis.

According to the above analysis, the flow chart of machine translation of English-Chinese long sentences under the constraint of the mixed corpus is shown in Figure 5.

4. Experimental Test Analysis

The programming language JAVA 1.5.4 is used as the programming software for multi-fuzzy semantic analysis and judgment programming in machine translation, and the development environment is Eclipse 3.4.2. Under the constraint of the tested mixed corpus, the machine translation of English-Chinese long sentences comes from the position-TAg1 text database, and the information attribute set consisting of A, B, C, and D4 semantic features is obtained by selecting $k = 4$. Set the dimension of spatial dynamic distribution of English-Chinese long sentences under the constraint of the mixed corpus as 12, the sample size of English-Chinese long sentences under the constraint of the mixed corpus as 1200, the training set as 100, the iteration times of fuzzy clustering of English-Chinese long sentences under the constraint of the mixed corpus as 300, and use ICTCLAS2015 machine translation software for batch Chinese-English machine translation. The number of characteristic words $(1'', 1') = (\{1, 4\}, \{a, r, f\})$, $(1'', 2') = (\{2\}, \{a, c, d, e\})$ in machine translation of English-Chinese long sentences under the constraint of the mixed corpus, and the concept tree of context semantic mapping of English-Chinese long sentences under the constraint of the mixed corpus is set. Parameter setting of concept tree of context semantic mapping is shown in Figure 6.

According to the above simulation environment and parameter settings, the multi-fuzzy semantic automatic judgment of machine translation of English-Chinese long sentences under the constraint of the mixed corpus is carried out, and the results of semantic linear fitting of English-Chinese long sentences under the constraint of the mixed corpus are shown in Figure 7.

The analysis of Figure 7 shows that, when limited by a mixed corpus, machine translation of English-Chinese long sentences has a strong semantic allocation capacity. The matching degree of English-Chinese language structure types is discovered using the results of the linear fitting of English-Chinese long sentences with the mixed corpus constraints. By using this method and traditional methods, the comparison results of semantic matching degree of machine translation of English-Chinese long sentences under the constraint of the mixed corpus are shown in Figure 8.

Analysis of the simulation results in Figure 8 shows that the semantic matching degree of English-Chinese long sentence machine translation under the constraint of the mixed corpus is high, and the feature registration rate of context semantic mapping is greatly improved, which is superior and improves the accuracy of English-Chinese long sentence machine translation under the constraint of the mixed corpus [20].

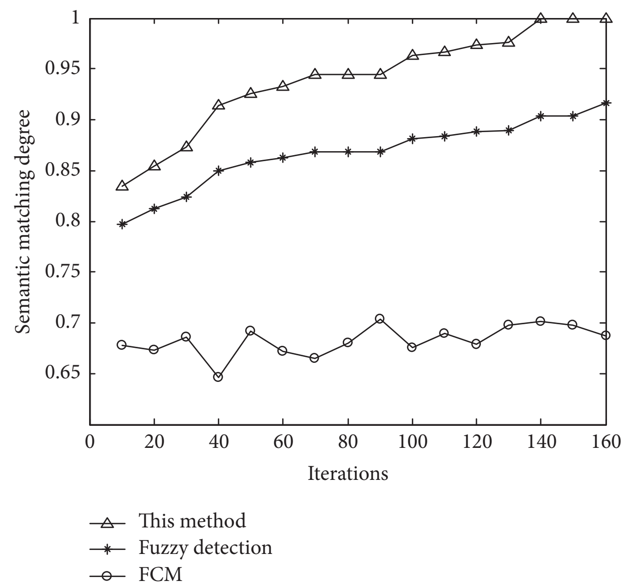

Figure 8: Comparison of feature registration rates of semantic mapping.

5. Conclusions

In this paper, a machine translation method of English-Chinese long sentences based on fuzzy semantic optimization is proposed. The English-Chinese long sentence subject list ontology mapping model is created under the restriction of the mixed corpus by using the method of systematic semantic configuration structure fusion. In the distribution structure space of English-Chinese long sentences under the constraint of the mixed corpus, the method of extracting the best semantic relevance and analyzing the concept tree of context semantic mapping is adopted to mine the dynamic features of the subject list of English-Chinese long sentence translation under the constraint of the mixed corpus. In the intrinsic distribution sequence of semantic mapping between concepts, the fuzzy semantic combination features of English-Chinese long sentences under the constraint of the mixed corpus are extracted, and according to the distribution results of word knowledge priority selection sequence, the automatic translation of English-Chinese long sentences under the constraint of the mixed corpus is realized by scheduling grammatical error correction task list. The simulation results show that the semantic configuration ability of machine translation of English-Chinese long sentences under the constraint of the mixed corpus is good, and the matching degree between English and Chinese language structures is high, which improves the automatic translation ability of English-Chinese long sentences under the constraint of the mixed corpus. This method has a good application value in the design of machine translation software for English-Chinese long sentences under the constraint of the mixed corpus [22].

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 there are no conflicts of interest.

References

[1] M. Singh, R. Kumar, and I. Chana, "Corpus based machine translation system with deep neural network for Sanskrit to Hindi translation," *Procedia Computer Science*, vol. 167, pp. 2534–2544, 2020.

[2] A. Fan, S. Bhosale, H. Schwenk, and C. B. Sinli, "Beyond English-centric multilingual machine translation," *Journal of Machine Learning Research*, vol. 22, no. 107, pp. 1–48, 2021.

[3] S. Lyons, "A review of Thai–English machine translation," *Machine Translation*, vol. 34, no. 2-3, pp. 197–230, 2020.

[4] V. Srinivasarao and U. Ghanekar, "Speech enhancement - an enhanced principal component analysis (EPCA) filter approach," *Computers & Electrical Engineering*, vol. 85, Article ID 106657, 2020.

[5] J. T. Camino, C. Artigues, L. Houssin, and S. Mourgues, "MILP formulation improvement with k-means clustering for the beam layout optimization in multibeam satellite systems," *Computers & Industrial Engineering*, vol. 158, Article ID 107228, 2021.

[6] M. Kundi and R. Chitchyan, "Use case elicitation with FrameNet frames," in *Proceedings of the In2017 IEEE 25th international requirements engineering conference workshops (REW)*, pp. 224–231, Lisbon, Portugal, September 2017.

[7] K. Shuang, Z. Zhang, J. Loo, and S. Su, "Convolution–deconvolution word embedding: an end-to-end multi-prototype fusion embedding method for natural language processing," *Information Fusion*, vol. 53, pp. 112–122, 2020.

[8] Z. Dhifaoui and J. M. Bardet, "Local correlation dimension of multidimensional stochastic process," *Statistics & Probability Letters*, vol. 181, Article ID 109262, 2022.

[9] G. Coletti and B. Bouchon-Meunier, "A study of similarity measures through the paradigm of measurement theory: the fuzzy case," *Soft Computing*, vol. 24, no. 15, pp. 11223–11250, 2020.

[10] S. Maruf, F. Saleh, and G. Haffari, "A survey on document-level neural machine translation: methods and evaluation," *ACM Computing Surveys*, vol. 54, no. 2, pp. 1–36, 2022.

[11] R. Kasperė, J. Horbačauskienė, J. Motiejūnienė, V. Liubiniene, I. Patasiene, and M. Patasius, "Towards sustainable use of machine translation: usability and perceived quality from the end-user perspective," *Sustainability*, vol. 13, no. 23, Article ID 13430, 2021.

[12] J. Roostaei, S. Colley, R. Mulhern, A. A. May, and J. M. Gibson, "Predicting the risk of GenX contamination in private well water using a machine-learned Bayesian network model," *Journal of Hazardous Materials*, vol. 411, Article ID 125075, 2021.

[13] D. Fawzy, S. Moussa, and N. Badr, "The spatiotemporal data fusion (STDF) approach: IoT-based data fusion using big data analytics," *Sensors*, vol. 21, no. 21, p. 7035, 2021.

[14] J. Sanz, M. Sesma-Sara, and H. Bustince, "A fuzzy association rule-based classifier for imbalanced classification problems," *Information Sciences*, vol. 577, pp. 265–279, 2021.

[15] M. Zishan Anwar, S. Bashir, and M. Shabir, "An efficient model for the approximation of intuitionistic fuzzy sets in terms of soft relations with applications in decision making," *Mathematical Problems in Engineering*, vol. 2021, Article ID 6238481, 19 pages, 2021.

[16] T. Hempel and A. Al-Hamadi, "An online semantic mapping system for extending and enhancing visual SLAM," *Engineering Applications of Artificial Intelligence*, vol. 111, Article ID 104830, 2022.

[17] F. Salehi, M. R. Keyvanpour, and A. Sharifi, "GT2-CFC: g," *Information Sciences*, vol. 578, pp. 297–322, 2021.

[18] E. Bas and E. Egrioglu, "A fuzzy regression functions approach based on Gustafson-Kessel clustering algorithm," *Information Sciences*, vol. 592, pp. 206–214, 2022.

[19] S. De Vito, G. Di Francia, E. Esposito, S. Ferlito, F. Formisano, and E. Massera, "Adaptive machine learning strategies for network calibration of IoT smart air quality monitoring devices," *Pattern Recognition Letters*, vol. 136, pp. 264–271, 2020.

[20] B. N. Davis and R. J. LeVeque, "Analysis and performance evaluation of adjoint-guided adaptive mesh refinement for linear hyperbolic PDEs using c," *ACM Transactions on Mathematical Software*, vol. 46, no. 3, pp. 1–28, 2020.

[21] R. Cantini, F. Marozzo, G. Bruno, and P. Trunfio, "Learning sentence-to-h semantic mapping for hrm," *ACM Transactions on Knowledge Discovery from Data*, vol. 16, no. 2, pp. 1–26, 2022.