The Application and Practice of Computer-Aided Translation in English Teaching

Youming Xu [1]

[1]Wuxi Institute of Technology, Wuxi, 214121, China

Corresponding author: Youming Xu, xu_youm1999@163.com

Abstract: In order to improve the improvement effect of English teaching, this paper applies computer aided translation to English teaching, and proposes a semantic similarity vector space model. Moreover, this paper applies the Cell-Type Membrane Computational Optimization Algorithm (CMCOA) to the topic crawler. CMCOA is used to optimize the weighting factor of the unvisited URLs priority value formula. In addition, this paper selects the optimal object through the communication rules and evolution rules of each membrane, proposes a topic crawler based on semantic understanding and intelligent learning to improve the semantic analysis effect of English translation, and builds a computer aided translation system. The research shows that the computer translation proposed in this paper has better application effect in English teaching and can effectively improve the quality of English teaching.

Keywords: computer; assisted translation; English teaching; application
DOI: https://doi.org/10.14733/cadaps.2023.S12.216-230

1. INTRODUCTION

Because the computer has a strong storage capacity, it can greatly reduce the labor required for translation and improve the efficiency of translation. Moreover, more and more translators have begun to pay attention to intelligent translation software, and are trying to apply artificial intelligence software to translation practice, such as Dao Translation and Baidu Translation.

Translation memory can greatly improve translation quality. Limited by the computer-aided translation environment of a single machine, it cannot provide users with the limitation of translation libraries. However, the computer-aided translation environment based on the cloud platform can provide users with large-capacity translation memory storage capacity in the network environment. The application of translation memory function allows users to establish consistent language expression habits, unified professional terminology and the expression form of proper names when translating the same type of corpus. In addition, the realization effect of some high-quality translation memory can also track and record the contextual environment and subject awareness of the corpus information to achieve higher-quality translation effects.

Computer-aided translation software has two core technologies, one of which is translation memory. When a translator translates a sentence, the translation memory of computer-assisted translation will automatically search for the sentence that has been translated and stored in the memory. If the original sentence or similar sentence is searched, the system will directly extract the translation result, or give it to Similar reference translation [2]. Translators can directly use the translation results provided by the system, or make adjustments according to the system prompts, and the newly added translation content will be stored in the translation memory for future reference. Therefore, the storage unit of computer-aided translation software computer-aided translation is a segment rather than a word or phrase, and compared with the previous machine translation, the syntactic and semantic levels have been greatly improved. According to investigation and research, the translation memory of computer-aided translation technology can save about 40% of the cost in just one year, and nearly 60% in three years [15].

Translation termbase is another core technology of computer-aided translation software computer-aided translation, which mainly uses SDLMultiterm to establish terminology [5]. What is stored in the translation termbase is not the common bilingual parallel phrase corpus, but the less common terminology. In the traditional translation mode, the translator needs to look up different specialized dictionaries or thesaurus to translate specialized terms, and each time he needs to repeat the lookup to ensure the accuracy, and after looking up, he needs to return to the translation interface [9]. Using computer-aided translation software, computer-aided translation, translators can directly use the translation terminology database to search for related terms in the translation operation interface during the translation process, and do not need to switch out, because the computer-aided translation is equipped with Multiterm terminology database [4] . Such an operating system greatly improves translation efficiency and greatly reduces the time for consulting professional vocabulary. In addition, termbases ensure uniformity of terminology. For example, when the workload of a translation project is large and the project needs to be assigned to several translators to translate at the same time, there will often be inconsistencies in the translation of professional terms. According to the survey and research, most of the wasted translation cost comes from rework. Using computer-aided translation software, project managers can publish a unified translation termbase and share it with all translators participating in the project. The translation termbase that the translators refer to during translation is the same, so the translation can also maintain consistency and professionalism [17] .

Master the basics of computer-assisted translation software. At this stage, these translation memory-related software are constantly developing and expanding. If you only master some translation software of lower quality, it is difficult to keep up with the trend of market development even if you are very proficient in using it. Therefore, students should focus on learning the basics of translation machines. Principles and how to use these translation software, and continue to study the concepts and ideas of these translation software design, students should improve their ability to use translation software for translation work, organically combine language services and translation skills, and master corpus, intelligent Editing and translation technology and other concepts, actively exploring the language habits and text styles of the translated sentences, and summarizing some universal translation rules through research on syntactic patterns, coherent forms, and lexical norms, so that they can be effectively selected to adapt to translation strategy [7]. This is enough to show that it is very important for students to master the necessary principles and skills of computer-assisted translation related software. The teaching mode should be student-centered [20]. The teaching goal of computer-assisted translation is to cultivate students' ability to learn independently and solve practical problems while cultivating students' ability to use various translation software and work [19]. In the process of teaching, teachers should fully pay attention to the subject status of students, and can allow students to practice on a certain translation topic in groups, so that in the process of actual translation exercises, students can not only exercise autonomous learning and solve practical problems. Moreover, students have fully played their main role and improved their translation skills through cooperation and communication. Not only that, group cooperation also enhances students' communication skills and teamwork awareness, which plays a positive role in cultivating comprehensive talents [11]. At the same time, teachers should also carefully select the template materials used in teaching, and should adjust the materials used appropriately according to the actual situation, and try to select template materials with relatively standardized format and content, and will be regularly updated and improved [16]. Reasonable use of Internet teaching resources. In our daily study, we often use network resources such as CNKI, Baidu Encyclopedia, and Interpretation Network. On these websites, students can search for materials related to translation in time [13]. For example, while providing students with the latest translation materials, Interpretation Network also has many test preparation resources, translation recording software, and translation-related note-taking methods. Cloud software is also a popular and convenient network resource in recent years. It can be used without a network. Cloud software can also provide huge computer resources, which can be recycled after the user completes the function. In addition, there are some cloud-based translation software that can translate in a remote form, which facilitates mutual communication during the translation process, thereby improving translation quality and translation efficiency [6].

The new technologies involved in translation activities are very wide, including word processing technology, format conversion technology, typesetting technology, query technology, machine translation technology, translation memory technology, corpus technology, terminology management technology, localization technology, translation project management technology, data Security technology, etc., even includes some basic computer hardware technology [18]. Faced with a dazzling array of software, teachers should not adopt a comprehensive and punctual teaching model, but should make an objective and fair comparison, identify the pros and cons, and choose a representative mainstream translation aid as the teaching content [12].

The application of computer technology can provide great help for translation, and on this basis, the research direction of computer-assisted translation is derived. Cloud platform is a computer technology that has developed very rapidly in recent years, and has made great progress in many application fields [14]. Using the cloud platform can solve the contradiction between computing resources, storage resources and implementation costs of a single network end user, and provide users with a high-quality computing and storage environment without adding too much cost to users. Applied to the field of translation [10].

An important technical support for computer-assisted translation is translation memory capability. In the traditional translation process, because the translator has translation memory ability, when translating the same type of corpus, as the translation content continues to increase, his grasp of the content of the corpus becomes more and more accurate, and the obtained translation The quality can also be gradually improved [1]. Using computer-assisted translation, it is also possible to create storage data of translation memory in the user's computer for reference in subsequent translations. However, the translation memory capability that a single computer can provide is very limited, and the representation, storage, organization and utilization of translation memory data are very complex, not simply stacking the words or sentences in translation. Using the powerful storage and computing power of the cloud platform, a large-capacity translation memory can be established at the back end of the cloud platform, and complex translation processing programs can be run to meet the needs of deep translation memory [3].

Translation memory can improve the quality of translation very well. Due to a stand -alone computer -assisted translation environment, it cannot provide users with restrictions on translation libraries. The computer -based computer -aided translation environment based on the cloud platform can provide users with large -capacity translation in the network environment Memory storage ability. Applying the function of translation memory allows users to translate the same type of corpus, establish a consistent language expression habits, unified professional terms, and the form of proprietary names before and after. Track and record the context and theme consciousness of corpus information to achieve higher quality translation effects.

In order to improve the effect of English teaching, this paper applies computer-aided translation to English teaching, and constructs an intelligent English teaching system to improve the effect of English teaching.

2. SEMANTIC TOPIC CRAWLER BASED ON MEMBRANE COMPUTING OPTIMIZATION ALGORITHM

Definition 1 is that the term space is a set of basic terms, which is defined as follows:

$$\text{TSpace} = \{ \text{term}_i | \text{term}_i \in \text{Lex}, 1 \le i \le |\text{Lex}| \} \tag{2.1}$$

In the formula, TSpace is the term space, Lex is the basic vocabulary set, $term_i$ is the terms in Lex, and $|Lex|$ is the total number of terms in the basic vocabulary set Lex.

Definition 2 is that the score of the term vector corresponds, which is defined as follows:

$$\text{TVector} = \left(\text{val}_1, \text{val}_2, \cdots, \text{val}_{|\text{Lext}} \right) \tag{2.2}$$

In the formula, TVector is the term vector.That is, the dimension of the term vector TVector is equal to the total number of terms $|Lex|$ in the term space TSpace.

Definition 3 is that the semantic space is a set of two-term items, which is defined as follows:

$$SSpace = \{ dbt_l | \ dbt_l = (term_i, term_j), term_i \in Lex, term_j \in Lex, \\ 1 \le i \le |Lex|, 1 \le j \le |Lex|, 1 \le l \le |Lex|^2 \} \tag{2.3}$$

In the formula, SSpace is the semantic space, and Lex is the same as the above, that is, the basic vocabulary set. dbt_l is a two-term tuple and corresponds to terms $term_i$、$term_j$, where $term_i$、$term_j$ are all elements in Lex. The SSpace dimension of the semantic space is $|Lex|^2$.

Definition 4 is that the scores of the semantic vector all correspond to the semantic space, which is defined as follows:

$$\text{SVector} = \left(\text{val}_1, \text{val}_2, \cdots, \text{val}_{|\text{Leet}|^2} \right) \tag{2.4}$$

In the formula, SVector is the semantic vector,That is, the dimension of the semantic vector SVector is equal to the total number of double-term items $|Lex|^2$ in the semantic space SSpace.

Dcuments and topics obtain document terms and topic terms through text processing operations such as stop word removal and stemming. The weight of each term can be calculated by the term frequency inverse document frequency (TF-IDF) equation, and these term weights can be converted into document and topic term vectors, as follows:

$$\text{DTS}_k = \{ \text{term}_1, \text{term}_2, \cdots, \text{term}_n \}, \quad \overrightarrow{d_k} = (w_{k1}, w_{k2}, \cdots, w_{kn}, 0, 0, \cdots, 0) \\ \text{TTS} = \{ \text{term}_1, \text{term}_2, \cdots, \text{term}_m \}, \quad \vec{t} = (w_{t1}, w_{t2}, \cdots, w_{tm}, 0, 0, \cdots, 0) \tag{2.5}$$

In the formula, DTS_k、TTS are the term sets of document d_k and topic t, respectively, $\overrightarrow{d_k}$、$\vec{t}$ are the term vector of document d_k and topic t, respectively, and $w_{ki}(1 \le i \le n)$ is the TF-IDF value of term i in document d_k. $w_{tj}(1 \le j \le m)$ is the TF-IDF value of the term j in topic t, n and m are the total number of terms in document d_k and topic t, and the 0 value in the term vector $\overrightarrow{d_k}$ corresponds to the value of the term in the term space TSpace but not in the term set DTS_k. The number of 0 values is the dimension of the term vector $\overrightarrow{d_k}$ minus the total number of terms.The number of 0 values is the dimension of the term vector $\vec{t}$ minus the total number of terms in topic t, which is $|Lex|-m$.

	$\overrightarrow{DSV}$	$\overrightarrow{TSV}$
book platform	1.128	0.066
book lesson	0.901	0.471
book document	1.449	0.069
review platform	0.535	0.066
review lesson	0.385	0.042
review document	0.602	0.06

$$\begin{array}{ccccc} & \text{term}_1 & \text{term}_2 & \cdots & \text{term}_m \\ \text{term}_1 & \text{Sem}_{11}^k & \text{Sem}_{12}^k & \cdots & \text{Sem}_{1m}^k \\ \text{term}_2 & \text{Sem}_{21}^k & \text{Sem}_{21}^k & \cdots & \text{Sem}_{2m}^k \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \text{term}_n & \text{Sem}_{n1}^k & \text{Sem}_{n2}^k & \cdots & \text{Sem}_{nm}^k \end{array} \quad (2.6)$$

The details are as follows:

$$\overrightarrow{DSV}_k = \left(\overrightarrow{dsv_1}, \overrightarrow{dsv_2}, \cdots, \overrightarrow{dsv_n}, \vec{0}\right) \quad \overrightarrow{dsv_i} = \left(dva1_{(i-1)^* m+1}, dva1_{(i-1)^* m+2}, \cdots, dval_{i+m}\right)$$
$$\overrightarrow{TSV}_k = \left(\overrightarrow{tsv_1}, \overrightarrow{tsv_2}, \cdots, \overrightarrow{tsv_n}, \vec{0}\right) \quad \overrightarrow{tsv_i} = \left(tva1_{(i-1)^{m+1},}, \ \text{tva}_{(i-1)^* m+2}, \cdots, \ \text{tval}_{i*m}\right)$$
$$dbTS_k = \left\{dbt_r \middle| \ dbt_r = \left(\text{term}_p, \text{term}_q\right), \text{term}_p \in DTS_k, \text{term}_q \in TTS, 1 \le p \le n, 1 \le q \le m\right\}$$
$$dbTS_{ki} = \left\{dbt_r \middle| \ dbt_r = \left(\text{term}_i, \text{term}_j\right), \text{term}_i \in DTS_k, \text{term}_j \in TTS, 1 \le j \le m\right\}$$
$$dbTS_{ki} \subseteq dbTS_k \quad \vec{0} = (0, 0, \cdots 0)_{1_{1\times\left(|Letl^2 - m^* n\right)}}$$
$$dva1_{(i-1)^* m+j} = w_{ki} * \text{Sem}_{ij}^k \quad tva1_{(i-1)^* m+j} = w_{tj}{}^* \text{Sem}_{ij}^k \quad (1 \le i \le n, 1 \le j \le m) \quad (2.7)$$

The similarity between d_k and topic t is the cosine value of the semantic vector $\overrightarrow{DSV}_k$ of document d_k and the semantic vector $\overrightarrow{TSV}_k$ of topic t, and the specific expression is as follows[20]:

$$\mathrm{Sim}\left(d_k,t\right)=\overrightarrow{DSV}_k\cdot\overrightarrow{TSV}_k=\frac{\sum_{i=1}^{n}\sum_{j=1}^{m}w_{ki}w_{tj}\left(\mathrm{Sem}_{ij}^{k}\right)^2}{\sqrt{\sum_{i=1}^{n}\left(w_{ki}^2\sum_{j=1}^{m}\left(\mathrm{Sem}_{ij}^{k}\right)^2\right)}\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m}w_{tj}^2\left(\mathrm{Sem}_{ij}^{k}\right)^2}} \qquad (2.8)$$

In Figure 1, the term vector and semantic vector of the above example document d_1 and topic t are represented in the three-dimensional subspace coordinate system of the term space and the semantic space, respectively. The correlation between document d_1 and topic t can be obtained by the above formula 8, that is, the specific expression is as follows:

$$\overrightarrow{DSV}_1=(0.83,0.448,0.933,0.739,0.394,0.844,0.243,0.162,0.243,0,0,\cdots,0)_{1\times|\mathrm{Lex}|^2}$$

$$\overrightarrow{TSV}_1=(0.066,0.032,0.06,0.074,0.035,0.069,0.06,0.035,0.048,0,0,\cdots,0)_{1\times|\mathrm{Lex}|^2} \qquad (2.9)$$

$$\mathrm{Sim}\left(d_1,t\right)=\overrightarrow{DSV}_1\cdot\overrightarrow{TSV}_1=0.964$$

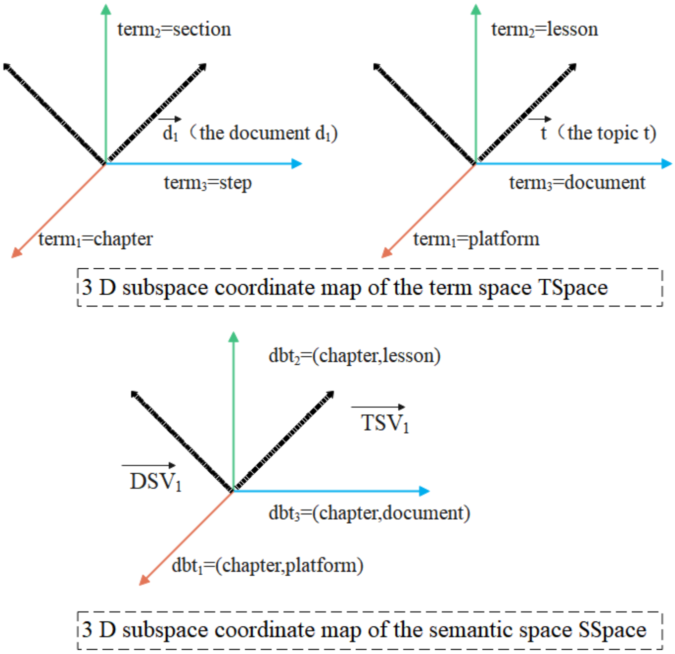

Figure 1: Three-dimensional subspace coordinate diagram of term space and semantic space.

The structure of CMCOA can be represented by "[]" nesting, each "[]" represents a membrane, and its specific representation is as follows:

$$\mu = [_1[_2[_4]_4[_5]_5[_6]_6]_2[_3[_7]_7[_8]_8]_3]_1 \tag{2.10}$$

In the formula, μ is the structure of CMCOA, and the number represents the number of the membrane. The structure μ of this CMCOA can be represented in Figure 2 below:

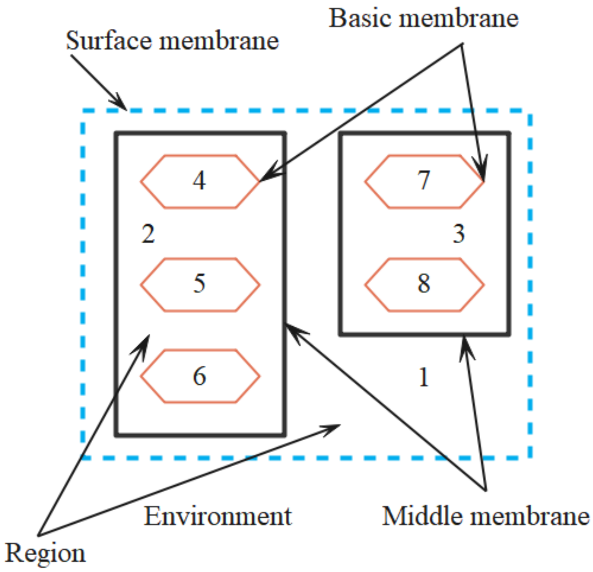

Figure 2: Structure of CMCOA.

In Figure 2 above, film 1 is the surface film, films 2 and 3 are intermediate films, films 4, 5, 6, 7, and 8 are basic films, the outside of the surface film 1 is the environment, and the inside of each film is the region.

The object and fitness function of CMCOA are specifically expressed as follows:

$$M_k = \text{random}\{o_{k1}, o_{k2}, \cdots, o_{kN_k}\} \quad k = 1, 2, \cdots, 8$$

$$o_{kj} = \left(\lambda_{j1}^k, \lambda_{j2}^k, \lambda_{j3}^k, \lambda_{j4}^k\right) \quad fit_j = \frac{1}{\sqrt{\sum_{i=1}^{N}\left(y_{it} - y_{ic}\right)^2}}$$

$$y_{ic} = \lambda_{j1}^k \text{Sim}_{i1} + \lambda_{j2}^k \text{Sim}_{i2} + \lambda_{j3}^k \text{Sim}_{i3} + \lambda_{j4}^k \text{Sim}_{i4} \quad (1 \le i \le N) \tag{2.11}$$

In the formula, o_{kj} is the weighting factor of the objects j, λ_{j1}^k、 λ_{j2}^k、 λ_{j3}^k、 λ_{j4}^k in the membrane k corresponding to the full text, pin text, title level, and context, respectively.

Sim_{i1}、Sim_{i2}、Sim_{i3}、Sim_{i4} are the training values of the topic relevance of full text, pin text, title, and context, respectively, and N is the total number of training data.
The selection rules of CMCOA are specifically expressed as follows:

$$FatherGeneration:\left[o_{kj}\right]_k \xrightarrow{f} NextGeneration:\left[o_{kj}\right]_k$$
$$f: P_{rand} \le P_{sj}, \quad P_{sj} = \frac{\sum_{r=1}^{j} fit_r}{\sum_{i=1}^{N_k} fit_i} \tag{2.12}$$

In the formula, o_{kj} is the object j in the membrane k, f is the condition for the object o_{kj} to be selected into the next generation, P_{rand} is a random value of [0, 1], and P_{sj} is the probability value of the object o_{ij} being selected. fit_r、 $fit_i\left(1 \le r \le j, 1 \le i \le N_k\right)$ are the fitness values of object o_{kr} and object o_{ki}, respectively, and N_k is the total number of objects in the membrane k.
The crossover rules and mutation rules of CMCOA are specifically expressed as follows:

$$\left[o_{ki}, o_{kj}\right]_k \xrightarrow{f_c} \left[o_{ki}^{'}, o_{kj}^{'}\right]_k \quad f_c: P_{crand} \le P_{ci}, P_{cj} \quad \left[o_{kh}\right]_k \xrightarrow{f_m} \left[o_{kh}^{'}\right]_k \quad f_{mi}: P_{mrand} \le P_{mh}$$
$$o_{ki}^{'} = \alpha o_{ki} + (1-\alpha) o_{kj} \quad o_{kj}^{'} = \alpha o_{kj} + (1-\alpha) o_{ki} \quad o_{kh}^{'} = o_{kh} + \beta(1,1,1,1) \tag{2.13}$$

Each object corresponds to a four-dimensional vector and the vector scores are in [0, 1], f_c、 f_m are the conditions for crossover and mutation of individuals in the membrane k, respectively. P_{ci}、 P_{cj} are the probability of crossover of object o_{ki} and object o_{kj} in membrane k, respectively, and P_{mh} is the probability of mutation of object o_{kh} in membrane k. P_{crand}、 P_{mrand} are randomly generated numbers and the value range is $[0,1]$, α、 β are all given constants and the value range is also within [0, 1].

The specific calculation expressions of the crossover probability P_{ci}、 P_{cj} and the mutation probability P_{mh} in the above formula (13) are as follows:

$$P_{cg}=\begin{cases} r_c\dfrac{fit_{\max}-fit_g}{fit_{\max}-fit_{\text{avg}}} & fit_g\geq fit_{\text{avg}} \\ r_c & fit_g<fit_{avg}\end{cases}\quad (g=i,j)\quad P_{mh}=\begin{cases} r_m\dfrac{fit_{\max}-fit_h}{fit_{\max}-fit_{\text{avg}}} & fit_h\geq fit_{\text{avg}} \\ r_m & fit_h<fit_{avg}\end{cases}$$

$$r_c=\frac{P_{c\max}-P_{c\min}}{(1+\exp(-aU(t)))}+P_{c\min}\quad r_m=\frac{P_{m\max}-P_{m\min}}{(1+\exp(-aU(t)))}+P_{m\min}\quad a=9.903438$$

$$U(t)=\frac{2H}{\log_2(1+t)\log_2 N_k}-1\quad H=-\sum_{d=1}^{M}P_d\log_2 P_d\quad P_d=\frac{n_d}{N_k}\left(M\leq N_k\right)\quad fit_{\text{avg}}=\frac{\sum_{l=1}^{N_k}fit_l}{N_k}\tag{2.14}$$

In the formula, P_{cg} is the crossover probability P_{ci} of the object o_{ki} in the membrane k or the crossover probability P_{cj} of the object o_{kj}, and the mutation probability of the object o_{kh} in the $P_{m,h}$ membrane k.

The communication rules of CMCOA are specifically expressed as follows:

$$\left[o_{kj}\right]_k\rightarrow[]_k o_{kj}\quad i=1,4,5,6,7,8\quad f_1: fit_j\geq fit_i\quad\left(1\leq i\leq N_k\right)$$

$$\left[o_{kj}\right]_k\xrightarrow{f_2}[]_k o_{kj}\quad i=2,3\quad f_2:\begin{cases} fit_j<fit_{old} & \text{and } P_{rand}\leq P_{rj} \\ fit_j\geq fit_{old}\end{cases} orfit\geq fit\quad\left(1\leq i\leq N_k\right)\tag{2.15}$$

This condition means that the top and base films only send the best objects (corresponding to the largest fitness value) into the external environment or outer film containing them. f_2 is the condition for the objects in the interlayer to communicate, P_{rand} is a randomly generated number in the range of [0,1], and P_{rj} is the communication probability of the object o_{kj} in the interlayer k, that is, the probability value of the object o_{kj} being sent.

In the above formula (15), the above three values correspond to fit_j、fit_{old}、P_{rj} respectively, the first two fitness values can be calculated according to the above formula (11), and P_{rj} is mainly calculated with reference to the energy distribution function in physics. In the literature, this energy distribution function is specifically expressed as follows:

$$P_r=\exp\left(\left(E_f-E\right)/\left(k_B T\right)\right)\tag{2.16}$$

In the formula, P_r is the probability of electrons occupying the energy E, E_f is the Fermi level, k_B is the Boltzmann constant, and T is the thermodynamic temperature.

Using the current evolutionary algebra to replace the thermodynamic temperature, the exchange probability P_{rj} of the object o_{kj} in the interlayer k can be obtained, and the specific expression is as follows:

$$P_{rj} = 1 - \exp\left(\left(ft_j - fit_{old}\right)/g\right) \tag{2.17}$$

In the formula, fit_j is the current fitness value of the object o_{kj} in the intermediate film k, fit_{old} is the maximum fitness value of all objects in the previous generation film k, and g is the current exchange algebra.
The algorithm flowchart of CMCOA is shown in Figure 3.

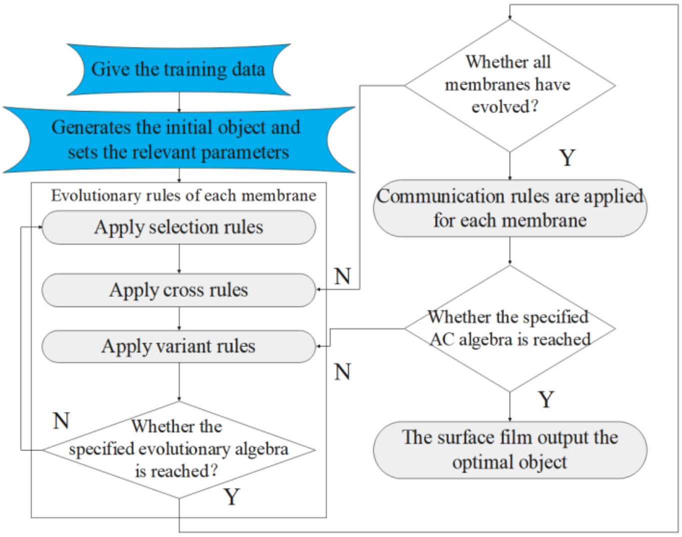

Figure 3: Algorithm flow chart of CMCOA.

A topic crawler based on semantic understanding and intelligent learning is proposed. This topic crawler uses the full text, anchor text, title, and context of the web page as hyperlinked documents, and uses SSVSM to calculate the similarity between the above four texts and the topic. At the same time, it uses CMCOA to optimize the calculation of four weighting factors in the hyperlink's topic relevancy equation, and linearly combines these four optimized weighting factors with the topic similarity of the corresponding four texts. Finally, this combined value is used as the topic relevance of the hyperlink, which is also called the ranking priority value of the hyperlink, as shown in Figure 4.

In Figure 4, the subject crawler system based on semantic understanding and intelligent learning is mainly divided into four modules: weighting factor optimization, related numerical calculation of terms, subject similarity calculation of documents, and sorting priority value calculation.The optimization weighting factor module mainly uses the evolution rules and communication rules of CMCOA to optimize four weighting factors, which correspond to the weighting factors of full text, anchor text, title and context respectively. The structure diagram of this module is shown in Figure 5 below. In Figure 5 below, the training data mainly includes training URLs related to the topic, its own topic similarity, and the topic similarity of its parent web page's full text, anchor text, title, and context. The relevant numerical module for obtaining terms is mainly to calculate the term TF-IDF weights of the four documents that have not visited URLs. At the same time, it calculates the semantic similarity between four document terms and topic terms, and obtains four semantic similarity matrices accordingly. The structure of this module is shown in Figure6 below. In Figure 6, document and topic preprocessing mainly includes stop word removal and stemming.

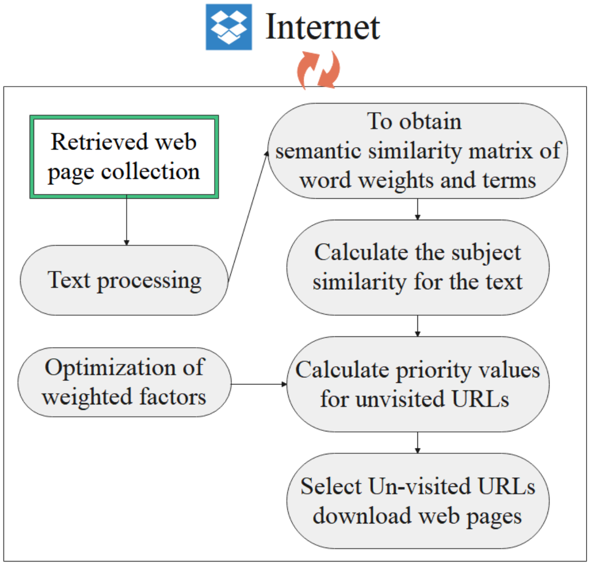

Figure 4: Structure diagram of crawler system based on semantic understanding and intelligent learning.

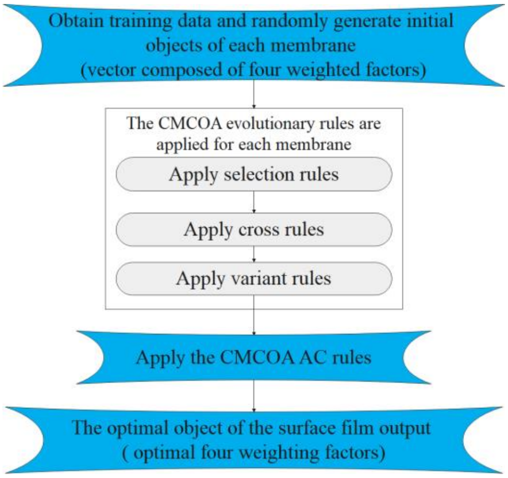

Figure 5: Structure diagram of optimization weighting factor module.

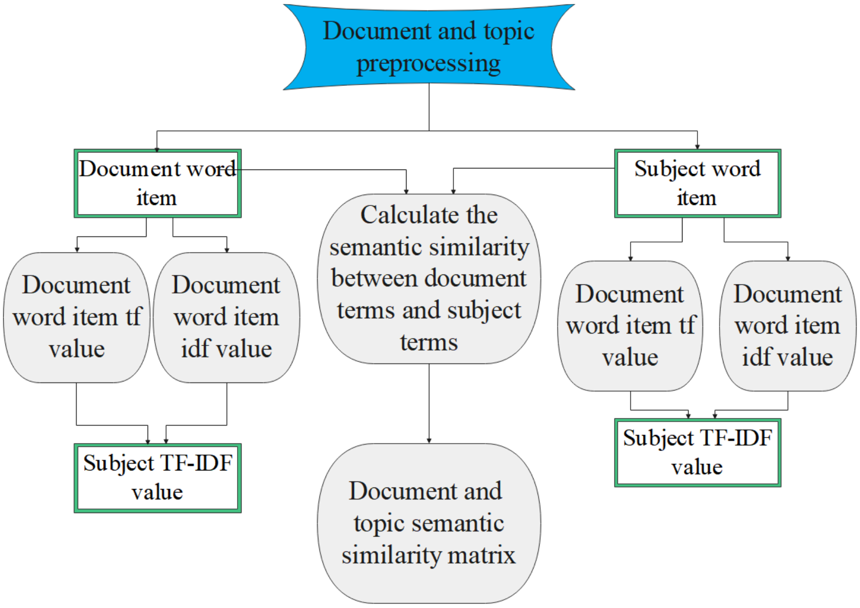

Figure 6: Structure diagram of related numerical modules for acquiring terms.

3. ENGLISH TEACHING BASED ON COMPUTER AIDED TRANSLATION

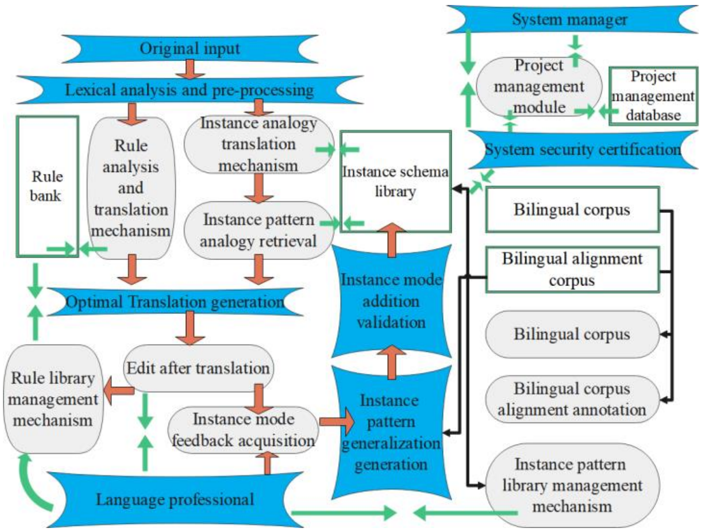

Figure 7: Overall design scheme.

The computer-aided translation system shown in Figure7 is constructed, the system is applied to English teaching, the effect of English teaching practice is counted, and the effect is verified by the simulation test method, and the results shown in Table 1 and Figure 8 are obtained.

Number	Teaching effect	Number	Teaching effect	Number	Teaching effect
1	75.80	19	75.23	37	73.47
2	70.18	20	70.15	38	70.62
3	74.20	21	74.64	39	83.37
4	84.82	22	78.84	40	75.14
5	80.52	23	83.45	41	83.48
6	82.93	24	74.48	42	85.24
7	81.34	25	77.30	43	77.81
8	73.12	26	79.53	44	72.39
9	83.80	27	70.71	45	80.61
10	82.31	28	70.18	46	80.47
11	72.38	29	81.99	47	74.37
12	85.74	30	73.15	48	82.28
13	77.75	31	79.85	49	71.70
14	81.69	32	83.21	50	72.05
15	74.45	33	75.93	51	70.24
16	70.11	34	76.25	52	83.92
17	74.79	35	71.75	53	85.51
18	75.67	36	78.10	54	71.81

Table 1: The role of computer aided translation in promoting English teaching.

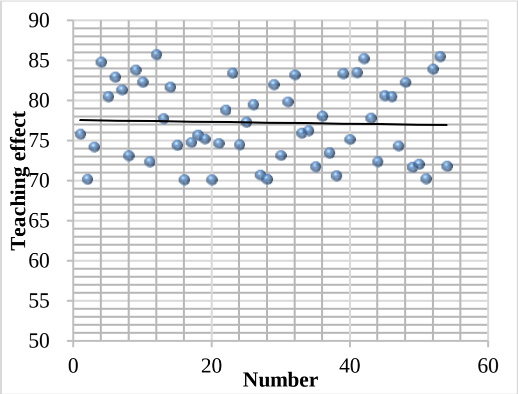

Figure 8: The application effect of computer-aided translation in English teaching.

From the above research, it can be seen that the computer translation proposed in this paper has a good application effect in English teaching, and can effectively improve the quality of English teaching.

4. CONCLUSION

Simple machine translation is basically based on the translation at the lexical level, and there are often major errors at the syntactic and semantic levels. Also, machine translation is not very good at converting two language systems. Language is not only the words itself, but also contains powerful cultural information and language expression habits. Therefore, machine translation often provides translations with little reference. When people realize that they cannot rely entirely on artificial intelligence translation, but should use intelligent translation to serve translators, computer-assisted translation was born. In order to improve the effect of English teaching, this paper applies computer-aided translation to English teaching and constructs an intelligent English teaching system. The research shows that the computer translation proposed in this paper has better application effect in English teaching and can effectively improve the quality of English teaching.

Conflict of interest

The author declare no competing interests.

Data Availability Statement

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

Youming Xu, https://orcid.org/0009-0006-5277-1511

REFERENCES

[1] Bowker, L.: Machine translation literacy: Academic libraries' role, Proceedings of the Association for Information Science and Technology, 56(1),2019,618-619. https://doi.org/10.1002/pra2.108
[2] Brynjolfsson, E.; Hui, X.; Liu, M.: Does machine translation affect international trade? Evidence from a large digital platform, Management Science, 65(12), 2019, 5449-5460. https://doi.org/10.1287/mnsc.2019.3388
[3] Clivaz, C.: Lost in translation? the odyssey of 'digital humanities'in French, Studia Universitatis Babeș-Bolyai Digitalia, 62(1), 2017, 26-41. https://doi.org/10.24193/subbdigitalia.2017.1.02
[4] Dinneen, C.: Students' use of digital translation and paraphrasing tools in written assignments on Direct Entry English Programs, English Australia Journal, 37(1), 2021 ,40-51.
[5] Du, Y.; Liu, W. ; Lv, X. ; Peng, G. ; An improved focused crawler based on semantic similarity vectorspace model. Applied Soft Computing, 2015 https://doi.org/10.1016/j.asoc.2015.07.026
[6] Fadele, A. A.; Kamsin, A.; Ahmad, K.; Rasheed, R. A.: A novel Hadith authentication mobile system in Arabic to Malay language translation for android and iOS Phones, International Journal of Information Technology, 13(4), 2021, 1683-1692. https://doi.org/10.1007/s41870-020-00542-5
[7] Farooque, M. M. J. F.; Jadad, B.: A comparative analysis of translation software using data science approach to arabic statements, International Journal of Advanced Computer Systems and Software Engineering, 2(3), 2021 , 13-19.
[8] Gurumurthy, K.; Anand, D.; Karthik, K.: Democratizing Translation Using Digital Methods, Language and Language Teaching, 10(1), 2021, 30-36.
[9] Karlgren, K.; Lakkala, M.; Toom, A.; Ilomäki, L.; Lahti-Nuuttila, P.; Muukkonen, H.: Assessing the learning of knowledge work competence in higher education–cross-cultural translation and adaptation of the Collaborative Knowledge Practices Questionnaire. Research Papers in Education, 35(1), 2020 , 8-22. https://doi.org/10.1080/02671522.2019.1677752
[10] Lin, L., Hu, P., Tong, X., Na, S., Cao, R., Yuan, X., ..., Wang, L. V.: High-speed three-dimensional photoacoustic computed tomography for preclinical research and clinical

translation, Naturecommunications, 12(1), 2021, 1-10. https://doi.org/10.1038/s41467-021-212321
[11] Minors, H.: Desblache, Lucile (2019): Music and Translation: New Mediations in the Digital Age. London: Macmillan, 407 p. Meta: journal des traducteurs/Meta: Translators' Journal, 66(1), 2021, 230-232. https://doi.org/10.7202/1079333ar
[12] Parsa, R. N.: Trends in e-tools and resources for translators and interpreters, Translation & Interpreting, 13(2), 2021, 183-186. https://doi.org/10.12807/ti.113202.2021.r01
[13] Rankin, Y. A.; Thomas, J. O.;Joseph, N. M.: Intersectionality in HCI: Lost in translation, Interactions, 27(5), 2020,68-71. https://doi.org/10.1145/3416498
[14] Salter, C.: Disturbance, translation, enculturation: necessary research in new media, technology, and the senses, Visual Anthropology Review, 34(1),2018,87-97. https://doi.org/10.1111/var.12156
[15] Schulte, R.: Expanding the horizon of translation studies, Translation Review, 106(1), 2020 , 1-3. https://doi.org/10.1080/07374836.2020.1734422
[16] Singh, J.: Lost in translation? The emergence of the digital Guru Granth Sahib, Sikh Formations, 14(3-4), 2018, 339-351. https://doi.org/10.1080/17448727.2018.1485355
[17] Sun, Y.; Li, D.: Digital Humanities Approaches to Literary Translation, Comparative Literature Studies, 57(4), 2020, 640-654. https://doi.org/10.5325/complitstudies.57.4.0640
[18] Tracey, P.; Dalpiaz, E.; Phillips, N.: Fish out of water: Translation, legitimation, and new venture creation, Academy of Management Journal, 61(5), 2018 1627-1666. https://doi.org/10.5465/amj.2015.0264
[19] Türkmen, B.; Utilising digital media as a second language (L2) support: A case study on Netflix with translation applications, Interdisciplinary Description of Complex Systems: INDECS, 18(4), 2020, 459-470. https://doi.org/10.7906/indecs.18.4.6
[20] Vandepitte, S.; Lefever, E.: Translation as a multilingual activity in the digital era, Revue francaise de linguistique appliquee, 23(2), 2018 , 59-72 https://doi.org/10.3917/rfla.232.0059