English Semantic Recognition Based on an Intelligent Algorithm

Na Deng
International Education School, Huanghuai University, Zhumadian, Henan 463000, China
E-mail: n79597@yeah.net

Keywords: intelligent algorithm, English translation, semantic recognition, transformer model, BLEU

Received: September 3, 2021

In the process of translation, semantic barriers have attracted extensive attention from researchers. Taking the translation between Chinese and English as an example, this paper used intelligent algorithms to recognize the semantic role of English, introduced the semantic role labeling, designed a semantic role encoder, integrated the encoder with the transformer model, and tested the translation performance of the system. The experimental results showed that the BLEU-4 score of the combined system was significantly higher than the baseline system and the traditional transformer system. The average BLEU-4 values of the three systems were 35.02, 35.78, and 36.9, respectively, and the score of the combined system was the highest. The specific analysis of several examples also found that the translation results of the combined system were more reliable. The experimental results verify the effectiveness of the combined system in machine translation and the importance of semantic recognition in translation.

Povzetek: Pri prevajanju med angleščino in kitajščino pomaga inteligentni semantični algoritem.

1 Introduction

Machine translation refers to translating one natural language into another through machine tools. With the development of technology, machine translation is also improving, which has made a great contribution to the communication of social culture and information, and more and more methods have been applied in machine translation. Sennrich et al. [1] encoded rare and unknown words into subword unit sequences, i.e., translate various categories of words by units smaller than words. Through the comparison on WMT 15 translation tasks, they found that subword models improved over a back-off dictionary baseline for the WMT 15 translation tasks English-German and English-Russian by 0.8 and 1.5 BLEU, respectively. Aiming at the problem of disambiguation in translation, Choi et al. [2] designed a method of contextualizing the word embedding vectors using a nonlinear bag-of-words representation of the source sentence and proposed to represent special tokens with using typed symbols. Experiments showed that the method could significantly improve the quality of translation. For neural machine translation, Luong et al. [3] trained a neural machine translation (NMT) system on data that was augmented by the output of a word alignment algorithm, allowing the NMT system to output, for each OOV word in the target sentence, its corresponding word in the source sentence. Experiments on WMT'14 English to French translation tasks showed that the method achieved an improvement of 2.8 BLEU. Lee et al. [4] mapped the source character sequence to the target character sequence without any segmentation. They adopted the character-level convolution network with maximum pooling on the encoder to improve the speed of model training. Through experiments, they found that the method had higher translation quality. Most of the current machine translations need to be adjusted manually to achieve high readability [5]. One of the major problems is the obstacle of semantic recognition in machine translation [6]. Almost all language behaviors of people are related to semantics. In translation activities, the size of the obstacles to semantic recognition depends on the translator's mastery of the source language and the target language. The higher the mastery degree is, the smaller the obstacles to semantic recognition in the process of translation are. This is for manual translation. In machine translation, the machine will not be affected by emotion, language, etc. If the relevant knowledge input into the system is perfect enough and the learning mechanism is intelligent enough, the obstacles of semantic recognition will be smaller. Based on semantic recognition, this study analyzed the method of Chinese-English machine translation and carried out experiments on the designed system to understand the reliability of the system in translation and make some contributions to the better development of machine translation.

2 Neural machine translation combined with semantic roles

2.1 Semantic role labeling

Machine translation, with the advantages of high speed and no need for manual work, has attracted extensive attention in translation work. Neural machine translation is the mainstream [7], but it can only learn from bilingual parallel corpus, ignoring linguistic knowledge, which leads to the poor quality of translation [8]. The natural language contains a lot of fuzziness, near meaning, and polysemy [9]. Therefore, semantic recognition is of great

value to NMT. In the current research, semantic annotation and recognition are of great help to disambiguation and role understanding [10]. Among them, semantic role annotation (SRL) [11] is an important content, which can realize shallow semantic analysis and identify the labels of argument, agent, patient, etc. in a sentence. SRL is helpful for the computer to better understand the true meaning of a sentence. An example is as follows.

[Alice] Agent [met] Predicate [Bob] Patient in the [church] Location [yesterday] time [evening] time.

In this sentence, [Alice] is the agent, [yesterday] and [evening] are the time, [church] is the location, [met] is the predicate, and [Bob] is the patient.

In the translation method designed in this paper, the AllenNLP tool [12] is used to realize the SRL of the source corpus. According to the number of predicates in a sentence, the sentence is divided into the corresponding number of lines. Each line includes the predicate label and its corresponding semantic role label. The prefix "BIO" is used for further modification. "B -" means that the current word is the first word within the scope of the semantic role, "I -" means that the current word is a middle word or a tail word, and "O -" means that the current word does not belong to any semantic role.

2.2 Transformer model combined with semantic role encoder

The characteristic of NMT is that it has an encoder and a decoder. The encoder reads the source sequence and outputs the vector; then, the decoder generates the correct translation according to the source vector. The specific methods include convolution neural networks [13], deep neural networks [14], etc. The transformer model is one kind of NMT [15], which only uses the attention mechanism to encode and decode and comprises an attention mechanism and a feedforward neural network. In one operation of the attention mechanism, variables Query, Key, and Value, i.e., Q, K, and V, are involved, and the operation process is:

$$\text{Attention}(\mathrm{Q},\mathrm{K},\mathrm{V}) = \text{softmax}(\frac{\mathrm{QK^T}}{\sqrt{\mathrm{d_k}}})\mathrm{V},$$

where d_k refers to the dimension of the model.

In order to make up for the loss of the input sequence order, the transformer model introduces the concept of position-coding:

$$\mathrm{PE}(\mathrm{pos},2\mathrm{i}) = \sin(\mathrm{pos}/10000^{2\mathrm{i}/\mathrm{d_{model}}},$$

$$\mathrm{PE}(\mathrm{pos},2\mathrm{i}+1) = \cos(\mathrm{pos}/10000^{2\mathrm{i}/\mathrm{d_{model}}},$$

where pos refers to the position of the input word, i is a bit of the position vector, and d_{model} is the dimension of the model. In the process of translation, the encoder learns different features through the attention module, and the decoder decodes the semantic vector. Through softmax operation, the target language words can be generated one by one.

A semantic role encoder is used to combine semantic recognition with the transformer model. For a source sentence, $\mathrm{X} = [\mathrm{x_1},\mathrm{x_2},\ldots,\mathrm{x_n}]$, there is a same semantic role label sequence: $\mathrm{L} = [\mathrm{l_1},\mathrm{l_2},\ldots,\mathrm{l_n}]$. The transformer encoder is composed of N_s same layers. Every layer includes two sublayers. The calculation formula of the first sublayer is:

$$A^n = MultiHead(S^{n-1},S^{n-1},S^{n-1}),$$

where S^{n-1} refers to the output of the n-1-th layer of the encoder. After residual connection and layer regularization processing, there is:

$$B^n = LayerNorm(A^n + S^{n-1}).$$

Then, the second sublayer is a bit-by-bit fully connected feedforward neural network, which is also processed by residual connection and layer regularization:

$$\mathrm{C^n} = \mathrm{FFN}(\mathrm{B^n}),$$

$$\mathrm{S^n} = \mathrm{LayerNorm}(\mathrm{B^n} + \mathrm{C^n}).$$

The encoder that fuses the semantic role has only one layer, and the calculation formula is:

$$A_{SRL} = MultiHead(L,L,S^{N_S}),$$

where L refers to the label sequence of the semantic role, $\mathrm{S^{Ns}}$ is the final output of the source encoder, and $\mathrm{A_{SRL}}$ refers to the final output of the semantic encoder. Finally, the two outputs are fused by the gating mechanism. The calculation formulas are:

$$\gamma = \sigma(\mathrm{W_{N_s}S^{Ns}} + \mathrm{W_{SRL}S_{SRL}}),$$

$$S_{mix} = \gamma \odot S^{N_S} + (1-\gamma) \odot S_{SRL},$$

where $\mathrm{W_{Ns}}$ and $\mathrm{W_{SRL}}$ are the matrices of gating parameters, and $\mathrm{S_{mix}}$ is the mixture of the original sentence and the semantic role information as the input of the decoder.

3 Experimental analysis

3.1 Experimental data

The Chinese-English bilingual corpus used in the experiment contains about four million parallel sentence pairs, which come from some subsets of the LDC corpus, including LDC2003E14, LDC2004E12, etc. The GIZA++ tool was used for word alignment, and then the grow-diag-final-and heuristic method was used to obtain the information of word alignment. The development set of the experiment was the evaluation corpus of NIST2005 (NIST05), and the test sets were the evaluation corpus of NIST2002, 2003, 2004, 2005, and 2008 (NIST02, NIST03, NIST04, NIST05, and NIST08). The hierarchical phrase translation system [16] was used as the baseline system and compared with the traditional transformer system and the transformer system combined with semantic role recognition.

3.2 Evaluation index

This paper used BLEU [17] to automatically evaluate the quality of machine translation. BLEU index was based on N-ary grammar. First, the maximum frequency of a word in different results was recorded; then, the frequency of each candidate word in translation was corrected. The calculation formula is:

$$P_n = \frac{\sum_{c\in\{candidates\}}\sum_{N-gram\in c} count_{clip}(N-gram)}{\sum_{c'\in\{candidates\}}\sum_{N-gram\in c'} count_{clip}(N-gram')},$$

where c is the total length of the candidate translation sentence. It is assumed that the total length of the reference translation is r. There is also a length penalty term in BLEU:

$$BP = \begin{cases} 1, c > r \\ e^{1-\frac{r}{c}}, c \le r \end{cases}.$$

Finally, the calculation formula of BLEU is:

$$BLEU = BP \times \exp(\sum_{n=1}^{N} w_n \log p_n),$$

where N refers to the maximum order of N-ary grammar and w_n is the weight coefficient. In the experiment, BLEU-4 was used for evaluation, i.e., $N = 4, w_n = \frac{1}{N}$.

3.3 Experimental results

On different test sets, the BLEU-4 values of the two systems are shown in Figure 1.

It was seen from Figure 1 that the transformer system performed better in translation than the baseline system, and the translation performance of the transformer system further improved after combining with semantic role recognition. On NIST 02, the BLEU-4 scores of the three systems were 37.2, 38.1, and 39.7, respectively, and the BLEU-4 value of the transformer system combined with semantic role recognition was 6.72 % higher than that of the baseline system and 4.20 % higher than that of the transformer system. On NIST 03, the BLEU-4 scores of the three systems were 36.6, 37.2, and 38.4, respectively, and the BLEU-4 value of the transformer system combined with semantic role recognition was 4.92 % higher than that of the baseline system and 3.23 % higher than that of the transformer system. On NIST 04, the BLEU-4 scores of the three systems were 38.1, 39.2, and 40.1, respectively, and the BLEU-4 value of the transformer system combined with semantic role recognition was 5.25 % higher than that of the baseline system and 2.30 % higher than that of the transformer system. On NIST 05, the BLEU-4 values of the three systems were 35.4, 36.1, and 37.2, respectively; the BLEU-4 value of the transformer system combined with semantic role recognition was 5.08% higher than that of the baseline system and 3.05% higher than that of the transformer system. On NIST08, the BLEU-4 values of the three systems were 27.8, 28.3, and 29.1, respectively; the BLEU-4 value of the transformer system combined with semantic role recognition was 4.68% higher than that of the baseline system and 2.83% higher than that of the transformer system.

The average BLEU-4 values of the three systems were calculated, and they were 35.02, 35.7, and 36.9, respectively. It was found that the transformer system combined with semantic role recognition had the highest BLEU-4 value and the highest translation quality in the process of the English translation. Finally, the translation performance of the system was analyzed taking several sentences as the example.

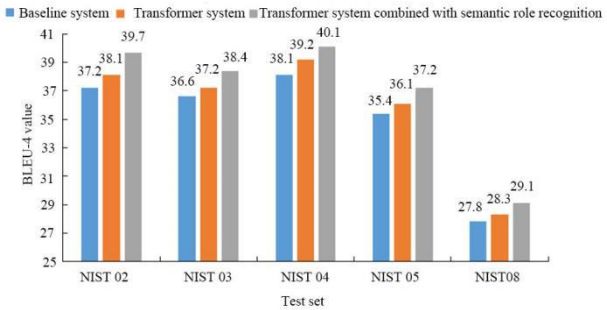

Figure 1: Comparison of bLEU-4 values.

System	Sentence
Original sentence	**Charges** against four other men were found not proven.
Baseline system	对四名其他男子的**收费**没有被证明。
Transformer system	对四名男子的**收费**被发现没有证实。
The transformed system combined with semantic role recognition	对四名男子的**指控**被发现没有被证实。
Original sentence	The data are stored in the computer's **memory**.
Baseline system	数据被存储在计算机的**记忆**中。
Transformer system	数据存储在计算机的**记忆**中。
The transformed system combined with semantic role recognition	数据存储在计算机的**内存**中。
Original sentence	After a full test, he submitted his **application**.
Baseline system	经过全面测试后，他提交了**申请**。
Transformer system	经过全面测试，他提交了他的**申请**。
The transformed system combined with semantic role recognition	经过全面测试，他提交了他的**应用程序**。

Table 1: Example sentence analysis.

As shown in Table 1, due to semantic differences, there are some differences in translation results. In example 1, "charges" has multiple meanings, such as fee collection, fee, accusation, rushing, etc.; in the output of the first two systems, it was translated as "fee collection", but in the transformer system combined with semantic role recognition, it was translated as "charge". Combined with context and semantics, the translation of the transformer system combined with semantic role recognition was correct. In example 2, "memory" means something that is remembered, internal storage, etc.; in this sentence, it should mean "internal storage". In the first two systems, it was wrongly translated as "something that is remembered" because the two systems did not recognize the semantics, which led to the wrong choice of candidate words. In example 3, "application" means application, complaint, use, application program, etc.; in this sentence, "application program" was correct.

It was found from Figure 1 and Table 1 that the transformer system combined with semantic role recognition showed better performance in the process of

translation and produced more accurate and reliable results.

4 Conclusion

For Chinese-English machine translation, this study designed a new system based on the encoder and transformer model, which combined semantic role recognition, and tested it. It was found that the designed system had better performance, the highest BLEU score, and more accurate translation results, compared with the baseline system and the traditional transformer system. This work contributes the further development of machine translation and also helps to improve the importance of semantic recognition in translation work.

References

[1] Rico Sennrich, Barry Haddow, and Alexandra Birch. Neural Machine Translation of Rare Words with Subword Units. *Computer Science*, 1715-1725, 2016.

[2] Heeyoul Choi, Kyunghyun Cho, and Yoshua Bengio. Context-Dependent Word Representation for Neural Machine Translation. *Computer Speech & Language*, 45(Sep.):149-160, 2017. https://doi.org/10.1016/j.csl.2017.01.007

[3] Minh-Thang Luong, Ilya Sutskever, Quoc V. Le, Oriol Vinyals, and Wojciech Zaremba. Addressing the Rare Word Problem in Neural Machine Translation. *Bulletin of University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca. Veterinary Medicine*, 27(2):82-86, 2015. https://doi.org/10.3115/v1/P15-1002

[4] Jason Lee, Kyunghyun Cho, and Thomas Hofmann. Fully Character-Level Neural Machine Translation without Explicit Segmentation. *Transactions of the Association for Computational Linguistics*, 5:365-378, 2017. https://doi.org/10.1162/tacl_a_00067

[5] Jernej Vicic. A fast implementation of rules based machine translation systems for similar natural languages. *Informatica*, 37(4):455-456, 2013.

[6] Mohamed El Bachir Menai. Word sense disambiguation using an evolutionary approach. *Informatica*, 38(2):155-169, 2014.

[7] Castilho Sheila, Moorkens Joss, Gaspari Federico, Calixto Iacer, Tinsley John, and Way Andy. Is Neural Machine Translation the New State of the Art?. *The Prague Bulletin of Mathematical Linguistics*, 108(108):109-120, 2017. https://doi.org/10.1515/pralin-2017-0013

[8] Philip Arthur, Graham Neubig, and Satoshi Nakamura. Incorporating Discrete Translation Lexicons into Neural Machine Translation. *Computer Science*, eprint arXiv:1606.02006, 2016. https://doi.org/10.18653/v1/D16-1162

[9] Chong Chai Chua, Tek Yong Lim, Lay-Ki Soon, Enya Kong Tang, and Bali Ranaivo-Malançon. Meaning preservation in Example-based Machine Translation with structural semantics. *Expert Systems with Applications*, 78(JUL.):242-258, 2017. https://doi.org/10.1016/j.eswa.2017.02.021

[10] Gang Zhang. Research on the efficiency of intelligent algorithm for english speech recognition and sentence translatione. Informatica, 45(2):309-314, 2021. https://doi.org/10.31449/inf.v45i2.3564

[11] Yoko Nakajima, Michal Ptaszynski, Hirotoshi HONMA, and Fumito Masui. A Method for Extraction of Future Reference Sentences Based on Semantic Role Labeling. *IEICE Transactions on Information & Systems*, E99.D(2):514-524, 2016. https://doi.org/10.1587/transinf.2015EDP7115

[12] Matt Gardner, Joel Grus, Mark Neumann, Oyvind Tafjord, Pradeep Dasigi, Nelson Liu, Matthew Peters, Michael Schmitz, and Luke Zettlemoyer. AllenNLP: A Deep Semantic Natural Language *Processing Platform. Proceedings of Workshop for NLP Open Source Software (NLP-OSS)*, 1-6, 2018. https://doi.org/10.18653/v1/W18-2501

[13] Taiki Watanabe, Akihiro Tamura, Takashi Ninomiya, and Teguh Bharata Adji. Neural Machine Translation with CKY-based Convolutional AttentionCKY に基づく畳み込みアテンション構造を用いたニューラル機械翻訳. *Journal of Natural Language Processing*, 26(1):207-230, 2019. https://doi.org/10.5715/jnlp.26.207

[14] Jiajun Zhang, and Chengqing Zong. Deep Neural Networks in Machine Translation: An Overview. *IEEE Intelligent Systems*, 30(5):16-25, 2015. https://doi.org/10.1109/MIS.2015.69

[15] Chao Su, Heyan Huang, Shumin Shi, Ping Jian, and Xuewen Shi. Neural machine translation with Gumbel Tree-LSTM based encoder. *Journal of Visual Communication and Image Representation*, 71:102811, 2020. https://doi.org/10.1016/j.jvcir.2020.102811

[16] Xiaoyin Fu, Wei Wei, Lichun Fan, Shixiang Lu, a nd Bo Xu. Nesting hierarchical phrase-based model for speech-to-speech translation. *2012 8th International Symposium on Chinese Spoken Language Processing (ISCSLP)*, 368-372, 2012. https://doi.org/10.1109/ISCSLP.2012.6423497

[17] Yaya Tian, Shaweta Khanna, Anton Pljonkin. Research on machine translation of deep neural network learning model based on ontology. *Informatica*, 45(5):643-649, 2021. https://doi.org/10.31449/inf.v45i5.3559