

SSAG-Net: Syntactic and Semantic Attention-Guided Machine Reading Comprehension

Chenxi Yu and Xin Li*

Department of Information Technology and Cyber Security People's Public Security University of China, Beijing, 102623, China

*Corresponding Author: Xin Li. Email: lixin@ppsuc.edu.cn

Received: 04 March 2022; Accepted: 19 April 2022

Abstract: Machine reading comprehension (MRC) is a task in natural language comprehension. It assesses machine reading comprehension based on text reading and answering questions. Traditional attention methods typically focus on one of syntax or semantics, or integrate syntax and semantics through a manual method, leaving the model unable to fully utilize syntax and semantics for MRC tasks. In order to better understand syntactic and semantic information and improve machine reading comprehension, our study uses syntactic and semantic attention to conduct text modeling for tasks. Based on the BERT model of Transformer encoder, we separate a text into two branches: syntax part and semantics part. In syntactic component, an attention model with explicit syntactic constraints is linked with a self-attention model of context. In semantics component, after the framework semantic parsing, the lexical unit attention model is utilized to process the text in the semantic part. Finally, the vectors of the two branches converge into a new vector. And it can make answer predictions based on different types of data. Thus, a syntactic and semantic attention-guided machine reading comprehension (SSAG-Net) is formed. To test the model's validity, we ran it through two MRC tasks on SQuAD 2.0 and MCTest, and the SSAG-Net model outperformed the baseline model in both.

Keywords: Machine reading comprehension; syntax dependency; frame semantics

1 Introduction

Machine Reading Comprehension task (MRC) requires machines to understand natural language information in the same way humans do and to answer questions about it [1]. This requires machines to try to infer the meaning of natural language text.

In the past, people constructed syntactic or semantic skills by hand definition to match the most appropriate answer. With the development of neural network models, the field of natural language processing began to model sentences using fine-tuned pre-trained language models, and then train the models in different tasks.

Although neural network models have made significant progress in the field of natural language processing, certain research has indicated that these models cannot truly comprehend the text as humans



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

do, and that natural language processing interpretability using neural network models alone is limited. This is a clear departure from the most basic definition of machine reading comprehension.

People began to consider using neural network models to model the syntax or semantics of text based on human reading habits. Training is done using a more accurate parser and a neural network model [2–4].

However, just a few research has combined syntax and semantics modeling. Syntax parsing and frame semantics are among the current research methodologies that have been integrated into the MRC's neural network. The input sequence is syntactically and semantically processed using integrated method. However, because this strategy ignores the expansion of the attention layer, the text is unable to consider significant terms selectively. Our study seeks to combine syntactic parsing model and frame semantics model into an end-to-end machine reading comprehension model with improved attention to solve machine reading comprehension tasks based on this.

The main contributions of this paper are as follows:

- 1) Based on syntactic and semantic attention, we explore the integration of these two models.
- 2) We propose a new reading comprehension model, in which syntactic parsing model and frame semantics model are added to the encoding layer and attention layer respectively, so that sentences in the text can be fully analyzed and more comprehensive expressions can be obtained.
- 3) We use two types of datasets for testing, and the experimental results show that our proposed method works well in MRC tasks.

2 Related Work

In natural language tasks, dependency syntactic parsing is a crucial task. Researchers debated how to represent word dependence in phrase structure at first. Schabes proposed the lexicalized tree join syntax (LTAG) in 1988 [5]. The combinatorial category grammar (CCG) [6] proposed by Steedman in 2000. The head-driven phrase structure grammar (HPSG) [7] proposed by Pollard and Sag in 1994. In 2016, syntactic dependent parsing trees were proved to have a catalytic effect in natural language processing [8]. With the continuous development of neural network, syntactic parsing has been tried to be applied [9–11], and many syntactic parsers with good accuracy have been innovated, such as, a novel Attentional NMT with Source dependency representation proposed by Chen in 2017 [2], and in 2019, Duan introduce syntax into Transformer for better machine translation [4].

The concept of frame semantics was first proposed by Fillmore [12], who pointed out the role of the concept of frame in cognition, and understanding. In 1998, Baker proposed the FrameNet project [13], which is devoted to the semantic and syntactic generalization of corpus evidence, and the semantic part makes use of the concept of frame semantics. In the past, MRC tasks were typically performed by manually defining semantic functions combined with machine learning. A heavy reliance on manually defined semantics leads to a model that is less portable and less applicable to other domains. With the continuous development of neural networks, frame semantics are also studied based on neural networks using pre-training models [14]. In 2018, Zhang used pre-training model to model semantic knowledge in neural network in MRC task [15]. In 2020 Guo [16] embedded the LU vector in the previous framework on the basis of framework semantics and improved the modeling of the framework as a whole. At the same time, multi-frame relationship integration is tried.

All the above studies have considered only one aspect of syntax and semantics. However, previous studies have attempted to fuse syntax and semantics. From 2015, relying on hand-defined Max-margin learning framework [17], in 2018, based on analytic tree and semantic framework of structured inference result question answering pair [18], to 2020, syntactic and semantic fusion of neural network model [19]. At present, many researches on the fusion of frame semantics and syntactic parsing simply connect

syntactic and semantic labels and words together [20], which is too rough and fails to properly represent labels and words. Based on this, we construct the MRC model by integrating syntax and frame semantics at the coding layer and attention layer.

3 Syntax Dependency and Frame Semantics

3.1 Syntactic Dependency

Based on the dependency structure, syntactic dependency describes a sentence’s structure by determining which parts of the sentence each word depends on [21]. If a word modifies another word, then that word is dependent on the other word in this context. Adjectives modify nouns, adverbs modify adjectives, and so on. We may build a syntactic structure tree of dependencies based on a sentence using these syntactic dependencies [22]. We created the syntactic dependency tree shown in Fig. 2 based on the passage and questions in Fig. 1. According to the dependency tree, we can analyze that ‘what’ in the question corresponds to the word ‘Formation’ in passage, so the answer is ‘Formation’.

Passage

On February 6, 2016, Beyonc released a new single exclusively on music streaming service Tidal called **Formation**.

Question

Beyonce's new single released before the super bowl was called what?

Answer

Formation

Figure 1: Examples of passage, questions and answers for machine reading comprehension

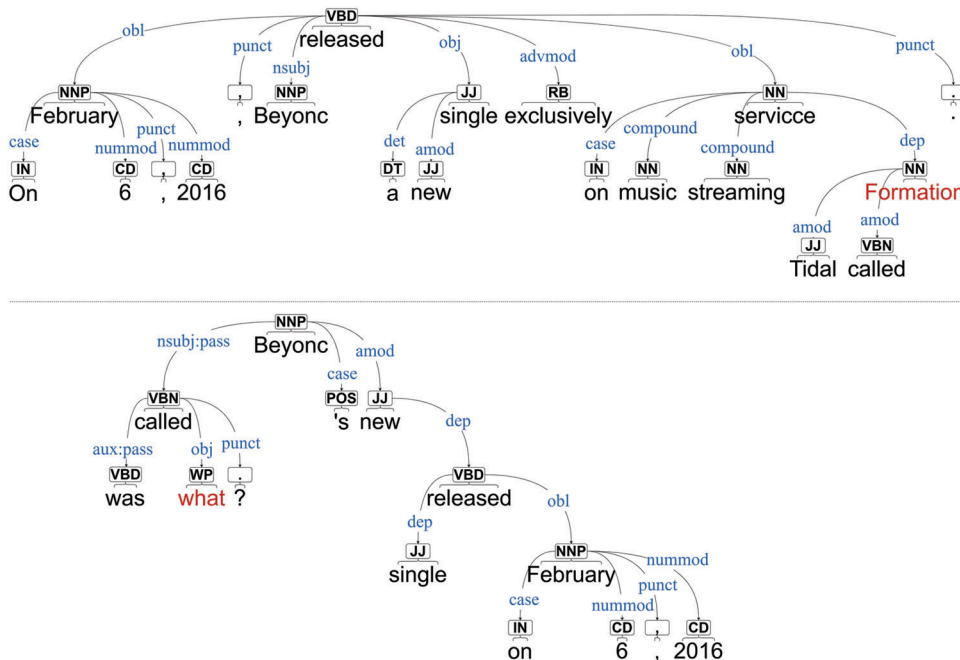


Figure 2: Example of syntactic dependency tree

3.2 Frame Semantics

The FrameNet project was the first to establish frame semantics, which has the advantage of allowing most phrases to be comprehended by breaking them up according to the defined framework. Frame semantics has four definitions: frame (F), frame elements (FEs), lexical units (LUs), and target (T). FrameNet's frames contain textual descriptions of what they represent, and they are schematics of situations involving multiple actors, items, and other conceptual feelings. Frames are made up of frame elements and lexical units. Frame element is a type of semantic role in frame, with core words, non-core words peripheral words. Lexical unit is the group of words that connect frame element to make frame. The term 'T' evokes up the frame for identifying the lexical unit in the actual sentence.

For example, the semantic framework of the above sentence is shown in Fig. 3. There are two semantic frames in this sentence as shown in Tab. 1, which are 'Publishing' and 'Referring_by_name' respectively. Where Target 'called' activates frame 'Referring_by_name', the frame element corresponding to it is FEs-core's 'Name', which is 'Formation'.

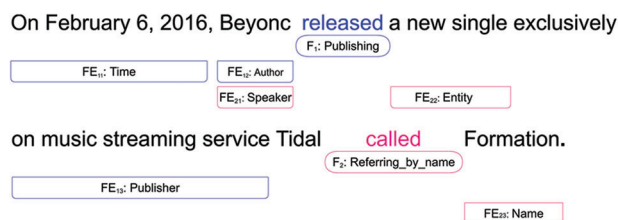


Figure 3: Graphical examples of semantic framework

Table 1: Example of semantic framework

F	FEs-Core	FEs-Non-Core	LUs
Publishing	author, publisher, work	audience, manner, means, medium, place, purpose, time	appear.v, publication.n, publish.v, put out.v, release [event].n, release.v
Referring_by_name	entity [ent], name [nam], speaker [spe]	descriptor [des], time [tim]	address.v, call.v, designation.n, name.n, refer.v

4 Model Overview

Input layer, encoding layer, attention layer, and answer prediction layer are four sections of a machine reading comprehension model.

The goal of this paper is to create a neural network model that incorporates both syntactic and semantic analysis. The entire model is divided into two branches in this research, which deal with syntactic parsing and frame semantics, respectively. Finally, the answer prediction layer combines the two vectors.

The model overview is shown in Fig. 4.

For the section of syntactic analysis processing, we use the BERT [14] model to get an encoding sentence, and then we use syntax to build a self-attention mechanism. Then, before answer prediction, we integrate the original coding with the encoding of the attention mechanism which is driven by syntax, and it can provide an encoding representation of the syntactic analysis phase.

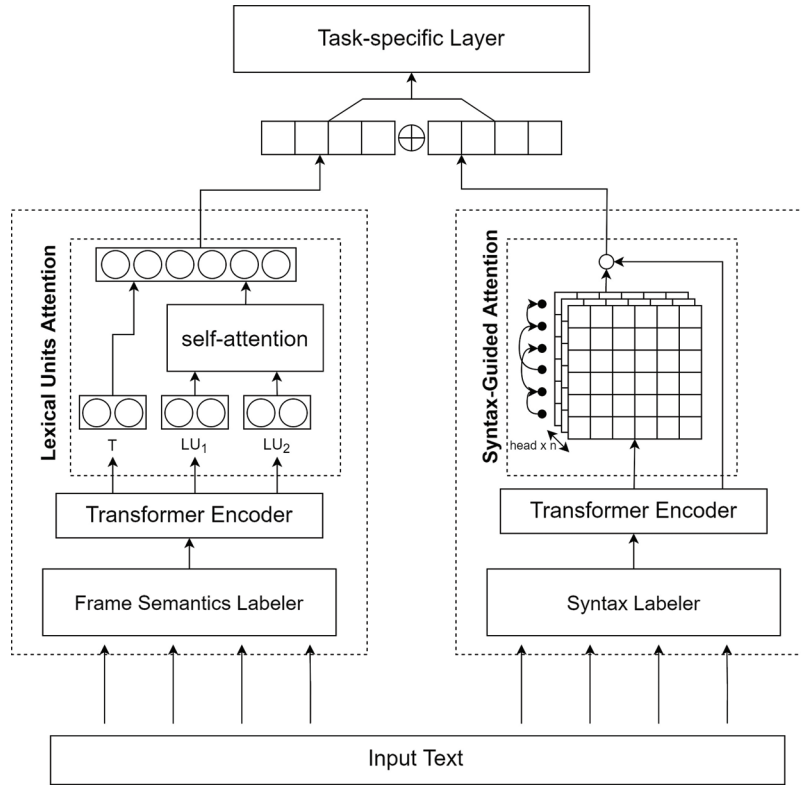


Figure 4: SSAG-Net model

We use SEMAFOR [23] to annotate the framework semantic text for the framework semantic component, and then we encode the annotated text using BERT [14] model of Transformer encoder. After passing through a TLUA attention layer, the coding representation of the framework semantic part is completed. Finally, the answer prediction layer is used to predict answers based on distinct tasks using syntactic parsing and framework semantic coding. The extraction and single choice are both tested in this study.

4.1 Attention

We use 2020 Zhang’s SG-Net [24] model for the syntax-guided attentional mechanism layer in this paper. The attention layer pre-trains and marks the dependency structure of each sentence using a syntactic dependency parser. The word’s relationships might help focus attention on the sentence’s most significant words.

Find the ancestor node corresponding to each word based on the sentence structure analyzed by the parser. If a word j is the ancestor node of the word i , then the Mask $M[i, j]$ of the word is marked as 1; otherwise, it is marked as 0. See Eq. (1)

$$M[i, j] = \begin{cases} 1, & \text{if } j \in P_i \text{ or } j = i \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

The grammar-guided attention layer is determined using multi-head self-attention computation once all of the masks have been annotated. The output vector of the syntactic processing section is the vector obtained.

We use the lexical unit attention model (TLUA) proposed by Guo in 2020 [16] for the semantic attention layer of the framework in this paper. This model solves the problem of not treating frame as a whole. By using the attention mechanism of target in the sentence, the weight of different LU is automatically determined and the weight of target is increased at the same time, which can reduce the noise interference caused by irrelevant LUs. After TLUA processing, the vectors of the semantic part of the framework are obtained. The formulas Eqs. (2) and (3) are as follows:

$$F_m = t^{F_m} + \sum_{\tilde{U}^{F_m}} att(u_n^{F_m}) \cdot u_n^{F_m} \quad (2)$$

$$att(u_n^{F_m}) = \frac{\exp(t^{F_m} \cdot u_n^{F_m})}{\sum_{u_k^{F_m} \in \tilde{U}^{F_m}} \exp(t^{F_m} \cdot u_k^{F_m})} \quad (3)$$

Here, \tilde{U}^{F_m} represents the LUs set of F_m which is not include t^{F_m} , and $\tilde{U}^{F_m} \in R^{H \cdot (n-1)}$.

The syntactic and semantic vectors are then combined with the Transformer-encoded context vectors. Finally, put them into the prediction layer to predict answers based on various task types.

4.2 Answer Prediction Layer

In this research, the model is applied to two types of machine reading comprehension tasks: extraction and single choice.

By analyzing the given text content and relevant queries, the aim of extraction is to predict the corresponding starting and ending places from the text. In the test, we used the softmax operation to predict the starting and ending positions of the encoded vectors in the linear layer [25].

$$s, e = \text{SoftMax}(\text{Linear}(\bar{H})) \quad (4)$$

And then selected the one with the highest probability score as the final answer output.

$$score_{has} = \max(s_k + e_l), 0 \leq k \leq l \leq n \quad (5)$$

The single choice task requires us to choose the correct answer from a set of k options that contain the correct answer based on the question and the information in the current paragraph. The response could be a single word, a phrase, or even a complete statement [26].

When there is no response, a threshold is set. There is no output if the sum of probability does not surpass this level. The model inputs the vector classifier for the single-choice task and selects the option with the highest probability as the prediction result for output during the test [27].

5 Experiment

5.1 Dataset and Evaluation Index

The experiment was based on SQuAD 2.0 and MCTest, which were used to test two MRC tasks.

5.1.1 SQuAD 2.0

The extractive MRC was tested on SQuAD 2.0. SQuAD 1.0 created a vast machine reading comprehension dataset with 100,000 questions, containing over 500 Wikipedia articles. In SQuAD 1.0, the answers form part of the corresponding text. SQuAD 2.0 adds over 50,000 human-designed unanswerable questions to SQuAD 1.0, so the system must account for unanswered questions in addition to enhancing the accuracy of the responses. The EM and F1 score are the assessment criteria for the

extractive MRC task of the model in this research. The percentage of projected ground truth answers is calculated using EM.

$$EM = \frac{N_{real}}{N_{all}} \quad (6)$$

F1 score is the coverage of test answers vs. correct answers, calculated by precision and recall.

$$P = \frac{N_{overlap}}{N_{all\ answer}} \quad (7)$$

$$R = \frac{N_{overlap}}{N_{truth\ answer}} \quad (8)$$

$$F1 = \frac{2 \times P \times R}{P + R} \quad (9)$$

5.1.2 MCTest

The MCTest is a single-choice MRC test. The MCTest dataset is a collection of seven-year-old-appropriate fairy tales, each containing one-of-four single-choice questions from the original text. Its correctness is its evaluation criterion. The accuracy rate is the percentage of right answers given by the model out of a total of M answers.

5.2 Implementation Details

For the syntactic analysis part, we adopt Zhang’s syntax-Guided network [21]. For the semantic part of the framework, SEMAFOR is used to annotate sentences automatically, and Guo’s lexical unit attention model is used to pay attention to key words. We use BERT [14] as input for all parts. Our initial learning rate is set as [8e-6, 1e-5, 2e-5, 3e-5], and Batch size is set as 10. And in the dual context aggregation, the weight α is 0.5. The maximum sequence length of SQuAD2.0 and MCTest is set to 382.

5.3 Experiments Results

For the experimental results, we designed the comparative experiment and model ablation experiment on SQuAD2.0 and MCTest.

Comparative Experiment

The comparative experiment is to compare the existing machine reading comprehension models which perform well in syntactic or semantic aspects. According to the different datasets applicable to the model, we divided the plan into two groups for experiments. The extraction machine reading comprehension model based on SQuAD2.0 dataset and the single choice machine reading comprehension model based on MCTest dataset were adopted respectively. As shown in [Tabs. 2](#) and [3](#):

Regular Track and BERT Track respectively refer to the effect of training each machine reading comprehension model after using traditional methods and using BERT pre-training model. BERT is the baseline model of this paper. BERT is a large pre-training model released by Google in 2018. In the field of MRC, adding an answer output layer on the BERT and fine-tuning the parameters of the pre-training part can perform well in MRC tasks. SSAG-Net is the final model of this paper.

[Tabs. 2](#) and [3](#) show the comparison between the above models and SSAG-Net model according to different evaluation indicators. It can be seen that the current models with good performance all use BERT as the pre-training model, indicating that the MRC model improved on the basis of BERT model has become the current trend and achieved good results. Although, it can be seen from [Tab. 2](#) that the

EM and F1 score of the model in this paper are improved to some extent after the addition of the attention mechanism model based on syntactic and semantic improvement, with EM being 85.8 and F1 score being 88.4. It can be seen from [Tab. 3](#) that the SSG-Net model also improves the test capability of the single-choice dataset to a certain extent, where the MCTest-160 value is 88.2 and the MCTest-160 value is 87.4.

Table 2: The performance comparison of different models on SQuAD 2.0 dataset

Model	Dev		Test	
	EM	F1	EM	F1
<i>Regular track</i>				
Joint SAN	69.3	72.2	68.7	71.4
U-Net	70.3	74.0	69.2	72.6
RMR + ELMo + Verifier	72.3	74.8	71.7	74.2
<i>BERT track</i>				
Human	-	-	86.8	89.5
SemBERT	-	-	84.8	87.9
BERT (LARGE)	-	-	82.1	84.8
SG-Net	85.1	87.9	-	-
Baseline	84.1	86.8	-	-
SSAG-Net	85.8	88.4	-	-

Table 3: The performance comparison of different models on two MCTest datasets

Model	MCTest-160 (100%)	MCTest-500 (100%)
<i>Regular track</i>		
Wang et al. (2015) + Syntax + Semantics	75.3	69.9
Li et al. (2018) + Syntax + Semantics	74.6	72.7
<i>BERT track</i>		
BERT + DCMN + (Zhang et al., 2019)	85.0	86.5
XLNet + DCMN + (Zhang et al., 2019)	86.2	86.6
FSR (Guo et al., 2020) + Semantics	86.1	84.2
SS-MRC	87.2	86.7
Baseline	82.5	80.9
SSAG-Net	88.2	87.4

Ablation Experiment

In this study, we found that syntactic and semantic processing of input text at the attention level can effectively improve the processing effect of machine reading comprehension tasks. In this paper the model is improved comprehensively based on SG-Net model and FSR model, and the two types of

datasets are trained and tested respectively. To verify the validity of our model, we performed an ablation experiment for two different MRC tasks. The ablation experiment used SQuAD 2.0 and MCTest data sets to compare the improved evaluation indicators in each part. The results are shown in [Tabs. 4 and 5](#).

Table 4: Extraction ablation study of SSAG-Net model

Method	EM	F1
SSAG-NET	85.8	88.4
-Syntax	85.1	87.9
-Frame semantics	84.8	87.4
-Syntax-frame semantics	84.1	86.8

Table 5: Single choice ablation study of SSAG-Net model

Method	MCTest-160	MCTest-500
SSAG-Net	88.2	87.4
-Syntax	87.6	86.4
-Frame semantics	86.2	84.8
-Syntax-frame semantics	82.5	80.9

- syntax-frame semantics is the BERT model of the baseline, meaning that Syntax and semantics are not added. In the SQuAD2.0 dataset, the EM and F1 score of the model in this paper are improved by 1.7% and 1.6% compared with this model. In the MCTest dataset, compared with this model, the MCTest-160 value is improved by 4.7% and the MCTest-500 value is improved by 6.5%.
- Frame Semantics adds syntactic semantics to baseline. In the SQuAD2.0 dataset, the EM and F1 score of this model are both improved by 1.0%. In the MCTest dataset, compared with this model, the value of MCTest-160 and MCTest-500 are improved by 2.0% and 2.6% respectively.
- Syntax adds a framework semantic part to the baseline. In the SQuAD2.0 dataset, the EM and F1 score of the model in this paper are improved by 0.7% and 0.5% compared with this model. In the MCTest dataset, compared with this model, the MCTest-160 and MCTest-500 values of the proposed model are improved by 0.6% and 1.0% respectively.

Research results demonstrated that the effect of our model on the extractive reading comprehension task and the single choice machine reading comprehension task was greatly enhanced by testing on the extractive MRC dataset and the single choice MRC dataset, as well as the ablation experiment. The experimental results show that adding syntactic and semantic modules improves the machine reading comprehension model significantly, and the frame semantics improve the machine reading comprehension model most significantly. It is also verified that this conclusion is similar to the framework semantic model proposed by Guo in 2020 [19], the grammar-guided machine reading comprehension model proposed by Zhang in 2020 [21], and the combined semantic and frame model proposed by Guo in 2020 [16].

6 Discussion

6.1 Effect of Different Questions Length

To explore the impact of different questions lengths in text on the model, we used questions from SQuAD2.0 to sort by question length, grouping questions within the same length range and same number of words. We used the baseline model and SSAG-Net to test the accuracy of each group, and the test results are shown in Figs. 5 and 6. We can see from the results in the figures, as the Question length gradually increases, the matching accuracy of the baseline decreases significantly, and our proposed SSAG-Net decreases slightly. However, our matching accuracy remains stable compared to the baseline. This shows that adding syntactic and semantic expressions can deal with sentences with longer problems better.

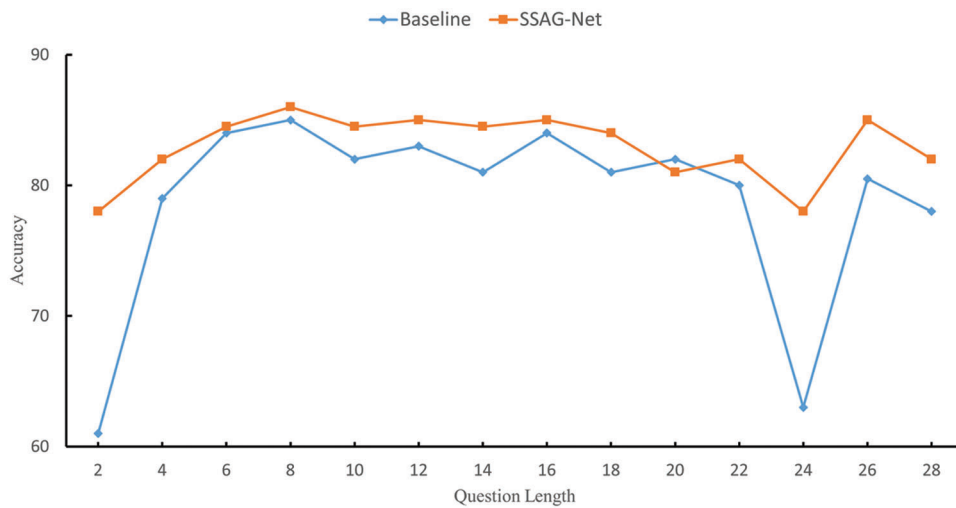


Figure 5: Accuracy of different questions' length ranges

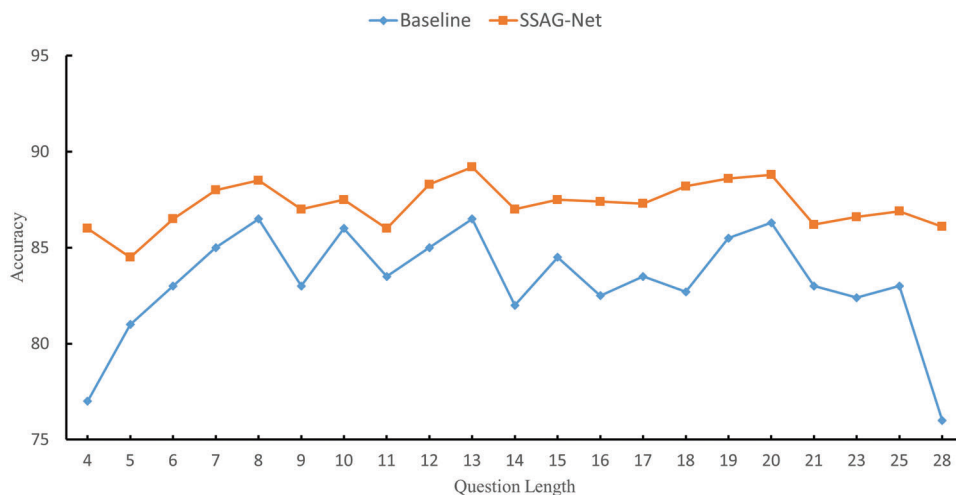


Figure 6: Accuracy of different questions' number of words

6.2 Limitations

In the present study, we propose a machine reading comprehension model based on neural network, which integrates syntax and semantics and explicitly considers the key parts of the reading text. In addition, through

the tests of the extractive MRC dataset and the single-choice MRC dataset, as well as the final ablation test and the accuracy test for different question lengths, we prove that the effect of our model on the extractive reading comprehension task and the multiple-choice machine reading comprehension task is significantly improved.

At the same time, there are some limitations for this study. First of all, our model is based on the fusion of syntactic model and semantic model, and the encoder part of the model adopts the original BERT model. Up to now, there have been many BERT models fine-tuned for language information, so there is room for improvement in the text coding. In addition, the semantic analysis of the framework has not taken into account the connection between frameworks, which makes the understanding of semantics of the model intermittent.

This paper will continue to solve the above problems in the future research, hoping to make more contributions to the field of machine reading comprehension in the future.

7 Conclusion

We offer a new model for reading comprehension in this research, based on the BERT model, to investigate the fusion of syntactic and semantic constraints included into the attentional process to guide text modeling. This allows for a more detailed rendering of the text. As a result, we developed SSAG-Net, a three-layer context architecture that combines primitive self-attention, syntactic guided attention, and lexical unit attention. The semantic framework must pre-analyze the source text that passes through the lexical unit attention model. BERT model with pre-training is used for encoding. According to the experimental results of two machine reading comprehension tests based on span and single choice model answer extraction. Our model performs well in both of these tasks, demonstrating that syntactic and semantic processing can effectively construct reading comprehension model.

Acknowledgement: We would like to thank teachers Yuan Deyu and Sun Haichun for their helpful discussion and feedback, as well as the team's technical support.

Funding Statement: This work was supported by Ministry of public security technology research program [Grant No.2020JSYJC22], Fundamental Research Funds for the Central Universities (No.2021JKF215), Open Research Fund of the Public Security Behavioral Science Laboratory, People's Public Security University of China (2020SYS03), and Police and people build/share a smart community (PJ13-201912-0525), Key Projects of the Technology Research Program of the Ministry of Public Security (2021JSZ09).

Conflicts of Interest: The authors declare that they have no conflicts of interest regarding the publication of this paper.

References

- [1] X. L. Xu, J. L. Zheng and Z. M. Yin, "Overview of technical studies on machine reading comprehension," *Journal of Chinese Computer Systems*, vol. 41, no. 3, pp. 464–470, 2020.
- [2] K. Chen, R. Wang, M. Utiyama, L. Liu, A. Tamura *et al.*, "Neural machine translation with source dependency representation," in *Proc. of the 2017 Conf. on Empirical Methods in Natural Language Processing*, Copenhagen, Denmark, pp. 2846–2852, 2017.
- [3] K. Chen, R. Wang, M. Utiyama, E. Sumita and T. Zhao, "Syntax-directed attention for neural machine translation," in *Proc. of the AAAI Conf. on Artificial Intelligence*, New Orleans, Louisiana, USA, vol. 32, no. 1, 2018.
- [4] S. Duan, H. Zhao, J. Zhou and R. Wang, "Syntax-aware transformer encoder for neural machine translation," in *2019 Int. Conf. on Asian Language Processing (IALP)*, Shanghai, China, IEEE, pp. 396–401, 2019.
- [5] Y. Schabes, A. Abeille and A. K. Joshi, "Parsing strategies with 'lexicalized' grammars: Application to tree adjoining grammars," in *Technical Reports (CIS)*, 1988.

- [6] M. Steedman, *The Syntactic Process*, Cambridge, MA: MIT press, vol. 24, 2000.
- [7] C. Pollard and I. A. Sag, *Head-driven Phrase Structure Grammar*, in University of Chicago Press, Chicago, Illinois, USA, 1994.
- [8] S. R. Bowman, J. Gauthier, A. Rastogi, R. Gupta, C. D. Manning *et al.*, “A fast unified model for parsing and sentence understanding,” ArXiv Preprint ArXiv:1603.06021, 2016.
- [9] Z. Zhang, H. Zhao and L. Qin, “Probabilistic graph-based dependency parsing with convolutional neural network,” in *Proc. of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Berlin, Germany, pp. 1382–1392, 2016.
- [10] Z. Li, J. Cai, S. He and H. Zhao, “Seq2seq dependency parsing,” in *Proc. of the 27th Int. Conf. on Computational Linguistics*, Santa Fe, New Mexico, USA, pp. 3203–3214, 2018.
- [11] X. Ma, Z. Hu, J. Liu, N. Peng, G. Neubig *et al.*, “Stack-pointer networks for dependency parsing,” ArXiv Preprint ArXiv:1805.01087, 2018.
- [12] C. J. Fillmore, “Frame semantics and the nature of language,” in *Annals of the New York Academy of Sciences: Conf. on the Origin and Development of Language and Speech*, New York, vol. 280, no. 1, pp. 20–32, 1976.
- [13] C. F. Baker, C. J. Fillmore and J. B. Lowe, “The Berkeley framenet project,” in *36th Annual Meeting of the Association for Computational Linguistics and 17th Int. Conf. on Computational Linguistics*, Montreal, Quebec, Canada, vol. 1, pp. 86–90, 1998.
- [14] J. Devlin, M. W. Chang, K. Lee and K. Toutanova, “Bert: Pre-training of deep bidirectional transformers for language understanding,” ArXiv Preprint ArXiv:1810.04805, 2018.
- [15] X. Zhang, X. Sun and H. Wang, “Duplicate question identification by integrating framenet with neural networks,” in *Proc. of the AAAI Conf. on Artificial Intelligence*, New Orleans, Louisiana, USA, vol. 32, no. 1, 2018.
- [16] S. Guo, R. Li, H. Tan, X. Li, Y. Guan *et al.*, “A frame-based sentence representation for machine reading comprehension,” in *Proc. of the 58th Annual Meeting of the Association for Computational Linguistics*, Online, pp. 891–896, 2020.
- [17] H. Wang, M. Bansal, K. Gimpel and D. McAllester, “Machine comprehension with syntax, frames, and semantics,” in *Proc. of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th Int. Joint Conf. on Natural Language Processing (Volume 2: Short Papers)*, Beijing, China, pp. 700–706, 2015.
- [18] C. Li, Y. Wu and M. Lan, “Inference on syntactic and semantic structures for machine comprehension,” in *Thirty-Second AAAI Conf. on Artificial Intelligence*, New Orleans, Louisiana, USA, 2018.
- [19] S. Guo, Y. Guan, R. Li, X. Li and H. Tan, “Incorporating syntax and frame semantics in neural network for machine reading comprehension,” in *Proc. of the 28th Int. Conf. on Computational Linguistics*, Barcelona, Spain (Online), pp. 2635–2641, 2020.
- [20] J. Kasai, D. Friedman, R. Frank, D. Radev and O. Rambow, “Syntax-aware neural semantic role labeling with supertags,” ArXiv Preprint ArXiv:1903.05260, 2019.
- [21] Z. G. Qu, H. R. Sun and M. Zheng, “An efficient quantum image steganography protocol based on improved EMD algorithm,” *Quantum Information Processing*, vol. 20, no. 53, pp. 1–29, 2021.
- [22] Z. G. Qu, Y. M. Huang and M. Zheng, “A novel coherence-based quantum steganalysis protocol,” *Quantum Information Processing*, vol. 19, no. 362, pp. 1–19, 2020.
- [23] D. Das, D. Chen, A. F. T. Martins, N. Schneider and R. Shah, “Frame-semantic parsing,” *Computational Linguistics*, vol. 40, no. 1, pp. 9–56, 2014.
- [24] Z. Zhang, Y. Wu, J. Zhou, S. Duan, H. Zhao *et al.*, “SG-Net: Syntax-guided machine reading comprehension,” in *Proc. of the AAAI Conf. on Artificial Intelligence*, New York, USA, vol. 34, no. 5, 2020.
- [25] Y. Peng, X. Li, J. Song, Y. Luo, S. Hu *et al.*, “Verification mechanism to obtain an elaborate answer span in machine reading comprehension,” *Neurocomputing*, vol. 446, pp. 80–91, 2021.
- [26] M. Shaheen and U. Abdullah, “Carm: Context based association rule mining for conventional data,” *Computers, Materials & Continua*, vol. 68, no. 3, pp. 3305–3322, 2021.
- [27] S. Sun, J. Zhou, J. Wen, Y. Wei and X. Wang, “A DQN-based cache strategy for mobile edge networks,” *Computers, Materials & Continua*, vol. 71, no. 2, pp. 3277–3291, 2022.