

# Incorporating Syntax and Frame Semantics in Neural Network for Machine Reading Comprehension

Shaoru Guo<sup>1</sup>, Yong Guan<sup>1</sup>, Ru Li<sup>1,2\*</sup>, Xiaoli Li<sup>3</sup>, Hongye Tan<sup>1,2</sup>

1. School of Computer and Information Technology, Shanxi University, Taiyuan, China

2. Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, Shanxi University, Taiyuan, China

3. Institute for Infocomm Research, A\*Star, Singapore

{guoshaoru0928, guanyong0130}@163.com

{liru, tanhongye}@sxu.edu.cn, xlli@i2r.a-star.edu.sg

## Abstract

Machine reading comprehension (MRC) is one of the most critical yet challenging tasks in natural language understanding (NLU), where both *syntax* and *semantics* information of text are essential components for text understanding. It is surprising that *jointly considering syntax and semantics in neural networks* was never formally reported in literature. This paper makes the first attempt by proposing a novel *Syntax and Frame Semantics model for Machine Reading Comprehension* (SS-MRC), which takes full advantage of syntax and frame semantics to get richer text representation. Our extensive experimental results demonstrate that SS-MRC performs better than ten state-of-the-art technologies on machine reading comprehension task.

## 1 Introduction

Machine Reading Comprehension (MRC) requires machines to read and understand a text passage, and answer relevant questions about it, where MRC systems must have the ability to infer the meanings of underlying natural language, where both syntax and semantics are critical for text understanding.

Traditional MRC methods are *feature-based* approaches, where they first manually generate syntactic or semantic features and subsequently apply a standard machine learning model to identify a best answer. For instance, Sachan and Xing (2016) use the Abstract Meaning Representation (AMR) formalism in a max-margin framework, which only focuses on semantic information. In addition, there are studies that take both syntax and semantics into account for MRC (Wang et al., 2015; Li et al., 2018). However, these methods heavily rely on manually defined features and are difficult to generalize to other tasks.

Recently, as large QA datasets become available, *Neural-based* methods have been proposed for MRC, where even without time consuming feature engineering, they can still achieve favorable results. Especially significant progress has been achieved on MRC tasks by fine-tuning a pre-trained general purpose language model (Devlin et al., 2018; Radford et al., 2018). Despite the success of those neural models, a number of studies have found there is a huge gap between MRC deep learning models and human beings (Wang and Jiang, 2019), as they might not really understand the natural language text (Mudrakarta et al., 2018). As such, some works employ semantic knowledge in neural network to facilitate sentence modelling (Guo et al., 2020; Zhang et al., 2018a; Zhang et al., 2018b). However, no work has been focused on integrating syntax and semantics into neural network for MRC.

Note that FrameNet (Fillmore, 1976; Baker et al., 1998), as a widely adopted knowledge base, provides rich schematic scenario representation that could be potentially leveraged to better understand sentences. In this paper, we proposed a *Syntax and Frame Semantics model for Machine Reading Comprehension* (SS-MRC), which fuses Syntax and Frame Semantics into an end-to-end neural model for addressing MRC task. The key contributions of this work are summarized as follows:

1. To our best knowledge, we are *the first to explore the schema of fusing syntax and frame semantics into an end-to-end neural network* for Machine Reading Comprehension (MRC) task.

---

\*Corresponding author: Ru Li.

This work is licensed under a Creative Commons Attribution 4.0 International Licence. Licence details: <http://creativecommons.org/licenses/by/4.0/>.

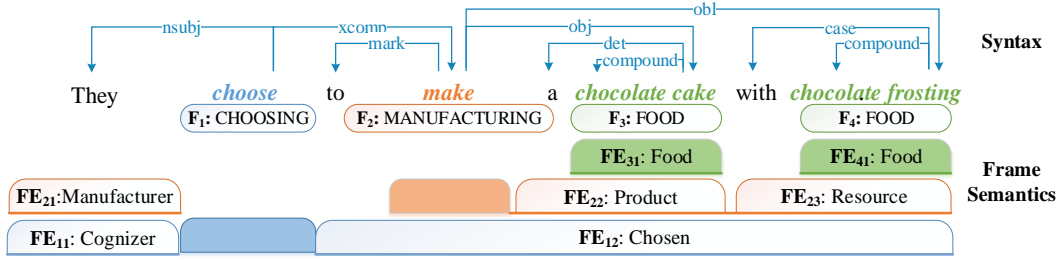


Figure 1: An Annotated Sentence with Syntax and Frame Semantics.

2. We propose a novel *Syntax and Frame Semantics for Machine Reading Comprehension* (SS-MRC) method, which takes full advantage of syntax and frame semantics for every token in the sentence/sequence to obtain richer and more comprehensive representation.
3. Our extensive experimental results demonstrate our proposed SS-MRC method is *significantly better than ten state-of-the-art methods across two benchmark datasets* for MRC task.

## 2 Syntax and Frame Semantics Labeling

We employ Stanford CoreNLP (Manning et al., 2014) to analyze the syntactic structure of every sentence in given text passage. Figure 1 shows dependency parse results (top part) of a example sentence. Words on the arrows are dependency labels, e.g., the arrow from word *choose* to word *they* indicates that *they* is the **subject** of *choose* and the dependency label of *they* is **nsubj** — a nominal syntactic subject.

In addition, we employ SEMAFOR (Das et al., 2014) to automatically process different sentences with multiple semantic annotations (Kshirsagar et al., 2015). In particular, a *Frame* (F) is defined as a composition of *Lexical Units* (LUs) and a set of *Frame Elements* (FEs). Given a sentence, if its certain word evokes a frame by matching a LU, then it is called *Target* (T). Figure 1 provides an example sentence (bottom part) with four T, namely *choose*, *make*, *chocolate cake* and *chocolate frosting*. Each T has its evoked semantic frame right below it, i.e.  $F_1, F_2, F_3, F_4$ . For each frame, its corresponding FEs are shown enclosed in the block. For example, T **choose** evokes the  $F_1$ :**Choosing** frame, and has two FEs *Cognizer*, *Chosen*, fulfilled by *They* and *to make a chocolate cake with chocolate frosting* respectively. In addition, evoked  $F_2$ :**Manufacturing** frame has three FEs, namely, *Manufacturer*, *Product*, *Resource*, fulfilled by *They*, *a chocolate cake* and *chocolate frosting*, respectively. From this example, it is very clear that both syntax and frame semantics information are very useful for MRC task.

## 3 Model Overview

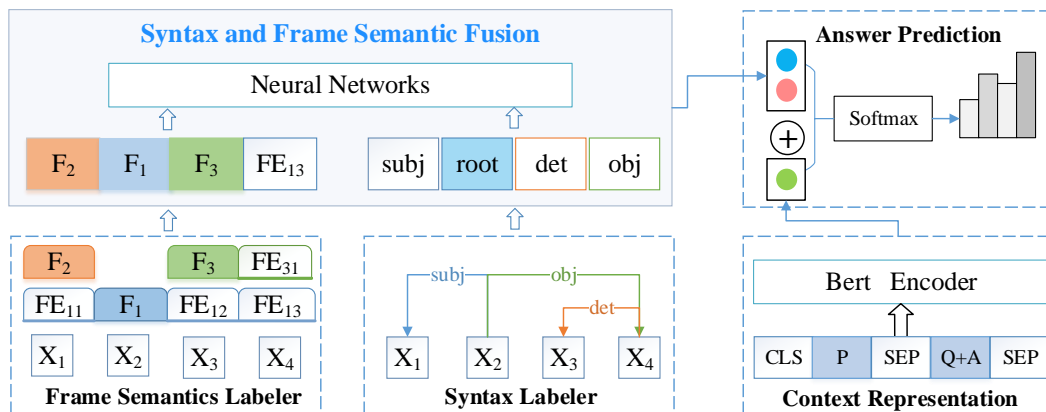


Figure 2: An Overview of the SS-MRC Model.

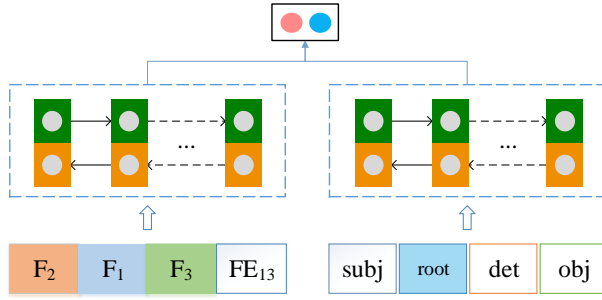


Figure 3: Siamese-based Fusion Method.

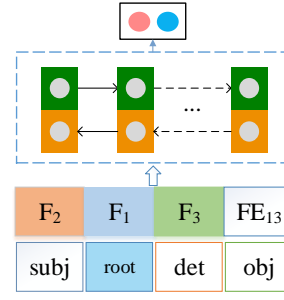


Figure 4: Location-wise Fusion Method.

Figure 2 shows our proposed SS-MRC model, consisting of five modules: three input modules (semantics, syntax, and context), one fusion module and one answer prediction module. We first pack passage, question and candidate answer into a sequence  $x = \{x_1, x_2, \dots, x_n\}$ . Input module takes in source context  $x$  and external feature text  $x^{sf}$ , i.e., syntactic context  $x^s$  and frame semantic context  $x^f$ . In particular, syntactic context  $x^s$  are produced by replacing words with their dependency labels, while frame semantic context  $x^f$  are produced by replacing words with frames and frame elements (Guo et al., 2020). Then Bert (Devlin et al., 2018) is employed to encode the source context  $x$  into a vector  $g^x$ . After that, Syntax and Frame Semantics Fusion module fuses syntactic context  $x^s$  and frame semantic context  $x^f$  into a feature vector  $g^{sf}$ , which will be elaborated in next subsection. Finally, Answer Prediction module predicts answers based on both source context representation  $g^x$  and overall feature representation  $g^{sf}$ .

## 4 Syntax and Frame Semantics Fusion Module

In this paper, we explore three different fusion methods to generate  $g^{sf}$  by integrating  $x^s$  and  $x^f$ . Note that our *Syntax and Frame Semantics Fusion Module* is backbone-free, which indicates that we can use any existing neural models, i.e., LSTM, GRU and Transformer. In this work, we use Bi-LSTM as our backbone model.

### 4.1 Siamese-based Fusion Method (SFM)

Siamese-based Fusion Method (SFM) is a straightforward idea. Its architecture, as shown in Figure 3, consists of two sub-networks (Bromley et al., 1993). We run Bi-LSTM on syntactic context  $x^s$  and frame semantic context  $x^f$  independently, and then aggregate their vectorized representations into a vector  $g^{sf}$ :

$$g^s = BiLSTM(x^s) \quad (1)$$

$$g^f = BiLSTM(x^f) \quad (2)$$

$$g^{sf} = f(g^s \oplus g^f) \quad (3)$$

Where  $\oplus$  is the concatenation of  $g^s$  and  $g^f$ , and  $f(\cdot)$  is a non-linear transformation.

### 4.2 Mixed-based Fusion Method (MFM)

While the Siamese structure is easy to train, there is no interaction between the two feature text  $x^s$  and  $x^f$  during the training process, which causes information loss (Wang et al., 2017). As such, we propose a Mixed-based Fusion Method (MFM), which directly concatenates the syntactic context  $x^s$  and frame semantic context  $x^f$  into a sequence  $x^{sf}$ , and then performs a single BiLSTM on  $x^{sf}$  to get vector  $g^{sf}$ .

$$g^{sf} = BiLSTM(Concat(x^s, x^f)) \quad (4)$$

### 4.3 Location-wise Fusion Method (LFM)

We observe that every token/word can concurrently have both syntax and frame semantics information. Thus, instead of simply mixing up them in sentence level, like above two methods, we design a novel Location-wise Fusion Method (LFM) to coherently integrate both syntax and frame semantic information at token level, obtaining a better sentence representation, shown in Figure 4.

Method	MCTest-160 (%)	MCTest-500 (%)
Sachan and Xing (2016) + <i>Semantics</i>	–	70.33
Wang et al. (2015) + <i>Syntax</i> + <i>Semantics</i>	75.27	69.94
Li et al. (2018) + <i>Syntax</i> + <i>Semantics</i>	74.58	72.67
Attentive Reader (Hermann et al., 2015)	46.3	41.9
Neural Reasoner (Peng et al., 2015)	47.6	45.6
Parallel-Hierarchical (Trischler et al., 2016)	74.58	71.00
Reading Strategies (Sun et al., 2018)	81.7	82.0
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) + <i>Semantics</i>	86.1	84.2
<b>SS-MRC</b>	<b>87.2</b>	<b>86.7</b>

Table 1: The Performance Comparison of 11 Different Models on Two MCTest Datasets.

Method	160 (%)	500 (%)	Method	160 (%)	500 (%)
Bert	82.5	80.9	SS-MRC	87.2	86.7
SS-MRC (SFM)	86.2	84.5	-Syntax	86.1	84.2
SS-MRC (MFM)	86.6	85.0	-Frame semantics	85.4	83.0
SS-MRC (LFM)	87.2	86.7	-Syntax -Frame semantics	82.5	80.9

Table 2: Performance Comparison with Three Different Fusion Models.

Table 3: Ablation Study of SS-MRC Model.

$$x_i^{sf} = f(x_i^s \oplus x_i^f) \quad (5)$$

$$g^{sf} = BiLSTM(x^{sf}) \quad (6)$$

For simplicity, we use  $x_i^s$  and  $x_i^f$  to represent the  $i$ -th location information of  $x^s$  and  $x^f$  respectively.  $x_i^{sf}$  is the concatenation ( $\oplus$ ) of  $x_i^s$  and  $x_i^f$ . Thus,  $x^{sf}$  is represented as  $\{x_1^{sf}, \dots, x_N^{sf}\}$ .

## 5 Experiments

### 5.1 Datasets for MRC

We employ MCTest (Richardson et al., 2013) to test the performance of different models for multiple-choice machine comprehension task. It consists of two data sets, namely MCTest-160 and MCTest-500.

### 5.2 Implementation Details

Our implementation is based on the PyTorch of BERT (Devlin et al., 2018) and Bi-LSTM (Zhang et al., 2018a). We have used a single GPU, Nvidia P100 with 16G memory, for training our models. Adam has been selected as our optimizer with a batch size of 8, and the initial learning rate is set as  $5e-5$ .

### 5.3 Experiment Results

Based on standard training-test setting of MCTest, Table 1 shows our SS-MRC model achieves 87.2% and 86.7% accuracy on MCTest-160 and MCTest-500 respectively, which is better than ten state-of-the-art methods consistently, including three feature-based models (the first block), six neural-based models without syntax and semantic (the second block), and 1 neural method with Frame semantic (FSR).

Recall in Section 4, we proposed three different methods, namely, SFM, MFM, LFM, to integrate syntax and semantic information. Table 2 shows their detailed results. We have the following observations:

(1) No matter which of the three fusion methods we choose, their performance are all better than standard BERT model, indicating both syntax and frame semantic information are valuable in helping language understanding, and thus they can boost reading comprehension performance.

Passage	...They chose to <b>make a <i>chocolate cake</i></b> with <i>chocolate frosting</i> . She helped measure the <i>flour</i> , the <i>sugar</i> ... They ate the chocolate cake at <i>Julia's</i> party with scoops of vanilla <i>ice cream</i> and fresh <i>strawberries</i> . Annie gave their dog, <i>Sunny</i> ...
Question	What did Annie and her mother <b>make</b> ?
Option	A) flour and sugar                      * B) cake and frosting C) ice cream and strawberries    D) Julia and Sunny
Frame Semantic	{flour, sugar, cake, frosting, ice cream, strawberries} ∈ Food {Julia, Sunny} ∉ Food <b>Make</b> in the given passage and question evokes the same Frame <b>Manufacturing</b> .
Syntax	<i>cake</i> and <i>frosting</i> are <i>obj</i> and <i>obl</i> of <b>make</b> in the passage, and <i>what</i> is <i>obj</i> of <b>make</b> in the question.

Table 4: A Case Study Example. Correct answer is marked with (\*).

(2) LFM performs better than SFM and MFM, signifying that location-wise fusion method is more effective. As a token within a sentence typically has both syntax and multi-semantic annotations/functions, token level integration can systematically integrate corresponding syntax and semantic information.

To evaluate the contributions of different key factors in our SS-MRC method, three ablation studies are performed. From their results in Table 3, we observe both syntax and frame semantics contribute to the overall performance of our model, with frame semantics contributing more significantly than syntax. Note that the performance of SS-MRC (without syntax) is exactly the same as FSR (Guo et al., 2020).

## 5.4 Case Study

For case study, Table 4 shows an example in MCTest, where our proposed model is able to answer it correctly. Note both *cake*, *frosting* et al. belong to the **Food** Frame, while *Julia* and *Sunny* evoke two different Frames **People** and **Animals** respectively. The target word **Make** in the given passage and question evokes Frame **Manufacturing**. As shown in Figure 1 and Figure 5, we know that *cake* and *frosting* are *obj* and *obl* of **make** in the passage, and *what* is *obj* of **make** in the question, so we can infer that 'What' in question refers to 'cake and frosting' in passage.

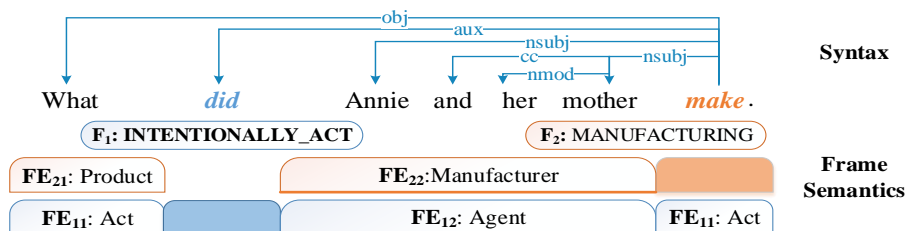


Figure 5: An Example from MCTest Demonstrates the Effectiveness of Frame Semantic and Syntax.

## 6 Conclusion

We propose a novel syntax and frame semantic fusion method for MRC in a neural network, which, to our best knowledge, is the first attempt in this area. Our extensive experimental results demonstrate it works better than ten state-of-the-art methods for the challenging machine reading comprehension task.

## Acknowledgments

We thank the anonymous reviewers for their helpful comments and suggestions. This work was supported by the National Key Research and Development Program of China (No.2018YFB1005103) and the National Natural Science Foundation of China (No.61772324, No.61936012).

## References

- Collin F. Baker, Charles J. Fillmore, and John B. Lowe. 1998. The berkeley framenet project. In *Proceedings of the 17th International Conference on Computational Linguistics*, COLING '98, pages 86–90, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Jane Bromley, Isabelle Guyon, Yann LeCun, Eduard Säckinger, and Roopak Shah. 1993. Signature verification using a siamese time delay neural network. In *Proceedings of the 6th International Conference on Neural Information Processing Systems*, NIPS93, page 737744, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.
- Dipanjan Das, Desai Chen, Andr F. T. Martins, Nathan Schneider, and Noah A. Smith. 2014. Frame-semantic parsing. *Computational Linguistics*, 40(1):9–56.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. BERT: pre-training of deep bidirectional transformers for language understanding. *CoRR*, abs/1810.04805.
- Charles J. Fillmore. 1976. Frame semantics and the nature of language. *Annals of the New York Academy of Sciences*, 280(1):20–32.
- Shaoru Guo, Ru Li, Hongye Tan, Xiaoli Li, Yong Guan, Hongyan Zhao, and Yueping Zhang. 2020. A frame-based sentence representation for machine reading comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 891–896, Online, July. Association for Computational Linguistics.
- Karl Moritz Hermann, Tomas Kocisky, Edward Grefenstette, Lasse Espeholt, Will Kay, Mustafa Suleyman, and Phil Blunsom. 2015. Teaching machines to read and comprehend. In C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems 28*, pages 1693–1701. Curran Associates, Inc.
- Meghana Kshirsagar, Sam Thomson, Nathan Schneider, Jaime Carbonell, Noah A. Smith, and Chris Dyer. 2015. Frame-semantic role labeling with heterogeneous annotations. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 218–224, Beijing, China, July. Association for Computational Linguistics.
- Chenrui Li, Yuanbin Wu, and Man Lan. 2018. Inference on syntactic and semantic structures for machine comprehension.
- Christopher D. Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, and David McClosky. 2014. The Stanford CoreNLP natural language processing toolkit. In *Association for Computational Linguistics (ACL) System Demonstrations*, pages 55–60.
- Pramod Kaushik Mudrakarta, Ankur Taly, Mukund Sundararajan, and Kedar Dhamdhere. 2018. Did the model understand the question? *CoRR*, abs/1805.05492.
- Baolin Peng, Zhengdong Lu, Hang Li, and Kam-Fai Wong. 2015. Towards neural network-based reasoning. *CoRR*, abs/1508.05508.
- Alec Radford, Karthik Narasimhan, Tim Salimans, and Ilya Sutskever. 2018. Improving language understanding by generative pre-training.
- Matthew Richardson, Christopher J.C. Burges, and Erin Renshaw. 2013. MCTest: A challenge dataset for the open-domain machine comprehension of text. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pages 193–203, Seattle, Washington, USA, October. Association for Computational Linguistics.
- Mrinmaya Sachan and Eric Xing. 2016. Machine comprehension using rich semantic representations. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 486–492, Berlin, Germany, August. Association for Computational Linguistics.
- Kai Sun, Dian Yu, Dong Yu, and Claire Cardie. 2018. Improving machine reading comprehension with general reading strategies. *CoRR*, abs/1810.13441.
- Adam Trischler, Zheng Ye, Xingdi Yuan, Jing He, Philip Bachman, and Kaheer Suleman. 2016. A parallel-hierarchical model for machine comprehension on sparse data. *CoRR*, abs/1603.08884.

- Chao Wang and Hui Jiang. 2019. Explicit utilization of general knowledge in machine reading comprehension. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 2263–2272, Florence, Italy, July. Association for Computational Linguistics.
- Hai Wang, Mohit Bansal, Kevin Gimpel, and David McAllester. 2015. Machine comprehension with syntax, frames, and semantics. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 700–706, Beijing, China, July. Association for Computational Linguistics.
- Zhiguo Wang, Wael Hamza, and Radu Florian. 2017. Bilateral multi-perspective matching for natural language sentences. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17*, pages 4144–4150.
- Xiaodong Zhang, Xu Sun, and Houfeng Wang. 2018a. Duplicate question identification by integrating framenet with neural networks. In *Thirty-Second AAAI Conference on Artificial Intelligence*.
- Zhuosheng Zhang, Yuwei Wu, Zuchao Li, Shexia He, Hai Zhao, Xi Zhou, and Xiang Zhou. 2018b. I know what you want: Semantic learning for text comprehension. *CoRR*, abs/1809.02794.
- Zhuosheng Zhang, Yuwei Wu, Hai Zhao, Zuchao Li, Shuailiang Zhang, Xi Zhou, and Xiang Zhou. 2019. Semantics-aware bert for language understanding.