A Frame-based Sentence Representation for Machine Reading Comprehension

Shaoru Guo[1], **Ru Li**[1,2*], **Hongye Tan**[1,2], **Xiaoli Li**[3], **Yong Guan**[1], **Hongyan Zhao**[1] **and Yueping Zhang**[1]
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,guanyong0130,13233587117}@163.com
{liru,tanhongye}@sxu.edu.cn, xlli@ntu.edu.sg
hongyanzhao@tyust.edu.cn

Abstract

Sentence representation (SR) is the most crucial and challenging task in Machine Reading Comprehension (MRC). MRC systems typically only utilize the information contained in the sentence itself, while human beings can leverage their semantic knowledge. To bridge the gap, we proposed a novel *Frame-based Sentence Representation* (FSR) method, which employs frame semantic knowledge to facilitate sentence modelling. Specifically, different from existing methods that only model lexical units (LUs), Frame Representation Models, which utilize both LUs in frame and Frame-to-Frame (F-to-F) relations, are designed to model frames and sentences with attention schema. Our proposed FSR method is able to integrate multiple-frame semantic information to get much better sentence representations. Our extensive experimental results show that it performs better than 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. Human beings can easily understand the meaning of a sentence based on their semantic knowledge. For instance, given a sentence *Katie bought some chocolate cookies*, people know *Katie* is a *buyer*, *chocolate cookies* are *goods* and belong to *Food* class etc. Existing machine learning approaches, however, face great challenges to address complicated MRC questions, as they do not have above semantic knowledge.

Nevertheless, FrameNet (Fillmore, 1976; Baker et al., 1998), as a knowledge base, provides schematic scenario representation that could be potentially leveraged to better understand sentences.

F	**Commerce_buy**
FEs	Buyer, Goods, ...
LUs	**buy.v**, buy.n, buyer.n, purchase.n,...
T	$[Katie]_{Buyer}$ **bought**$_{\textbf{Commerce\_buy}}$ $[some\ chocolate\ cookies]_{Goods}$
F-to-F	Commerce_buy—-Shopping—Seeking—Locating

Table 1: Example of F, FEs, LUs, T and F-to-F.

It enables the development of wide-coverage frame parsers (Gildea and Jurafsky, 2002; Das et al., 2014), as well as various real-world applications, ranging form event recognition (Liu et al., 2016), textual entailment (Burchardt et al., 2009), question answering (Ofoghi et al., 2009), narrative schemas (Chambers and Jurafsky, 2010) and paraphrase identification (Zhang et al., 2018), etc. In particular, *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). It is worth mentioning that FrameNet arranges different relevant frames into a network by defining Frame-to-Frame (F-to-F) relations. Table 1 provides an example of F, FEs, LUs, T and F-to-F, where target word **bought** in sentence *Katie bought some chocolate cookies* evokes a frame ***Commerce_buy*** as it matches with a LU **buy**. Note target word **chocolate cookies** evokes a different frame ***Food***.

How to utilize semantic knowledge from FrameNet? We observe the existing works mainly focus on LU vector embedding within a frame (Hermann and Blunsom, 2014; Bojanowski et al., 2017; Glavas et al., 2019), without *modeling a frame as a whole*. In addition, many sentences could have more than one target words that will evoke multiple frames, but there is less existing method to

*Corresponding author: Ru Li.

Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pages 891–896
July 5 - 10, 2020.

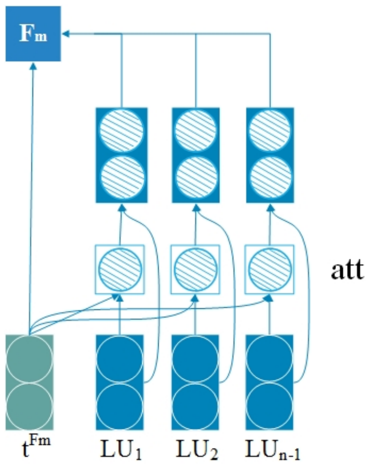

Figure 1: Lexical Units Attention Model.

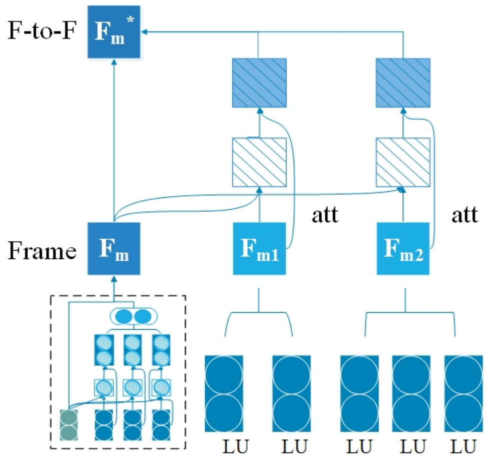

Figure 2: Frame Relation Attention Model.

integrate rich multi-frame relations from FrameNet together. To address the above problems, in this paper, we proposed a novel *Frame-based Sentence Representation* (FSR) method, which leverages rich frame semantic knowledge, including both generalizations of LUs and F-to-F relations, to better model sentences. The key contributions of this work are summarized as follows:

1. We propose novel attention-based frame representation models, which take full advantage of LUs and F-to-F relations to model frames with attention schema.

2. We propose a new *Frame-based Sentence Representation* (FSR) method that integrates multi-frame semantic information to obtain richer semantic aggregation for better sentence representation.

3. Our experimental results demonstrate our proposed frame-based sentence representation (FSR) method is very effective on Machine Reading Comprehension (MRC) task.

2 Frame Representation Model

In this section, we present our *Frame Representation Model*, considering both LUs and F-to-F.

Let $F = \{F_1, F_2, \ldots, F_m, \ldots\}$ represents a set of all frames in FrameNet, where $F_m \in \mathcal{R}^H$ is the representation of *m*-th frame of F. Let $U^{F_m} = \{u_1^{F_m}, u_2^{F_m}, \ldots, u_n^{F_m}, \ldots\}$ be the LUs set of F_m, where $U^{F_m} \in \mathcal{R}^{(H \cdot N)}$, N stands for the total number of LUs in F_m, and $u_n^{F_m}$ be the *n*-th LU of F_m. t^{F_m} is a target word, matching a LU in F_m. We proposed 3 different frame representation models.

2.1 Lexical Units Aggregation Model (LUA)

Lexical Units Aggregation Model (LUA) is a straightforward idea. Given a frame F_m, it averages all its underlying LU representation $u_n^{F_m}$ ($u_n^{F_m} \in U^{F_m}$) to represent the frame entirely:

$$F_m = \frac{1}{N} \sum_{U^{F_m}} u_n^{F_m} \tag{1}$$

2.2 Lexical Units Attention Model (TLUA)

Each frame in above LUA model has the same representation for different sentences, as they do not distinguish the importance of each LU in the frame. To address this issue, we propose TLUA model, utilizing an attention scheme to automatically weight different LUs for the frame, according to target word T in the given sentence, shown in Figure 1.

More specifically, we compute the weighted sum of target word T's representation and other LUs' representations based on their importance wrt T. In other words, we emphasize T as it occurs in the given sentence, which can reduce the potential noise introduced by irrelevant LUs in the same frame. It should be noted that we encode multiple word target by averaging of all words representations in it.

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

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

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

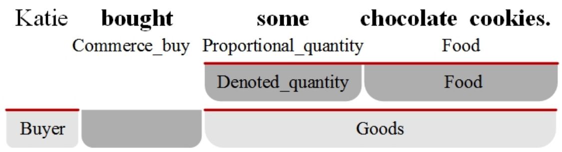

Figure 3: A sentence of FrameNet annotations.

2.3 Frame Relation Attention Model (FRA)

The key problem in MRC is to analyze semantic relations among multiple sentences. As such, we propose a novel FRA model, which takes advantage of F-to-F relations to get much richer semantic information, shown in Figure 2.

Given frame F_m, $F_m^+ = \{F_{m,1}, \dots, F_{m,w}, \dots\}$ represents its expanded frames, including all the frames that can be linked to F_m through F-to-F relation chains in FrameNet, with no more than 3 hops to only keep close relations. Note attention schemes have been designed for both *intra-frame* and *inter-frames*. Particularly, intra-frame attention focuses on relevant LUs, while inter-frames attention emphasizes relevant frames, avoiding the influence from less relevant but linked frames.

$$F_m^* = F_m + \sum_{w=1}^{W} att(F_{m,w}) \cdot F_{m,w} \tag{4}$$

$$att(F_{m,w}) = \frac{exp(F_m \cdot F_{m,w})}{\sum_{k=1}^{W} exp(F_m \cdot F_{m,k})} \tag{5}$$

3 Frame-based Sentence Representation

Given a sentence $s = \{x_1, x_2, \dots, x_k, \dots\}$ where each x_k is a word, let T_k be the k-th frame-evoking target of s, and T_k evokes F_k frame. FE_{ki} denotes the i-th *frame element* of F_k, and P_{ki} denotes the i-th span fulfilling FE_{ki}. We define a frame semantic quadruple $c_k =< T_k, F_k, FE_{kn}, P_{kn} >$, where c_k represents the k-th quadruple of s.

3.1 Sentence Semantic Annotations with Multiple Frames

In this paper, we employ SEMAFOR (Das et al., 2014) to automatically process sentences with multiple semantic annotations (Kshirsagar et al., 2015).

Figure 3 provides an example sentence with three T, namely *bought*, *some*, *chocolate cookies*. Each T has its evoked semantic frame right below it. For each frame, its FE are shown enclosed in the block where dark grey indicates the corresponding T, and the words fulfilling the FEs are connected to the corresponding text. For example, T **bought** evokes the ***Commerce_buy*** frame, and has the *Buyer*, *Goods* FEs fulfilled by *Katie* and *some chocolate cookies*.

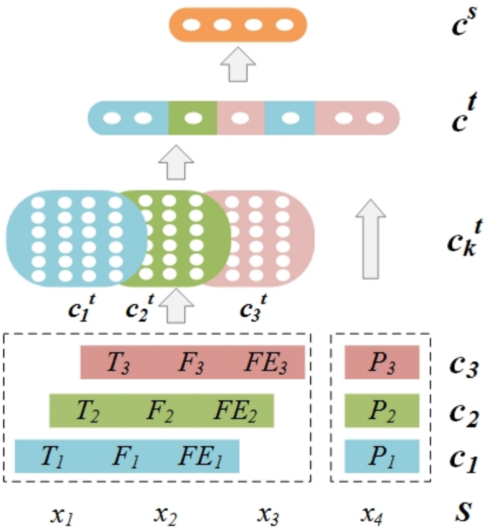

Figure 4: Frame Integration Representation Model.

The sentence s in Figure 3 has three quadruples:

1. c_1= <bought, *Commerce_buy*, [Buyer, Goods], [Katie, chocolate cookies]>

2. c_2= < some, *Proportional_quantity*, [Denoted_quantity], [some]>

3. c_3= <chocolate cookies, *Food*, [Food], [chocolate cookies]>

3.2 Frame Integration Representation

In Figure 4, c_k (k=1, 2, 3) is the input. We first compute its matrix representation c_k^t, with columns denoting different semantic information. Then, we formalize sentence representation as follows:

$$c^s = \mathcal{N}(c^t) \tag{6}$$

$$c^t = \phi(c_k^t, P_k) \quad (k = 1, \dots, K) \tag{7}$$

Where K represents the total number of quadruples in the sentence. $\phi(c_k^t, P_k)$ is aggregate operation, used to form frame set representation c^t based on the information of P and T in the sequence. Finally, we encode sentence information by neural network models.

4 Experiments

4.1 Models for MRC

To better analyze the performance of our proposed method on MRC, we apply both BERT (Devlin et al., 2018) and LSTM (Hochreiter and Schmidhuber, 1997) as our neural models. Also, we construct the input as: the passage as sequence A, and the

Method	MCTest-160 (%)	MCTest-500 (%)
Richardson et al. (2013)	69.16	63.33
Wang et al. (2015)	75.27	69.94
Li et al. (2018)	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 (Zhang et al., 2019)	73.8	80.4
BERT+DCMN+ (Zhang et al., 2019)	85.0	86.5
FSR	86.1	84.2

Table 2: The Performance Comparison of 10 Different Models on Two MCTest Datasets.

Method	160 (%)	500 (%)
Bert (Zhang et al., 2019)	73.8	80.4
Bert (Our implementation)	82.5	80.9
Bert+LUA	82.7	79.5
Bert+TLUA	84.6	82.7
Bert+FRA	86.1	84.2
bi-LSTM	54.2	49.5
bi-LSTM+LUA	59.4	57.5
bi-LSTM+TLUA	61.5	58.2
bi-LSTM+FRA	62.7	59.6

Table 3: Performance Comparison with Three Different Frame Representation Models.

Passage	Katie went to the store...She looked around for the *flowers*. She wanted cookies not *chips*. She **found** some ***chocolate cookies***. Katie then looked for a *bow*. ...
Question	What snack did Katie **buy**?
Option	A) Chips B) Chocolate cookies C) Flowers D) Bows
Answer	B
Frame Semantic	$\{$Chips , Chocolate cookies$\}$ $\in$ Food $\{$Flowers , Bows$\}$$\notin$Food **Found** and **Buy** have relations, as their frames are connected.

Table 4: A Case Study Example.

concatenation of question and one choice of answer as sequence B.

In addition, we apply a linear layer and a softmax layer on the final hidden state, and maximize the log-probability of correct labels during training.

4.2 Datasets for MRC

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

4.3 Experiment Results

Table 2 shows our FSR model achieves 86.1% accuracy on MCTest-160, which is significantly better than all the nine state-of-the-art methods. In addition, it also achieves very competitive results on MCTest-500, i,e, much better than eight existing methods, slightly worse than BERT+DCMN+ model. This is encouraging, as our model is much simpler than BERT+DCMN+, which uses much more sophisticated architecture.

Recall in Section 2, we proposed three different methods, namely, LUA, TLUA, FRA, for frame representation. Table 3 shows their detailed results:

(1) No matter for BERT or bi-LSTM, if we add frame semantic information, the performance improves by several percents, indicating frame information is valuable in semantic understanding.

(2) Comparing TLUA with LUA, TLUA performs better, signifying attention scheme in TLUA can capture semantic information more accurately.

(3) Finally, FRA further improves LUA and TLUA's performance, as sentences within a passage typically have semantic connections with each other, and it is thus necessary to take advantage of F-to-F relations to enrich semantic information.

4.4 Case Study

For case study, Table 4 shows an example in MCTest, where we are able to answer it correctly. Both *Chips*, *Chocolate cookies* belong to the ***Food*** frame, while *Flowers* and *Bows* evoke two different frames ***Plants*** and ***Accoutrements*** respectively. The target words **Found** and **Buy** in the given passage/question evoking different frames ***Locating*** and ***Commerce_buy*** — note in FrameNet they are connected due to their semantic relations, facilitating us to find answer B) Chocolate cookies.

5 Conclusion

We propose a novel Frame-based Sentence Representation method, which integrates multi-frame semantic information to facilitate sentence modelling. Our extensive experimental results demonstrate it works very well 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.61936012, No.61772324).

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.

Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. 2017. Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, 5:135–146.

Aljoscha Burchardt, Marco Pennacchiotti, Stefan Thater, and Manfered Pinkal. 2009. Assessing the impact of frame semantics on textual entailment. *Natural Language Engineering*, 15(4):527550.

Nathanael Chambers and Dan Jurafsky. 2010. A database of narrative schemas. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta. European Language Resources Association (ELRA).

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.

Daniel Gildea and Daniel Jurafsky. 2002. Automatic labeling of semantic roles. *Computational Linguistics*, 28(3):245–288.

Goran Glavas, Robert Litschko, Sebastian Ruder, and Ivan Vulic. 2019. How to (properly) evaluate cross-lingual word embeddings: On strong baselines, comparative analyses, and some misconceptions. *CoRR*, abs/1902.00508.

Karl Moritz Hermann and Phil Blunsom. 2014. Multilingual models for compositional distributed semantics. *CoRR*, abs/1404.4641.

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.

Sepp Hochreiter and Jrgen Schmidhuber. 1997. Long short-term memory. *Neural Computation*, 9(8):1735–1780.

Meghana Kshirsagar, Sam Thomson, Nathan Schneider, Jaime G 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.

Chenrui Li, Yuanbin Wu, and Man Lan. 2018. Inference on syntactic and semantic structures for machine comprehension. In *Thirty-Second AAAI Conference on Artificial Intelligence*.

Shulin Liu, Yubo Chen, Shizhu He, Kang Liu, and Jun Zhao. 2016. Leveraging FrameNet to improve automatic event detection. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2134–2143, Berlin, Germany. Association for Computational Linguistics.

Bahadorreza Ofoghi, John Yearwood, and Liping Ma. 2009. The impact of frame semantic annotation levels, frame-alignment techniques, and fusion methods on factoid answer processing. *Journal of the American Society for Information Science and Technology*, 60(2):247–263.

Baolin Peng, Zhengdong Lu, Hang Li, and Kam-Fai Wong. 2015. Towards neural network-based reasoning. *CoRR*, abs/1508.05508.

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. 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.

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. Association for Computational Linguistics.

Shuailiang Zhang, Hai Zhao, Yuwei Wu, Zhuosheng Zhang, Xi Zhou, and Xiang Zhou. 2019. Dcmn+: Dual co-matching network for multi-choice reading comprehension. *arXiv preprint arXiv:1908.11511*.

Xiaodong Zhang, Xu Sun, and Houfeng Wang. 2018. Duplicate question identification by integrating framenet with neural networks. In *Thirty-Second AAAI Conference on Artificial Intelligence*.