JBNU at MRP 2020: AMR Parsing Using a Joint State Model for Graph-Sequence Iterative Inference

Jinwoo Min[†], Seung-Hoon Na[†], Jong-Hun Shin[‡] and **Young-Kil Kim[‡]**
[†]Computer Science and Engineering, Jeonbuk National University, South Korea
[‡]Electronics and Telecommunication Research Institute (ETRI), South Korea
`jinwoomin4488@gmail.com, nash@jbnu.ac.kr`
`{jhshin82,kimyk}@etri.re.kr`

Abstract

This paper describes the Jeonbuk National University (JBNU) system for the 2020 shared task on Cross-Framework Meaning Representation Parsing at the Conference on Computational Natural Language Learning. Among the five frameworks, we address only the abstract meaning representation framework and propose a *joint state* model for the graph-sequence iterative inference of (Cai and Lam, 2020) for a simplified graph-sequence inference. In our joint state model, we update only a single joint state vector during the graph-sequence inference process instead of keeping the dual state vectors, and all other components are exactly the same as in (Cai and Lam, 2020).

1 Introduction

Recent studies on meaning representation parsing (MRP) have focused on different semantic graph frameworks (Oepen et al., 2019) such as bilexical semantic dependency graphs (Peng et al., 2017; Wang et al., 2018; Peng et al., 2018; Dozat and Manning, 2018; Na et al., 2019), a universal conceptual cognitive annotation (Hershcovich et al., 2017, 2018), abstract meaning representation (Wang and Xue, 2017; Guo and Lu, 2018; Song et al., 2019; Zhang et al., 2019; Cai and Lam, 2019, 2020; Zhou et al., 2020), and a discourse representation structure (Abzianidze et al., 2019; Liu et al., 2018; van Noord et al., 2018; Liu et al., 2019; Evang, 2019; Liu et al., 2020). To jointly address various semantic graphs, the aim of the Cross-Framework MRP task at the 2020 Conference (MRP 2020) on Computational Natural Language Learning (CoNLL) is to develop semantic graph parsing across the following five frameworks (Oepen et al., 2020): 1) **EDS**: Elementary Dependency Structures (Oepen and Lønning, 2006), 2) **PTG**: Prague Tectogrammatical Graphs

(Hajič et al., 2012), 3) **UCCA**: Universal Conceptual Cognitive Annotation (Abend and Rappoport, 2013), 4) **AMR**: Abstract Meaning Representation (Banarescu et al., 2013), and 5) **DRG**: Discourse Representation Graphs (Abzianidze et al., 2017).

For MRP 2020, we address only the AMR framework and present a *joint state* model for graph-sequence iterative inference, as a simple extension of (Cai and Lam, 2020). The graph-sequence iterative model of (Cai and Lam, 2020) incrementally constructs an AMR graph starting from an empty graph G_0 by alternatively applying two modules: 1) *Concept Solver*, which uses a previous graph hypothesis G_i to predict a new concept, and 2) *Relation Solver*, which uses a previous concept hypothesis to predict relations for the new concept.

The *dual-state* model of (Cai and Lam, 2020) deploys two state vectors x_t and y_t for the graph-sequence iterative inference, which refers to the t-th sequence hypothesis and t-th graph hypothesis, respectively. Unlike the dual state model, we instead maintain a *joint state* vector z_t, which encodes both sequence and graph hypotheses to apply a graph-sequence iterative inference in a simple and unified manner. During the iterative inference stage, we take the current joint state vector as a query vector and update the next joint state vector by applying attention mechanisms both to the text (i.e., sequence memory) and graph (i.e., graph memory) parts separately. The final joint state vector is then passed to the concept and relation solvers, which predict new concepts and their relations, respectively, as with the dual state model by (Cai and Lam, 2020).

We submitted the results of our AMR parsing model during the post-evaluation stage and ranked between 3rd and 4th place among the participants in the official results under the cross-framework

metric [1].

The remainder of this paper is organized as follows: Section 2 presents the detailed architecture of our system. Section 3 describes the detailed process used for training biaffine attention models. Section 4 provides the official results of MRP 2020. Finally, some concluding remarks and a description of future research are given in Section 5.

2 Model

Figure 1 shows the neural architecture based on the joint state model for a graph-sequence iterative inference.

The neural architecture consists of five components: a 1) *sequence encoder*, 2) *graph encoder*, 3) *concept solver*, 4) *relation solver*, and 5) *joint state model* for a graph-sequence iterative inference.

In the following, we briefly summarize the first four components of our model, which are almost the same as those of (Cai and Lam, 2020), where only the joint state vector is used as the unified query vector for the concept and relation solvers. We then present our joint state model in more details.

2.1 Sequence encoder: Multi-layer transformer

Following the notations of (Cai and Lam, 2020), let W be the input sentence consisting of $w_1, \cdots, w_n$. The sequence encoder is based on a multi-layer transformer architecture (Vaswani et al., 2017), where inputs at the bottom layer combine the character-level features, POS tag, named entity tags, and BERT-based features. Roughly, the sequence encoder takes W as input and generates the sequence of hidden states as follows:

$$h_0, h_1, \cdots, h_n = \text{SequenceEncoder}((BOS, w_1, \cdots, w_N))$$

where h_0 corresponds to the special token BOS.

2.2 Graph encoder: Multi-layer Transformer

Suppose that G_i is the current graph consisting of i nodes, $c_1, \cdots, c_i$. The graph encoder is based on a multi-layer transformer encoder with masked self-attention and source-attention. Roughly, the

graph encoder takes G and produces the following hidden states of the concept nodes:

$$s_0, s_1, \cdots, s_i = \text{GraphEncoder}(G = \{c_1, \cdots, c_i\})$$

where s_0 corresponds to the special token BOG.

2.3 Concept solver: Attention over words

Suppose that z_t is the current joint state vector that encodes the t-th sequence and graph hypotheses. The concept solver takes z_t and generates a new concept.

$$
\begin{aligned}
q_t &= W^Q z_t \\
k_{1:n} &= W^K h_{1:n} \\
v_{1:n} &= W^V h_{1:n} \\
[\alpha_t, r_t] &= Attention(q_t, k_{1:n}, v_{1:n}) \\
z_t' &= z_t + r_t
\end{aligned}
\tag{1}
$$

where $Attention(q, k, v)$ is the attention module that takes q, k, and v as the query vector, keys, and values, respectively, and returns the attention probabilities α_t and attentive representation r_t. Given z_t', the concept generation process equipped with the copying mechanism generates a new concept as follows:

$$
\begin{aligned}
P^{(vocab)} &= softmax\left(W^{(vocab)} z_t' + b^{(vocab)}\right) \\
[p_1, p_2, p_3] &= softmax\left(W^{(switch)} z_t'\right) \\
P(c) &= p_0 \cdot P^{(vocab)}(c) + \\
&\quad p_1 \cdot \left(\sum_{i \in L(c)} \alpha_t[i]\right) + \\
&\quad p_2 \cdot \left(\sum_{i \in T(c)} \alpha_t[i]\right)
\end{aligned}
$$

where $L(c)$ and $T(c)$ are index sets of lemmas and tokens respectively, which have the surface form as a concept c defined by (Cai and Lam, 2020).

2.4 Relation solver: Multi-head attention over graph nodes

The relation solver is based on the multi-head attention over graph nodes. Suppose that z_t is the current joint state vector, G_i is the current graph, and H is the number of heads for the multi-head attention. For each head h, the relation solver first

---

[1]In other words, when our submission is included in the official ranking, our AMR parsing system ranks 4th and 3rd on the full test set and the 100-sentence LPPS sub-set, respectively.
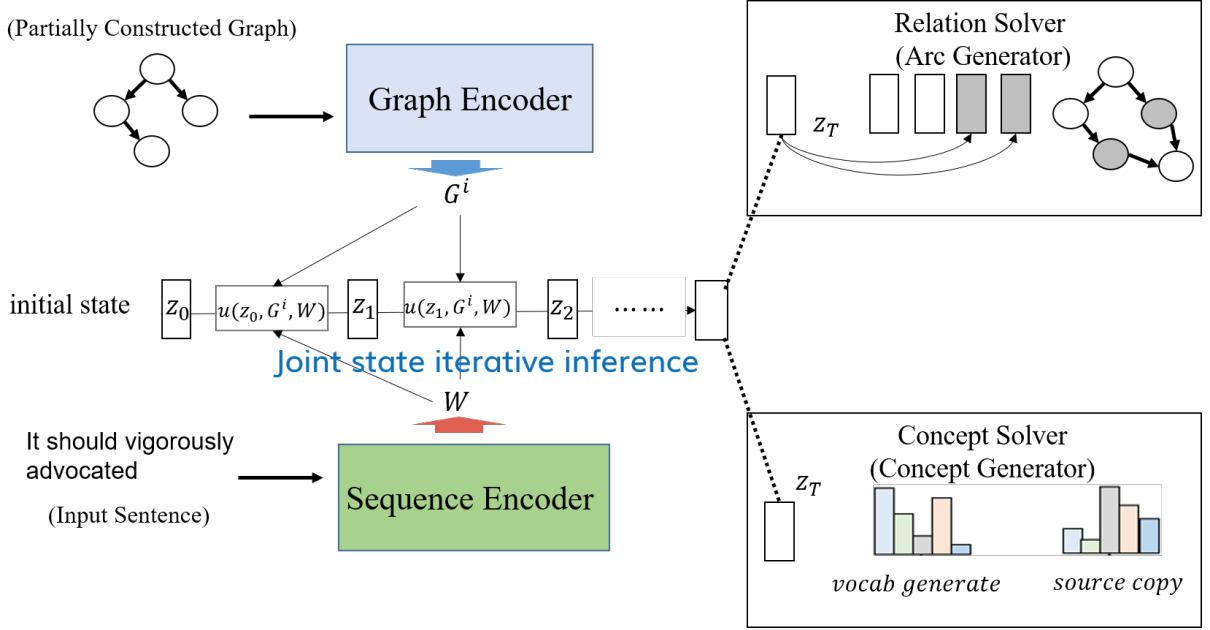
Figure 1: The neural architecture of the joint state model for a graph-sequence iterative inference.

applies the attention over graph nodes $s_0, \cdots, s_i$.

$$
\begin{aligned}
q_t^h &= W_h^Q z_t \\
k_{0:i}^h &= W_h^K s_{0:i} \\
v_{0:i}^h &= W_h^V s_{0:i} \\
\left[\beta_t^h, r_t^h \right] &= Attention(q_t^h, k_{0:i}^h, v_{0:i}^h)
\end{aligned}
\tag{2}
$$

The final edge probabilities are obtained by taking the maximum over the multi-attention probabilities:

$$
\beta_t[i] = max_{h=1}^H \beta_t^h[i] \tag{3}
$$

2.5 Iterative inference: Joint state model

The remaining part is to obtain the new joint state vector z_{t+1} from z_t. Suppose again that z_t is the current joint state vector and G_i is the current graph. The joint state model for an iterative inference is formulated as follows:

$$
\begin{aligned}
z_0 &= fusion(h_0, s_i) \\
g_t &= \sigma(W_g z_t + b_g) \\
\left[., z_t^{seq} \right] &= Attention(z_t, h_{1:n}, h_{1:n}) \\
\left[., z_t^{graph} \right] &= Attention(z_t, s_{0:i}, s_{0:i}) \\
z_{t+1} &= z_t + (1 - g_t) z_t^{seq} + g_t z_t^{graph}
\end{aligned}
$$

where W_g and b_g are the parameters for the affine transformation of z_t and $fusion(\mathbf{x}, \mathbf{y})$ is the fusion

Encoder	
lemma emb dim	300
char emb dim	50
cnn filter size	100
pos emb dim	50
ner emb dim	50
encoder layers	3
encoder dropout	0.33
Decoder	
concept emb dim	300
concept char emb dim	50
cnn filter size	100
decoder layers	3
decoder dropout	0.33

Table 1: Hyper-parameter settings

function defined by (Hu et al., 2018) [2].

3 Training

3.1 Hyperparameters

We used the Adam optimizer (Kingma and Ba, 2015) to train our AMR parsing model. Table 1 summarizes the hyper-parameters used for training these models

4 Official Results

The official results of our AMR parsing based on the joint state model, referred to as "Graph-Sequence+Joint", are summarized in Table 2,

---

[2]For simplicity, we obtain z_0 based on a function of h_0 and s_i. More generally, however, z_0 could be initialized as a function of $h_{1:n}$ and $s_{0:i}$.

which compares the results of the top-2 systems (Hitachi and ÚFAL). Overall, our system ranks between 3rd and 4th place among all participants which submitted to the AMR framework.

5 Summary and Conclusion

In this paper, we presented the Jeonbuk National University system based on a joint state model for a graph-sequence iterative inference on the AMR framework at MRP 2020 task. However, one limitation of the current joint state model is that the iterative inference process is separately formulated from the concept and relation solvers, without a tight coupling. In a future study, we plan to further elaborate the joint state models by reformulating an iterative inference based on attention results from the concept and relation solvers.

Acknowledgments

References

Omri Abend and Ari Rappoport. 2013. Universal conceptual cognitive annotation (UCCA). In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, ACL '13, pages 228–238. Association for Computational Linguistics.

Lasha Abzianidze, Johannes Bjerva, Kilian Evang, Hessel Haagsma, Rik van Noord, Pierre Ludmann, Duc-Duy Nguyen, and Johan Bos. 2017. The Parallel Meaning Bank: Towards a multilingual corpus of translations annotated with compositional meaning representations. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics*, pages 242–247.

Lasha Abzianidze, Rik van Noord, Hessel Haagsma, and Johan Bos. 2019. The first shared task on discourse representation structure parsing. In *Proceedings of the IWCS Shared Task on Semantic Parsing*.

Laura Banarescu, Claire Bonial, Shu Cai, Madalina Georgescu, Kira Griffitt, Ulf Hermjakob, Kevin Knight, Philipp Koehn, Martha Palmer, and Nathan Schneider. 2013. Abstract meaning representation for sembanking. In *Proceedings of the 7th Linguistic Annotation Workshop and Interoperability with Discourse*, pages 178–186.

Deng Cai and Wai Lam. 2019. Core semantic first: A top-down approach for AMR parsing. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 3799–3809.

Deng Cai and Wai Lam. 2020. AMR parsing via graph-sequence iterative inference. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 1290–1301.

Timothy Dozat and Christopher D. Manning. 2018. Simpler but more accurate semantic dependency parsing. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, ACL'18, pages 484–490.

Kilian Evang. 2019. Transition-based DRS parsing using stack-LSTMs. In *Proceedings of the IWCS Shared Task on Semantic Parsing*.

Zhijiang Guo and Wei Lu. 2018. Better transition-based AMR parsing with a refined search space. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, EMNLP '18, pages 1712–1722.

Jan Hajič, Eva Hajičová, Jarmila Panevová, Petr Sgall, Ondřej Bojar, Silvie Cinková, Eva Fučíková, Marie Mikulová, Petr Pajas, Jan Popelka, Jiří Semecký, Jana Šindlerová, Jan Štěpánek, Josef Toman, Zdeňka Urešová, and Zdeněk Žabokrtský. 2012. Announcing Prague Czech-English dependency treebank 2.0. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC-2012)*, LREC-2012, pages 3153–3160.

Daniel Hershcovich, Omri Abend, and Ari Rappoport. 2017. A transition-based directed acyclic graph parser for UCCA. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, ACL'17, pages 1127–1138.

Daniel Hershcovich, Omri Abend, and Ari Rappoport. 2018. Multitask parsing across semantic representations. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, pages 373–385.

Minghao Hu, Yuxing Peng, Zhen Huang, Xipeng Qiu, Furu Wei, and Ming Zhou. 2018. Reinforced mnemonic reader for machine reading comprehension. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, IJCAI '18, pages 4099–4106.

Diederick P Kingma and Jimmy Ba. 2015. Adam: A method for stochastic optimization. In *International Conference on Learning Representations*, ICLR '13.

Jiangming Liu, Shay B. Cohen, and Mirella Lapata. 2018. Discourse representation structure parsing. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 429–439.

Jiangming Liu, Shay B. Cohen, and Mirella Lapata. 2019. Discourse representation parsing for sentences and documents. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 6248–6262.

Jiangming Liu, Shay B. Cohen, and Mirella Lapata. 2020. Dscorer: A fast evaluation metric for discourse representation structure parsing. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 4547–4554. Association for Computational Linguistics.

Seung-Hoon Na, Jinwoo Min, Kwanghyeon Park, Jong-Hun Shin, and Young-Kil Kim. 2019. Jbnu at MRP 2019: Multilevel biaffine attention for semantic dependency parsing. In *Proceedings of the Shared Task on Cross-Framework*

method		tops			labels			properties			edges			all		
		P	R	F	P	R	F	P	R	F	P	R	F	P	R	F
Hitachi	lpps	0.84	0.84	0.84	0.83	0.85	0.84	0.86	0.77	0.81	0.71	0.73	0.72	0.78	0.79	0.79
	all	0.86	0.86	0.86	0.88	0.86	0.87	0.83	0.81	0.82	0.77	0.74	0.76	0.83	0.80	0.82
ÚFAL	lpps	0.86	0.86	0.86	0.85	0.87	0.86	0.78	0.71	0.75	0.69	0.71	0.70	0.77	0.79	0.78
	all	0.84	0.84	0.84	0.88	0.87	0.87	0.86	0.85	0.85	0.73	0.70	0.71	0.81	0.79	0.80
Graph-Sequence+Joint	lpps	0.86	0.86	0.86	0.79	0.80	0.79	0.54	0.45	0.49	0.68	0.67	0.68	0.74	0.73	0.74
	all	0.84	0.84	0.84	0.79	0.73	0.76	0.68	0.39	0.50	0.61	0.54	0.57	0.71	0.62	0.66

Table 2: The official results of the MRP metrics on the AMR framework, comparing the top-2 systems (Hitachi and ÚFAL) with our system using the joint state iterative inference (Graph-Sequence+Joint).

Meaning Representation Parsing at the 2019 Conference on Computational Natural Language Learning, pages 95 – 103, Hong Kong, China.

Rik van Noord, Lasha Abzianidze, Antonio Toral, and Johan Bos. 2018. Exploring neural methods for parsing discourse representation structures. *Transactions of the Association for Computational Linguistics*, 6:619–633.

Stephan Oepen, Omri Abend, Lasha Abzianidze, Johan Bos, Jan Hajič, Daniel Hershcovich, Bin Li, Tim O'Gorman, Nianwen Xue, and Daniel Zeman. 2020. MRP 2020: The Second Shared Task on Cross-framework and Cross-Lingual Meaning Representation Parsing. In *Proceedings of the CoNLL 2020 Shared Task: Cross-Framework Meaning Representation Parsing*, pages 1 – 22, Online.

Stephan Oepen, Omri Abend, Jan Hajič, Daniel Hershcovich, Marco Kuhlmann, Tim O'Gorman, Nianwen Xue, Jayeol Chun, Milan Straka, and Zdeňka Urešová. 2019. MRP 2019: Cross-framework Meaning Representation Parsing. In *Proceedings of the Shared Task on Cross-Framework Meaning Representation Parsing at the 2019 Conference on Computational Natural Language Learning*, pages 1 – 27, Hong Kong, China.

Stephan Oepen and Jan Tore Lønning. 2006. Discriminant-based MRS banking. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation*, LREC'06.

Hao Peng, Sam Thomson, and Noah A. Smith. 2017. Deep multitask learning for semantic dependency parsing. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, ACL '17, pages 2037–2048.

Hao Peng, Sam Thomson, Swabha Swayamdipta, and Noah A. Smith. 2018. Learning joint semantic parsers from disjoint data. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, NAACL '18, pages 1492–1502.

Linfeng Song, Daniel Gildea, Yue Zhang, Zhiguo Wang, and Jinsong Su. 2019. Semantic neural machine translation using AMR. *Transactions of the Association for Computational Linguistics*, 7:19–31.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Ł ukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in Neural Information Processing Systems 30*, pages 5998–6008.

Chuan Wang and Nianwen Xue. 2017. Getting the most out of AMR parsing. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, EMNLP '17, pages 1257–1268.

Yuxuan Wang, Wanxiang Che, Jiang Guo, and Ting Liu. 2018. A neural transition-based approach for semantic dependency graph parsing. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*, AAAI '18, pages 5561–5568.

Sheng Zhang, Xutai Ma, Kevin Duh, and Benjamin Van Durme. 2019. AMR parsing as sequence-to-graph transduction. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, ACL '19, pages 80–94.

Qiji Zhou, Yue Zhang, Donghong Ji, and Hao Tang. 2020. AMR parsing with latent structural information. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 4306–4319. Association for Computational Linguistics.