

HIT-SCIR at MRP 2019:

A Unified Pipeline for Meaning Representation Parsing
via Efficient Training and Effective Encoding

Wanxiang Che, Longxu Dou, Yang Xu, Yuxuan Wang, Yijia Liu, Ting Liu
Research Center for Social Computing and Information Retrieval
Harbin Institute of Technology

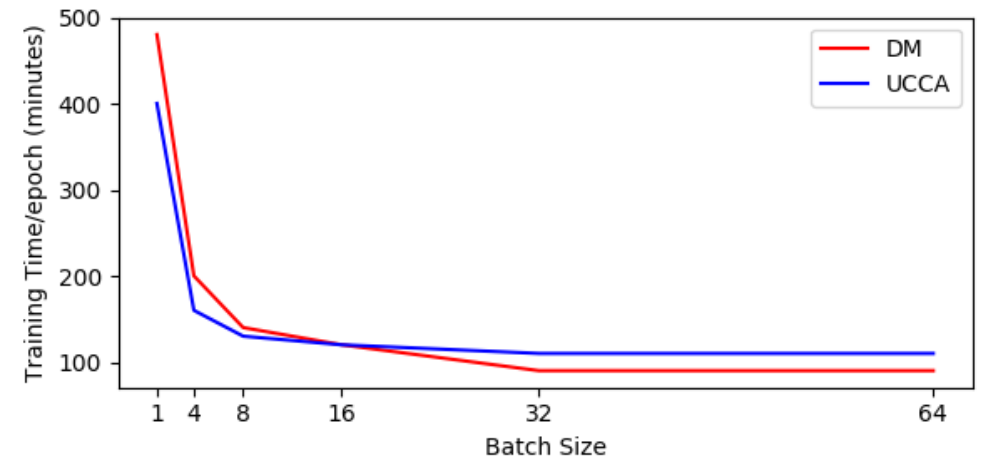
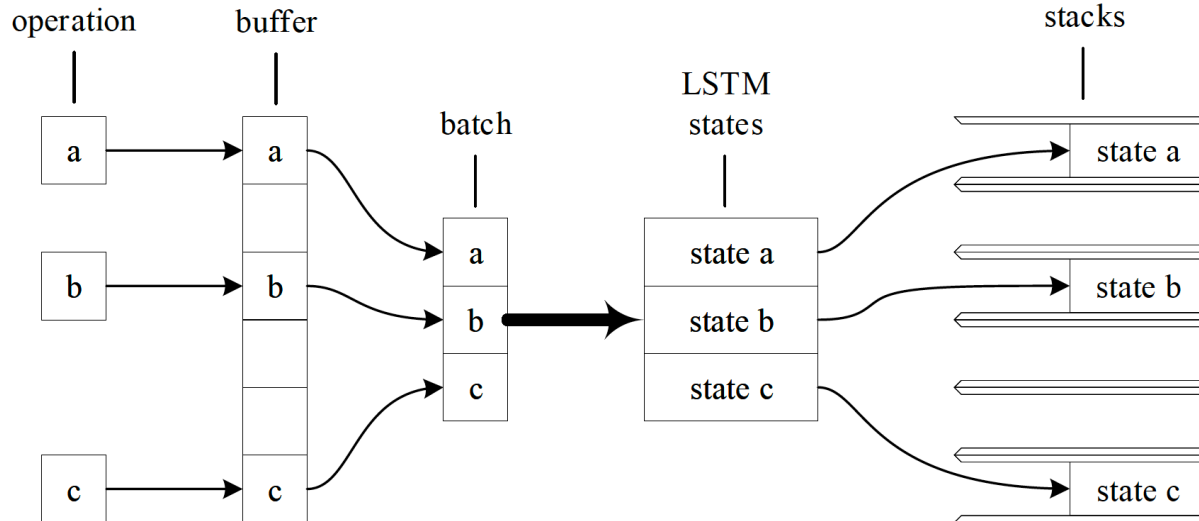
Overview of Our Techniques

- Rank 1st according to ALL-F1
- Baseline model: Transition-based Parser with Stack LSTM (Dyer et al., 2015)
- Our Extensions:
 - Efficient Training of Stack LSTM via parallel training
 - Effective Encoding via adopting BERT

System	DM	PSD	EDS	UCCA	AMR	ALL-F1
HIT-SCIR	95.08	90.55	90.75	<u>81.67</u>	72.94	<u>86.2</u>
SJTU-NICT	<u>95.50</u>	91.19	89.90	77.80	71.97	85.3
Suda-Alibaba	92.26	85.56	<u>91.85</u>	78.43	71.72	84.0
Saarland	94.69	<u>91.28</u>	89.10	67.55	66.72	81.9
Hitachi	91.02	91.21	83.74	70.36	43.86	76.0
Amazon	93.26	89.98	-	-	<u>73.38</u>	-

Parallel Training Stack-LSTM

- Aligning the similar operations in Stack-LSTM within a batch
- Computing them simultaneously



- Conduct experiments with GloVe
- 5.3x on DM
- 2.7x on UCCA

BERT is Amazing!

- We fine-tune the BERT
- Layer-wise scalar weighed BERT is adopted

Feature	DM	PSD	EDS	UCCA	AMR	Avg
GloVe	87.1	74.1	88.2	87.5	65.3	80.4
BERT(base)	94.3	83.6	91.5	92.8	71.4	86.7

- Metric: ALL-F1 based on mtool
- Dataset: Splited training data on 8:1:1 proportion

Structure vs Representation

- Transition-based Parser achieves comparable results with Graph-based Parser
- Kulmizev et al. (2019) found similar conclusion in PTB

Model	Feature	DM		PAS		PSD	
		id	ood	id	ood	id	ood
Wang et al	word2vec	89.3	83.2	91.4	87.2	76.1	73.2
Dozat et al	Glove+char	92.7	87.8	94.0	90.6	80.5	78.6
Transition	GloVe+char	86.1	79.2	89.8	85.2	72.8	68.5
Graph	GloVe+char	91.6	86.1	93.1	89.6	77.4	73.0
Transition	BERT	92.9	89.2	94.4	92.4	<u>81.6</u>	<u>81.0</u>
Graph	BERT	<u>94.1</u>	<u>90.8</u>	<u>94.8</u>	<u>92.9</u>	80.7	79.5

Wang et al: <A Neural Transition-Based Approach for Semantic Dependency Graph Parsing>

Dozat et al: <Simpler but More Accurate Semantic Dependency Parsing>

Kulmizev et al: <Deep Contextualized Word Embeddings in Transition-Based and Graph-Based Dependency Parsing – A Tale of Two Parsers Revisited>

Model Ensemble

- In follow up experiment, we obtain further improvement on lpps dataset
- Ensemble model consists of 5 single model

Systems	DM	PSD	EDS	UCCA	AMR	Avg
Single	93.98	87.41	89.83	82.61	69.03	84.57
Ensemble	94.00	87.79	89.57	83.41	71.35	85.16

Conclusion

- **Our Contribution:**
 - Efficient Training of Stack LSTM via parallel training
 - Effective Encoding through adopting BERT
- The performance gap between Graph and Transition on SDP is almost eliminated under BERT
- Our code: <https://github.com/HIT-SCIR/HIT-SCIR-CoNLL2019>