RESEARCH STATEMENT

Zhanming (Allan) Jie
ByteDance AI Lab
allan@bytedance.com

ABSTRACT

We have been working extensively on building efficient structured models for fundamental natural language processing tasks such as named entity recognition and semantic parsing. We proposed various kinds of hypergraph structures and develop efficient dynamic-programming algorithms for inference. Our research efforts demonstrated that the hypergraph-based models have a strong capability and expressivity in extracting structures from texts. We also developed a structured prediction framework that helps us perform efficient On top of that, we are currently focusing on developing structured models that are capable to perform symbolic reasoning and help us achieve the *system 2* ability (Kahneman, 2011; Bengio et al., 2021). In this research statement, we summarize our previous research works in structured prediction and layout the future direction towards symbolic reasoning.

1 STRUCTURED PREDICTION

We built structured models for the task of named entity recognition (NER) (Tjong Kim Sang & De Meulder, 2003) and semantic parsing (Zelle & Mooney, 1996). We focused on these two tasks as both of them require semantic-level understanding from natural language. Motivated by the shallow semantics conveyed by the syntactic dependency trees, we built *dependency-guided* models for NER. The property that named entities often form sub-tree structures in syntactic dependencies allows us to efficiently identify the possible candidates of named entities. We design a *dependency-guided model* (DGM) (Jie et al., 2017) based on the semi-Markov CRF (Lafferty et al., 2001; Sarawagi & Cohen, 2004) model with linear complexity on average while achieving comparable or better performance than the semi-Markov CRF. However, the above approach does not fully make use of the dependency structures such as the dependency label and high-order dependencies. Thus, we proposed the *dependency-guided LSTM-CRF* (Jie & Lu, 2019; Xu et al., 2021) to fully exploit above properties from dependency trees. Given the strong expressiveness of hypergraph-based models (Lu, 2017), we proposed several graphical models and design efficient algorithms for the NER and semantic parsing tasks (Jie & Lu, 2014; Jie et al., 2017; Jie & Lu, 2018; Jie et al., 2019; Jie & Lu, 2019; Xu et al., 2021).

1.1 EFFICIENT HYPERGRAPH INFERENCE FRAMEWORK

Different designs of hypergraphs require implementations of the particular inference algorithm. Such algorithms usually cost a significant amount of efforts to implement. In order to make the inference more generic, we have developed and continue to work on a structured prediction framework (Lu, 2017)[1] that is able to help us perform inference on almost any arbitrary structures. The underlying inference algorithm is a generic version of the *inside-outside* algorithm (Baker, 1979) where we develop the extension and apply to general hypergraph structures. As the linear-chain CRF is the most commonly used architecture in NLP, we also develop the efficient framework[2] for the community to use. Our framework is able to improve the complexity from $\mathcal{O}(N)$ to $\mathcal{O}(\log N)$ following parallel-scan algorithm (Blelloch, 1990; Rush, 2020). Along this line of work, we are interested in building more efficient algorithms for general hypergraph structures. For example, typical constituency parsing or other structured parsing tasks require efficient inference algorithm for

[1] https://github.com/sutd-statnlp/statnlp-neural
[2] https://github.com/allanj/pytorch_neural_crf

tree-based structures. How to efficiently incorporate the inference procedure into our frameworks remains as a technical challenge. The Torch-Struct (Rush, 2020) framework implemented many architectures but the underlying codebase is specific to some particular structures. Our idea is to allows developers or researchers to only focus on the structure design rather than the inference procedure.

2 Structured Symbolic Reasoning

Besides working on structured prediction tasks, we also focus on the reasoning ability of structured models (Jie et al., 2022). Specifically, we try to answer the question that "*How can we make use of the expressiveness in graphical models to perform symbolic reasoning?*". Our goal is to build an interpretable model that gives explainable predictions under the scope of modern neural architectures. We proposed a deductive system that allows us iteratively construct partial structures and perform reasoning over these structures. We adopt the math word problem (MWP) solving as the downstream task as it essentially evaluate the reasoning ability of our proposed model. To make iterative and explainable predictions, we carefully design neural module networks (Andreas et al., 2016; Gupta et al., 2020) to perform reasoning (such as *addition* and *subtraction*) among the mathematical quantities. Our experiments demonstrate that the underlying deductive system presents explainable step during the MWP solving procedure.

The underlying deductive system design not only applies to MWP solving but also other tasks that consider multi-step reasoning, such as HotpotQA (Yang et al., 2018). The reasoning happens between states (e.g., partial structures) where we apply differentiable actions to evolve the previous state to a new state. The resulting reasoning process is similar to the concept of "*reasoning chain*" (Jhamtani & Clark, 2020; Saha et al., 2020). For example, we are interested in how models exploit the relations among the rationales (i.e., relevant sentences) in reading comprehension. Such relations could be latent and we need to build latent-variable models (Jie & Lu, 2018) to find such relations. Besides, we can also apply the deductive system to dependency parsing where the tree generation happens in a top-down and step-by-step manner.

3 Self-Supervised Learning for Structures

The success of Transformers (Vaswani et al., 2017) architecture brings our attention to self-supervised learning. Transformer-based models with self-supervised pre-training are able to achieve state-of-the-art performance in downstream NLU tasks. In the recent remark by Meta Research[3]:

> "*We believe that self-supervised learning is one of the most promising ways to build such background knowledge and approximate a form of common sense in AI systems.*"

A huge amount of research efforts have demonstrated that the current language models with standard random masking pre-training strategy has a strong capability in understanding the natural language as well as the linguistic structures behind it. However, the secret behind remains a mystery to us as we cannot really explicitly extract or model the meaning or the linguistic structures through self-supervised learning. Such phenomena inspire us to explore the direction that extracting structures from the self-supervised perspective. For example, Hu et al. (2021) directly perform grammar induction after pre-training with a parsing specific objective function. The essential idea is to design an elegant pre-training objective function to incorporate the structured prior knowledge of the downstream task, such as parsing and named entity recognition.

References

Jacob Andreas, Marcus Rohrbach, Trevor Darrell, and Dan Klein. Neural module networks. In *Proceedings of CVPR*, 2016.

James K Baker. Trainable grammars for speech recognition. *The Journal of the Acoustical Society of America*, 65(S1):S132–S132, 1979.

[3]https://ai.facebook.com/blog/self-supervised-learning-the-dark-matter-of-intelligence/

Yoshua Bengio, Yann Lecun, and Geoffrey Hinton. Deep learning for ai. *Communications of the ACM*, 64(7):58–65, 2021.

Guy E Blelloch. Prefix sums and their applications. 1990.

Nitish Gupta, Kevin Lin, Dan Roth, Sameer Singh, and Matt Gardner. Neural module networks for reasoning over text. In *Proceedings of ICLR*, 2020.

Xiang Hu, Haitao Mi, Zujie Wen, Yafang Wang, Yi Su, Jing Zheng, and Gerard de Melo. R2d2: Recursive transformer based on differentiable tree for interpretable hierarchical language modeling. In *Proceedings of ACL*, 2021.

Harsh Jhamtani and Peter Clark. Learning to explain: Datasets and models for identifying valid reasoning chains in multihop question-answering. In *Proceedings of EMNLP*, 2020.

Zhanming Jie and Wei Lu. Multilingual semantic parsing : Parsing multiple languages into semantic representations. In *Proceedings of COLING*, 2014.

Zhanming Jie and Wei Lu. Dependency-based hybrid trees for semantic parsing. In *Proceedings of EMNLP*, 2018.

Zhanming Jie and Wei Lu. Dependency-guided lstm-crf for named entity recognition. In *Proceedings of EMNLP*, 2019.

Zhanming Jie, Aldrian Obaja Muis, and Wei Lu. Efficient dependency-guided named entity recognition. In *Proceedings of AAAi*, 2017.

Zhanming Jie, Pengjun Xie, Wei Lu, Ruixue Ding, and Linlin Li. Better modeling of incomplete annotations for named entity recognition. In *Proceedings of NAACL*, 2019.

Zhanming Jie, Jierui Li, and Wei Lu. Learning to reason deductively: Math word problem solving as complex relation extraction. In *Proceedings of ACL*, 2022.

Daniel Kahneman. *Thinking, fast and slow*. Macmillan, 2011.

John Lafferty, Andrew McCallum, and Fernando CN Pereira. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *Proceedings of ICML*, pp. 282–289, 2001.

Wei Lu. A unified framework for structured prediction: From theory to practice. In *Proceedings of EMNLP: Tutorial Abstracts*, 2017.

Alexander M Rush. Torch-struct: Deep structured prediction library. In *Proceedings of ACL: System Demonstrations*, pp. 335–342, 2020.

Swarnadeep Saha, Sayan Ghosh, Shashank Srivastava, and Mohit Bansal. Prover: Proof generation for interpretable reasoning over rules. In *Proceedings of EMNLP*, 2020.

Sunita Sarawagi and William W Cohen. Semi-markov conditional random fields for information extraction. In *Proceedings of NeurIPS*, pp. 1185–1192, 2004.

Erik F. Tjong Kim Sang and Fien De Meulder. Introduction to the CoNLL-2003 shared task: Language-independent named entity recognition. In *Proceedings of NAACL*, 2003.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Proceedings of NeurIPS*, 2017.

Lu Xu, Zhanming Jie, Wei Lu, and Lidong Bing. Better feature integration for named entity recognition. In *Proceedings of NAACL*, pp. 3457–3469, 2021.

Zhilin Yang, Peng Qi, Saizheng Zhang, Yoshua Bengio, William Cohen, Ruslan Salakhutdinov, and Christopher D Manning. Hotpotqa: A dataset for diverse, explainable multi-hop question answering. In *Proceedings of EMNLP*, 2018.

John M Zelle and Raymond J Mooney. Learning to parse database queries using inductive logic programming. In *Proceedings of the national conference on artificial intelligence*, 1996.