Ph.D. subject: Multitask Deep Learning for Joint Syntactic and Semantic Structure Prediction

Thierry Charnois, Joseph Le Roux, Nadi Tomeh

Deadline for application: August 25 2017

In this study we would like to explore multitask deep learning and structured prediction for joint syntactic and semantic structure analysis.

1 Scientific Context

In recent work on semantic parsing, [Peng et al., 2017] showed how to use deep learning and structured prediction to parse sentences into three semantic dependency graph formalisms with a single model trained from three different corpora using multitask learning. This model outperforms previous approaches which relied on separate syntactic analysis and hand-engineered features to extract information about local contexts to perform semantic tasks.

We would like to explore ways in which this approach could be extended to other NLP semantic tasks such as named-entity recognition (shallow structure) and discourse relation extraction (deep structure).

For multitask modeling, while [Peng et al., 2017] used parameters sharing across formalism and higher-order structure to predict the dependency graphs jointly, we propose to extend structured prediction energy networks (SPENs) [Belanger and McCallum, 2015] to jointly model several tasks. A deep architecture is used in SPENs to define an energy function of candidate labels in the output structure, and then predictions are produced by using backpropagation to iteratively optimize the energy with respect to the labels. This deep architecture captures dependencies between labels that would lead to intractable graphical models, and performs structure learning by automatically learning discriminative features of the structured output.

Another line of research is the exploration of structured attention models [Kim et al., 2017] for semantic parsing whether dynamic programming is possible or not.

For learning we propose to explore new models, from unrolled optimization to adversarial learning methods [Goodfellow et al., 2014] which have proved successful in vision in recent years and have been rarely applied to language related tasks [Gulrajani et al., 2017].

2 Administrative Context

This proposal is intended for master students with the following profile:

- proficient in NLP and Computational Linguistics (CL) eager to discover deep learning, or
- proficient in machine learning with a **strong** motivation to learn NLP and CL.

The hired candidate will be a Ph.D. candidate at LIPN, Universit Paris 13, in team RCLN (NLP and Knowledge Representation), member of LABEX EFL. S/he will possibly interact with other teams (machine learning, formal languages, optimization, etc).

This work will be supervised by:

Thierry Charnois Full Professor

Joseph Le Roux Assistant Professor

Nadi Tomeh Assistant Professor

To apply, and for any additional information, please contact leroux@lipn.fr, tomeh@lipn.fr, charnois@lipn.fr,

Applications must contain: a CV, a copy of Master's grades (M1 and M2 in France),a cover letter and optionally reference letters (recommended).

References

David Belanger and Andrew McCallum. Structured prediction energy networks. *CoRR*, abs/1511.06350, 2015.

Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron C. Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in Neural Information Processing Systems 27: Annual Conference on Neural Information Processing Systems 2014, December 8-13 2014, Montreal, Quebec, Canada*, pages 2672–2680, 2014.

I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville. Improved Training of Wasserstein GANs. *ArXiv e-prints*, March 2017.

Yoon Kim, Carl Denton, Luong Hoang, and Alexander M. Rush. Structured attention networks. *CoRR*, abs/1702.00887, 2017. URL http://arxiv.org/abs/1702.00887.

Hao Peng, Sam Thomson, and Noah A. Smith. Deep multitask learning for semantic dependency parsing. *CoRR*, abs/1704.06855, 2017.