Bethe Free Energy Minimization in MRFs for Combinatorial Optimization

Joseph Le Roux

November 28, 2024

1 Context

Machine Learning (ML) for Combinatorial Optimization (CO) is a fast growing field with many applications [BLP18]. In particular the probabilistic modelling of CO instances as Markov Random Fields (MRFs) has been used in Computer Vision and Natural Language Processing for tasks that require highly structured and constrained output [Kom11]. They are especially attractive since they not only encode the problem of finding optimal solutions as *Maximum A Posteriori* inference (MAP), but also allows the generation of near optimal solutions as sampling.

However recent ML applications to CO tend to drop the structural constraints and favor the perceptive power of highly expressive feature extractors. For instance, in approaches such as [Nai+21] the MRF is simply reduced to its factor graph encoded with a graph neural network (GNN). Variable node vectors are then used to parametrize classifiers which return the expected values of the corresponding variables. The interactions between variables represented as edges in the MRF are not directly accessible to the classifiers, only through GNN features.

For this internship we propose to extend this approach and explicitly incorporate the MRF structure, *i.e.* variable correlations, during the prediction. MAP inference is intractable in this setting and the internship work focuses on efficient approximate methods designed for neural networks.

2 Internship Description

We already have developed an approach that computes MRF probabilities on top of a GNN feature extractor. While promising, this method relies on a version of loopy belief propagation (LBP), a well known algorithm from the MRF literature. LBP suffers from several drawbacks: *(i)* it is slow to converge, and in some cases does not converge at all, *(ii)* it is intrinsically an iterative algorithm which makes it inefficient on current parallel architectures such as GPUs.

However LBP, when it converges, returns a local optimum of the Bethe Free Energy from which probabilities can be easily computed. Recently, Wiseman and Yoon [WK19] and Kuck and al. [Kuc+20] proposed methods to tackle these issues, either by computing directly the Bethe Energy of a MRF in the first case or to unroll LBP in the second.

The internship consists in incorporating these methods in our model for CO and study their usefulness in this context. We plan two phases:

1. Adaptation of the two algorithms to our frameworks, for training and inference;

2. Evaluation and comparison of the two algorithms and LBF on several CO problems, *i.e.* the Maximal Independence Set for which we have a large dataset of instances.

3 Application

We are looking for a candidate with either CO background (master level) with very good knowledge of Machine Learning methods for CO, or with a strong ML background willing to adapt recent models to CO tasks. We expect proficiency with python and deep learning libraries such as pytorch.

The internship work will be carried out on site (no remote) at LIPN, Université Sorbonne Paris Nord, with possibilities of extension to a three-year Ph.D. funding (2025-2028). This internship is funded by the ANR Project SEMIAMOR (2024-2028).

For additional information, please contact leroux@lipn.fr. If you are interested please attach to your application email a CV, a cover letter and a transcript of your Master level marks.

References

[BLP18] Yoshua Bengio, Andrea Lodi, and Antoine Prouvost. "Machine Learning for Combinatorial Optimization: a Methodological Tour d'Horizon". In: *CoRR* abs/1811.06128 (2018). arXiv: `1811.06128`. URL: `http://arxiv.org/abs/1811.06128`.

[Kom11] Nikos Komodakis. "Efficient training for pairwise or higher order CRFs via dual decomposition". In: *CVPR 2011*. 2011, pp. 1841–1848. DOI: `10.1109/CVPR.2011.5995375`.

[Kuc+20] Jonathan Kuck, Shuvam Chakraborty, Hao Tang, Rachel Luo, Jiaming Song, Ashish Sabharwal, and Stefano Ermon. "Belief Propagation Neural Networks". In: *Advances in Neural Information Processing Systems*. Ed. by H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin. Vol. 33. Curran Associates, Inc., 2020, pp. 667–678. URL: `https://proceedings.neurips.cc/paper_files/paper/2020/file/07217414eb3fbe24d4e5b6cafb91ca18-Paper.pdf`.

[Nai+21] Vinod Nair, Sergey Bartunov, Felix Gimeno, Ingrid von Glehn, Pawel Lichocki, Ivan Lobov, Brendan O'Donoghue, Nicolas Sonnerat, Christian Tjandraatmadja, Pengming Wang, Ravichandra Addanki, Tharindi Hapuarachchi, Thomas Keck, James Keeling, Pushmeet Kohli, Ira Ktena, Yujia Li, Oriol Vinyals, and Yori Zwols. "Solving Mixed Integer Programs Using Neural Networks". In: *arXiv:2012.13349 [cs, math]* (July 2021). arXiv: 2012.13349. URL: `http://arxiv.org/abs/2012.13349`.

[WK19] Sam Wiseman and Yoon Kim. "Amortized Bethe Free Energy Minimization for Learning MRFs". In: *Advances in Neural Information Processing Systems*. Ed. by H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett. Vol. 32. Curran Associates, Inc., 2019. URL: `https://proceedings.neurips.cc/paper_files/paper/2019/file/dc554706afe4c72a60a25314cbaece80-Paper.pdf`.