Neuro-Symbolic AI Systems Combining Neural Networks and Marginals-Augmented Constraint Programming

Gilles Pesant
Polytechnique Montréal, Montreal, Canada
gilles.pesant@polymtl.ca

Constraint Programming (CP) models a combinatorial problem using constraints that each represent a large portion of it and apply sophisticated dedicated search-space reduction algorithms on each, whose results are combined through *constraint propagation* by dynamically reducing the domain of each variable. These algorithms essentially ask whether a given variable-value pair $\langle x_i, d\rangle$ can appear (meaning variable x_i being assigned value d) in solutions to a given constraint. If it cannot, that variable-value assignment can certainly be removed from consideration since it must satisfy every constraint. Such *filtering algorithms*, applied throughout the solving process, can greatly reduce the search space to be explored.

One can view the set of solutions to a constraint or indeed to the whole problem as a multivariate discrete distribution. In that light, the filtering algorithms mentioned above seek to identify the *support* of that distribution with respect to each variable, i.e. the set of values that have non-zero frequency in the corresponding *marginal distribution*. Conventional constraint propagation essentially exchanges the supports of marginal distributions — I recently proposed to exchange whole marginal distributions (Pesant 2019). This required changes to the core of CP software systems: maintaining not only a domain of possible values for each variable but also a marginal distribution over that domain; a new operation on constraints, computing marginal distributions through weighted counting; replacing the propagation fixpoint algorithm by a sum-product message-passing algorithm, i.e. iterated *belief propagation* (BP). I implemented the resulting *CP-BP framework* in the publicly-available research prototype MiniCPBP [`github.com/PesantGilles/MiniCPBP`].

I recently led a two-year project funded by an IVADO [`ivado.ca/en`] fundamental research grant investigating the combination of CP and Machine Learning for tasks involving hard structure. We considered expressing such structure as constraints in a CP model and using the marginal distributions computed in the CP-BP framework either during the training or inference phases.

In the context of a constrained sequence generation task: For training, reinforcement learning was successfully used to combine in its reward signal the loss function from a pretrained RNN with the marginals and constraint violations from CP, in order to adjust the weights of the RNN (Lafleur, Chandar, and Pesant 2022); for inference when we wish to enforce long-term structure, the learned probabilities of the sequence model were mixed with the marginal probabilities of the CP model describing that structure and the next predicted token was then sampled from the resulting probability distribution (Manibod and Pesant 2022).

In the context of a structured prediction task: We compared against recent efforts from the CP research community to improve training by incorporating into the loss function a term related to the support of each variable (Silvestri, Lombardi, and Milano 2021), and to perform inference by solving an optimization problem over a CSP according to class probabilities (Mulamba et al. 2020). We showed that our ability to work with marginal distributions both greatly increases post-training accuracy by replacing coarse supports with finer-grained marginals and accelerates prediction inference by more simply solving a satisfaction problem fed with the class probabilities as marginals (Sargordi et al. 2022).

I would give a 30- to 60-minute talk on the CP-BP framework and its application to neuro-symbolic AI systems.

References

[Lafleur, Chandar, and Pesant 2022] Lafleur, D.; Chandar, S.; and Pesant, G. 2022. Combining reinforcement learning and constraint programming for sequence-generation tasks with hard constraints. In Solnon, C., ed., *28th International Conference on Principles and Practice of Constraint Programming, CP 2022, July 31 to August 8, 2022, Haifa, Israel*, volume 235 of *LIPIcs*, 30:1–30:16. Schloss Dagstuhl - Leibniz-Zentrum für Informatik.

[Manibod and Pesant 2022] Manibod, V., and Pesant, G. 2022. Long-term constrained sequential inference using constraint programming. Under evaluation.

[Mulamba et al. 2020] Mulamba, M.; Mandi, J.; Canoy, R.; and Guns, T. 2020. Hybrid classification and reasoning for image-based constraint solving. In Hebrard, E., and Musliu, N., eds., *Integration of Constraint Programming, Artificial Intelligence, and Operations Research - 17th International Conference, CPAIOR 2020, Vienna, Austria, September 21-24, 2020, Proceedings*, volume 12296 of *Lecture Notes in Computer Science*, 364–380. Springer.

[Pesant 2019] Pesant, G. 2019. From support propagation to belief propagation in constraint programming. *J. Artif. Intell. Res.* 66:123–150.

[Sargordi et al. 2022] Sargordi, M.; Verhaeghe, H.; Chandar, S.; and Pesant, G. 2022. Combining Neural Networks and Marginal Probabilities Derived from Constraint Programming Models to Perform Structured Prediction Tasks. Working paper.

[Silvestri, Lombardi, and Milano 2021] Silvestri, M.; Lombardi, M.; and Milano, M. 2021. Injecting domain knowledge in neural networks: A controlled experiment on a constrained problem. In Stuckey, P. J., ed., *Integration of Constraint Programming, Artificial Intelligence, and Operations Research - 18th International Conference, CPAIOR 2021, Vienna, Austria, July 5-8, 2021, Proceedings*, volume 12735 of *Lecture Notes in Computer Science*, 266–282. Springer.