Hyperbolic deep reinforcement learning for repeated exact combinatorial optimization

Paul Strang[1,2] Zacharie Ales[2] Come Bissuel[1] Olivier Juan[1] Safia Kedad-Sidhoum[2]

[1]EDF R&D, France, name.surname@edf.fr

[2]Conservatoire national des arts et metiers, France, name.surname@cnam.fr

Keywords: Mixed-integer programming, Branch-and-bound, Reinforcement learning, Hyperbolic geometry

Mixed-integer linear programs, or MILPs, are combinatorial optimization problems defined such as $P : \{\min c^T x \mid Ax \leq b \; ; \; x \in \mathbb{N}^{|\mathcal{I}|} \times \mathbb{R}^{n-|\mathcal{I}|}\}$ with $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, n the number of variables, m the number of linear constraints and $\mathcal{I}$ the indices of integer variables. MILPs are thus linear programs (LPs) with additional integrity constraints on a subset of variables. These integrity constraints make the problem non-convex, in fact, it is $\mathcal{NP}$-hard in the general case. We are interested in repeated problems of fixed dimensions $\{P_i = (A_i, b_i, c_i)\}_{i\in N}$, which are understood as realizations of a random variable $\mathcal{P}$ following an unknown distribution $\mathcal{D} : \Omega \to \mathbb{R}^{m\times n} \times \mathbb{R}^m \times \mathbb{R}^n$.

Mixed-integer programming solvers developed over the last decades have relied on the Branch and Bound (B&B) algorithm [Wol21] to efficiently explore the space of solutions while guaranteeing the optimality of the returned solution. These solvers are based on complex heuristics fine-tuned by experts on large MIP datasets to obtain the best average performance. In the context of real-world applications, in which similar instances with slightly varying inputs are solved on a regular basis, there is a huge incentive to reduce the solving time by learning efficient tailor-made heuristics.

Directly extending the work of Marc Etheve [Eth21], this thesis proposes to discover new B&B heuristics outperforming existing solvers using reinforcement learning (RL) algorithms. In fact, many contributions have sought to reformulate the task of learning optimal branching policies as a tree Markovian decision process (MDP) [Eth+20; Sun+22; Sca+22; PLB22]. However, tree MDPs represent a challenging setup for RL algorithms: credit assignement problem, large state-action spaces yielding complex exploration and partial observability are as many predicament hindering RL agents' performance on high dimensional MIP instances. We argue that learning better state representations acknowledging the tree structure of successive tree MDP states is key to mitigate these challenges.

Contributions adapting models from Riemannian geometry and hyperbolic spaces to traditional deep learning architectures have florished over the last five years [Pen+21]. Owing to their high capacity to model data exhibiting hierarchical (tree-like) structures, hyperbolic neural networks have been applied to improve performance in computer vision, natural language processing and other tasks involving graph embeddings. In turn, we propose to harness the inductive bias of hyperbolic neural networks [GBH18; SMH21] to learn better representations of tree Markovian decision processes.

References

[GBH18] Octavian-Eugen Ganea, Gary BÃ©cigneul, and Thomas Hofmann. "Hyperbolic Neural Networks". In: (2018).

[Eth+20] Marc Etheve et al. "Reinforcement Learning for Variable Selection in a Branch and Bound Algorithm". In: *Integration of Constraint Programming, Artificial Intelligence, and Operations Research*. Ed. by Emmanuel Hebrard and Nysret Musliu. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2020, pp. 176–185. ISBN: 978-3-030-58942-4. DOI: `10.1007/978-3-030-58942-4_12`.

[Eth21] Marc Etheve. "Solving repeated optimization problems by Machine Learning". PhD thesis. HESAM UniversitÃ©, Dec. 3, 2021. URL: `https://theses.hal.science/tel-03675471` (visited on 12/08/2022).

[Pen+21] Wei Peng et al. *Hyperbolic Deep Neural Networks: A Survey*. Feb. 17, 2021. DOI: `10.48550/arXiv.2101.04562`. arXiv: `2101.04562[cs]`. URL: `http://arxiv.org/abs/2101.04562` (visited on 05/30/2023).

[SMH21] Ryohei Shimizu, Yusuke Mukuta, and Tatsuya Harada. *Hyperbolic Neural Networks++*. Mar. 17, 2021. DOI: `10.48550/arXiv.2006.08210`. arXiv: `2006.08210[cs,stat]`. URL: `http://arxiv.org/abs/2006.08210` (visited on 07/25/2023).

[Wol21] Laurence A. Wolsey. *Integer programming*. Second edition. Hoboken, NJ: Wiley, 2021. ISBN: 978-1-119-60653-6.

[PLB22] Christopher W. F. Parsonson, Alexandre Laterre, and Thomas D. Barrett. *Reinforcement Learning for Branch-and-Bound Optimisation using Retrospective Trajectories*. Dec. 5, 2022. DOI: `10.48550/arXiv.2205.14345`. arXiv: `2205.14345[cs]`. URL: `http://arxiv.org/abs/2205.14345` (visited on 02/25/2023).

[Sca+22] Lara Scavuzzo et al. *Learning to branch with Tree MDPs*. Oct. 13, 2022. DOI: `10.48550/arXiv.2205.11107`. arXiv: `2205.11107[cs,math]`. URL: `http://arxiv.org/abs/2205.11107` (visited on 12/21/2022).

[Sun+22] Haoran Sun et al. "Improving Learning to Branch via Reinforcement Learning". In: Learning Meets Combinatorial Algorithms at NeurIPS2020. July 6, 2022. URL: `https://openreview.net/forum?id=z4D7-PTxTb` (visited on 12/14/2022).