Reinforcement Learning for Pricing Problem Optimization in Column Generation.

Abdo Abouelrous[1][0000−0003−3670−5136], Laurens Bliek[1][0000−0002−3853−4708], Adriana F. Gabor[2][0000−0003−4327−8048], Yaoxin Wu[1][0000−0002−3625−6599], and Yingqian Zhang[1][0000−0002−5073−0787]

[1] Eindhoven University of Technology, 5612 AZ Eindhoven, The Netherlands.
[2] Department of Mathematics, College of Computing and Mathematical Sciences, Khalifa University, United Arab Emirates.

Many decision-making problems such as planning, vehicle routing, scheduling problems can be solved by employing Combinatorial Optimization (CO) techniques. Often, these problems can be modelled by large integer programs and solved by Column Generation (CG) based algorithms, such as branch and price algorithms, that solve the LP relaxation at each node in a branch- and- bound procedure by a CG approach [3].

The idea of the CG approach is to start with an LP problem with a small number of variables, called the Restricted Master Problem (RMP) and add columns with negative reduced costs successively. To find the columns with negative reduced costs, a Pricing problem (PP) needs to be solved repeatedly until no columns with negative reduced costs can be added. The art of designing efficient CG based algorithms is in reformulating the problem such that the Pricing problem can be solved efficiently [7]. However, this remains a complex task, as for many practical problems such as vehicle routing, the PP is NP-hard.

In general, we can distinguish two paradigms to solve the Pricing Problem: Operations Research (OR) methods and Machine Learning (ML) methods. OR methods have been used most frequently. For some problems, such as variants of vehicle routing, the OR algorithms to solve the PP are computationally demanding in terms of both speed and memory. In contrast, ML techniques can be seen as an alternative to making decisions in a principled way, with methods such as deep learning being well-suited to problems characterized by a high-dimensional space [1]. These methods make use of data from the operational environment by extracting information from the solution structure which heuristisc do not exploit [6]. This helps to define a more guided search procedure.

Another advantage of ML algorithms is that they can learn from the collective expert knowledge and build on it to extend to other problems [1]. For instance, different problems can be repeatedly solved by learning common solution structures [4]. Although the training of ML methods – which normally happens offline – is not trivial, the methods can be easily applied to solve the problem on different instances once training concludes [8].

For CG, however, ML methods have been previously used mainly for pre- and post - processing the PP and not to solve the PP itself. Although ML methods can speed up a single CG iteration, they often terminate prematurely when they fail to find columns with negative reduced costs.

In this paper, we propose the first ML-based framework that directly solves the PP in CG using Reinforcement Learning (RL) based on the attention-mechanism architecture. More precisely, we leverage an end-to-end model, called POMO-CG, that optimizes the PP. Our model iteratively constructs columns by adding one component after another until a feasible column is generated. The objective is to generate columns with minimum reduced cost at every CG iteration - given a minimization objective. This is in contrast to other work like [2] where the ML agent does not generate the columns but deals more with the selection of pre-generated columns.

Although finding the columns with most negative reduced cost is conventionally not seen as necessary, it is important since the reduced cost of a variable represents the partial derivative of the objective function with respect to it. Thus, lower reduced costs translate to a larger potential to reduce the objective function and faster convergence [5]. This pursuit has been largely ignored due to the associated computational burden. However, as we are able to speed up the search for the minimum negative reduced cost through ML, we can carry this pursuit further.

To showcase the added value of our model, we propose a series of Vehicle Routing Problems with Capacity constraints and Time Windows for which we solve the linear relaxation. We test our method on instances of different sizes and compare the convergence speed and objective gap with another Dynamic Programming-based heuristic that solves the PP. Our results show that our method converges much faster by a factor of up to 30x and maintains a relatively small objective gap that is no more than 10%. Furthermore, our model can generalize to much larger instances than the ones it was trained on, reinforcing its added value in solving computationally demanding problems. Lastly, we provide a discussion on the advantages of our method with respect to other established CG techniques from the literature in terms of practicalities.

References

1. Bengio, Y., Lodi, A., Prouvost, A.: Machine learning for combinatorial optimization: a methodological tour d'horizon. European Journal of Operational Research **290**(2), 405–421 (2021)
2. Chi, C., Aboussalah, A., Khalil, E., Wang, J., Sherkat-Masoumi, Z.: A deep reinforcement learning framework for column generation. Advances in Neural Information Processing Systems **35**, 9633–9644 (2022)
3. Desaulniers, G., Desrosiers, J., Solomon, M.M.: Column generation, vol. 5. Springer Science & Business Media (2006)
4. Khalil, E., Dai, H., Zhang, Y., Dilkina, B., Song, L.: Learning combinatorial optimization algorithms over graphs. Advances in neural information processing systems **30** (2017)
5. Lübbecke, M.E., Desrosiers, J.: Selected topics in column generation. Operations research **53**(6), 1007–1023 (2005)
6. d O Costa, P.R., Rhuggenaath, J., Zhang, Y., Akcay, A.: Learning 2-opt heuristics for the traveling salesman problem via deep reinforcement learning. In: Asian conference on machine learning. pp. 465–480. PMLR (2020)

7. Václavík, R., Novák, A., Šcha, P., Hanzálek, Z.: Accelerating the branch-and-price algorithm using machine learning. European Journal of Operational Research **271**(3), 1055–1069 (2018)
8. Zhang, Z., Liu, H., Zhou, M., Wang, J.: Solving dynamic traveling salesman problems with deep reinforcement learning. IEEE Transactions on Neural Networks and Learning Systems **34**(4), 2119–2132 (2021)