Research Statement

CAO Zhiguang
School of Computing and Information Systems, Singapore Management University
Tel: (65) 6826 1341; Email: zgcao@smu.edu.sg
25-12-2023

Background

In recent years, there is a growing trend towards applying ***neural heuristics*** based on **deep (reinforcement) learning** against the traditional heuristics to solve **combinatorial optimization problems** (COPs), inspired by the remarkable success of deep neural networks in other domains. In general, all of them falls into the scope of **Learning to Optimize** (*L2Opt*). The rationale behind it comes from three aspects: (1) many COPs could be interpretated as optimizing a sequence (of nodes or elements), which is close to the NLP (Natural Language Processing) task in AI; (2) a class of COP instances may share similar structures and differ only in data which follows a distribution, such as the vehicle routing problems (VRPs) in logistics; and (3) neural heuristics based on deep models can discover the underlying patterns of a given COP class, which could be used to generate alternative algorithms that are better than hand-crafted ones in traditional heuristics.

Similar to the traditional heuristics, the neural heuristics are also mainly categorized into two types: neural ***construction*** heuristics and neural ***improvement*** heuristics. Taking solving VRP as an example, the former usually start with an empty solution, and sequentially add a node (customer) to the current (partial) solution until a complete route is formed. The latter usually start with an initial but complete solution, and iteratively select nodes or operators to perform certain local operations to generate a new solution with potentially improved quality.

Research Areas

Focusing on **Learning to Optimize** (*L2Opt*), we have developed a number of advanced neural heuristics based on deep models for solving various COPs such as vehicle routing problem (VRP), job shop scheduling problem (JSSP), bin packing problem (BPP), integer programming (IP), and constraint satisfaction problem (CSP).

A. **Neural Solvers for Vehicle Routing Problem** (TSP, CVRP, PDP)

It aims to find a shortest path for a (fleet of) vehicle(s) that departs from the depot to serve customers at different locations with different demands while complying certain constraints such as capacity limit, and the vehicle finally returns to the depot.

A.1 ***Multi-Decoder Attention Model*** (*Construction*): TSP, CVRP

Existing neural construction heuristics suffer from a major limitation, i.e., the generated solutions are not diverse enough. Intuitively, a more diverse set of solutions could potentially lead to better ones. This is because for VRP and many other COPs, multiple optimal solutions exist and trying to find different ones will increase the chance of finding at least one. To address this issue, we propose the Multi-Decoder Attention Model (MDAM) to train multiple construction policies. It employs a Transformer to encode the node information, and multiple identical attention decoders with unshared parameters to sample different trajectories. During training, each of the decoders learns distinct solution patterns, and is regularized by a Kullback-Leibler divergence loss to force the decoders to output dissimilar probability distribution of selecting nodes.

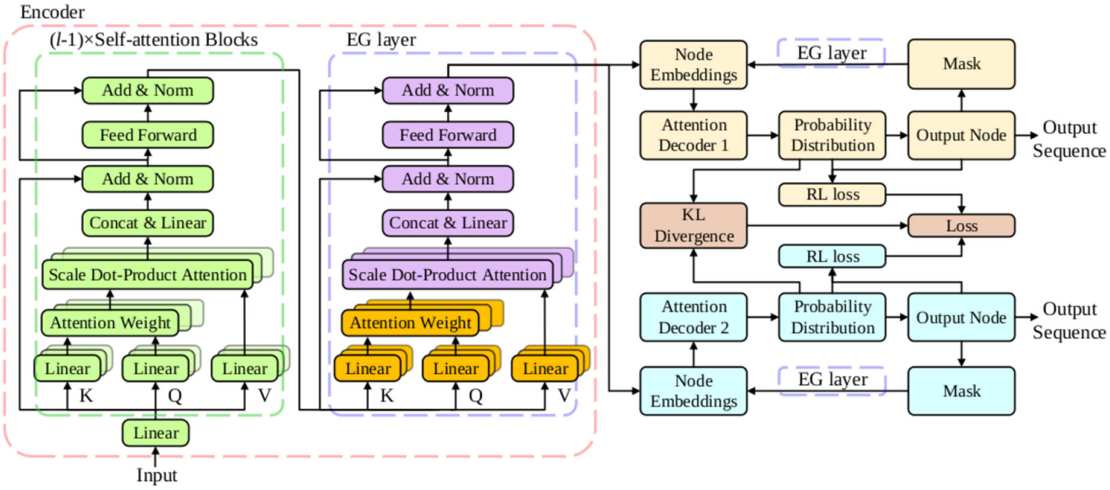


Figure 1. The overall architecture of our MDAM

A.2 ***Dual-Aspect Collaborative Transformer*** (*Improvement*): TSP, CVRP

The classic transformer is less effective in learning *improvement* models for VRP because its positional encoding (PE) method is not suitable in representing VRP solutions. To address this issue, we propose a novel Dual-Aspect Collaborative Transformer (DACT) to learn the embedding for the node and positional features separately, instead of fusing them together as done in existing ones, so as to avoid potential noises and incompatible correlations. Moreover, in our DACT, the positional features are embedded through a novel cyclic positional encoding (CPE) method to allow Transformer to effectively capture the circularity and symmetry of VRP solutions (i.e., cyclic sequences).

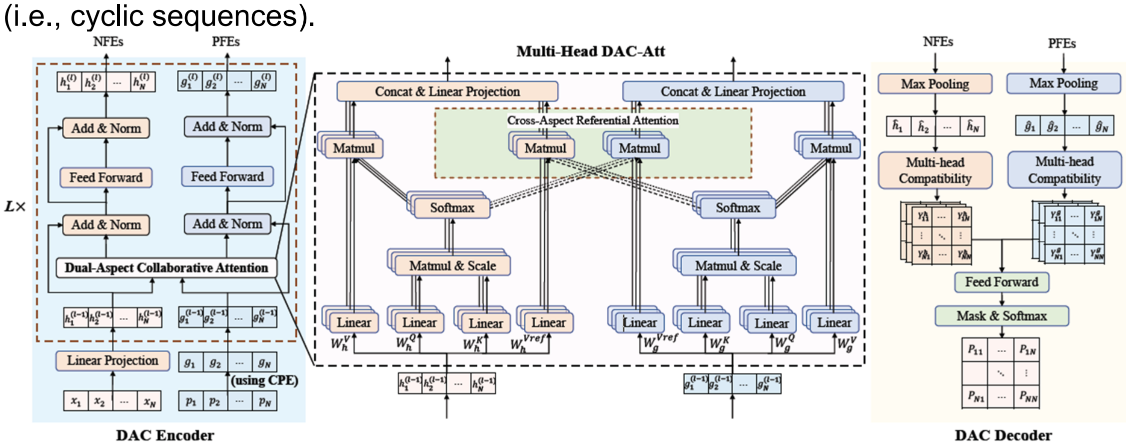


Figure 2. The overall architecture of our DACT

A.3 ***Efficient Neural Neighborhood Search*** (*Improvement*): PDP

The prevailing neural methods mainly focus on travelling salesman problem (TSP) or capacitated vehicle routing problem (CVRP), where efficient solvers for pickup and delivery problems (PDPs) are rarely studied. To bridge this gap, we present an efficient Neural Neighborhood Search (N2S) approach for PDPs. In specific, we design a powerful Synthesis Attention that allows the vanilla self-attention to synthesize various types of features regarding a route solution. We also exploit two customized decoders that automatically learn to perform removal and reinsertion of a pickup-delivery node pair to tackle the precedence constraint.

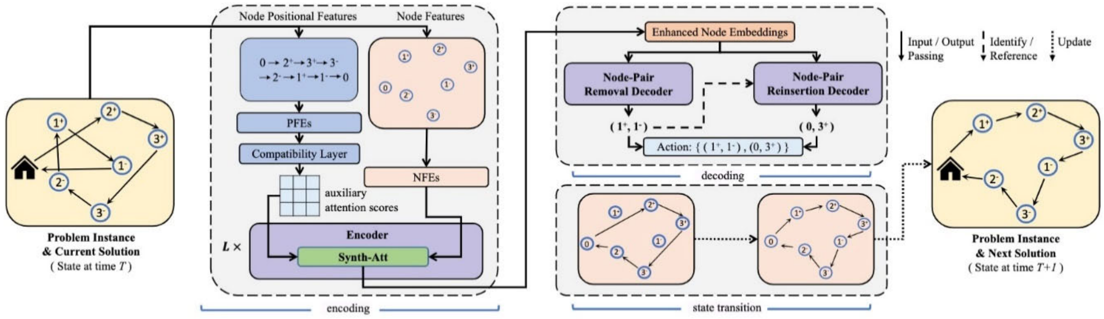
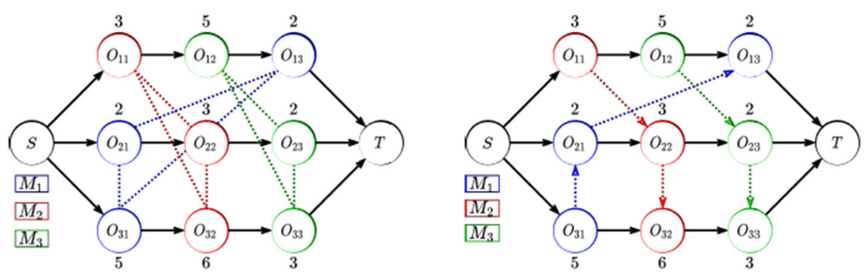


Figure 3. The overall architecture of our N2S

B. **Neural Solvers for Job Shop Scheduling Problem** (JSSP)

Learning to Dispatch (L2D, Construction). Priority dispatching rule (PDR) is widely used for solving real-world job-shop scheduling problem (JSSP). However, the design of effective PDRs is a tedious task, requiring a myriad of specialized knowledge and often delivering limited performance. To address this issue, we propose to automatically learn PDRs via an end-to-end deep reinforcement learning agent. We first cast the decision making (i.e., which job should go to which machine at each step) as determining the arc direction in a disjunctive graph, and then propose a Graph Neural Network based scheme to embed the states encountered during solving. The resulting policy network is size-agnostic, effectively enabling generalization on large-scale instances.



(a) JSSP instance (b) Complete solution

Figure 4. The representation as a disjunctive graph allows Graph Neural Network based embedding

C. **Neural Solvers for Bin Packing Problem** (BPP)

DRL with Multimodal Encoder (Construction). Due to the relatively less informative yet computationally heavy encoder, and considerably large action space inherent to the 3-D BPP, existing DRL methods are only able to handle up to 50 boxes.

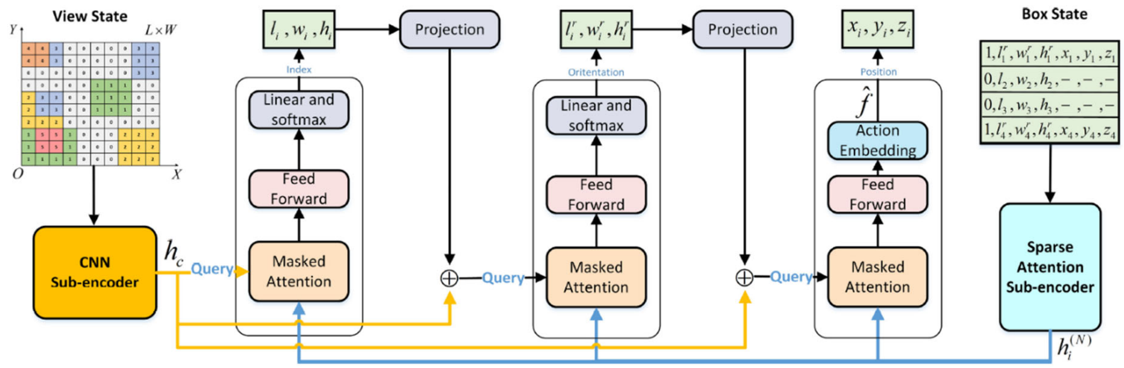


Figure 5. The overall architecture of our DRL with multimodal encoder

To address this issue, we propose a DRL agent, which sequentially addresses three subtasks of sequence, orientation, and position, respectively. Specifically, we exploit a multimodal encoder, where a sparse attention subencoder embeds the box state to mitigate the computation while learning the packing policy, and a convolutional neural network subencoder embeds the view state to produce auxiliary spatial representation. We also leverage an action representation learning in the decoder to cope with the large action space of the position subtask.

D. **Neural Solvers for Integer Programming** (IP)

DRL guided Large Neighborhood Search (Improvement). To tackle the issue that *how to improve a solver from externals such that it can find high-quality solutions more quickly*? we propose a high-level, learning based LNS method to solve general IP problems. Based on deep reinforcement learning (RL), we train a policy network as the destroy operator in LNS, which decides a subset of variables in the current solution for reoptimization. Then we use a solver as the repair operator, which solves sub-IPs to reoptimize the destroyed variables. Despite being heuristic, our method can effectively handle the large-scale IP by solving a series of smaller sub-IPs.

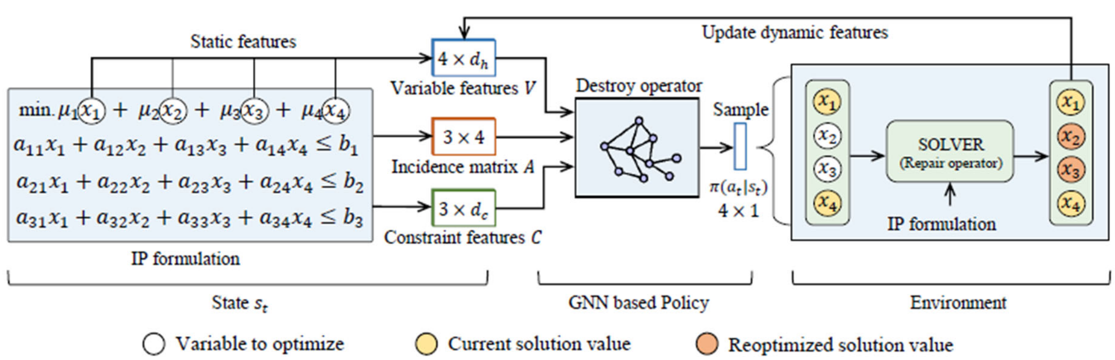


Figure 6. The overall architecture of our DRL guided large neighborhood search

E. **Neural Solvers for Constraint Satisfaction Problems** (CSP)

DRL guided Backtracking Search. Backtracking search algorithms are often used to solve the Constraint Satisfaction Problem (CSP). Its efficiency depends greatly on the variable ordering heuristics. Currently, the most commonly used heuristics are hand-crafted based on expert knowledge. We propose a deep reinforcement learning based approach to automatically discover new variable ordering heuristics that are better adapted for a given class of CSP instances, without the need of relying on hand-crafted features and heuristics. To capture the complex relations among the variables and constraints, we design a representation scheme based on Graph Neural Network that can process CSP instances with different sizes and constraint arities.

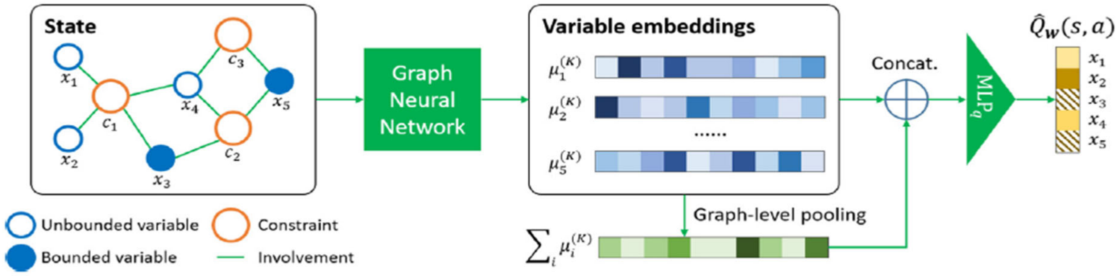


Figure 7. The overall architecture of our DRL guided backtracking search

Ongoing and Future Research Topics

Different from the computer vision and natural language processing in AI, it is still challenging for deep learning to solve combinatorial optimization problems like the vehicle routing problem (VRP), since they are NP-hard and computationally expensive. On the other hand, the neural heuristics based on deep models just emerged recently, while VRPs (and other COPs) have been studied by traditional methods (either heuristics or exact ones) for decades, and it would be hard for neural heuristics to outperform them with **arbitrary settings**.

In fact, most of existing neural heuristics always train and test the deep models using randomly generated VPR instances, which satisfy that, 1) the node locations are following a specific distribution, i.e., uniform, 2) the amount of nodes in an instance is fixed and small, i.e., 20, 50, and 100; 3) the constraints are relatively simple such as the feasibility of a node in TSP or the vehicle capacity limit in CVRP; and 4) the distance between nodes is usually measured in Euclidean space. Under those settings, the neural heuristics outstrip many traditional heuristics. However, the VRP instances in reality may follow various distributions and sizes, have more complex constraints, and are measured in real road networks which significantly differ from the Euclidean one, where the neural heuristics perform inferiorly to the traditional methods, and thus severely hinder their applications to solve the real-world problems. Therefore, how to strengthen the **generalization of *cross-distribution*, *cross-size*, *cross-constraints, and cross-metrics***, is currently my research focus, which aims to build generalizable and robust neural solvers for VRPs, as well as other COPs.

Selected Publications and Outputs

[1] Jianan Zhou, Yaoxin Wu, Wen Song, *Zhiguang Cao*, and Jie Zhang. Towards Omni-generalizable Neural Methods for Vehicle Routing Problems, International Conference on Machine Learning (**ICML**), 2023.

[2] Yuan Jiang, *Zhiguang Cao*, Yaoxin Wu, Wen Song, and Jie Zhang. Ensemble-based Deep Reinforcement Learning for Vehicle Routing Problems, Advances in Neural Information Processing Systems (**NeurIPS**), 2023.

[3] Yining Ma, *Zhiguang Cao*, and Yew Meng Chee. Learning to Search Feasible and Infeasible Regions of Routing Problems, Advances in Neural Information Processing Systems (**NeurIPS**), 2023.

[4] Jieyi Bi, Yining Ma, Jiahai Wang, *Zhiguang Cao*, Jinbiao Chen, Yuan Sun, and Yeow Meng Chee. Learning Generalizable Models for Vehicle Routing Problems via Knowledge Distillation, Advances in Neural Information Processing Systems (**NeurIPS**), 2022.

[5] Yining Ma, Jingwen Li, *Zhiguang Cao*, Wen Song, Le Zhang, Zhenghua Chen, and Jing Tang. Learning to Iteratively Solve Routing Problems with Dual-Aspect Collaborative Transformer, Advances in Neural Information Processing Systems (**NeurIPS**), 2021.

[6] Liang Xin, Wen Song, *Zhiguang Cao* and Jie Zhang, NeuroLKH: Combining Deep Learning Model with Lin-Kernighan-Helsgaun Heuristic for Solving the Traveling Salesman Problem, Advances in Neural Information Processing Systems (**NeurIPS**), 2021.

[7] Jinbiao Chen, Zizhen Zhang, *Zhiguang Cao*, Yaoxin Wu, Yining Ma, Te Ye, and Jiahai Wang. Neural Multi-Objective Combinatorial Optimization with Diversity Enhancement, Advances in Neural Information Processing Systems (**NeurIPS**), 2023.

[8] Jinbiao Chen, Zizhen Zhang, Te Ye, *Zhiguang Cao*, Siyuan Chen, and Jiahai Wang. Efficient Meta Neural Heuristic for Multi-Objective Combinatorial Optimization, Advances in Neural Information Processing Systems (**NeurIPS**), 2023.

[9] Haoran Ye, Jiarui Wang, *Zhiguang Cao*, Helan Liang, and Yong Li. DeepACO: Neural-enhanced Ant Systems for Combinatorial Optimization, Advances in Neural Information Processing Systems (**NeurIPS**), 2023.

[10] Cong Zhang, Wen Song, *Zhiguang Cao*, Jie Zhang, Puay Siew Tan and Chi Xu, Learning to Dispatch for Job Shop Scheduling via Deep Reinforcement Learning, Advances in Neural Information Processing Systems (**NeurIPS**), 2020.

[11] Yaoxin Wu, Wen Song, *Zhiguang Cao*, and Jie Zhang. Learning Large Neighborhood Search Policy for Integer Programming, Advances in Neural Information Processing Systems (**NeurIPS**), 2021.

[12] Zeyuan Ma, Hongshu Guo, Jiacheng Chen, Zhenrui Li, Guojun Peng, Yue-Jiao Gong, Yining Ma, *Zhiguang Cao*. MetaBox: A Benchmark Platform for Meta-Black-Box Optimization with Reinforcement Learning, Advances in Neural Information Processing Systems (**NeurIPS**), 2023.

[13] Yaoxin Wu, Wen Song, *Zhiguang Cao*, Jie Zhang, Abhishek Gupta, and Mingyan Simon Lin. Graph Learning Assisted Multi-Objective Integer Programming, Advances in Neural Information Processing Systems (**NeurIPS**), 2022.

[14] Yaoxin Wu, Wen Song, *Zhiguang Cao*, and Jie Zhang. Learning Scenario Representation for Solving Two-stage Stochastic Integer Programs, International Conference on Learning Representations (**ICLR**), 2022.