Research Statement

CAO Zhiguang
School of Computing and Information Systems, Singapore Management University
Tel: (65) 68261341; Email: zgcao@smu.edu.sg
10-12-2024

Background

In recent years, there has been a growing interest in applying neural heuristics based on deep (reinforcement) learning to solve combinatorial optimization problems (COPs), a field commonly known as **Learning to Optimize** (**L2Opt**). This trend is inspired by the success of deep neural networks in other domains. The rationale behind L2Opt is threefold: (1) many COPs can be interpreted as the optimization of a sequence (of nodes or elements), which bears similarity to tasks in natural language processing (NLP); (2) certain classes of COP instances share structural similarities, differing only in their data, such as vehicle routing problems (VRPs) in logistics; and (3) neural heuristics based on deep learning models can discover underlying patterns in COP classes, often generating algorithms that outperform traditional hand-crafted heuristics.

Like traditional heuristics, neural heuristics are typically categorized into two types: neural *construction* heuristics and neural *improvement* heuristics. For example, in solving VRPs, neural construction heuristics sequentially build a solution by adding nodes (customers) to the route until a complete solution is formed. In contrast, neural improvement heuristics begin with an initial complete solution and iteratively refine it by selecting nodes or operators to improve the solution's quality.

Phase I Research: *Specialized Neural Solvers*

Focusing on **L2Opt**, we have developed a range of specialized neural heuristic solvers, utilizing deep learning models, to address various COPs, including the vehicle routing problem (VRP), job shop scheduling problem (JSSP), bin packing problem (BPP), integer programming (IP), constraint satisfaction problem (CSP), and multi-objective COP (MOCOP). Notably, many of these contributions are considered pioneering within their respective domains.

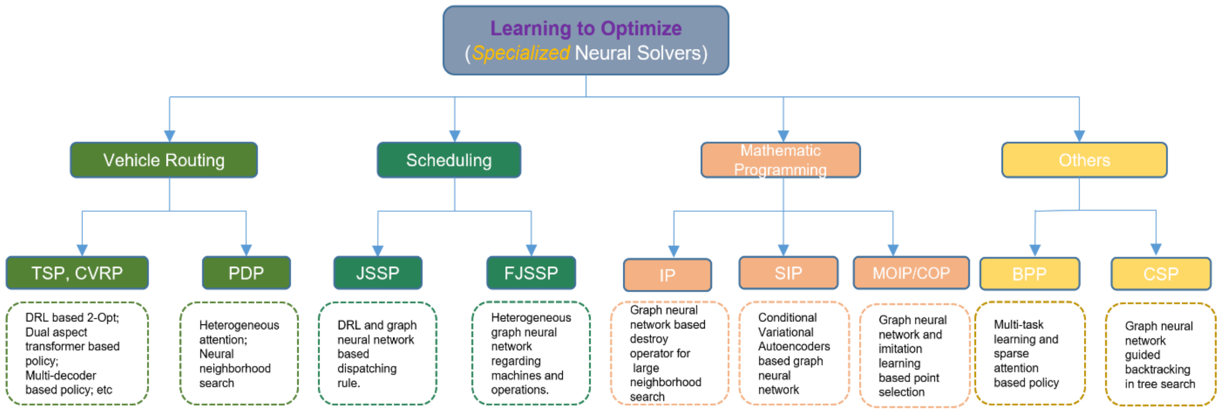

Figure 1. Phase I research focus – **Specialized** neural solvers for different COPs.

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

The goal of VRPs is to determine the shortest path for a fleet of vehicles, which depart from a depot to serve customers at different locations with varying demands, while adhering to constraints such as vehicle capacity. After serving all customers, the vehicles return to the depot. The classic VRP variants include the travelling salesman problem (TSP), capacitated vehicle routing problem (CVRP), and pickup and delivery problem (PDP).

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

A key limitation of existing neural construction heuristics is the lack of diversity in the generated solutions. A more diverse set of solutions increases the likelihood of finding optimal ones, especially since VRPs and other COPs often have multiple optimal solutions. To address this, we propose the Multi-Decoder Attention Model (MDAM), which trains multiple construction policies simultaneously. Using a Transformer to encode node information, MDAM employs multiple identical attention decoders with unshared parameters to sample different solution trajectories. During training, each decoder learns distinct solution patterns and is regularized using a Kullback-Leibler divergence loss, ensuring the decoders output diverse probability distributions for node selection in the solution construction process.

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

Standard transformers are less effective for learning improvement models for VRPs due to limitations in positional encoding (PE), which struggles to represent VRP solutions accurately. To overcome this, we propose the Dual-Aspect Collaborative Transformer (DACT), which independently learns embeddings for node and positional features, avoiding noise and incompatible correlations that arise from their fusion. Additionally, DACT incorporates a novel cyclic positional encoding (CPE) method, allowing it to effectively capture the circularity and symmetry inherent in VRP solutions (i.e., cyclic sequences).

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

Most neural methods primarily focus on the TSP or CVRP, with limited research on efficient solvers for pickup and delivery problems (PDPs). To fill this gap, we introduce the Neural Neighborhood Search (N2S) approach, specifically designed for PDPs. N2S leverages a powerful Synthesis Attention mechanism, enabling the vanilla self-attention to integrate various route features. We also develop two customized decoders that autonomously learn how to remove and reinsert pickup-delivery node pairs, effectively addressing the precedence constraints of PDPs.

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

Learning to Dispatch (L2D, Construction). Priority dispatching rules (PDRs) are commonly used to solve real-world JSSPs. However, designing effective PDRs is a complex and time-consuming task that requires specialized knowledge and often yields suboptimal results. To address this, we propose an end-to-end deep reinforcement learning (DRL) agent that automatically learns PDRs. We model the decision-making process—determining which job should be assigned to which machine at each step—as selecting arc directions in a disjunctive graph. A graph neural network is then employed to embed the problem states during solving, resulting in a size-agnostic policy network capable of generalizing to large-scale instances.

Learning to Improve (L2I, Improvement). While the L2D approach can be effective, its performance may still fall short of optimality due to the limitations of its graph representation in modeling partial solutions during the construction process. To overcome this, we propose a novel DRL-guided improvement heuristic for JSSPs. This approach employs a graph neural network-based representation scheme with two modules, designed to capture the dynamic topology and node types in the problem's graph. Additionally, we introduce a message-passing mechanism that enables the simultaneous evaluation of multiple solutions, allowing our method to scale linearly with problem size.

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

DRL with Multimodal Encoder (Construction). Existing DRL methods for 3D bin packing problems (BPPs) are limited by computationally heavy encoders and large action spaces, which restrict them to handling only small instances (up to 50 boxes). To address these limitations, we propose a DRL agent that sequentially solves three subtasks: sequencing, orientation, and positioning of the boxes. Specifically, we use a multimodal encoder, where a sparse attention sub-encoder embeds the box states to reduce computational costs, while a convolutional neural network sub-encoder captures spatial information. We also incorporate action representation learning in the decoder to manage the large action space of the positioning subtask.

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

DRL guided Large Neighborhood Search (Improvement). To accelerate the solution process for integer programming (IP) problems, we propose a high-level, learning-based large neighborhood search (LNS) method. Using DRL, we train a policy network to act as the destroy operator in LNS, identifying a subset of variables in the current solution for re-optimization. A solver is then used as the repair operator to solve sub-IP problems and reoptimize the selected variables. Although heuristic in nature, our method effectively handles large-scale IP problems by breaking them down into smaller, more manageable sub-problems.

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

DRL guided Backtracking Search. Backtracking search algorithms are widely used for solving constraint satisfaction problems (CSPs), with their efficiency heavily dependent on variable ordering heuristics. Traditionally, these heuristics are hand-crafted based on expert knowledge. We propose a DRL-based approach to automatically discover variable ordering heuristics that are better suited to specific CSP instances, eliminating the need for hand-crafted rules. Our approach leverages a graph neural network to process CSP instances of varying sizes and constraint arities, capturing the complex relationships between variables and constraints.

F. **Multi-objective Combinatorial Optimization Problems** (MOCOP)

DRL guided Diversity Enhancement. Existing neural methods for multi-objective combinatorial optimization problems (MOCOPs) rely heavily on decomposition, often resulting in repetitive solutions for individual subproblems and yielding a limited Pareto set. To improve diversity, we propose a novel neural heuristic with diversity enhancement (NHDE) that generates more Pareto solutions from two perspectives. First, to prevent redundant solutions, we introduce an indicator-enhanced DRL method that guides the model, alongside a heterogeneous graph attention mechanism to capture the relationships between the instance graph and the Pareto front. Second, we implement a multiple Pareto optima strategy to explore and preserve desirable solutions in each subproblem's neighborhood. Our NHDE framework is generic and can be applied across different neural methods for multi-objective combinatorial optimization.

Phase II Research: *Generalizable Neural Solvers*

Unlike fields such as computer vision and natural language processing, where deep learning has made significant strides, solving combinatorial optimization problems like the VRP remains challenging due to their NP-hard nature and computational complexity. While neural heuristics based on deep models are relatively new, traditional methods, including heuristics and exact algorithms, have been developed and refined for decades. As a result, it is difficult for specialized neural heuristics to consistently outperform these well-established methods under arbitrary conditions.

Building on the foundation of our earlier work and insights from Phase I, my current research focuses on improving the **generalization capabilities** of neural solvers for combinatorial optimization problems. Although previous specialized neural solvers have shown promising results, their performance is often confined to idealized, synthetic problem instances. In contrast, real-world applications demand solvers that can generalize across varying *distributions*, instance *sizes*, *constraints*, and *metrics*.

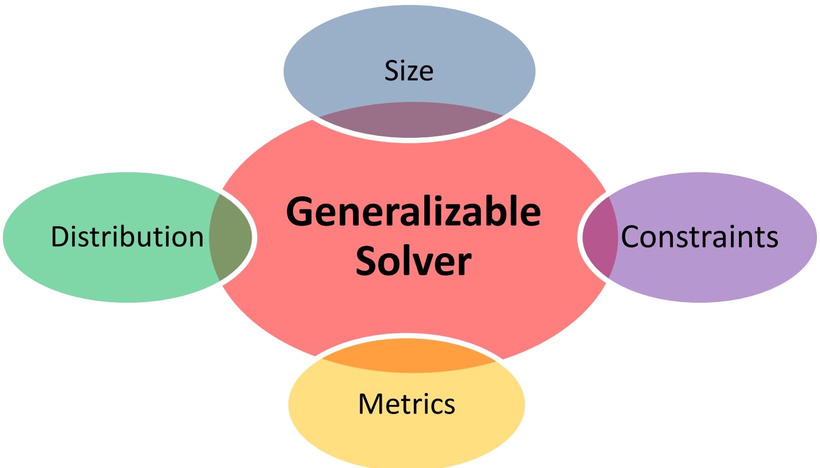

Figure 2. Phase II research focus – **Generalizable** neural solvers.

To address these challenges, using VRPs as an example, my focus is on developing **generalizable neural solvers** by improving cross-domain generalization through several approaches:

(1) **Cross-distribution generalization:** Existing neural solvers are typically trained on instances with a single, fixed distribution of node locations, such as uniform

distribution. Their performance degrades significantly when applied to instances with different distributions. To enhance cross-distribution generalization,

- We introduce an Adaptive Multi-Distribution Knowledge Distillation (AMDKD) scheme to train more generalizable deep models. AMDKD leverages knowledge from multiple teacher models trained on exemplar distributions to create a lightweight yet robust student model. An adaptive strategy is employed to allow the student model to focus on more challenging distributions, absorbing difficult-to-master knowledge more effectively. AMDKD achieves competitive results on both in-distribution and out-of-distribution instances, including synthetic and benchmark datasets (e.g., TSPLIB and CVRPLIB). Additionally, it requires fewer computational resources for inference.
- We also propose an ensemble-based deep reinforcement learning method for VRPs, which learns a diverse set of sub-policies to handle different instance distributions. To prevent parameter convergence, we enforce diversity across sub-policies using Bootstrap with random initialization, and apply regularization during training to further promote sub-policy inequality. This method outperforms state-of-the-art neural baselines on randomly generated distributions and generalizes well to benchmark datasets from TSPLib and CVRPLib.

(2) **Cross-size generalization**: Neural solvers are often trained on small, fixed-sized instances (e.g., 20, 50, or 100 nodes), resulting in poor performance when applied to larger instances. To improve cross-size generalization,

- We propose a continual learning framework that incrementally trains models on instances of increasing size. An inter-task regularization scheme retains knowledge from smaller problem sizes, while an intra-task regularization consolidates model performance by mimicking desirable behaviors from earlier training. We also use experience replay to revisit previously trained instances, mitigating catastrophic forgetting. Experimental results demonstrate superior performance across various problem sizes, including both seen and unseen instances, compared to state-of-the-art models.
- We also propose GLOP (Global and Local Optimization Policies), a unified hierarchical framework that scales efficiently to large routing problems. GLOP partitions large problems into Travelling Salesman Problems and further into Shortest Hamiltonian Path Problems, combining non-autoregressive neural heuristics for coarse-grained partitions and autoregressive heuristics for fine-grained route construction. GLOP achieves state-of-the-art real-time performance on large-scale routing problems, including TSP, ATSP, CVRP, and PCTSP.

(3) **Cross-constraint generalization**: Neural solvers are usually trained for specific problem constraints or tasks, requiring re-training from scratch for different VRP variants. To enhance cross-constraint generalization,

- We propose a unified neural solver for VRP variants, based on a multi-task vehicle routing solver with a mixture-of-experts (MVMoE) architecture. This approach enhances model capacity without proportional increases in computation, using a hierarchical gating mechanism to balance performance and complexity. Our method significantly improves zero-shot generalization on 10 unseen VRP variants and delivers strong results in both few-shot settings and on real-world benchmark instances. The extensive studies on MVMoE configurations also show the superiority of hierarchical gating for out-of-distribution data.

- We also introduce cross-problem learning, which leverages transferable knowledge across different VRP variants. We modularize the neural architecture, separating the backbone Transformer for TSP from lightweight modules that handle problem-specific features in complex VRPs. The backbone Transformer is pre-trained for TSP and fine-tuned for each target VRP variant. Extensive experiments demonstrate that fully fine-tuning the backbone Transformer outperforms training from scratch, while an adapter-based approach offers comparable performance with improved parameter efficiency. This method enhances cross-distribution applicability and model versatility.

(4) **Cross-metric generalization**: Neural solvers are often trained on instances with Euclidean coordinates, yet real-world problems frequently lack coordinate data or use non-Euclidean distance metrics. To improve cross-metric generalization,

- We propose a novel lifelong learning framework that incrementally trains neural solvers to manage VRPs in various contexts. Our framework utilizes a Transformer-based lifelong learner (LL) with an inter-context self-attention mechanism to transfer knowledge from previously solved VRPs to new ones. A dynamic context scheduler (DCS) with cross-context experience replay further enhances the model's ability to recall past policies. Extensive results on synthetic and benchmark instances show that this lifelong learner discovers effective policies for handling generic VRPs across different metrics, outperforming other neural solvers and achieving superior performance on most VRPs.

Selected Publications and Outputs

[1] Detian Kong, Yining Ma, *Zhiguang Cao*, Tianshu Yu, and Jianhua Xiao. Efficient Neural Collaborative Search for Pickup and Delivery Problems, IEEE Transactions on Pattern Analysis and Machine Intelligence (**T-PAMI**), 2024.

[2] Haoran Ye, Jiarui, Wang, *Zhiguang Cao*, Jinkyoo Park, and Guojie Song. ReEvo: Large Language Models as Hyper-heuristics with Reflective Evolution, Advances in Neural Information Processing Systems (**NeurIPS**), 2024.

[3] Jieyi Bi, Yining Ma, Jianan Zhou, Wen Song, *Zhiguang Cao*, Yaoxin Wu, and Jie Zhang. Learning to Handle Complex Constraints for Vehicle Routing Problems, Advances in Neural Information Processing Systems (**NeurIPS**), 2024.

[4] Jianan Zhou, Yaoxin Wu, *Zhiguang Cao*, Wen Song, Jie Zhang, and Zhiqi Shen. Collaboration! Towards Robust Neural Methods for Vehicle Routing Problems, Advances in Neural Information Processing Systems (**NeurIPS**), 2024.

[5] Jianan Zhou, *Zhiguang Cao*, Yaoxin Wu, Wen Song, Yining Ma, and Jie Zhang. MVMoE: Multi-Task Vehicle Routing Solver with Mixture-of-Experts, International Conference on Machine Learning (**ICML**), 2024.

[6] Yunzhuang Shen, Yuan Sun, Xiaodong Li, *Zhiguang Cao*, Andrew Eberhard, and Guangquan Zhang. Adaptive Stabilization Based on Machine Learning for Column Generation, International Conference on Machine Learning (**ICML**), 2024.

[7] Cong Zhang, *Zhiguang Cao*, Wen Song, Yaoxin Wu, and Jie Zhang. Deep Reinforcement Learning Guided Improvement Heuristic for Job Shop Scheduling, International Conference on Learning Representations (**ICLR**), 2024.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.
[17] 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.

[18] 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.

[19] 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.

[20] 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.

[21] 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.