GOAL: A Generalist Combinatorial Optimization Agent Learner

Darko Drakulić, Sofia Michel and Jean-Marc Andreoli
firstname.lastname@naverlabs.com

NAVER LABS Europe

ICLR 2025

NAVER LABS

Introduction

- Neural Combinatorial Optimization (NCO) - learn heuristics from data instead of manually devise them.
- Different approaches: transformer, GCN, diffusion -based.
- Current state-of-the-art models are problem specific

    Cannot solve multiple classes of CO problems (routing, scheduling, packing, graph, location) without changing the model and/or retraining from scratch.

Introduction

- Neural Combinatorial Optimization (NCO) - learn heuristics from data instead of manually devise them.
- Different approaches: transformer, GCN, diffusion -based.
- Current state-of-the-art models are problem specific

  Cannot solve multiple classes of CO problems (routing, scheduling, packing, graph, location) without changing the model and/or retraining from scratch.

- **Research question**: *Can we design and train a single model so that it can solve a variety of CO problems and efficiently fine-tuned to downstream tasks?*

Introduction

- Neural Combinatorial Optimization (NCO) - learn heuristics from data instead of manually devise them.
- Different approaches: transformer, GCN, diffusion -based.
- Current state-of-the-art models are problem specific

  Cannot solve multiple classes of CO problems (routing, scheduling, packing, graph, location) without changing the model and/or retraining from scratch.

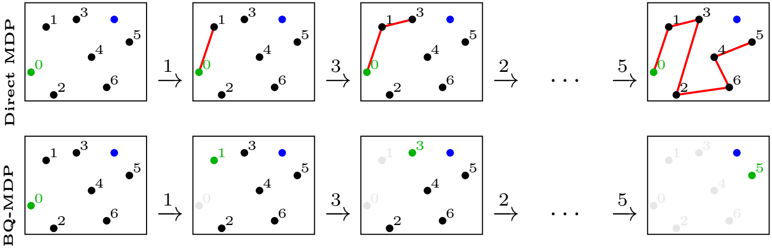
- **Research question**: *Can we design and train a single model so that it can solve a variety of CO problems and efficiently fine-tuned to downstream tasks?*
- This approach is investigated in other domains (language, vision, Atari, proprioceptive) (Reed et al., 2022), but it is novel in NCO domain.

Constructive CO and task representation

Solving a CO problem instance can be formulated as a sequential decision procedure and modeled as an MDP.

We use the BQ-MDPs (Drakulic et al., 2023), which are defined for “tail-recursive” problems.

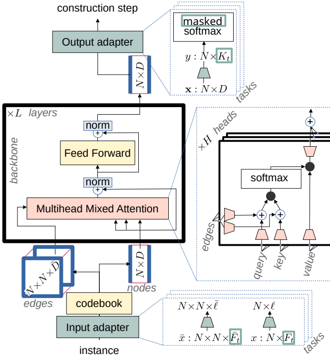
- Action space: construction steps.
- State space:
  - **traditional**: $\langle$ instance$^{\text{fixed}}$, partial solution$^{\text{changing}}\rangle$ pair
  - **BQ-MDP**: instance$^{\text{changing}}$ only
- Transitions: (illustrated on path-TSP)



Architecture of the model

Each instance of size N of a task (CO problem) is represented by:

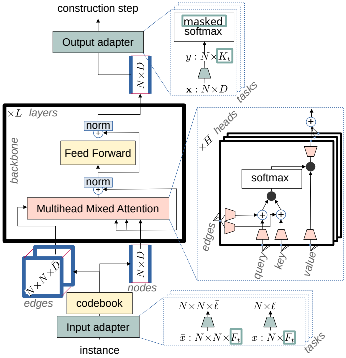
- a set of N abstract "nodes", together with features which can be attached at node
- a set of relations between nodes, "edges", at most $N \times N$, and
- instance features



Architecture of the model

Each instance of size N of a task (CO problem) is represented by:

- a set of N abstract “nodes”, together with features which can be attached at node
- a set of relations between nodes, ”edges”, at most $N \times N$, and
- instance features

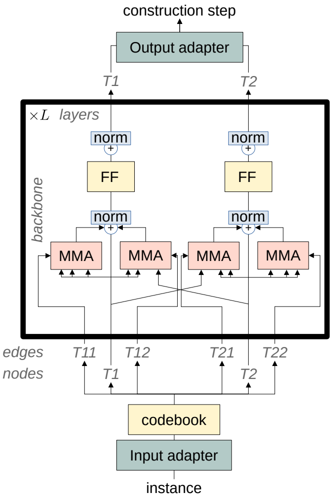


$$Q_n^{(h)} = Q_n \mathbf{W}_Q^{(h)}, \qquad K_m^{(h)} = K_m \mathbf{W}_K^{(h)}, \qquad V_m^{(h)} = V_m \mathbf{W}_V^{(h)}$$

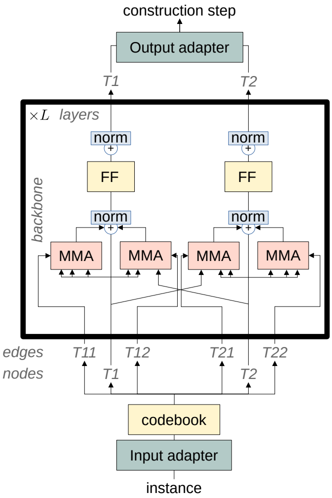
$$Q'^{(h)}_{mn} = E_{mn} \mathbf{W}'^{(h)}_Q, \qquad K'^{(h)}_{mn} = E_{mn} \mathbf{W}'^{(h)}_K.$$

$$r = \sum_h \mathrm{softmax}_{\mathrm{col}}(S^{(h)} + \mathcal{M})^\top V^{(h)} \mathbf{W}_O^{(h)\top} \quad \text{with} \quad S^{(h)}_{mn} = (Q_n^{(h)} + Q'^{(h)}_{mn})(K_m^{(h)} + K'^{(h)}_{mn})^\top$$

Mixed Type Attention



Mixed Type Attention



- All blocks **share** the same parameters!
  ⇒ we can use the same backbone for single and multi-type problems

Experiments: One Model to Solve Them All

- We train GOAL by imitation learning on eight classic and varied CO problems.
  1. Asymmetric Traveling Salesman Problem (ATSP)
  2. Capacitated Vehicle Routing Problem (CVRP)
  3. CVRP with Time Windows (CVRPTW)
  4. Orienteering Problem (OP)
  5. Job Shop Scheduling Problem (JSSP)
  6. Unrelated Machine Scheduling Problem (UMSP)
  7. Knapsack Problem (KP)
  8. Minimum Vertex Covering Problem (MVC)
- Training instances of “small” size (100 nodes)
- Good-quality solutions provided by specialized heuristics

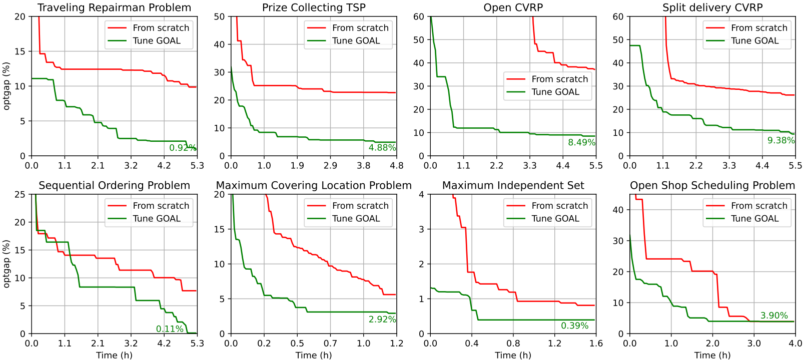
Performance on the training tasks

	ATSP100		CVRP100		CVRPTW100		OP100	
	gap	time	gap	time	gap	time	gap	time
Oracle solver	0.00%	29s	0.00%	12m	0.00%	10m	0.00%	1.1m
MDAM greedy	-		4.84%	1s	-		2.88%	16s
POMO no aug	-		**1.21**%	1s	-		-	
Sym-NCO greedy	-		3.33%	1s	-		2.03%	2s
BQ-NCO greedy	1.27%	2s	2.79%	2s	-		0.22%	2s
MVMoE/4E	-		1.65%	1s	4.90%	1s	-	
RouteFinder-TE	-		1.50%	1s	3.19%	1s	-	
MatNet greedy	0.93%	1s	-		-		-	
GOAL SINGLE-TASK greedy	**0.30**%	10s	2.34%	10s	**2.61**%	10s	**-0.04**%	3s
GOAL MULTI-TASK greedy	0.91%	10s	3.16%	10s	3.82%	10s	0.43%	3s

	KP100		MVC100		UMSP100x20		JSSP10x10	
	gap	time	gap	time	gap	time	gap	time
Oracle solver	0.00%	1s	0.00%	2m	0.00%	2m	0.00%	47s
POMO no aug	0.19%		-		-		-	
BQ-NCO greedy	**0.10**%		-		-		-	
S2V-DQN	-		0.97%	2s	-		-	
COMPASS	-		-		-		4.70%	3h
Gumbeldore greedy	-		-		-		3.17%	9s
GOAL SINGLE-TASK greedy	**0.10**%	3s	**0.23**%	3s	**2.82**%	7s	**2.73**%	15s
GOAL MULTI-TASK greedy	0.12%	3s	0.37%	3s	3.84%	7s	4.13%	15s

Fine-tuning to new tasks by self-improvement

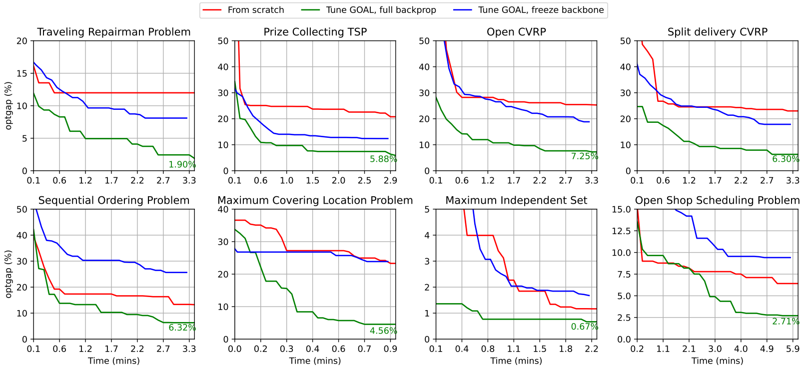
NAVER LABS



Fine-tuning to new tasks by imitation

NAVER LABS

Finetuning of 128 solved instances.



Conclusion

We propose a generalist neural CO model (GOAL) with a novel architecture which can solve multiple CO problems that can be represented as a graph.

1. Single-task specialized model provides state-of-the-art results on 7 of 8 test benchmarks,
2. A multi-task model successfully solves various CO problems with the same backbone, and
3. A pre-trained model can be efficiently fine-tuned to new problems in a supervised or unsupervised way

Thank you for your attention

References

Drakulic, D., Michel, S., Mai, F., Sors, A., & Andreoli, J.-M. (2023).BQ-NCO: Bisimulation Quotienting for Efficient Neural Combinatorial Optimization. *Advances in Neural Information Processing Systems*, *36*, 77416–77429.

Reed, S., Zolna, K., Parisotto, E., Colmenarejo, S. G., Novikov, A., Barth-Maron, G., Gimenez, M., Sulsky, Y., Kay, J., Springenberg, J. T., Eccles, T., Bruce, J., Razavi, A., Edwards, A., Heess, N., Chen, Y., Hadsell, R., Vinyals, O., Bordbar, M., & de Freitas, N. (2022, November). A Generalist Agent.