Tackling the Long Tail of Transportation Optimization with Machine Learning

Cathy Wu

Associate Professor | MIT LIDS, CEE, & IDSS

Swiss Transport Research Conference (STRC)

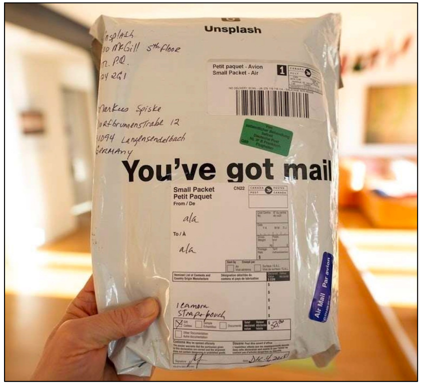
Data Science for Transportation Systems

An **AI for Engineering** Research Group

Focus: Design AI & optimization methods to make it easier to understand and improve transportation systems

Motivating example: Vehicle Routing Problems (VRP)

- One of the most widely studied in Combinatorial Optimization:
  - ≈10,000 works published in 2025 alone (Google Scholar), mostly heuristics
  - Direct application in real systems that distribute goods and provide services
  - Last-mile delivery, waste collection, field service, ride-pooling, on-demand transit

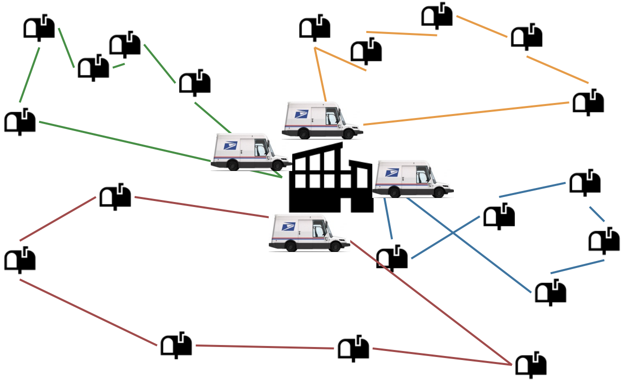


~10,000	**$11T+**	**65M+**
papers in 2025	global logistics costs, 2023	US parcel deliveries / day

Dantzig, G. B., & Ramser, J. H. "The Truck Dispatching Problem." Management Science, 1959.
Stats: Statista/Research & Markets (2024); ShipMatrix US parcel report (2024).

Long Tail of Transportation Optimization Problems

- Definition: **Vehicle Routing Problems (VRPs)**
- Given
  - Depot
  - N customers (location, demand)
  - Vehicle fleet (capacity)
  - … **Other constraints** …
- Find the lowest cost set of routes that satisfies all customers



Classical dimensions

- Vehicle capacities
- Time windows
- Heterogeneous fleets
- Multiple depots
- Pickup and delivery, backhauling
- Split delivery
- Arc routing (e.g., garbage collection)

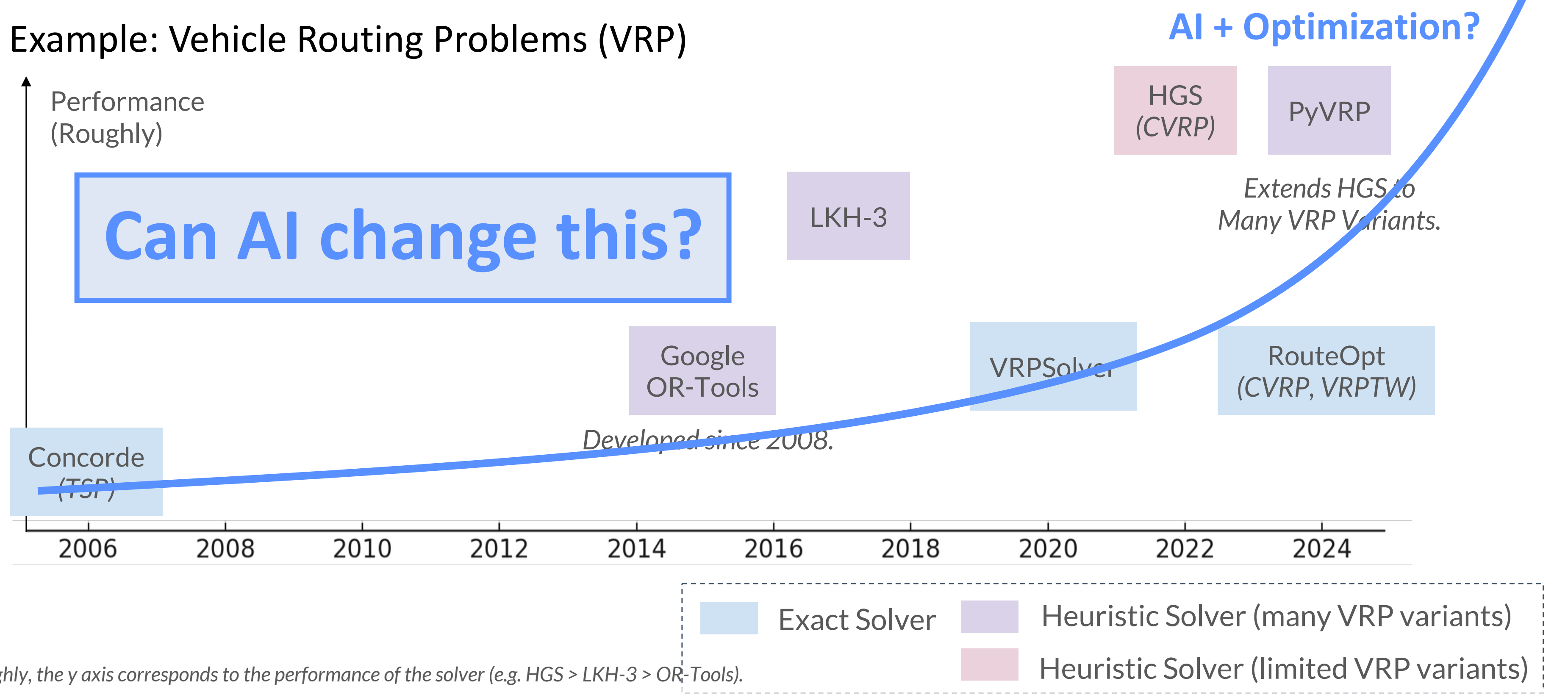
Emerging variants

- Electric vehicle routing (charging decisions)
- Drone routing (3D, endurance limits)
- Warehouse routing (pickers, sortation)
- Multi-modal logistics (truck + drone, rail + truck)

Laporte, G. "Fifty Years of Vehicle Routing." Transportation Science, 2009.
Dantzig, G. B., & Ramser, J. H. "The Truck Dispatching Problem." Management Science, 1959.

Solving the Long Tail is Slow and Laborious

Example: Vehicle Routing Problems (VRP)



*Roughly, the y axis corresponds to the performance of the solver (e.g. HGS > LKH-3 > OR-Tools).

Laporte, G. "Fifty Years of Vehicle Routing." Transportation Science, 2009.
Dantzig, G. B., & Ramser, J. H. "The Truck Dispatching Problem." Management Science, 1959.


WAYMO

Policy Implications of Connected and Automated Vehicles?

CAV affordance: Eco-driving at signalized intersections

Video credits: Audi

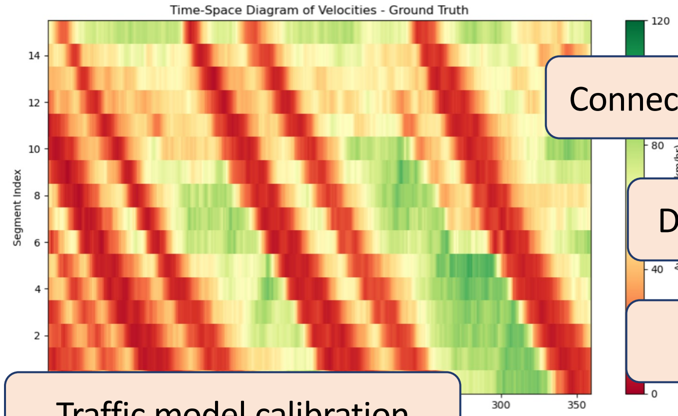
From ‘how to eco-drive’ (1000+ articles) to ‘should we eco-drive’ (≈0 articles)?

Research Question: Would intelligent control of vehicles to reduce CO_2 emissions at intersections at the city scale move the needle on climate change mitigation goals?

* Estimated using Google Scholar

[1] Mintsis et al. *Dynamic eco-driving near signalized intersections: Systematic review and future research directions*. Journal of Transportation Engineering, Part A, 2020

The Long Tail of Optimization is Cross-Cutting



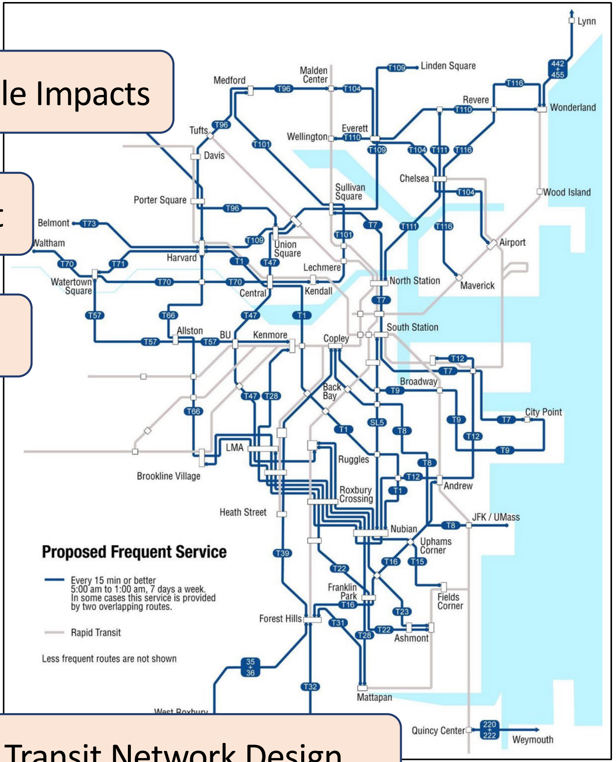
Connected and Automated Vehicle Impacts

Dynamic traffic assignment

Network optimization

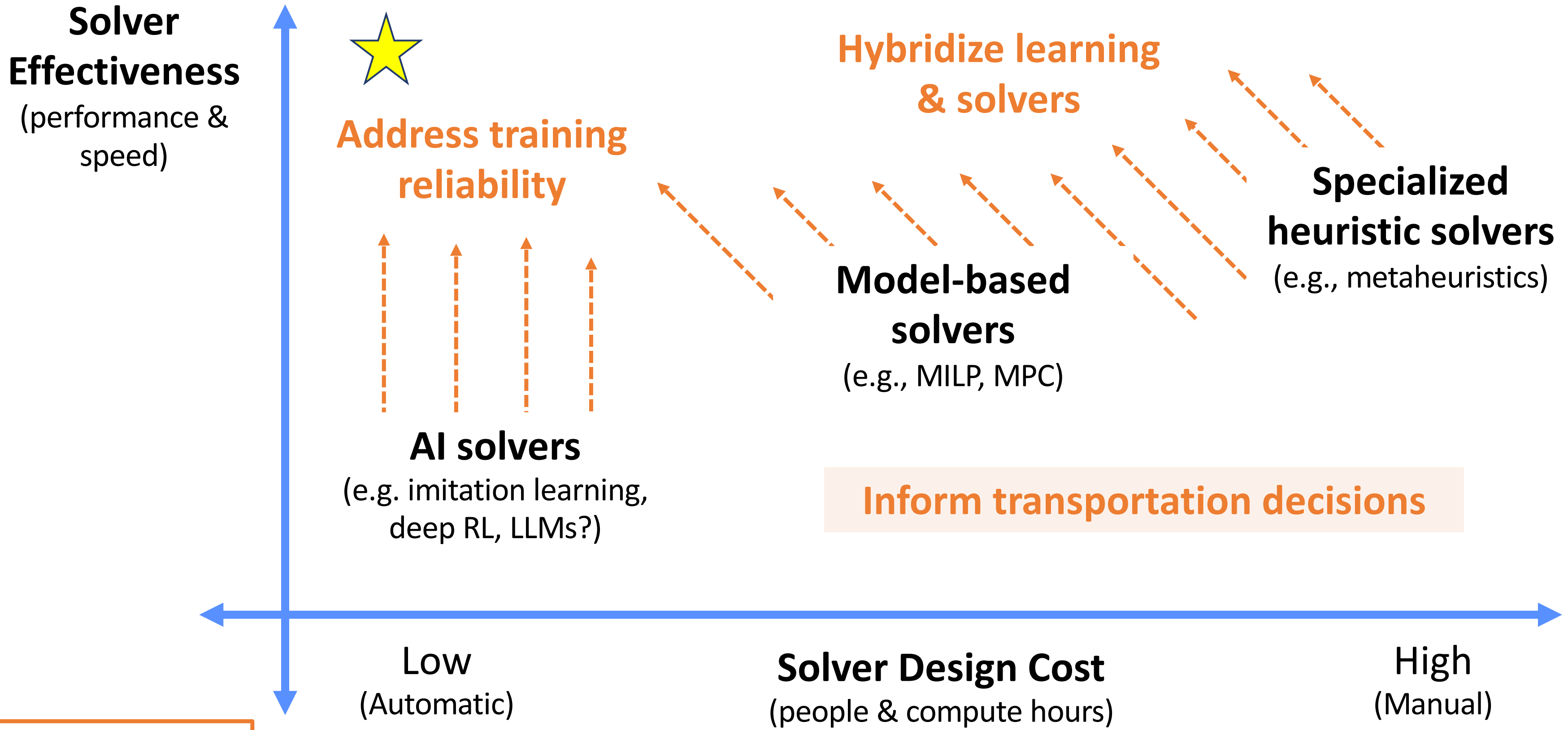
Traffic model calibration

Warehouse operations



Transit Network Design

Optimization Solver Landscape



My research lab

Research Overview

Address training reliability in contextual reinforcement learning

Identify RL non-robustness
[NeurIPS'22]

Model-based transfer learning
[**NeurIPS'24**, TRO'25, AAAI'26, L4DC'26, arXiv]

Residual policy learning
[ICRA'24, arXiv]

Hybridize learning & solvers for combinatorial optimization

Decomposition methods
[NeurIPS'21 **Spotlight (3%)**, **ICLR'24**]

Mixed Integer Linear Programming
[NeurIPS'23, ICLR'24, ICML'26]

Permutation methods
[ICLR'25 **Oral (1%)**, ICLR'25]

Inform transportation decisions for traffic and multi-agent coordination

Cooperative eco-driving
[ICLR'25, TCNS'25, **TRC'25**, TRC'26]

Automated warehousing
[ICAPS'24, ICRA'24, ICLR'24, JAIR'26]

Driving advisories
[TCNS'23, JATS'24, ITSM'25, T-ITS'25]

Traffic digital twins
[TRBAM'23, AAP'25, ISTTT'26, IVS'26, arXiv]

Routing
[KDD'25 **Oral (1%)**, TCNS'25, T-ITS'26]

Mixed autonomy traffic [TASE'22, TRO'23, TRO'24, T-ITS'24, IROS'24, IROS'24]

Outline

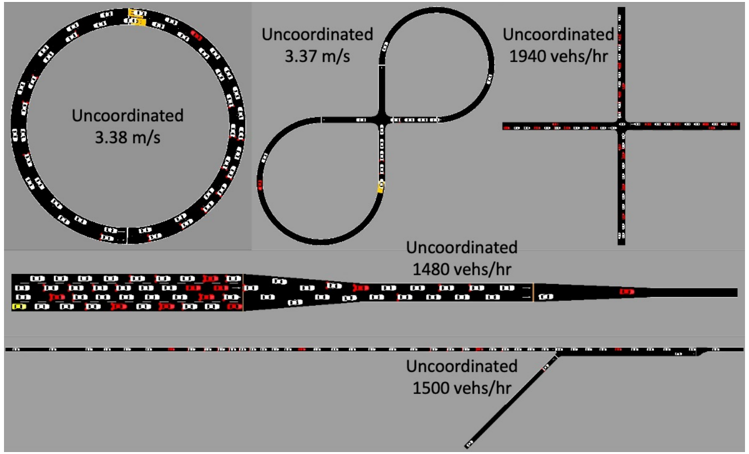
1. **Addressing Training Reliability in Contextual RL**
2. **Hybridizing Learning & Solvers in Combinatorial Optimization**
3. **Applications: Informing Transportation Decisions**

Outline

1. **Addressing Training Reliability in Contextual RL**
2. **Hybridizing Learning & Solvers in Combinatorial Optimization**
3. **Applications: Informing Transportation Decisions**

Summary of my first two years at MIT

Success during PhD [1-3]

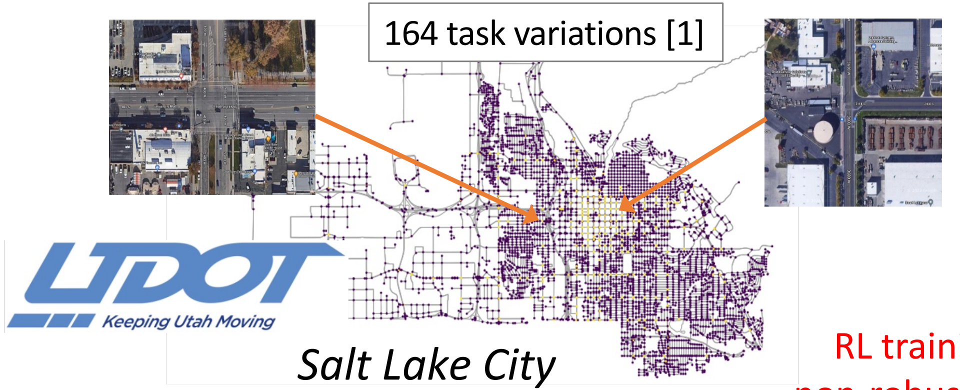


≈10 **unsuccessful** applications of RL

Success with enough engineering

[1] **Wu**, Kreidieh, Vinitsky, Bayen, "Emergent behaviors in mixed-autonomy traffic," in *CoRL*, 2017.
[2] **Wu**, Kreidieh, Parvate, Vinitsky, Bayen, "Flow: A modular learning framework for mixed autonomy traffic," *IEEE T-RO*, 2021.
[3] Yan, Kreidieh, Vinitsky, Bayen, **Wu**, "Unified automatic control of vehicular systems with reinforcement learning," *IEEE T-ASE*, 2022.

Finding: RL non-robustness to task variation



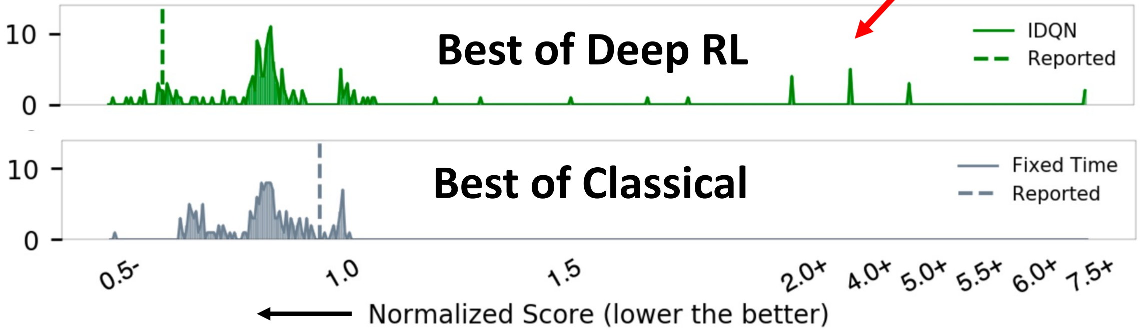
Methods [2]	Relative Delay (↓)
Fixed Time	**0.82**
Max Pressure	0.98
IDQN	1.06
IPPO	1.69
MPLight	1.11
MPLight*	1.49

Classical

Deep RL

Worse than fixed time

RL training non-robustness

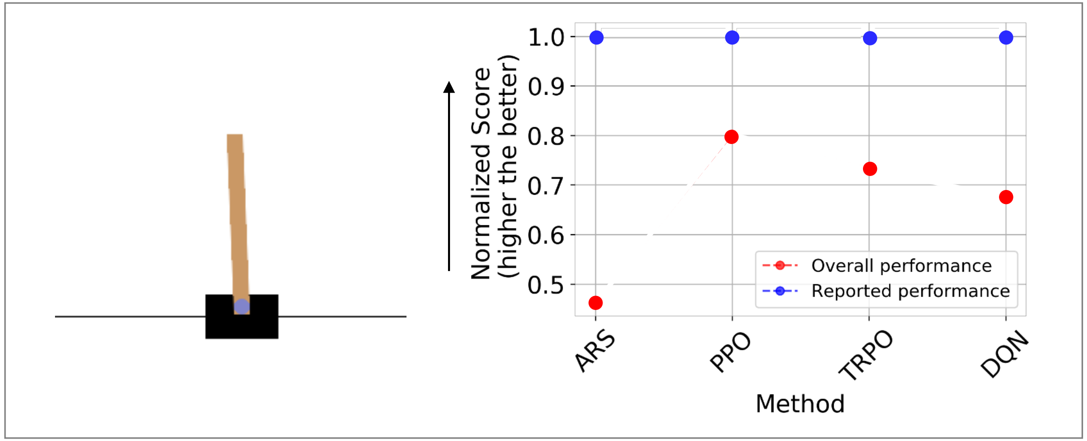


Deep RL methods are *not robust* to variations in traffic signal control

[1] Qu*, Valiveru*, Tang, Jayawardana, Freydt, **Wu**. "What is a Typical Signalized Intersection in a City? A Pipeline for Intersection Data Imputation from OpenStreetMap." TRB, 2023.
[2] Ault, Sharon. "Reinforcement learning benchmarks for traffic signal control." Advances in Neural Information Processing Systems (NeurIPS), 2021.
Jayawardana, Tang, Li, Suo, **Wu**. "The Impact of Task Underspecification in Evaluating Deep Reinforcement Learning." Advances in Neural Information Processing Systems (NeurIPS), 2022.

Finding: RL non-robustness to task variation

Similar findings with general deep RL methods & standard control benchmarks



Deep RL methods are **not robust** to variations in Cartpole

Deep RL methods are highly sensitive to task variation.

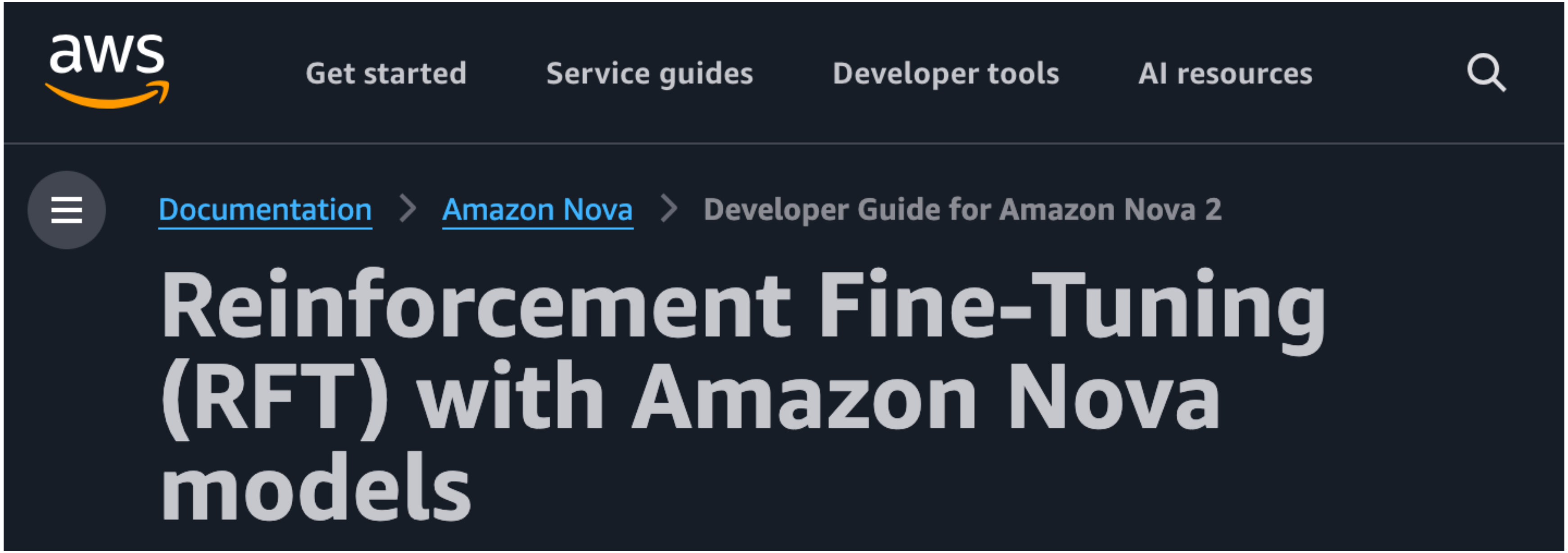
Yet, task variation is everywhere in real applications.

*Reported performance is reproduced from common benchmark task specification.

Jayawardana, Tang, Li, Suo, **Wu**. "The Impact of Task Underspecification in Evaluating Deep Reinforcement Learning." Advances in Neural Information Processing Systems (NeurIPS), 2022.

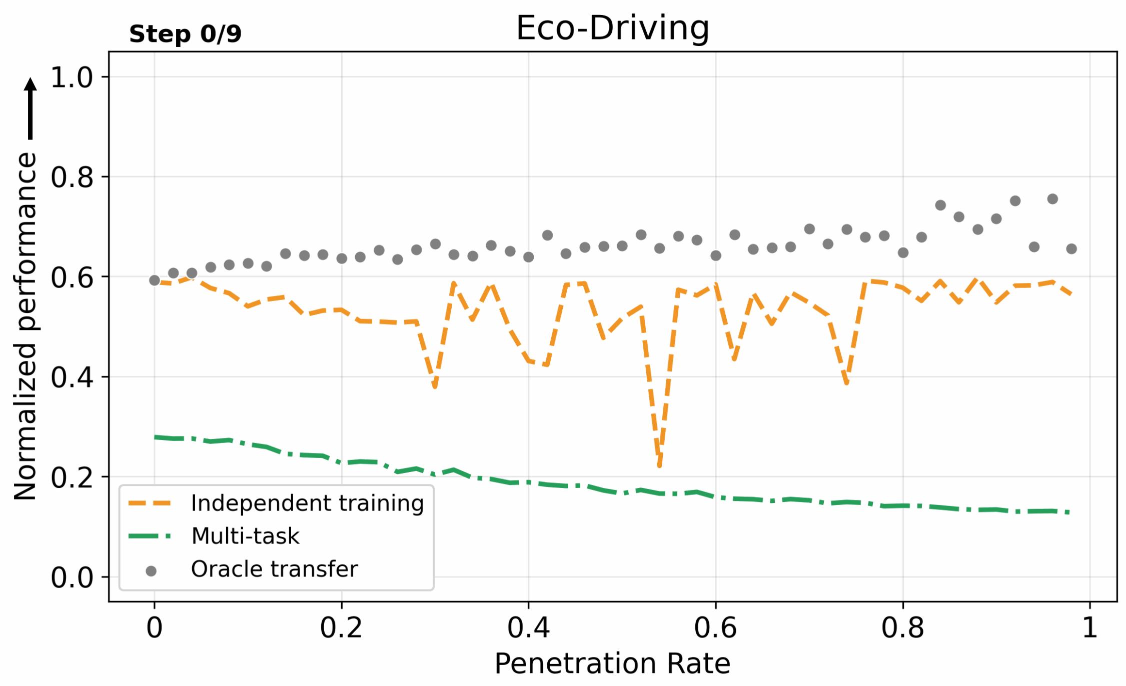
Similar challenge appears in RL fine-tuning of LLMs

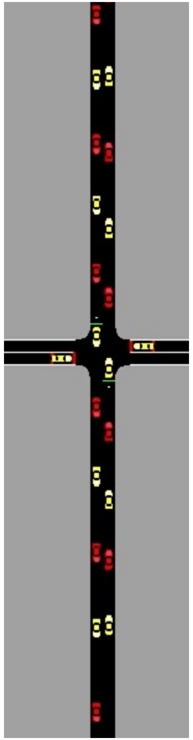
- Transportation-driven research!



H. Zheng *et al.*, **C. Wu**, "Internalizing Curriculum Judgment for LLM Reinforcement Fine-Tuning," arXiv. doi: [10.48550/arXiv.2605.11235](https://doi.org/10.48550/arXiv.2605.11235).

Task Variations as a Contextual MDP



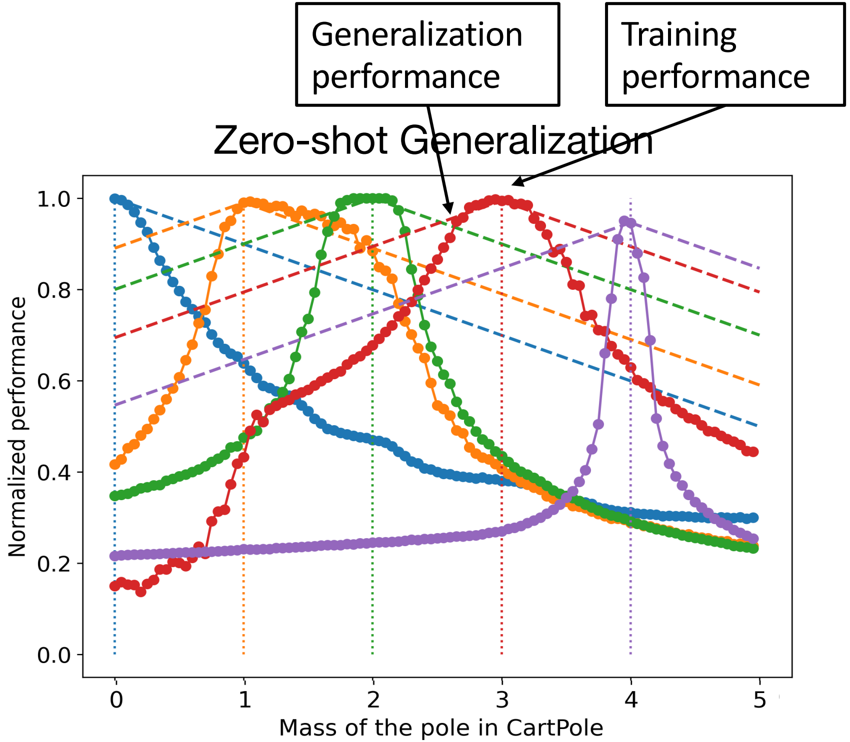


- Contextual MDP $\mathcal{M}_\Phi$: Family of related Markov Decision Problems (MDPs), with transitions/rewards conditioned on a context vector $\phi \in \Phi$ (context space)

J.-H. Cho, V. Jayawardana, S. Li, and C. **Wu**, "Model-Based Transfer Learning for Contextual Reinforcement Learning," in NeurIPS, 2024.

Model-based Transfer Learning

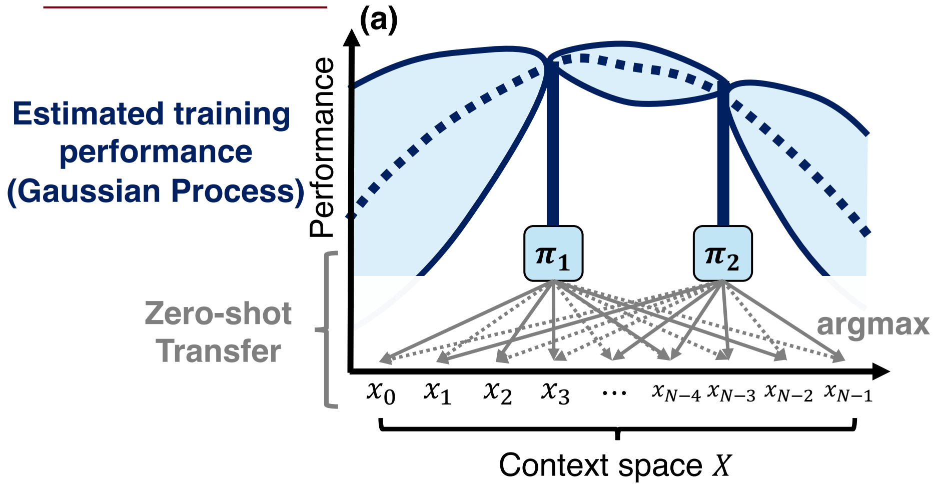
- *Key idea*: Success on task ϕ can come from either:
  1. Directly training on task ϕ
  2. Transferring from some task $\phi' \neq \phi$
- Training is expective, transfer is cheap
- Where to train?
- Approach: Active learning with Bayesian optimization
- Performance models
  - Training performance: Gaussian process
  - Transfer performance: Linear in Euclidean context distance



J.-H. Cho, V. Jayawardana, S. Li, and C. **Wu**, "Model-Based Transfer Learning for Contextual Reinforcement Learning," in NeurIPS, 2024.

Model-Based Transfer Learning

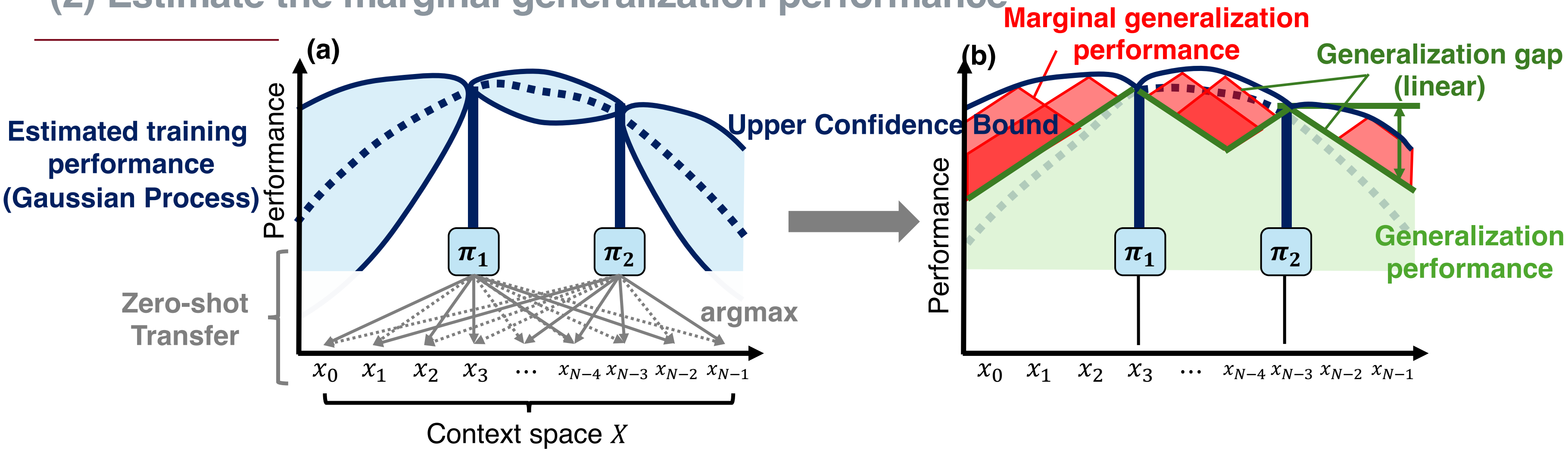
(1) Estimate training performance with Gaussian Process



Wu

Model-Based Transfer Learning

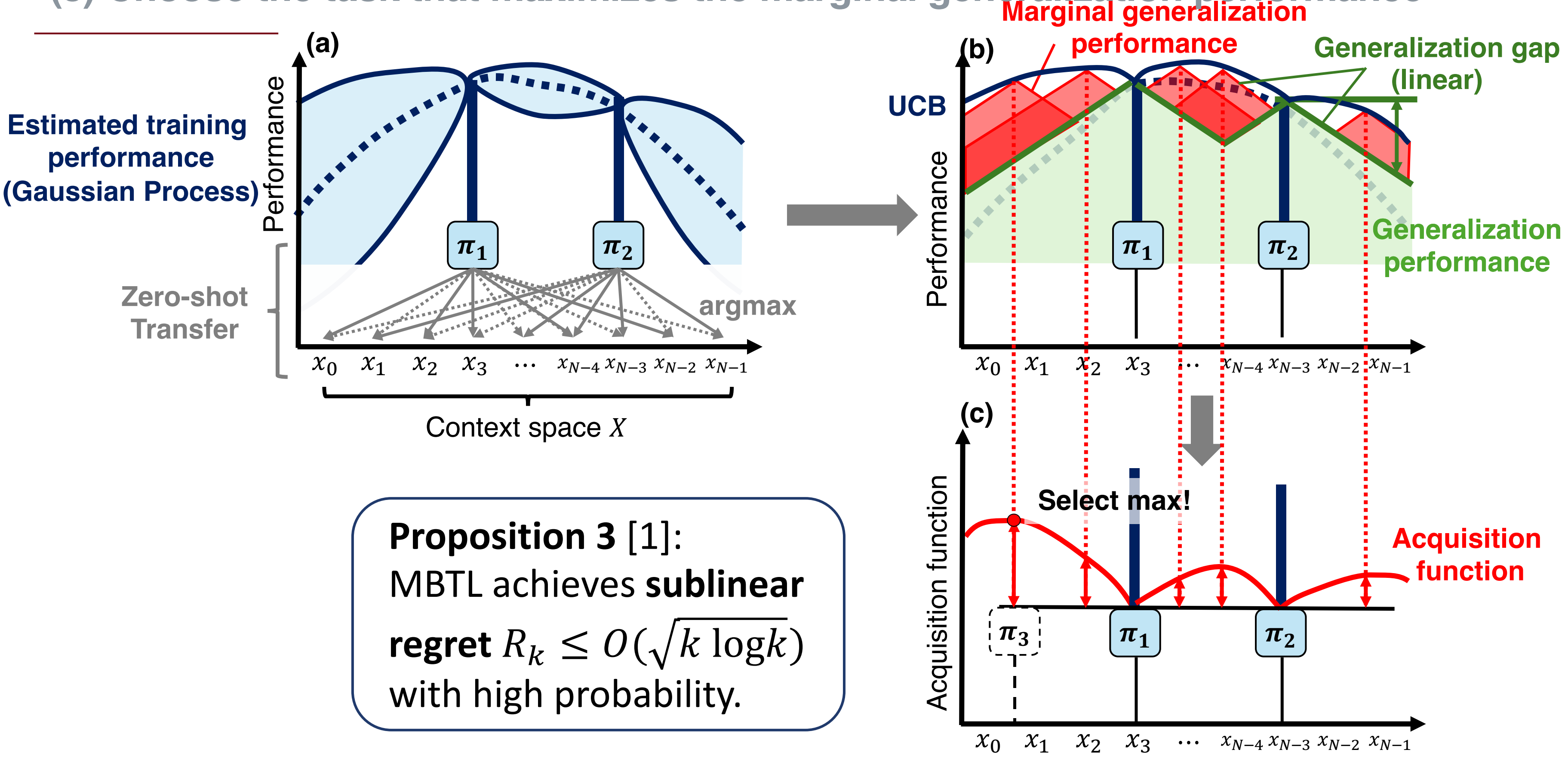
(2) Estimate the marginal generalization performance



Wu

Model-Based Transfer Learning

(3) Choose the task that maximizes the marginal generalization performance

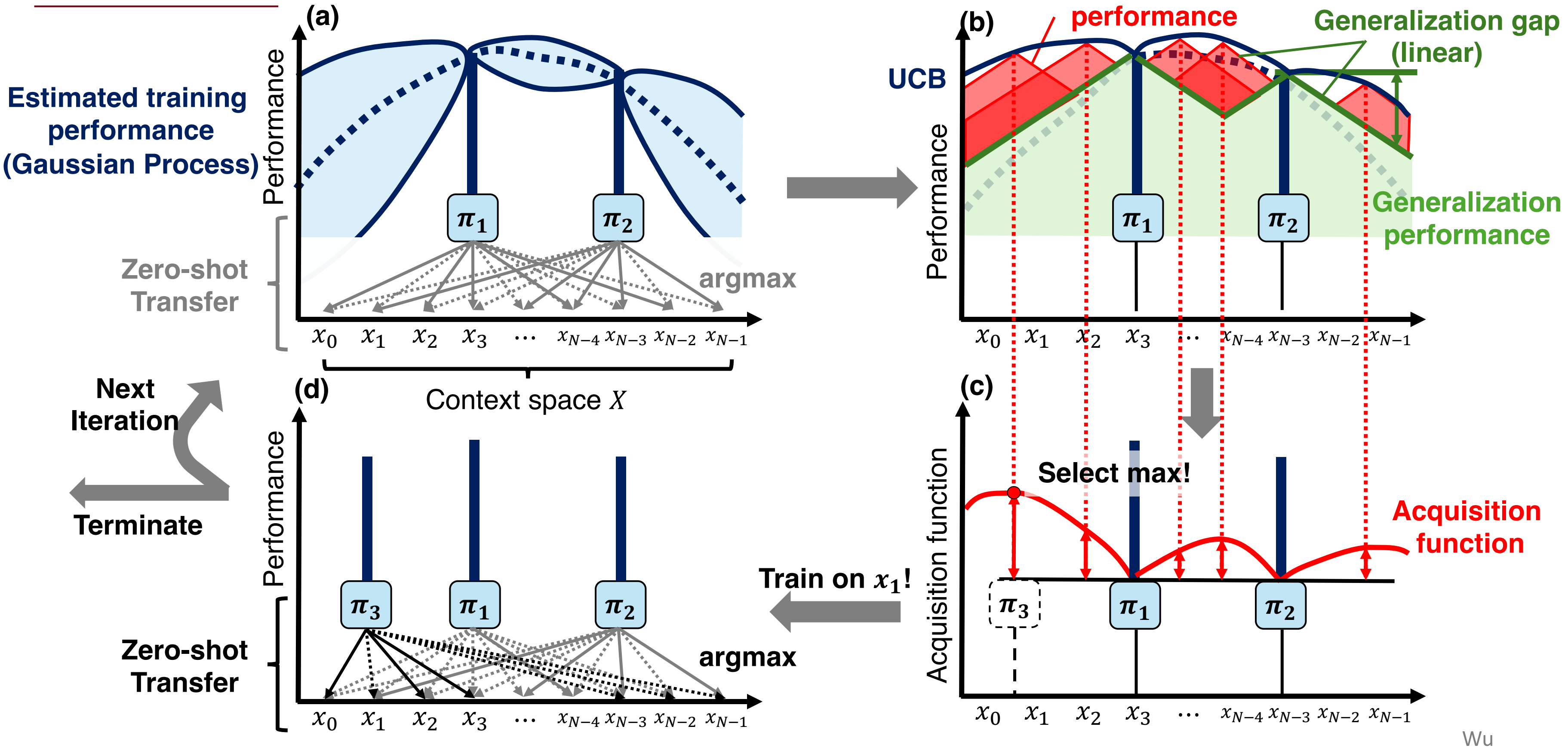


Proposition 3 [1]:
MBTL achieves **sublinear regret** $R_k \leq O(\sqrt{k \log k})$ with high probability.

[1] N. Srinivas, A. Krause, S. M. Kakade, and M. W. Seeger, "Information-Theoretic Regret Bounds for Gaussian Process Optimization in the Bandit Setting," IEEE Trans. Inform. Theory, 2012.

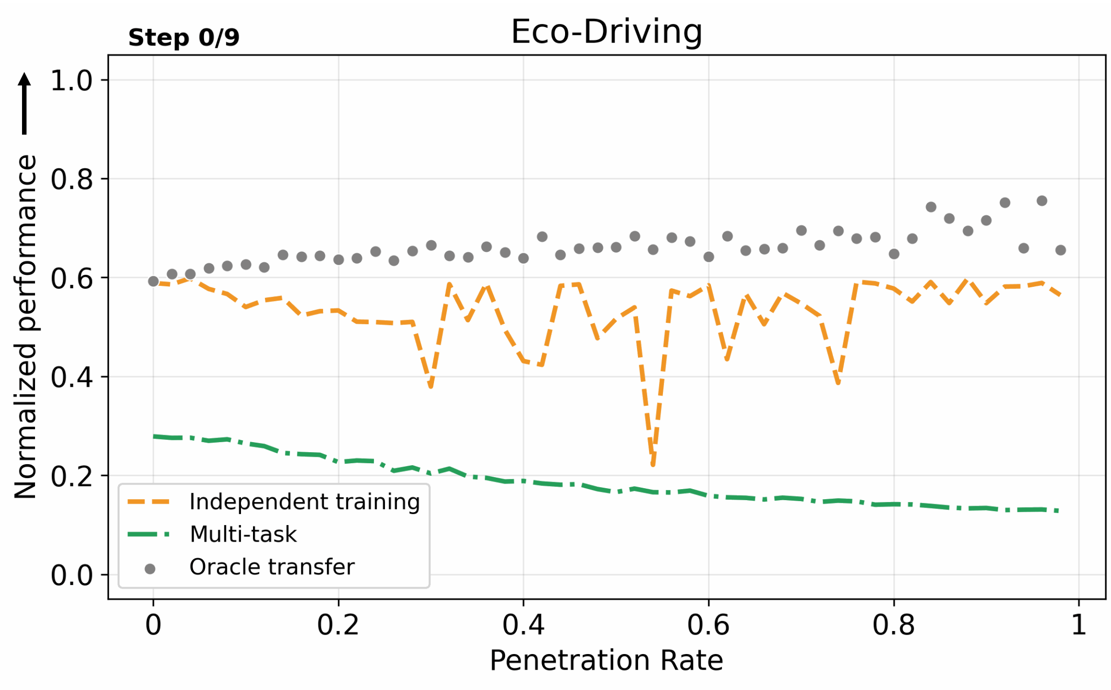
Model-Based Transfer Learning

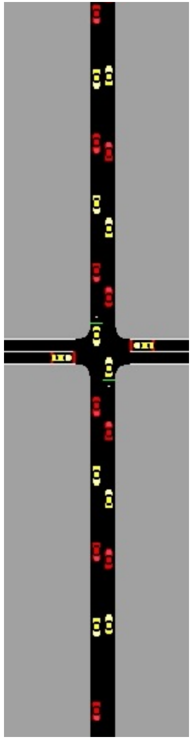
(4) Calculate generalization performance using zero-shot transfer



Wu

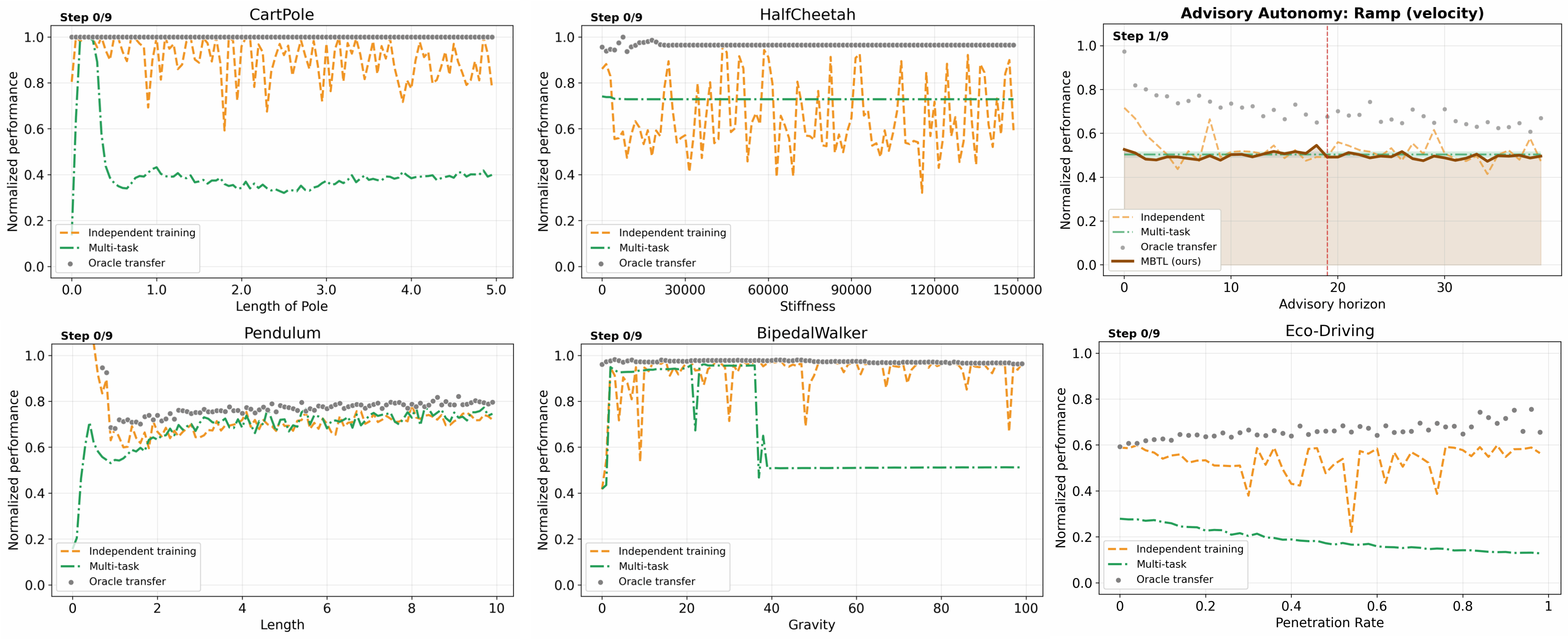
Bayesian optimization for task selection → 10-30x RL efficiency

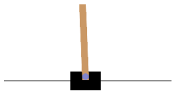


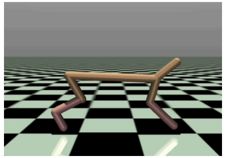


J.-H. Cho, V. Jayawardana, S. Li, and C. **Wu**, “Model-Based Transfer Learning for Contextual Reinforcement Learning,” in NeurIPS, 2024.

Bayesian optimization for task selection → 10-30x RL efficiency

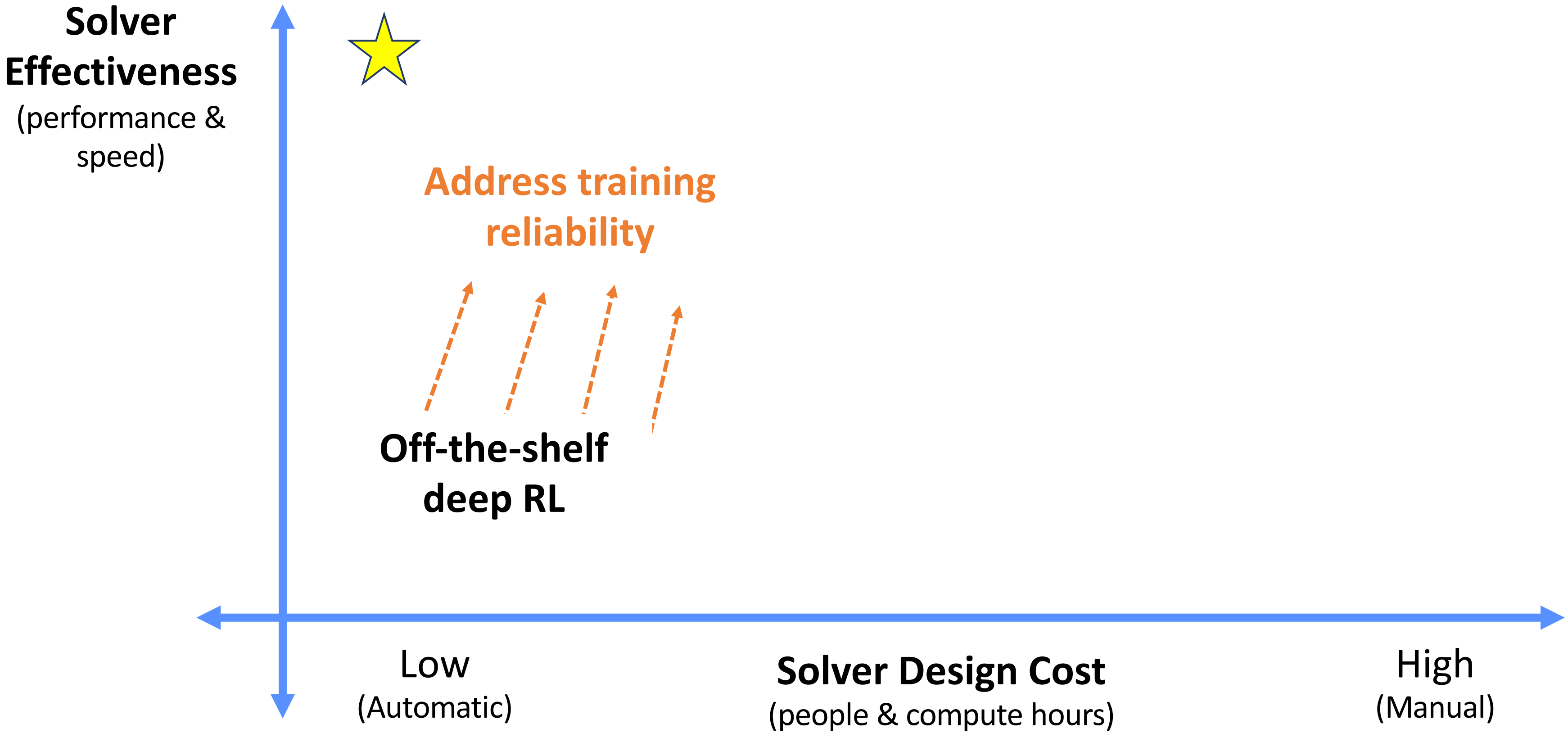






J.-H. Cho, V. Jayawardana, S. Li, and C. **Wu**, "Model-Based Transfer Learning for Contextual Reinforcement Learning," in NeurIPS, 2024.

AI for Optimization Landscape

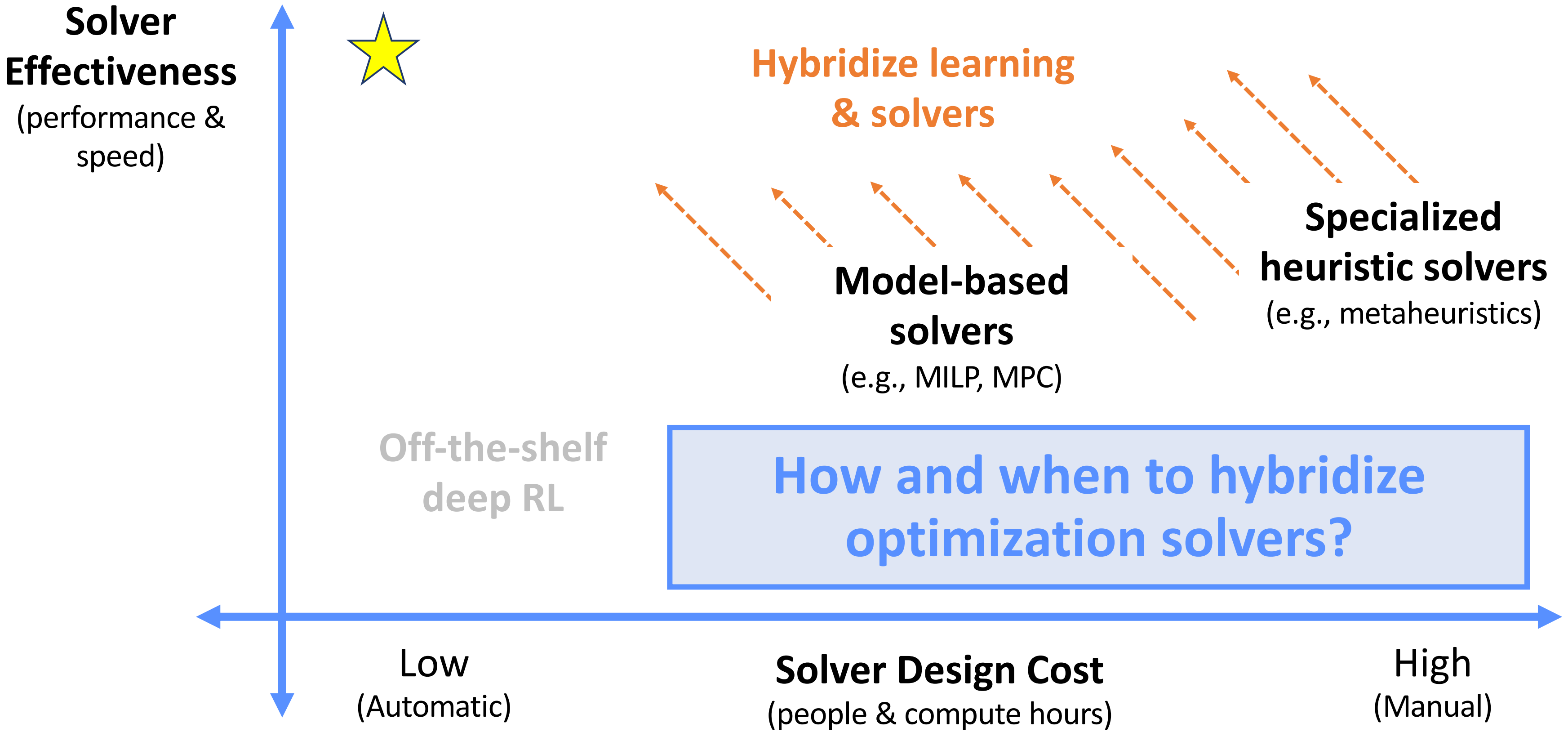


[1] T. Zhou, J.-H. Cho, and **C. Wu**, “Structure Detection for Contextual Reinforcement Learning,” in *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, 2026.

Outline

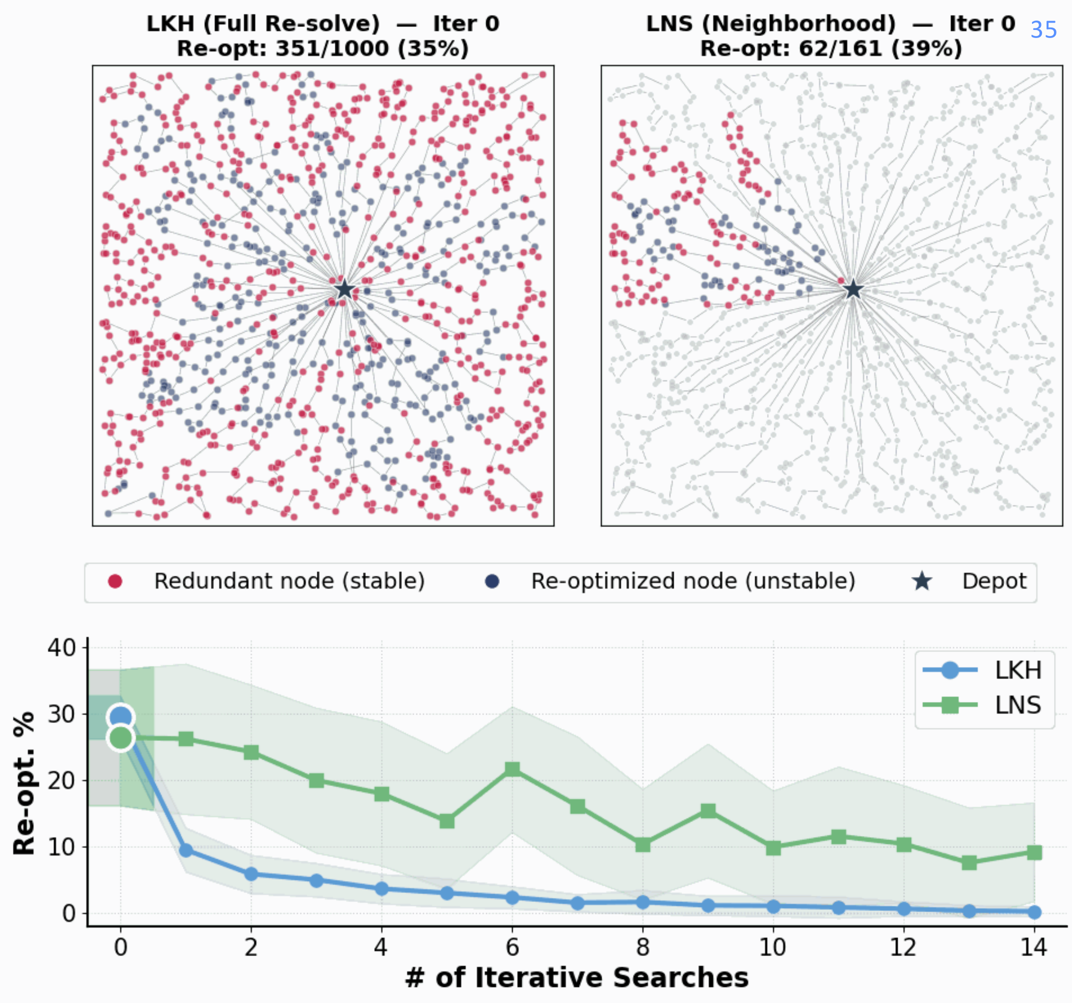
1. Addressing Training Reliability in Contextual RL
2. **Hybridizing Learning & Solvers in Combinatorial Optimization**
3. Applications: Informing Transportation Decisions

AI for Optimization Landscape



Solver Redundancy

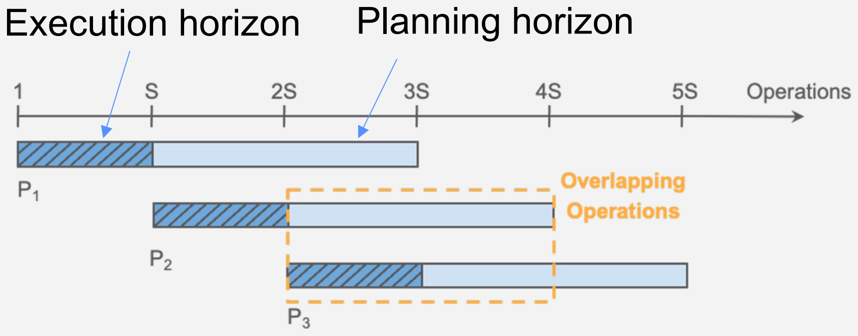
- Example: Vehicle Routing (Solvers: LKH, LNS + LKH)



W. Ouyang, S. Li, Y. Ma, and C. Wu, "Learning to Segment for Vehicle Routing Problems," International Conference on Learning Representations (ICLR), 2026. **Oral (1%)**.

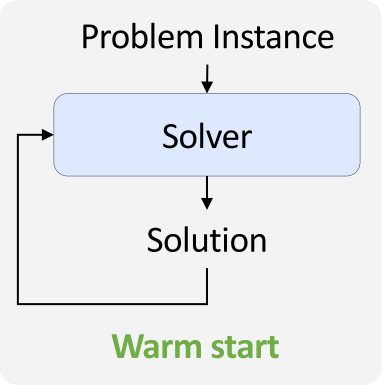
Learn to Reuse: Avoid Redundant Calculations

Setting: Rolling Horizon Optimization (RHO)

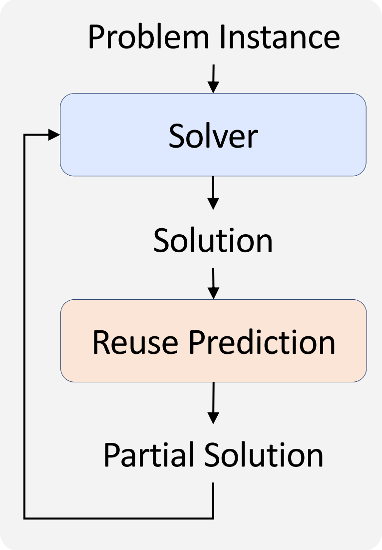


L-RHO: Up to 50% faster by *learning to reuse solutions in consecutive RHO subproblems* [1]

Warm start RHO



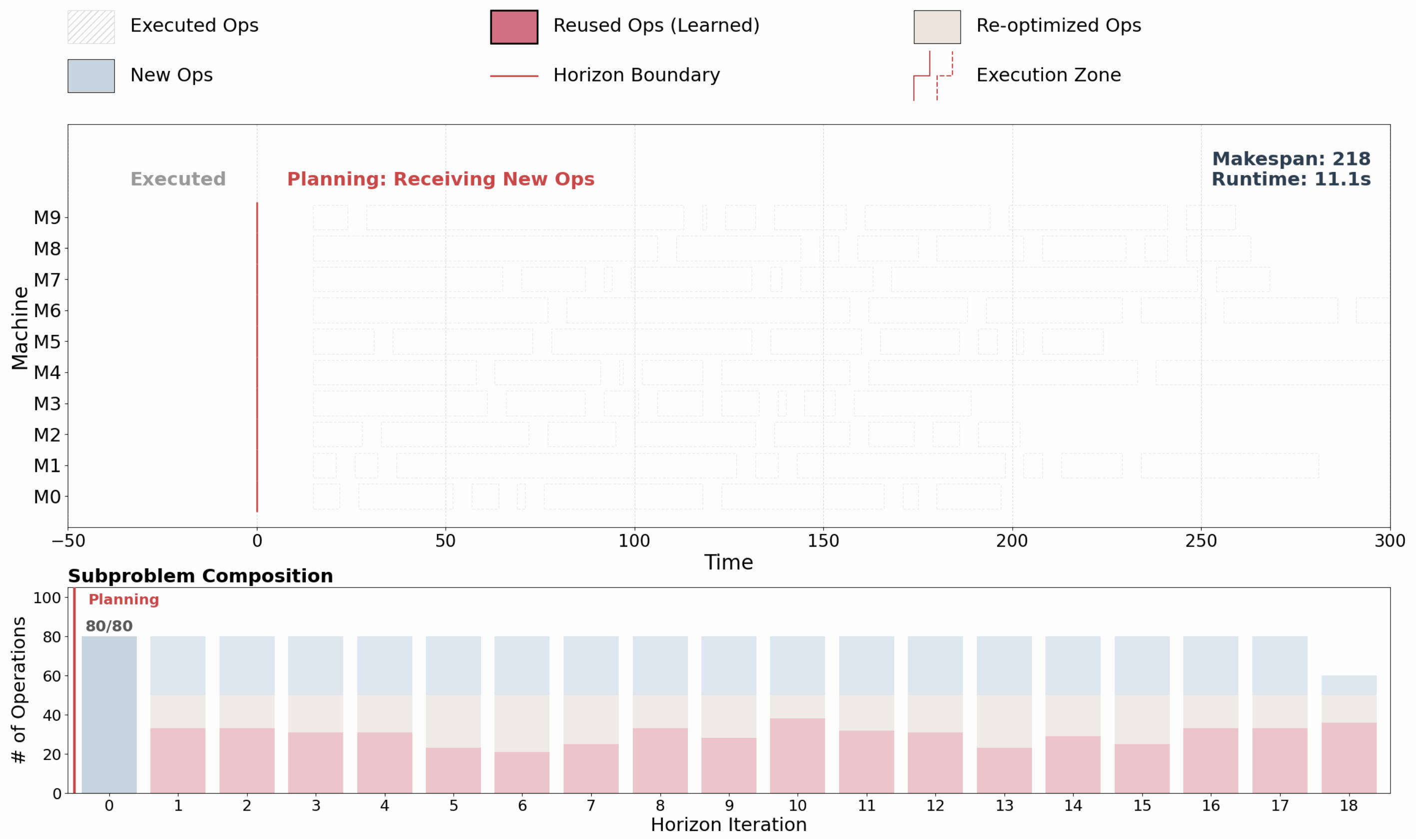
L-RHO (Ours)



- Method: Train a deep neural network to predict if a solution variable will change from one subproblem to the next
- Training Labels: Look-ahead oracle

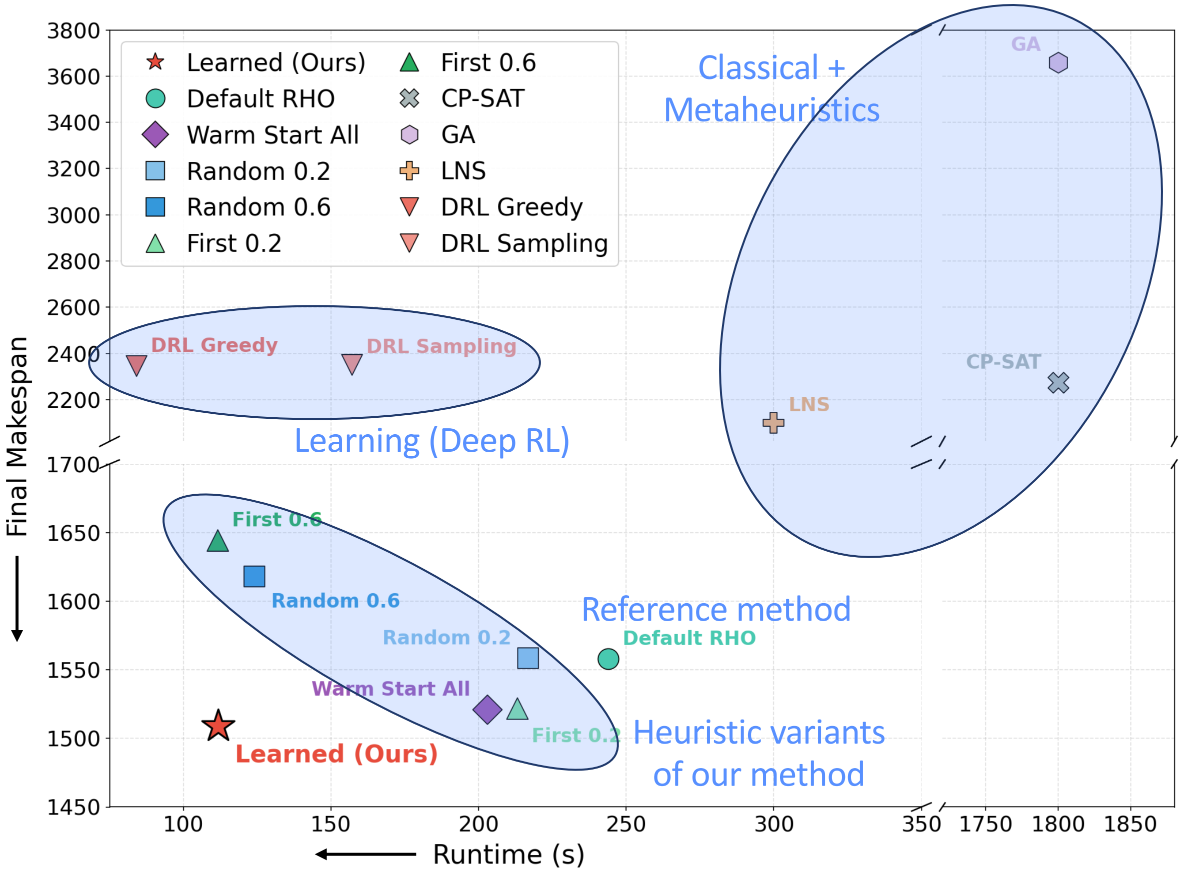
S. Li, W. Ouyang, Y. Ma, and **C. Wu**, "Learning-Guided Rolling Horizon Optimization for Long-Horizon Flexible Job Shop Scheduling," in ICLR, 2025.

L-RHO for Flexible Job Shop Scheduling



S. Li, W. Ouyang, Y. Ma, and **C. Wu**, "Learning-Guided Rolling Horizon Optimization for Long-Horizon Flexible Job Shop Scheduling," in ICLR, 2025.

L-RHO Performance Evaluation

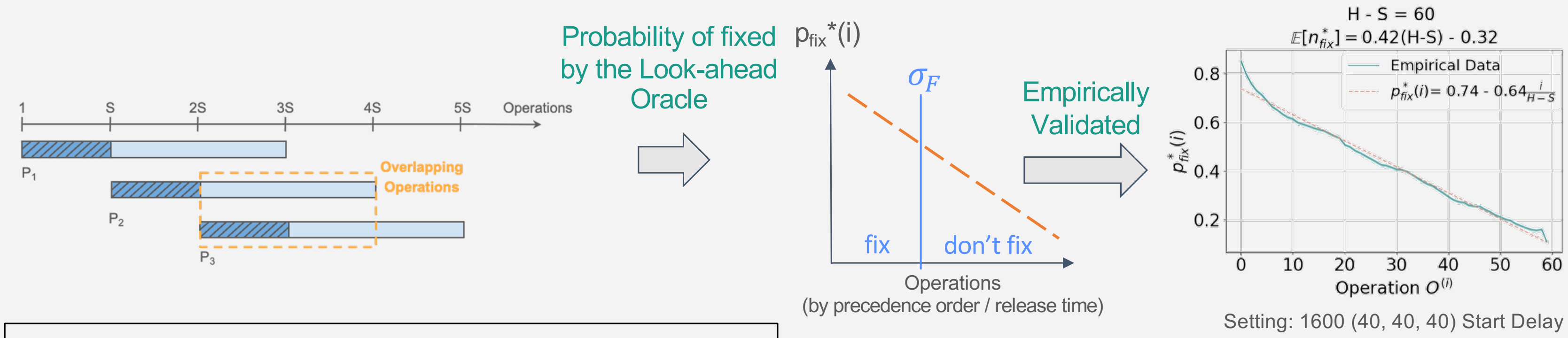


Result: Best trade-off in objective and speed

S. Li, W. Ouyang, Y. Ma, and **C. Wu**, “Learning-Guided Rolling Horizon Optimization for Long-Horizon Flexible Job Shop Scheduling,” in ICLR, 2025.

FP = False Positive; FN = False Negative; TNR = True Negative Rate

Pre-learning Diagnostics: Avoid “Try-and-See”



Proposition. Under a probabilistic model derived from FJSP data, the FP/FN rates for L-RHO, Random, and First can be given analytically in closed form.

→ **Training effort & cost vs Expected payoff**

Among the first results on **when to use** learning-guided optimization.

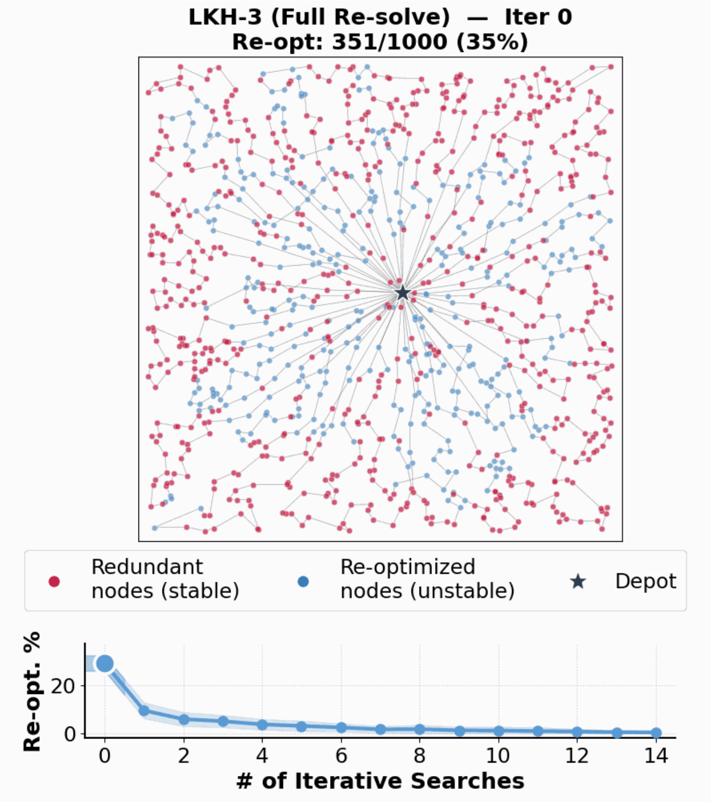
Gap Analysis for Flexible Job Shop Scheduling (FJSP)

Category	Recall	TNR	Speed	Obj
Oracle	100%	100%	+57%	+4.9%
Best heuristic (First 20%)	21%	83%	+13%	+2.3%
Learned (L-RHO)	77%	70%	+54%	+3.1%

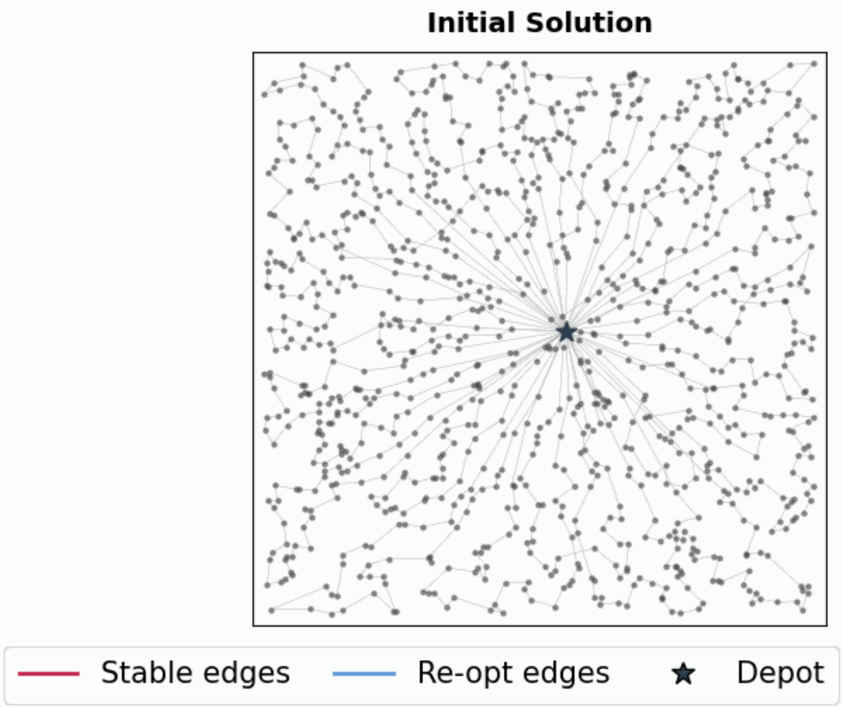
S. Li, W. Ouyang, Y. Ma, and **C. Wu**, “Learning-Guided Rolling Horizon Optimization for Long-Horizon Flexible Job Shop Scheduling,” in ICLR, 2025.

Learning to Reuse for Large Vehicle Routing

State-of-the-art Solver LKH (metaheuristic)



Learning-to-Segment (ours): Transformer-based prediction of stable edges



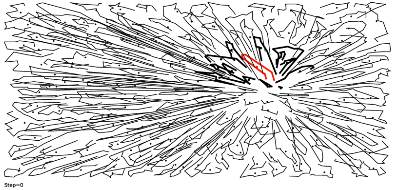
Results: 7x faster on CVRP with 2000 customers

W. Ouyang, S. Li, Y. Ma, and **C. Wu**, "Learning to Segment for Vehicle Routing Problems," International Conference on Learning Representations (ICLR), 2026. **Oral (1%).**

Learning-guided combinatorial optimization

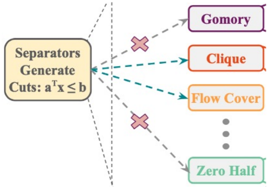
Contribution: Deep learning can accelerate combinatorial optimization solvers by 2-10x.

Learning to Focus and Prioritize Search

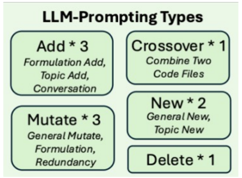


NeurIPS21 **Spotlight (<3%)** ICLR24 JAIR26

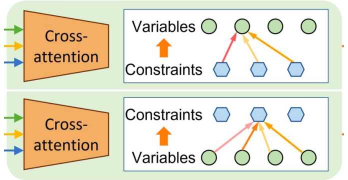
Learning for Mixed-Integer Linear Programming



NeurIPS23

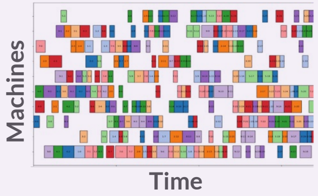


ICLR25

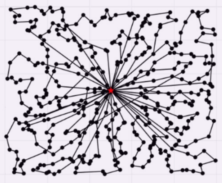


ICML26

Learning to Reuse Across Iterations

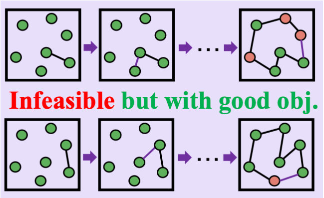


ICLR25



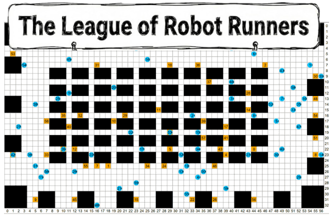
ICLR26 **Oral (1%)**

Learning to Repair Solutions

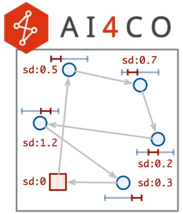


ICLR26

Benchmarks & Competitions



ICAPS24



KDD25 **Oral (9%)**

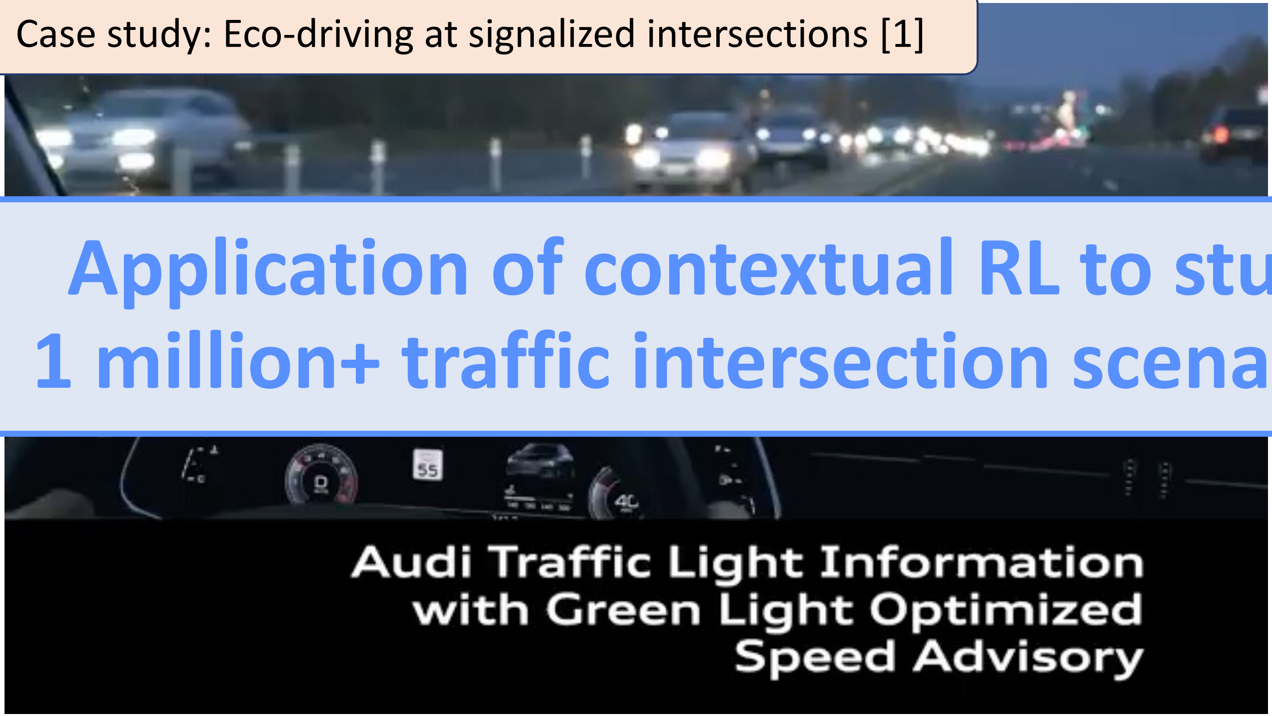
Modest data and compute (1,000 instances, 1 day on V100); **Fairly generalizable**.

Outline

1. Addressing Training Reliability in Contextual RL
2. Hybridizing Learning & Solvers in Combinatorial Optimization
3. **Applications: Informing Transportation Decisions**
   a. Policy Implications: Eco-driving Case Study
   b. Congestion Mitigation in Automated Warehousing

Application: Policy implications of automated vehicles

Case study: Eco-driving at signalized intersections [1]



Application of contextual RL to study 1 million+ traffic intersection scenarios

Video credits: Audi

Research Question: Would intelligent control of vehicles to reduce CO_2 emissions at intersections at the city scale move the needle on climate change mitigation goals?

[1] Mintsis et al. *Dynamic eco-driving near signalized intersections: Systematic review and future research directions*. Journal of Transportation Engineering, Part A, 2020
V. Jayawardana, *et al.*, **C. Wu**, "Mitigating Metropolitan Carbon Emissions with Dynamic Eco-driving at Scale," *Transportation Research Part C: Emerging Technologies*, 2025.

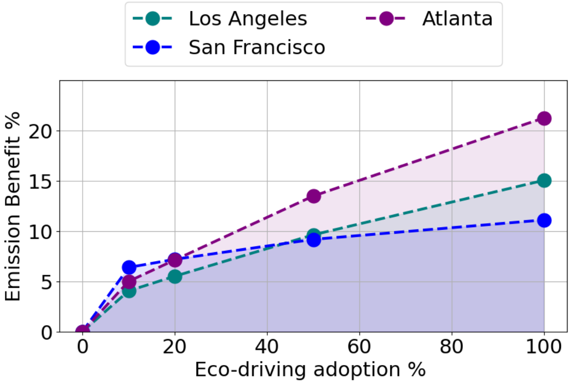
Findings and Policy Implications

Emissions impact:
11-22% of intersection CO_2 emissions

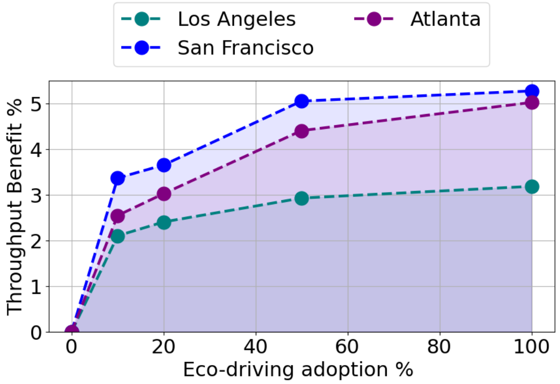
Near-term potential:
Majority of impact results from a minority of drivers and intersections

Long-term relevance:
Impact is robust through 2050 due to fleet impacts & travel demand

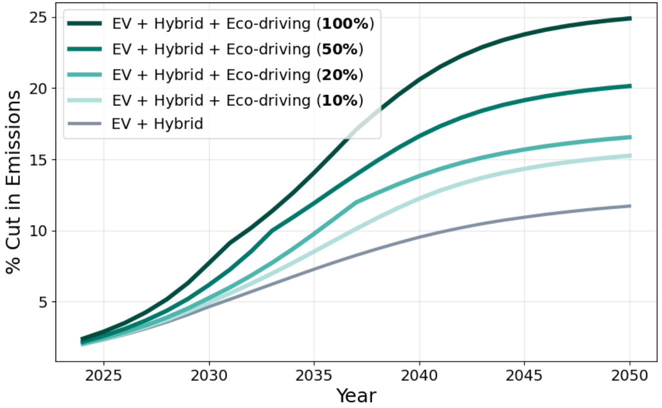
Emission reduction



Throughput increase



Projection through 2050 (LA)

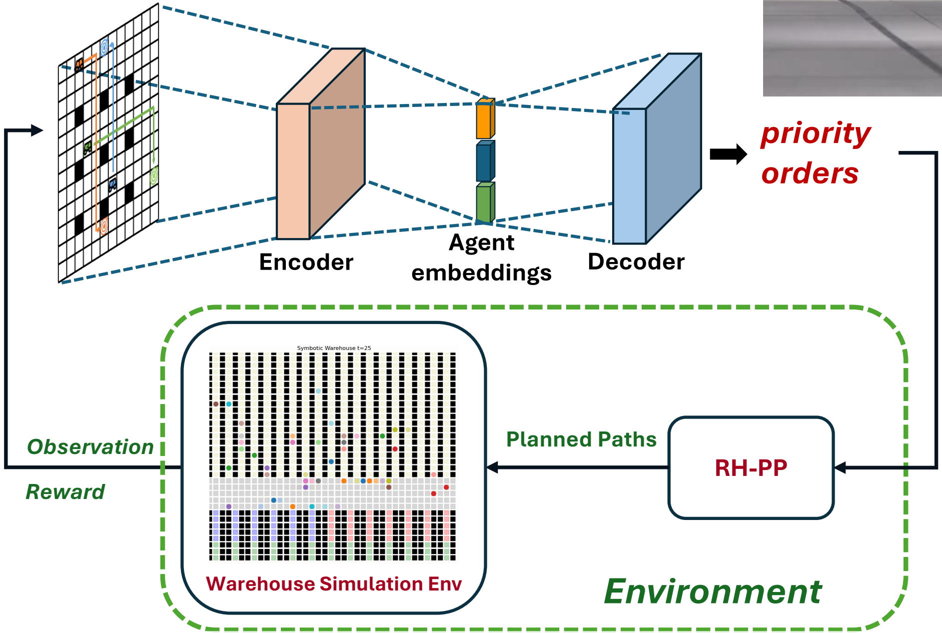


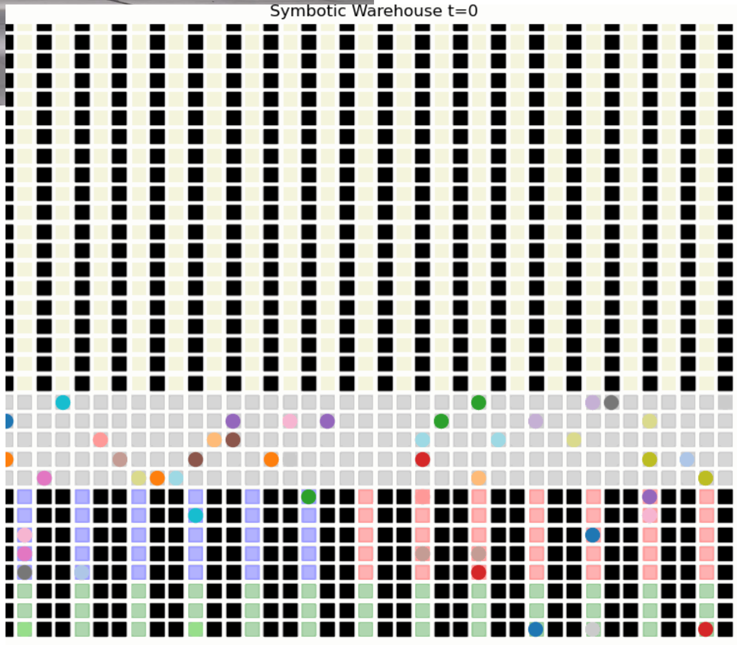
High-impact, relatively **low cost**, and **complements** existing climate interventions

V. Jayawardana, *et al.*, ***C. Wu***, "Mitigating Metropolitan Carbon Emissions with Dynamic Eco-driving at Scale," *Transportation Research Part C: Emerging Technologies*, 2025.

Congestion Mitigation in Automated Warehousing

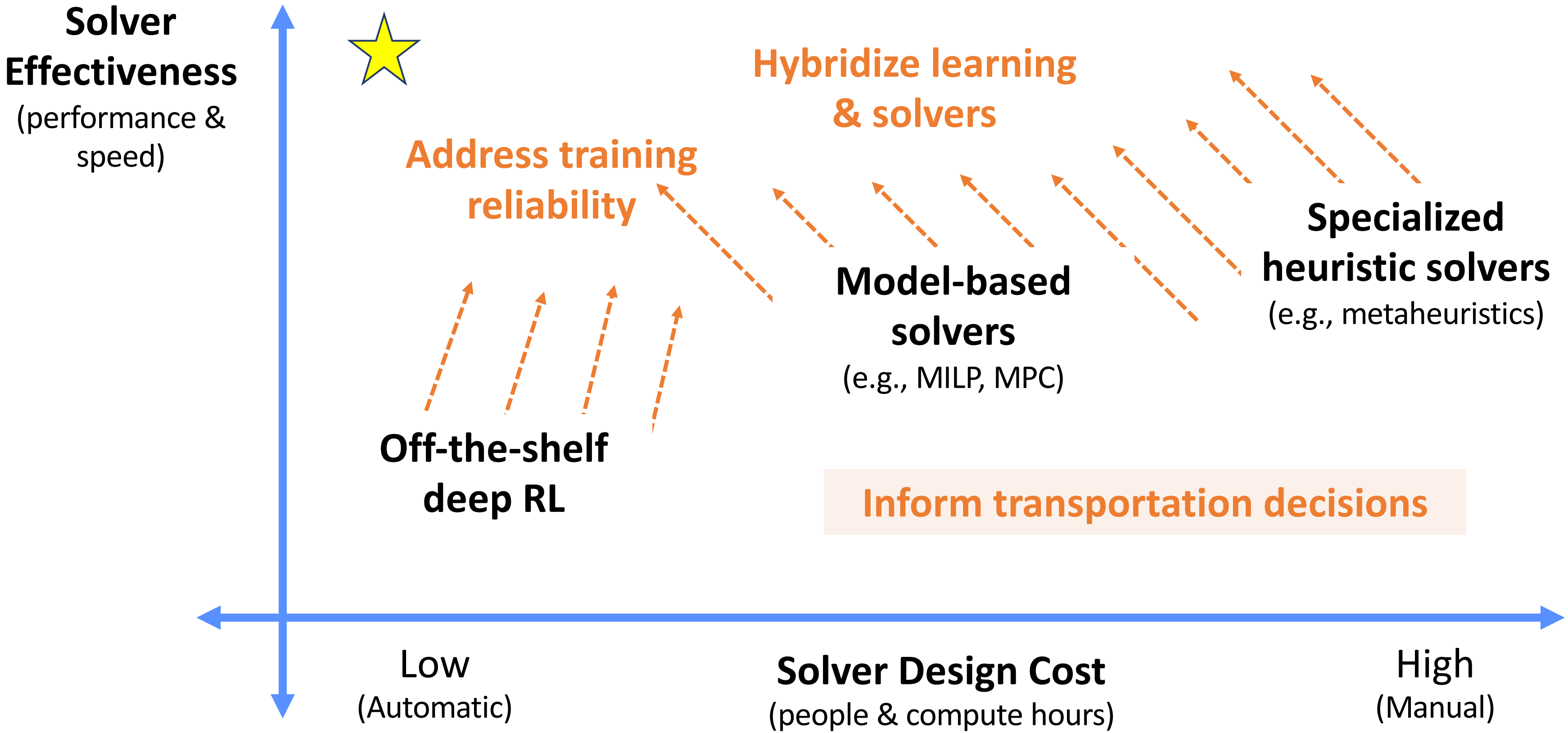
- Learn **priority order** of agents
- Method: Deep RL-guided prioritized planning
- **Results: +20% throughput on Symbotic-style warehouses**



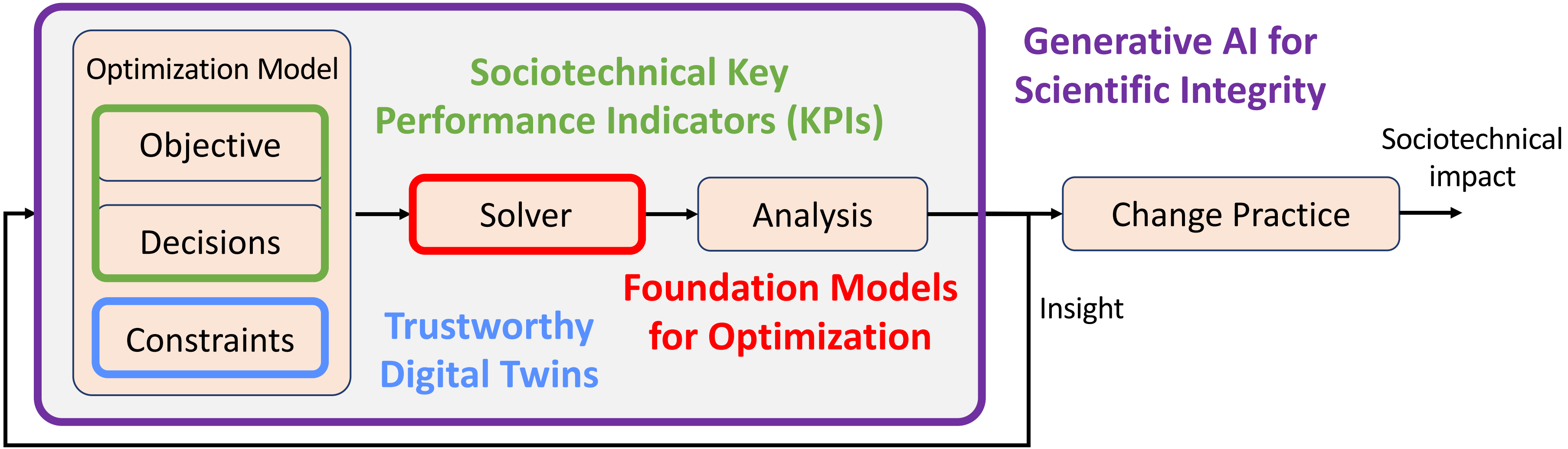


H. Zheng, Y. Ma, B. Araki, J. Chen, and **C. Wu**, "Learning-guided Prioritized Planning for Lifelong Multi-Agent Path Finding in Warehouse Automation," *Journal of Artificial Intelligence Research (JAIR)*, 2026.

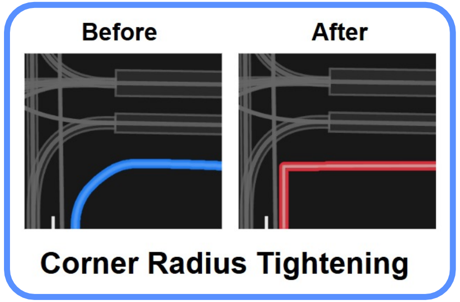
Optimization Solver Landscape



Future Research: Scaling Sociotechnical Decisions



[1]



[2]

5% of papers that use code share code

4% of papers that use data share data

[1] R. H. Lewis, C. Hickert, Y. Xie, E. Vinitsky, and C. **Wu**, "CrossRoads: A Robustness Benchmark for Driving Foundation Models on Roadway Counterfactuals". Under review.
[2] J. Ji* *et al.*, C. **Wu**, "Measuring the State of Open Science in Transportation Using Large Language Models," arXiv. doi: 10.48550/arXiv.2601.14429. Under review.

REproducible Research In Transportation Engineering (RERITE) Working Group

RERITE

Cathy Wu
Associate Professor, MIT
Chair, RERITE

Tutorial website

Challenge and opportunity

Advancing reproducibility is possibly the ***single most important action*** *each of us can take to improve the impact of transportation research.*

Do you want to join the action?

Join our Google group & take on your favorite RR task:

https://www.rerite.org/getinvolved.html

Join our LinkedIn group for the latest news:

https://www.linkedin.com/groups/10084451/

Research Support

amazon robotics

NSF

LIDS
LABORATORY FOR INFORMATION AND DECISION SYSTEMS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Civil and Environmental Engineering

IDSS
MIT INSTITUTE FOR DATA, SYSTEMS, AND SOCIETY

MathWorks

Microsoft

Cintra

U.S. Department of Transportation
Federal Highway Administration

symbotic

MIT Energy Initiative

amazon science

UDOT
Keeping Utah Moving

MIT Mobility Initiative

IBM Watson

MIT Research Support Committee

MIT Indonesia Seed Fund

MISTI
MIT Global Experiences

aws

MIT Quest for Intelligence