Combinatorial optimization augmented machine learning: Algorithms and applications in scheduling

Axel Parmentier

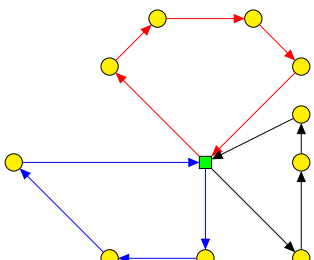
Cermics École Nationale des Ponts et Chaussées

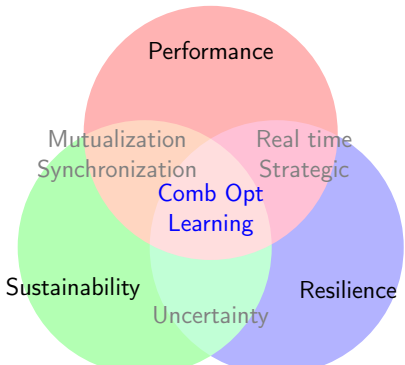
October 16th, 2025, Grand Séminaire MACS 2025

How to handle uncertainty in Operations Research (OR)

Operations Research

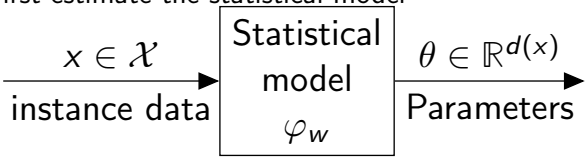
- allocation of resources
- optimization of industrial processes





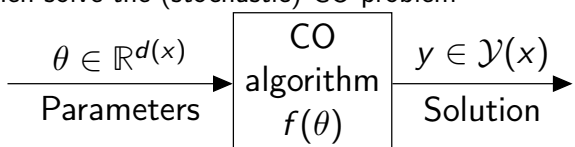
Leveraging data in operations research

First estimate the statistical model



Many OR problems use statistical models

Then solve the (stochastic) CO problem



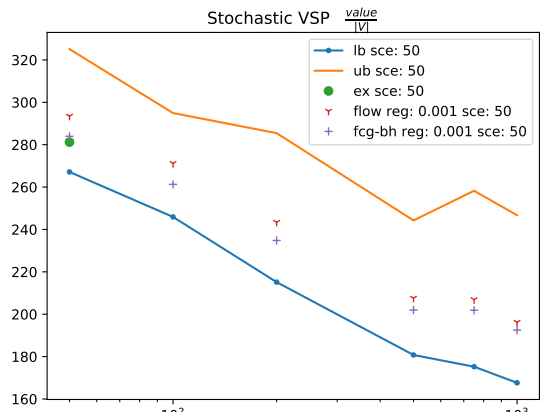
Researchers traditionally focus on the CO algorithm

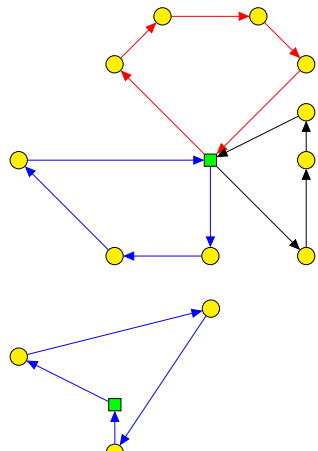
Learning problem:

Training set $(x_1, \bar{\theta}_1), \ldots, (x_n, \bar{\theta}_n)$
Loss $\mathcal{L}(\theta, \bar{\theta})$

$$\min_w \frac{1}{n} \sum_{i=1}^{n} \mathcal{L}(\varphi_w(x_i), \bar{\theta}_i)$$

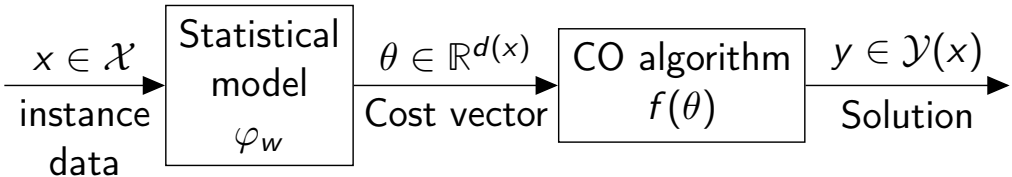
Main source of performance in OR ?





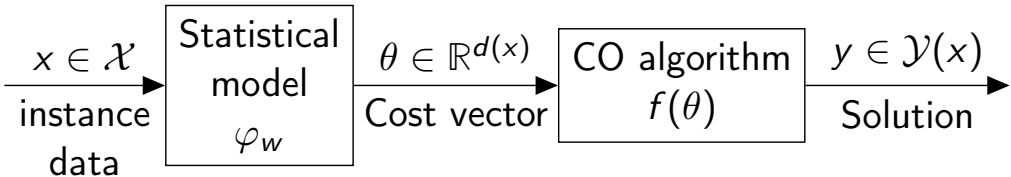
Augmenting the combinatorics reduces the marginal cost

Combinatorial Optimization Augmented Machine Learning

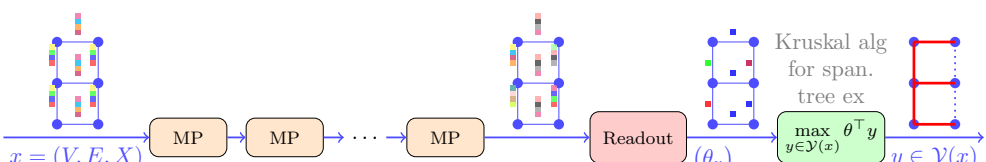


Trained by decision focused learning $\min_w \frac{1}{n} \sum_{i=1}^{n} \mathcal{L}(\varphi_w(x_i), \bar{y}_i)$.

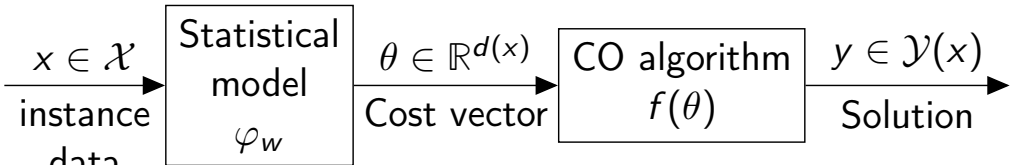
Combinatorial Optimization Augmented Machine Learning



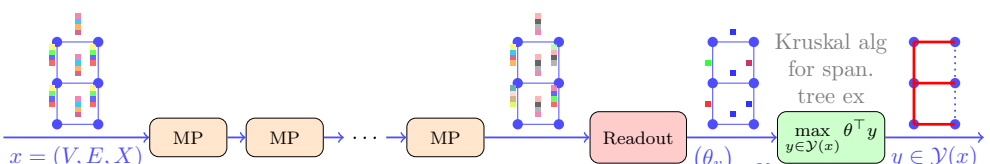
Trained by decision focused learning $\min_w \frac{1}{n} \sum_{i=1}^{n} \mathcal{L}(\varphi_w(x_i), \bar{y}_i)$.



Combinatorial Optimization Augmented Machine Learning



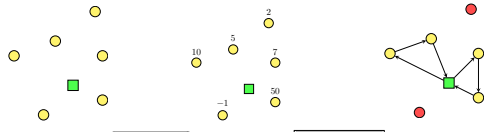
Trained by decision focused learning $\min_w \frac{1}{n} \sum_{i=1}^{n} \mathcal{L}(\varphi_w(x_i), \bar{y}_i)$.



Neural combinatorial optimization does decision focused learning without the CO layer

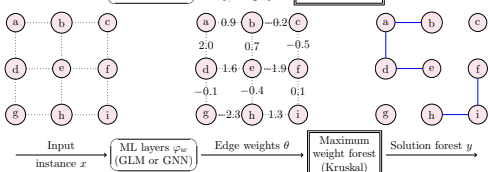
Settings and architectures

Multistage stochastic optimization



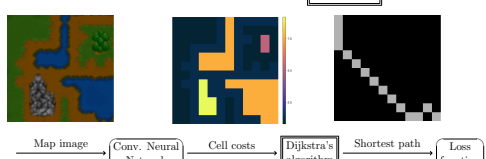
EURO NeurIPS challenge 2022. Baty et al. 2024; Greif et al. 2024

Contextual stochastic optimization



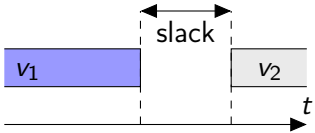
Donti, Amos, and Kolter 2017; Dalle et al. 2022

Data-driven optimization



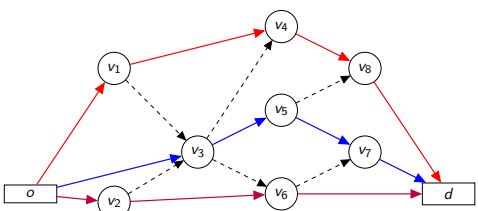
Pogančić et al. 2019; Berthet et al. 2020

Resilience: Stochastic Vehicle Scheduling Problem



$$c_p = \text{vehicle cost} + \mathbb{E}(\text{propagated delay cost})$$
$$= c^{\text{veh}} + \frac{1}{|\Omega|} \sum_{\omega \in \Omega} \sum_{v \in P} \xi_v^P(\omega)$$

Reduce costs dues to delay propagation along rotations



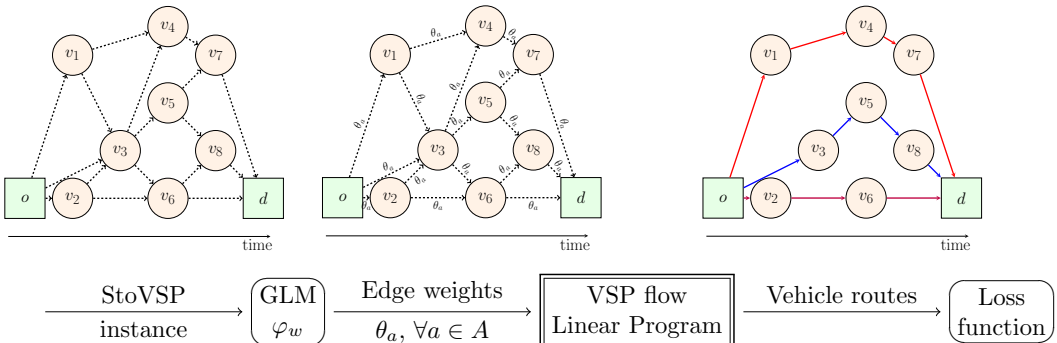
$$\min \sum_{P \in \mathcal{P}} c_P z_y$$
$$\sum_{P \ni v} y_P = 1 \quad \forall v$$
$$y_P \in \{0,1\}$$

Challenges

Even with simplest delay models

- No tractable moment formulation
- SAA does not scale (exact $|V| \leq 80$, heuristics $|V| \leq 400$)
- Cannot afford more than a single deterministic resolution

Decision aware learning for Contextual Stochastic VSP



Excellent performance on large scale instances[1]

Enables being contextual

[1] A. P. (Apr. 2021). "Learning to Approximate Industrial Problems by Operations Research Classic Problems". In: *Operations Research*; Guillaume Dalle et al. (July 2022). *Learning with Combinatorial Optimization Layers: A Probabilistic Approach*. eprint: 2207.13513.

Contextual stochastic combinatorial optimization

Contextual stochastic optimization problem[2]

$$\min_{\pi \in \mathcal{H}} \mathcal{R}(\pi) \quad \text{where} \quad \mathcal{R}(\pi) = \mathbb{E}_{(\boldsymbol{x}, \xi),\, \boldsymbol{y} \sim \pi(\cdot|\boldsymbol{x})} \left[c\left(\boldsymbol{x}, \boldsymbol{y}, \xi \right) \right]$$

noise correlated with $\boldsymbol{x}$

context in $\mathcal{X}$

decision in $\mathcal{Y}(x)$

Assumptions:

- we have an efficient algorithm to solve

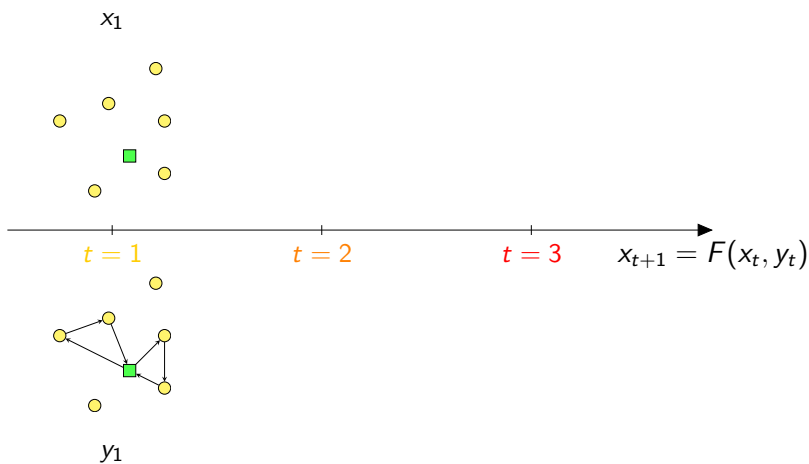
$$\min_{y \in \mathcal{Y}(x)} c\big(x(\omega), y, \xi(\omega)\big) + \langle \theta | y \rangle$$

- $\mathcal{Y}(x)$ is finite (but exponentially large)
- we have access to a dataset $\mathcal{D} = (x_i, \xi_i)_{i \in [N]}$

[2] Utsav Sadana et al. (Mar. 2024). "A Survey of Contextual Optimization Methods for Decision-Making under Uncertainty". In: *European Journal of Operational Research*. issn: 0377-2217. doi: 10.1016/j.ejor.2024.03.020. (Visited on 07/12/2024).

Dynamic Vehicle Routing Problem

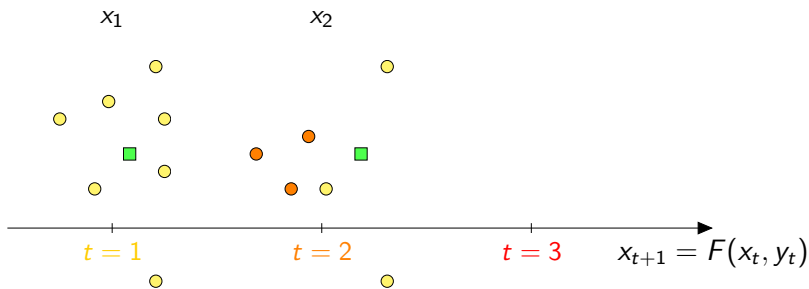
State
$x_t \in \mathcal{X}$
set of customers



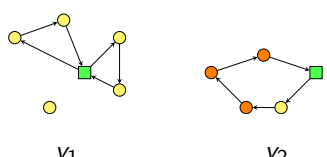
Decision
$y_t \in \mathcal{Y}(x_t)$
set of routes

Dynamic Vehicle Routing Problem

State
$x_t \in \mathcal{X}$
set of customers

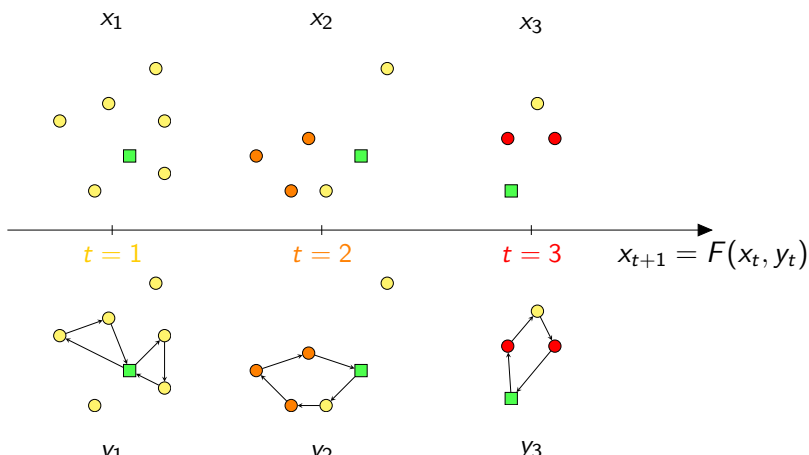


Decision
$y_t \in \mathcal{Y}(x_t)$
set of routes



Dynamic Vehicle Routing Problem

State
$x_t \in \mathcal{X}$
set of customers



Decision
$y_t \in \mathcal{Y}(x_t)$
set of routes

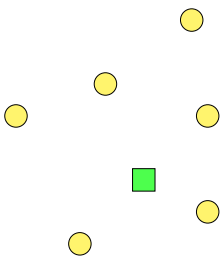
A solution of this problem is a **policy**:

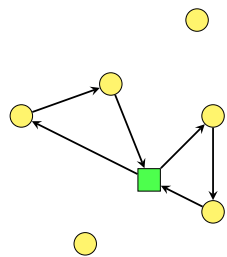
$$\pi : \mathcal{X} \to \mathcal{Y}$$
$$x_t \mapsto y_t$$

Objective: find $\pi^\star$, serving all customers before end of horizon, and minimizing total cost

$$\pi^\star = \arg\min_{\pi} \mathbb{E}\left[\sum_{\text{epochs } t} \text{total cost of routes in decision } y_t = \pi(x_t)\right]$$

Policy that won the EURO-NeurIPS challenge[3]





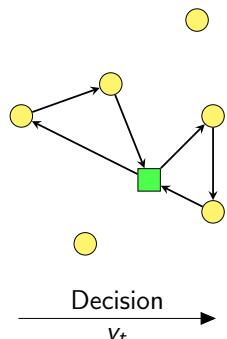
[3] Léo Baty et al. (Feb. 2024). "Combinatorial Optimization-Enriched Machine Learning to Solve the Dynamic Vehicle Routing Problem with Time Windows". In: *Transportation Science*. issn: 0041-1655. doi: 10.1287/trsc.2023.0107. (Visited on 07/18/2024).

Policy that won the EURO-NeurIPS challenge[3]

Epoch decisions can be seen as the solution of a Prize Collecting VRPTW:

- Serving customers is optional
- Serving customer v gives prize θ_v
- **Objective**: max profit = prize − route*cost*

$$\max_{y \in \mathcal{Y}(x_t)} \underbrace{\sum_{(u,v) \in x_t^2} \theta_v y_{u,v}}_{\text{total prize}} - \underbrace{\sum_{(u,v) \in x_t^2} c_{u,v} y_{u,v}}_{\text{total routes cost}} .$$



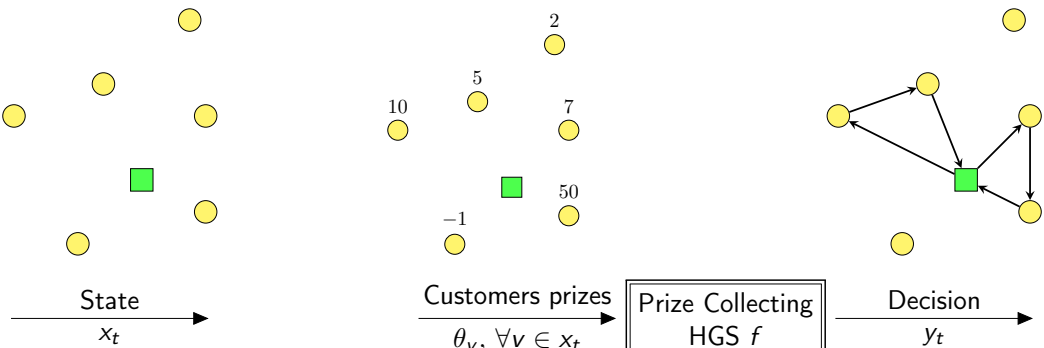
- **Algorithm**: Prize Collecting Hybrid Genetic Search

$\Rightarrow$ Combinatorial Optimization layer f

[3] Léo Baty et al. (Feb. 2024). "Combinatorial Optimization-Enriched Machine Learning to Solve the Dynamic Vehicle Routing Problem with Time Windows". In: *Transportation Science*. issn: 0041-1655. doi: 10.1287/trsc.2023.0107. (Visited on 07/18/2024).

Policy that won the EURO-NeurIPS challenge[3]

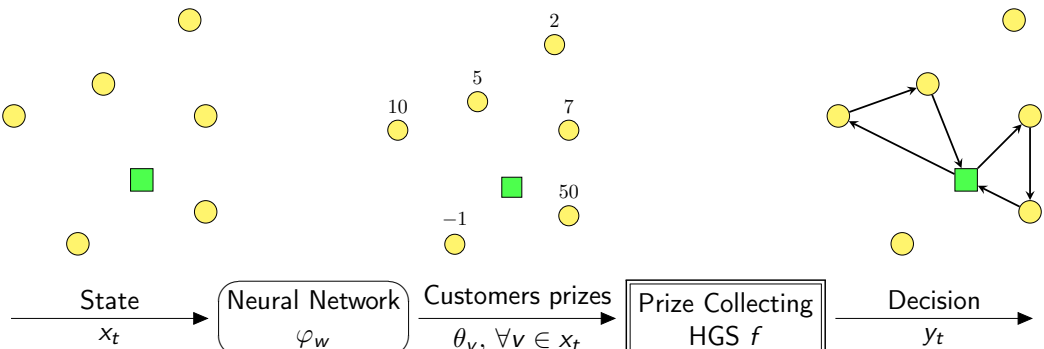
Difficulty: no natural way of computing meaningful prizes



[3] Léo Baty et al. (Feb. 2024). "Combinatorial Optimization-Enriched Machine Learning to Solve the Dynamic Vehicle Routing Problem with Time Windows". In: *Transportation Science*. issn: 0041-1655. doi: 10.1287/trsc.2023.0107. (Visited on 07/18/2024).

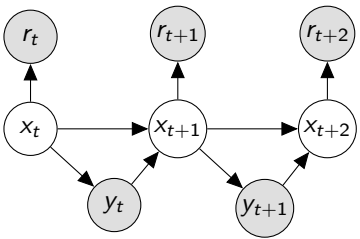
Policy that won the EURO-NeurIPS challenge[3]

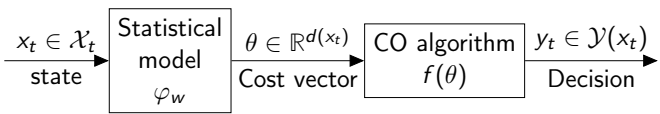
Solution: use a neural network to predict request prizes $\theta = \varphi_w(x_t)$



[3] Léo Baty et al. (Feb. 2024). "Combinatorial Optimization-Enriched Machine Learning to Solve the Dynamic Vehicle Routing Problem with Time Windows". In: *Transportation Science*. issn: 0041-1655. doi: 10.1287/trsc.2023.0107. (Visited on 07/18/2024).

Policy for multistage stochastic optimization





Neural network with a CO layer: policy for MDPs with large state *and decision* spaces.

$$\min_w \mathbb{E}_\pi \sum_t r_t \qquad \text{with} \qquad \pi_{w,t} : \mathcal{X}_t \to \mathcal{Y}_t$$

Policy encoded by neural networks with CO layers

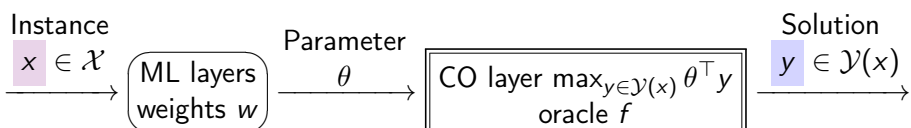
Goal: find a policy π that minimizes

$$\min_{\pi \in \mathcal{H}} \mathbb{E}_{\boldsymbol{x} \sim \mathbb{P}_{\boldsymbol{x}},\, \boldsymbol{y} \sim \pi(\cdot|\boldsymbol{x})} \left[c^0 \left(\boldsymbol{x} ; \boldsymbol{y} \right) \right]$$

cost function; instance in $\mathcal{X}$; decision in $\mathcal{Y}(x)$

$\mathbb{P}_{\boldsymbol{x}}$ unknown but access to $x_1, \ldots, x_n$.

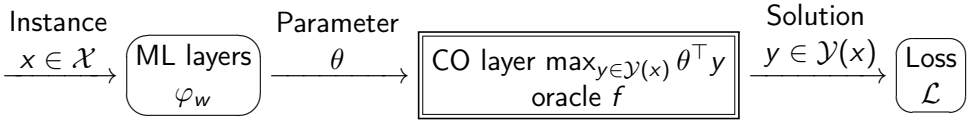
Model choice: we restrict ourselves to policies π_w based on



We thus seek weights w that minimize the risk

$$\min_{w} \mathbb{E}_{\boldsymbol{x} \sim \mathbb{P}_{\boldsymbol{x}},\, \boldsymbol{y} \sim \pi_w(\cdot|\boldsymbol{x})} \left[c^0 \left(\boldsymbol{x} ; \boldsymbol{y} \right) \right]$$

End-to-end learning: two paradigms



Empirical risk minimization

Dataset: $\mathcal{D} = (x_i)_{i \in [N]}$
Learning problem:

$$\min_w \frac{1}{N} \sum_{i=1}^{N} c^0\Big(x_i; f\big(\varphi_w(x_i)\big)\Big)$$

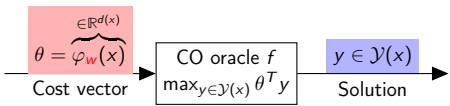
Supervised learning

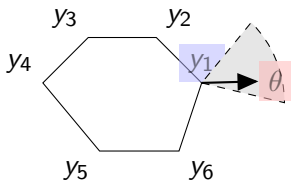
Dataset: $\mathcal{D} = (x_i, \bar{y}_i)_{i \in [N]}$
Learning problem:

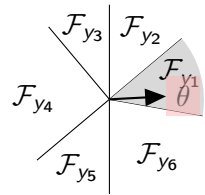
$$\min_w \frac{1}{N} \sum_{i=1}^{N} \mathcal{L}\Big(x_i; f\big(\varphi_w(x_i)\big), \bar{y}_i\Big)$$

$\rightarrow$ We would like both to rely on stochastic gradient descent (SGD)

Lack of informative derivatives



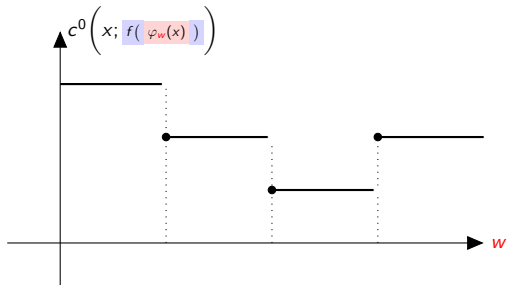




Piecewise-constant learning problem

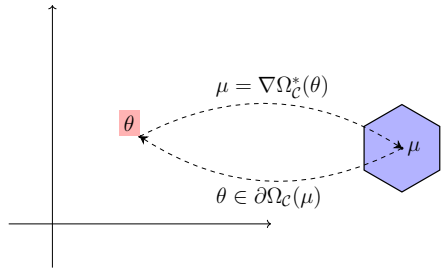
$$\frac{1}{N}\sum_{i=1}^{N} c^0\Big(x_i;\ f\big(\ \varphi_w(x_i)\ \big)\Big)$$

$$\frac{1}{N}\sum_{i=1}^{N} \mathcal{L}\Big(x_i;\ f\big(\ \varphi_w(x_i)\ \big), \bar{y}_i\Big)$$



Smoothing by regularization or perturbation

$$\max_{\mu \in \mathcal{C}(x)} \theta^T \mu - \Omega(\mu), \quad \mathcal{C}(x) = \text{conv}\left(\mathcal{Y}(x)\right)$$



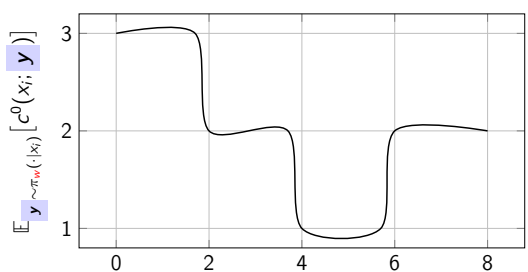
Ex. 1: $\Omega(\mu) = ||\mu||_2^2 + \mathbb{I}_{\mathcal{C}(x)}(\mu)$

Ex. 2: $\Omega^*(\theta) = \mathbb{E}_{\boldsymbol{Z}}[\max_{\mu \in \mathcal{C}(x)}(\theta + \varepsilon \boldsymbol{Z})^\top \mu]$

Smoothed learning problem

$$\frac{1}{N} \sum_{i=1}^{N} \mathbb{E}_{\boldsymbol{y} \sim \pi_w(\cdot|x_i)} \left[c^0(x_i; \boldsymbol{y}) \right]$$

$$\frac{1}{N} \sum_{i=1}^{N} \mathbb{E}_{\boldsymbol{y} \sim \pi_w(\cdot|x_i)} \left[\mathcal{L}(x_i; \boldsymbol{y}, \bar{y}_i) \right]$$



Supervised learning: Fenchel-Young losses[4]

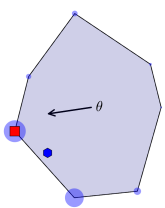
Non-optimality of $\bar{y}$ as a solution of the regularized prediction problem

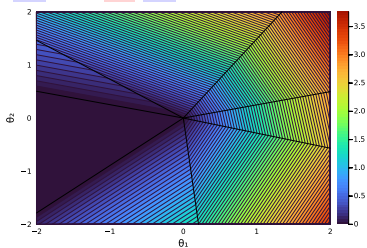
$$\mathcal{L}_\Omega(\theta ; \bar{y}) = \max_{y \in \mathcal{C}(x)} \big(\langle \theta \,|y\rangle - \Omega(y)\big) - \big(\langle \theta \,|\, \bar{y} \,\rangle - \Omega(\bar{y})\big)$$

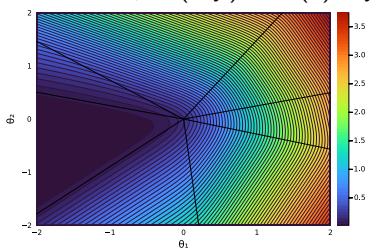
$$\mathcal{L}_\Omega(\theta ; \bar{y}) = \Omega^*(\theta) + \Omega(\bar{y}) - \langle \theta \,|\, \bar{y} \,\rangle$$

Properties that make SGD tractable

- $\mathcal{L}_\Omega(\theta; \bar{y}) \geq 0$
- $\mathcal{L}_\Omega(\theta; \bar{y}) = 0 \Leftrightarrow \bar{y} = \nabla\Omega^*(\theta)$
- Convex in θ
- $\nabla_\theta \mathcal{L}_\Omega(\theta; \bar{y}) = \widehat{f}_\Omega(\theta) - \bar{y}$

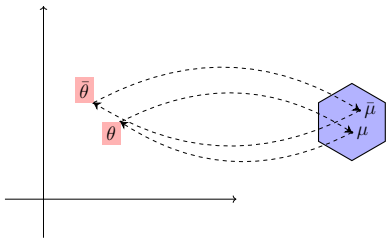


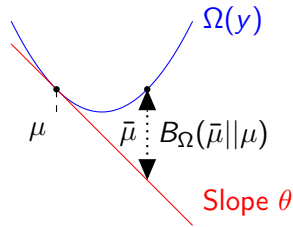




[4] Blondel, Martins, and Niculae 2020.

Fenchel-Young loss as a primal-dual Bregman divergence





$$B_\Omega(\bar{\mu}||\mu) = \Omega(\bar{\mu}) - \Omega(\mu) - \langle \nabla\Omega(\mu)|\bar{\mu} - \mu\rangle \quad \text{and} \quad B_\Omega(\bar{\mu}||\mu) = \mathcal{L}_\Omega(\theta;\bar{\mu}) = B_{\Omega^*}(\theta||\bar{\theta})$$

$$\min_{\mu\in\mathcal{C}} \frac{1}{N}\sum_{i=1}^{N} B_\Omega(\bar{\mu}_i||\mu) \Leftrightarrow \min_{\theta\in\mathbb{R}^d} \frac{1}{N}\sum_{i=1}^{N} \mathcal{L}_\Omega(\theta;\bar{\mu}_i)$$

Solves supervised learning? Research gap remains for empirical risk minimization
Why ERM? Take the cost function into account, get guarantees (Aubin-Frankowski et al. 2024), but hard in practice

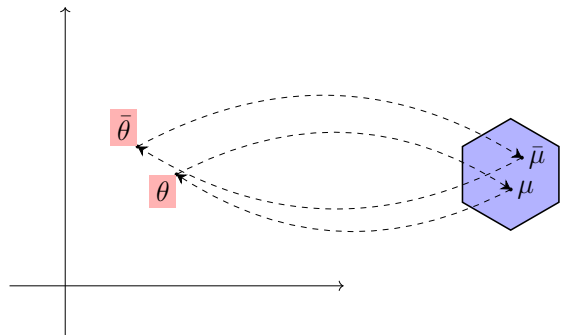
Regularization by perturbation

$$F_{\varepsilon,\mathcal{C}}(\theta) = \mathbb{E}[\max_{y\in\mathcal{Y}}(\theta + \varepsilon \boldsymbol{Z})^\top y]$$
$$= \mathbb{E}[\max_{y\in\mathcal{C}}(\theta + \varepsilon \boldsymbol{Z})^\top y]$$

$\rightarrow \Omega_{\varepsilon,\mathcal{C}} = F^*_{\varepsilon,\mathcal{C}}$

Proposition Berthet et al. 2020

- defined over $\mathbb{R}^d$
- strict convexity
- $\nabla_\theta F_{\varepsilon,\mathcal{C}}(\theta) = \mathbb{E}[\arg\max_{y\in\mathcal{C}}(\theta + \varepsilon \boldsymbol{Z})^\top y]$
- $\mathrm{dom}(F^*_{\varepsilon,\mathcal{C}}) = \mathcal{C}$
- $F^*_{\varepsilon,\mathcal{C}}$ Legendre-type



A distributional point of view:

$$\mathbb{P}(y|\theta) = \mathbb{P}_Z\Big(y = \arg\max_{y\in\mathcal{Y}}(\theta + Z)^\top y\Big)$$

Structured Reinforcement Learning

Reinforcement learning setting

$$\min_{\pi \in \mathcal{H}} \mathcal{R}(\pi) \quad \text{where} \quad \mathcal{R}(\pi) = \mathbb{E}_{(\boldsymbol{x},\xi),\, \boldsymbol{y}\sim\pi(\cdot|\boldsymbol{x})}\left[c(\boldsymbol{x},\boldsymbol{y},\xi)\right]$$

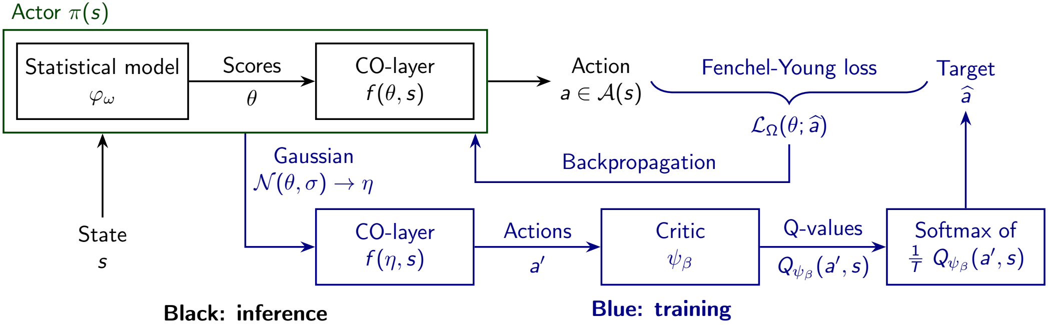
noise (not observed) → ξ

context in $\mathcal{X}$ → $\boldsymbol{x}$

decision in $\mathcal{Y}(x)$ → $\boldsymbol{y}$

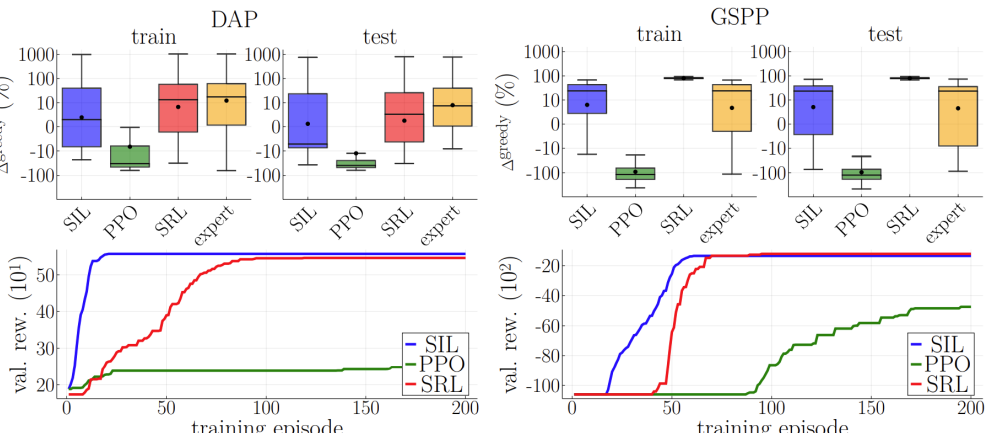
Access to evaluation oracle for $c(x, y, \xi)$.

No optimization oracle for $\min_{y\in\mathcal{Y}(x)} c(x, y, \xi)$ or $\min_{y\in\mathcal{Y}(x)} \mathbb{E}_\xi\left[c(x, y, \xi)\right]$



Evaluation oracle as critic ψ_β for single step problems.

Numerical results



Approximate version of alternating minimization scheme

Structure RL is an approximation of alternating minimization algorithm

$$\mu_i^{(t+1)} = \mathbb{E}_{\boldsymbol{Z}}\Big[\operatorname*{arg\,min}_{y_i \in \mathcal{Y}_i(x_i)} c(x_i, y_i, \xi_i) - \kappa(\varphi_{\bar{w}^{(t)}}(x_i) + \varepsilon \boldsymbol{Z})^\top y_i\Big] \qquad \text{(decomposition)}$$

$$\bar{w}^{(t+1)} \in \operatorname*{arg\,min}_{w \in \mathcal{W}} \frac{1}{N} \sum_{i=1}^{N} \mathcal{L}_\Omega\Big(\varphi_w(x_i); \mu_i^{(t+1)}\Big) \qquad \text{(coordination)}$$

for the surrogate problem, which is a relaxation of single step reinforcement learning problem

$$\min_{w, q_\otimes} \mathcal{S}_N(s_w; q_\otimes) := \min_{w, q_\otimes} \frac{1}{N} \sum_{i=1}^{N} \mathbb{E}_{\boldsymbol{y} \sim q_i}\Big[c(x_i, \boldsymbol{y}, \xi_i)\Big] + \kappa \mathcal{L}_{\Omega_{\Delta(x_i)}}\Big(Y(x_i)^\top \varphi_w(x_i); q_i\Big)$$

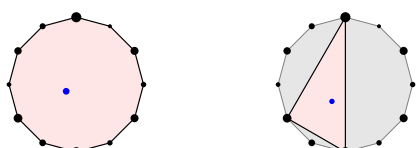
which requires optimization oracle for $\min_{y \in \mathcal{Y}(x)} c(x, y, \xi)$ or $\min_{y \in \mathcal{Y}(x)} \mathbb{E}_\xi\big[c(x, y, \xi)\big]$.

Sample to avoir optimization oracle[5]

Replace $\mathcal{Y}(x_i)$ by $\hat{\mathcal{Y}}_k^{(t)}(x_i)$: k atoms sampled from $p(\boldsymbol{y}|x_i, w^{(t)})$

$$\mu_i^{(t+1)} = \mathbb{E}_{\boldsymbol{Z}}\Big[\underbrace{\operatorname*{arg\,min}_{y_i \in \hat{\mathcal{Y}}_k^{(t)}(x_i)}}\ c(x_i, y_i, \xi_i) - \kappa(\varphi_{\bar{W}^{(t)}}(x_i) + \varepsilon \boldsymbol{Z})^\top y_i\Big] \qquad \text{(decomposition)}$$

Tractable by evaluation of the k elements of $\hat{\mathcal{Y}}_k^{(t)}(x_i)$



[5] Heiko Hoppe et al. (2025). *Structured Reinforcement Learning for Combinatorial Decision-Making*. arXiv: 2505.19053 [cs.LG]. url: https://arxiv.org/abs/2505.19053.

Convergence results

Given some technical assumptions / settings restrictions

Theorem *Convergence to surrogate optimum* Bouvier, Prunet, Leclère, P., 2025

The (average) iterates $q_{\otimes}^{(t)}$ concide with those of mirror descent, and, provided some technical assumptions, converge to $\min_{q_\otimes} \min_{s_\otimes} \mathcal{S}_N(s_\otimes, q_\otimes)$

Proposition *Empirical risk bound*, Bouvier, Prunet, Leclère, P., 2025

$$\theta_{\mathcal{S},N} \in \arg\min_{\theta} \min_{q_\otimes} \mathcal{S}_N(s_\theta, q_\otimes) \quad \Longrightarrow \quad \mathcal{R}_N(\theta_{\mathcal{S},N}) - \min_{\theta} \mathcal{R}_N(\theta) \leq \dots$$

Theorem *Generalization bounds*, Aubin-Frankowski, De Castro, P., Rudi, 2024

In the large data regime, $\mathcal{R}(\theta_{\mathcal{S},N}) - \min_\theta \mathcal{R}(\theta) \leq \dots$

Conclusion

Neural network with combinatorial optimization layers improve state of the art

- Contextual Stochastic Optimization (tactic, strategic)
- Dynamic problems (operations)

in combinatorial settings.

Structured Reinforcement learning

- Deep learning compatible
- No need for experts to imitate
- Leads to practically better policies
- Convergence to minimum of empirical risk minimization problem
- Generalization guarantees (approximation ratio in probability)

`https://github.com/JuliaDecisionFocusedLearning`

Combinatorial, convex, stochastic optimization, statistical learning.

References I

- Aubin-Frankowski, Pierre-Cyril et al. (July 2024). *Generalization Bounds of Surrogate Policies for Combinatorial Optimization Problems*. doi: 10.48550/arXiv.2407.17200. arXiv: 2407.17200 [stat]. (Visited on 12/10/2024).
- Baty, Léo et al. (Feb. 2024). "Combinatorial Optimization-Enriched Machine Learning to Solve the Dynamic Vehicle Routing Problem with Time Windows". In: *Transportation Science*. issn: 0041-1655. doi: 10.1287/trsc.2023.0107. (Visited on 07/18/2024).
- Berthet, Quentin et al. (2020). "Learning with differentiable pertubed optimizers". In: *Advances in neural information processing systems* 33, pp. 9508–9519.
- Blondel, Mathieu, André F. T. Martins, and Vlad Niculae (2020). "Learning with Fenchel-Young Losses". In: *Journal of Machine Learning Research* 21.35, pp. 1–69.
- Dalle, Guillaume et al. (July 2022). *Learning with Combinatorial Optimization Layers: A Probabilistic Approach*. eprint: 2207.13513.

References II

- Donti, Priya, Brandon Amos, and J. Zico Kolter (2017). "Task-Based End-to-end Model Learning in Stochastic Optimization". In: *Advances in Neural Information Processing Systems*. Vol. 30. Curran Associates, Inc. (Visited on 07/24/2024).
- Greif, Toni et al. (Feb. 2024). *Combinatorial Optimization and Machine Learning for Dynamic Inventory Routing*. arXiv: 2402.04463 [math]. (Visited on 03/04/2024).
- Hoppe, Heiko et al. (2025). *Structured Reinforcement Learning for Combinatorial Decision-Making*. arXiv: 2505.19053 [cs.LG]. url: https://arxiv.org/abs/2505.19053.
- P., A. (Apr. 2021). "Learning to Approximate Industrial Problems by Operations Research Classic Problems". In: *Operations Research*.
- Pogančić, Marin Vlastelica et al. (Sept. 2019). "Differentiation of Blackbox Combinatorial Solvers". In: *International Conference on Learning Representations*. (Visited on 07/22/2024).

References III

Sadana, Utsav et al. (Mar. 2024). "A Survey of Contextual Optimization Methods for Decision-Making under Uncertainty". In: *European Journal of Operational Research*. issn: 0377-2217. doi: 10.1016/j.ejor.2024.03.020. (Visited on 07/12/2024).