

Importance-aware Co-teaching for Offline Model-based Optimization

Ye Yuan ^{*1}, Can (Sam) Chen ^{*†1,2},
Zixuan Liu ³, Willie Neiswanger ⁴, Xue Liu ¹



¹
McGill
UNIVERSITY



²
Mila

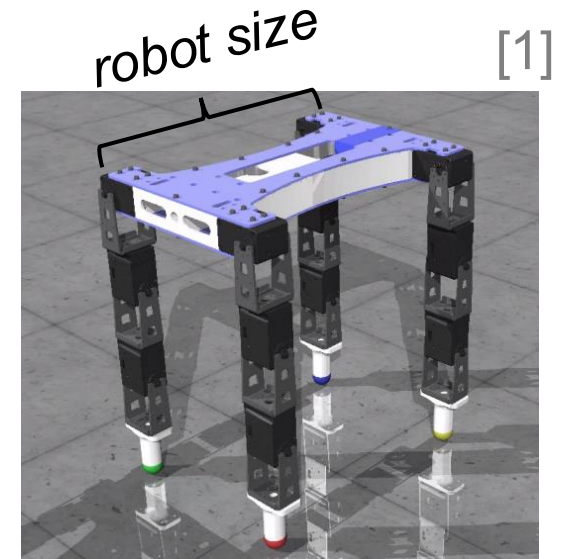
³
W
UNIVERSITY of
WASHINGTON



⁴
Stanford
University

Problem Background

- Design objects with specific desired properties.
 - For example: Design a new robot to run faster.
- Evaluation can be expensive, so assume access only to an offline dataset of designs and their property scores.
 - For example: some pairs of robot size and running speed.
- **Offline Model-based Optimization (MBO):** find a design (robot size) to maximize its property (speed) with the offline dataset only.



Problem Formulation

$$\arg \max_{\mathcal{A}} [\mathbb{P}(\{f(\mathbf{x}^*) : \mathbf{x}^* \in \mathcal{A}(\mathcal{D}, K)\}, n)].$$

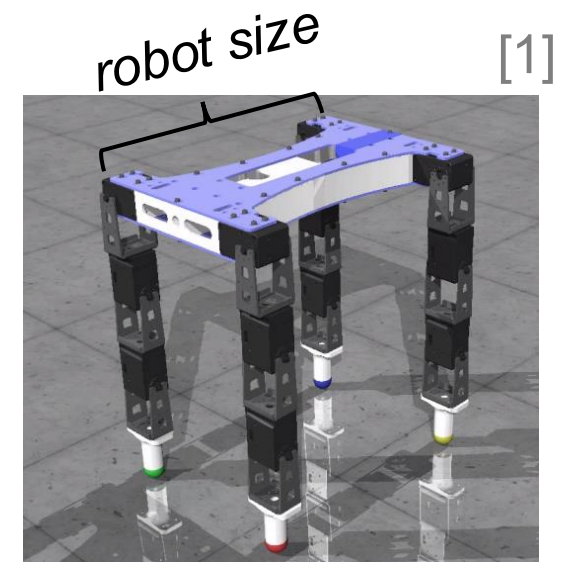
\mathbf{x}_i : some design (robot size);

$y_i = f(\mathbf{x}_i)$: some property (robot speed);

$\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^N$: an offline dataset;

\mathcal{A} : some algorithm outputs K candidates

$\mathbb{P}(S, n)$: the n^{th} percentile of S .



Related Work

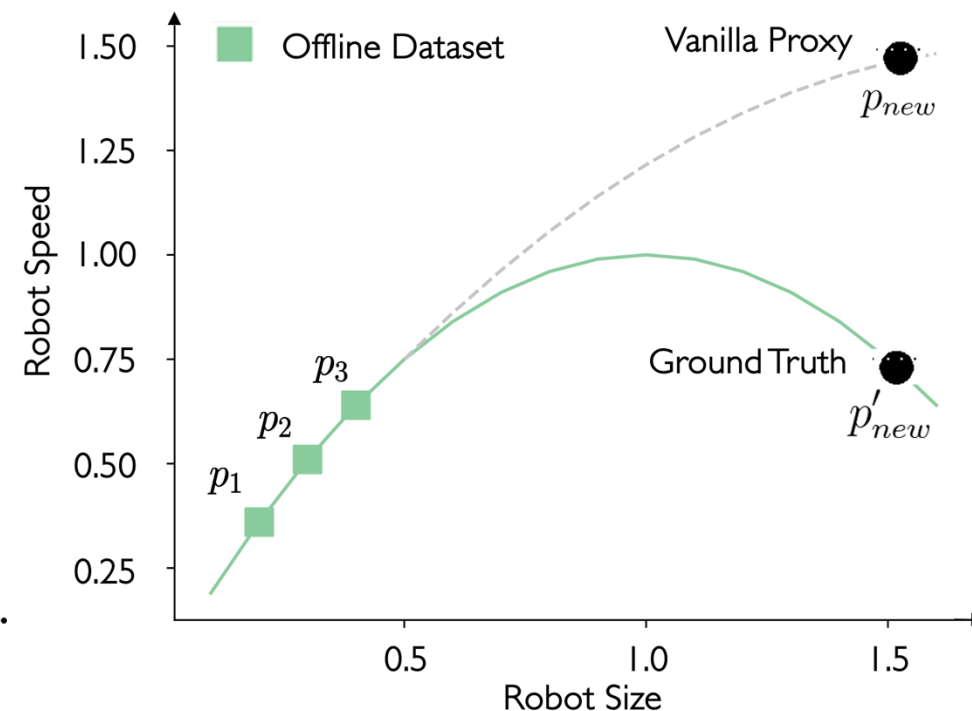
A common approach consists of 2 steps:

- 1) Fit a DNN proxy $f_{\theta}(\cdot)$ to \mathcal{D} .

$$\theta^* = \arg \min_{\theta} \frac{1}{N} \sum_{i=1}^N (f_{\theta}(\mathbf{x}_i) - y_i)^2.$$

- 2) Perform gradient ascent:

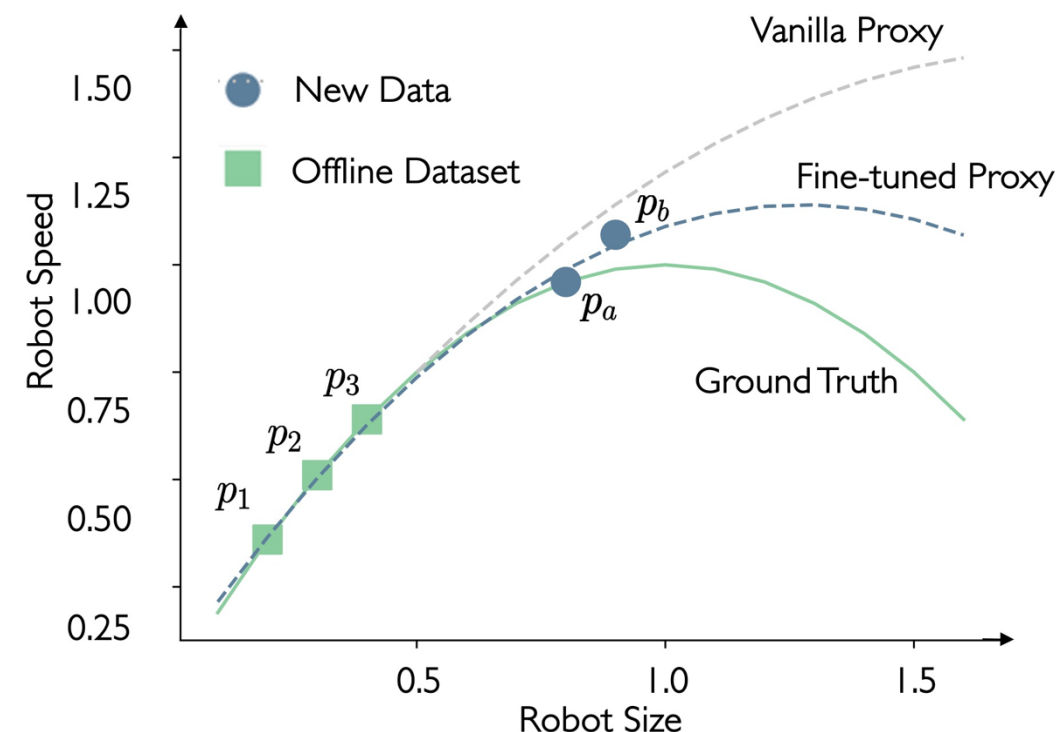
$$\mathbf{x}_t = \mathbf{x}_{t-1} + \eta \nabla_{\mathbf{x}} f_{\theta}(\mathbf{x}) \Big|_{\mathbf{x}=\mathbf{x}_t}, \quad \text{for } t \in [1, T].$$



Out-of-distribution (OOD) issue: The proxy overestimates the ground truth objective function, and the seemingly high-scoring design p_{new} obtained by gradient ascent has a low ground truth score.

Motivation

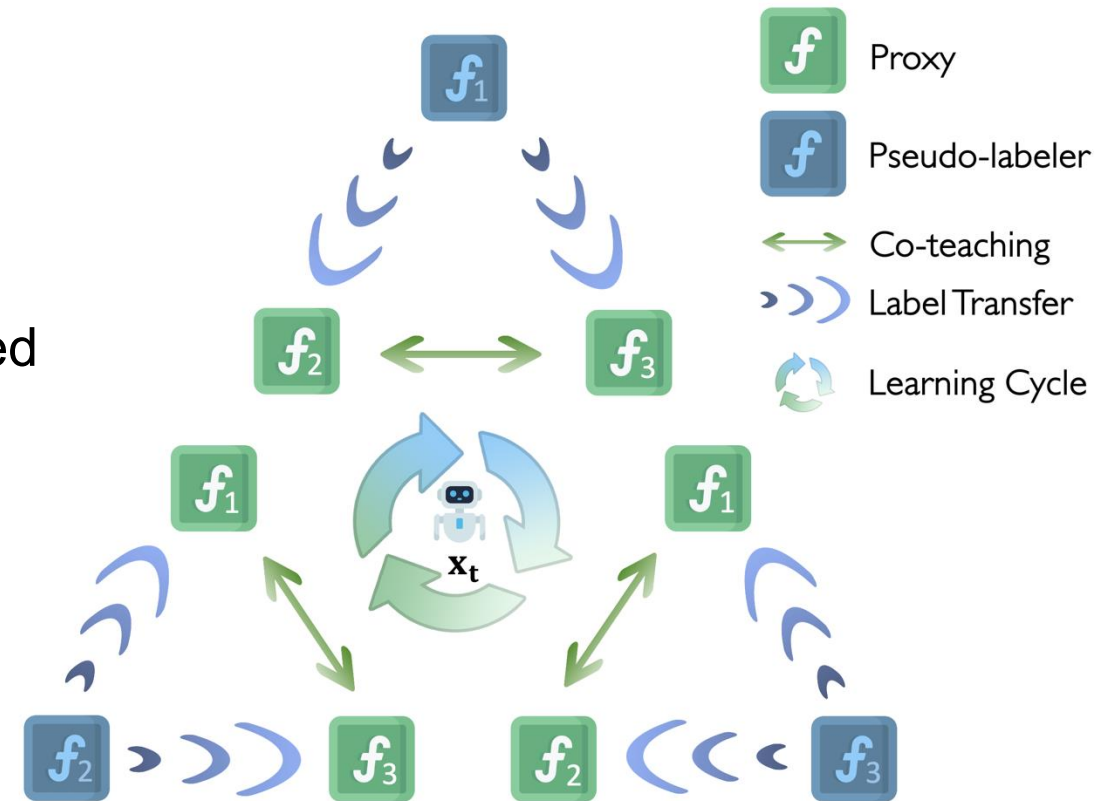
- What if we have more data points?
 - May train a better proxy!
- How to obtain these new data points?
 - Sample a set of points and use one proxy to pseudo-label them.
- How to identify the more accurate data points?
 - Let another two proxies co-teach each other to exchange valuable data.





Methodology: Pseudo-label-driven Co-teaching

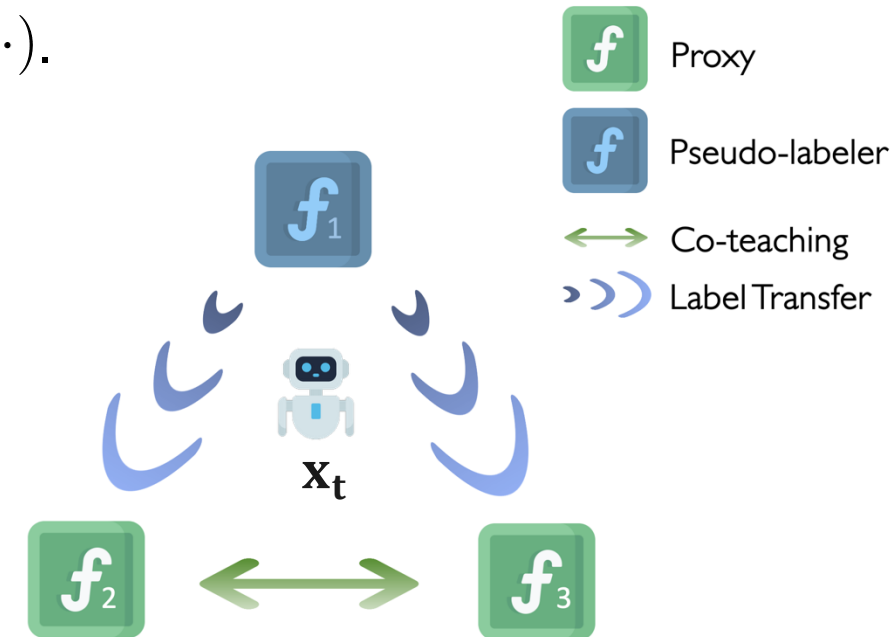
- Maintain three symmetric proxies and use their mean ensemble as the final proxy.
- Select one proxy as the pseudo-labeler, followed by a co-teaching process to enable knowledge sharing between the other two proxies.
- Repeat this process three times with different proxies as the pseudo-labeler in turn.





Methodology: Pseudo-label-driven Co-teaching

- Maintain three symmetric proxies, $f_{\theta_1}(\cdot)$, $f_{\theta_2}(\cdot)$, and $f_{\theta_3}(\cdot)$.
- Generate a set of points near the current point x_t and use $f_{\theta_1}(\cdot)$ to pseudo-label it.
- $f_{\theta_2}(\cdot)$ and $f_{\theta_3}(\cdot)$ co-teach each other by exchanging the small-loss samples in the pseudo-labeled dataset.



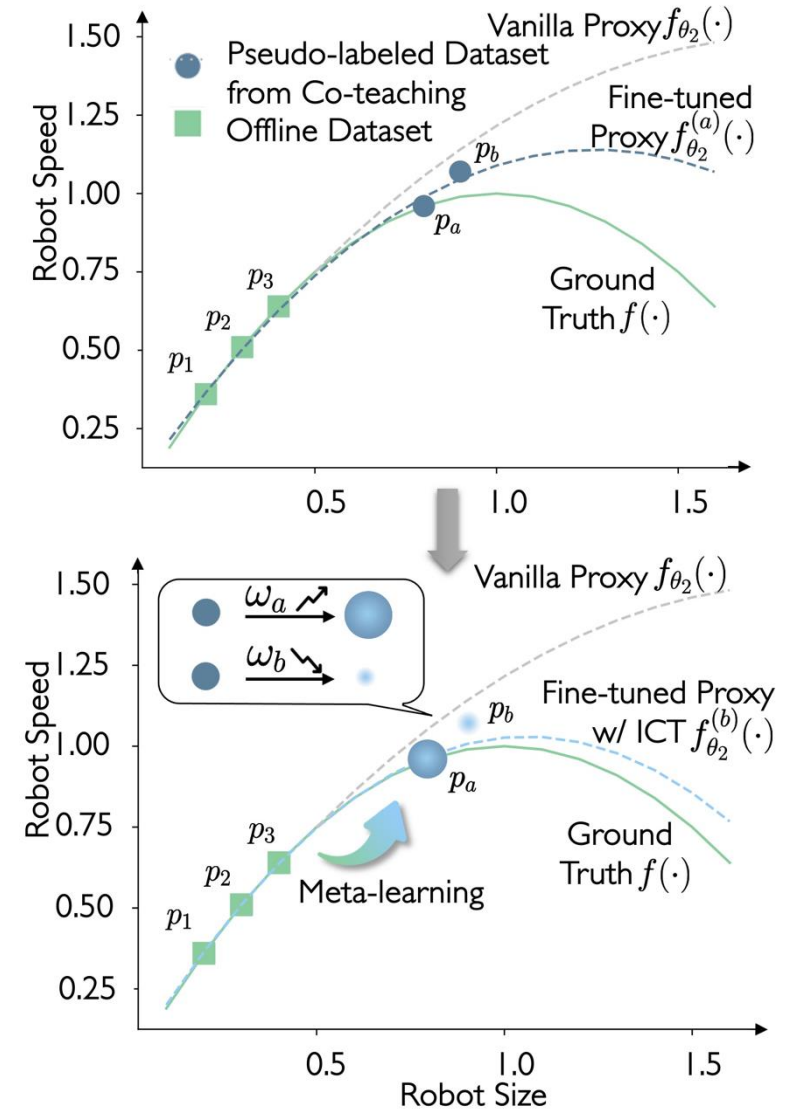


Methodology: Meta-learning-based Sample Reweighting

- Assign an importance weight ω_i to the i^{th} selected sample.
- Leverage the supervision signals from the offline dataset to update the weight:

$$\begin{aligned}\omega'_i &= \omega_i - \beta \frac{\partial \mathcal{L}(\theta^*(\omega))}{\partial \theta} \frac{\partial \theta^*(\omega)}{\partial \omega_i} \\ &= \omega_i + \frac{\alpha \beta}{K} \frac{\partial \mathcal{L}(\theta^*(\omega))}{\partial \theta} \frac{\partial (f_{\theta}(x_i^s) - \bar{y}_i^s)^2}{\partial \theta^\top},\end{aligned}$$

where $\mathcal{L}(\theta^*(\omega)) = \arg \min_{\omega} \frac{1}{N} \sum_{i=1}^N (f_{\theta^*(\omega)}(x_i) - y_i)^2$ is the loss on the offline data set.





Experimental Results: Continuous Tasks

Table 1: Experimental results on continuous tasks for comparison.

Method	Superconductor	Ant Morphology	D’Kitty Morphology	Hopper Controller
$\mathcal{D}(\text{best})$	0.399	0.565	0.884	1.0
BO-qEI	0.402 ± 0.034	0.819 ± 0.000	0.896 ± 0.000	0.550 ± 0.018
CMA-ES	0.465 ± 0.024	1.214 ± 0.732	0.724 ± 0.001	0.604 ± 0.215
REINFORCE	0.481 ± 0.013	0.266 ± 0.032	0.562 ± 0.196	-0.020 ± 0.067
CbAS	0.503 ± 0.069	0.876 ± 0.031	0.892 ± 0.008	0.141 ± 0.012
Auto.CbAS	0.421 ± 0.045	0.882 ± 0.045	0.906 ± 0.006	0.137 ± 0.005
MIN	0.499 ± 0.017	0.445 ± 0.080	0.892 ± 0.011	0.424 ± 0.166
Grad	0.483 ± 0.025	0.920 ± 0.044	0.954 ± 0.010	1.791 ± 0.182
Mean	0.497 ± 0.011	0.943 ± 0.012	0.961 ± 0.012	1.815 ± 0.111
Min	0.505 ± 0.017	0.910 ± 0.038	0.936 ± 0.006	0.543 ± 0.010
COMs	0.472 ± 0.024	0.828 ± 0.034	0.913 ± 0.023	0.658 ± 0.217
ROMA	0.510 ± 0.015	0.917 ± 0.030	0.927 ± 0.013	1.740 ± 0.188
NEMO	0.502 ± 0.002	0.952 ± 0.002	0.950 ± 0.001	0.483 ± 0.005
BDI	0.513 ± 0.000	0.906 ± 0.000	0.919 ± 0.000	1.993 ± 0.000
IOM	0.520 ± 0.018	0.918 ± 0.031	0.945 ± 0.012	1.176 ± 0.452
ICT_(ours)	0.503 ± 0.017	0.961 ± 0.007	0.968 ± 0.020	2.104 ± 0.357

Our method achieves top performance on all four continuous tasks.



Experimental Results: Discrete Tasks and Rankings

Table 2: Experimental results on discrete tasks, and ranking on all tasks for comparison.

Method	TF Bind 8	TF Bind 10	NAS	Rank Mean	Rank Median
$\mathcal{D}(\text{best})$	0.439	0.467	0.436		
BO-qEI	0.798 ± 0.083	0.652 ± 0.038	1.079 ± 0.059	9.9/15	11/15
CMA-ES	0.953 ± 0.022	0.670 ± 0.023	0.985 ± 0.079	6.1/15	3/15
REINFORCE	0.948 ± 0.028	0.663 ± 0.034	-1.895 ± 0.000	11.3/15	15/15
CbAS	0.927 ± 0.051	0.651 ± 0.060	0.683 ± 0.079	9.1/15	9/15
Auto.CbAS	0.910 ± 0.044	0.630 ± 0.045	0.506 ± 0.074	11.6/15	12/15
MIN	0.905 ± 0.052	0.616 ± 0.021	0.717 ± 0.046	11.0/15	12/15
Grad	0.906 ± 0.024	0.635 ± 0.022	0.598 ± 0.034	7.7/15	9/15
Mean	0.899 ± 0.025	0.652 ± 0.020	0.666 ± 0.062	6.6/15	6/15
Min	0.939 ± 0.013	0.638 ± 0.029	0.705 ± 0.011	7.3/15	8/15
COMs	0.452 ± 0.040	0.624 ± 0.008	0.810 ± 0.029	10.3/15	12/15
ROMA	0.924 ± 0.040	0.666 ± 0.035	0.941 ± 0.020	5.1/15	5/15
NEMO	0.941 ± 0.000	0.705 ± 0.000	0.734 ± 0.015	5.0/15	4/15
BDI	0.870 ± 0.000	0.605 ± 0.000	0.722 ± 0.000	7.9/15	8/15
IOM	0.878 ± 0.069	0.648 ± 0.023	0.274 ± 0.021	7.6/15	6/15
ICT_(ours)	0.958 ± 0.008	0.691 ± 0.023	0.667 ± 0.091	3.1/15	2/15

Our method achieves top performance on 2/3 discrete tasks.

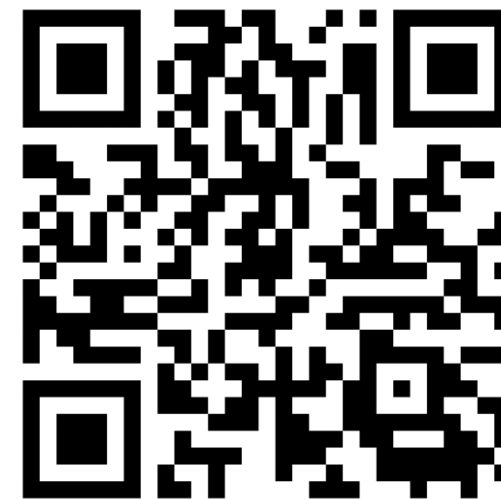
Thanks for your attention!



Paper



Code



Correspondence



McGill
UNIVERSITY



Mila



UNIVERSITY of
WASHINGTON



Stanford
University