NEURAL INFORMATION PROCESSING SYSTEMS

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

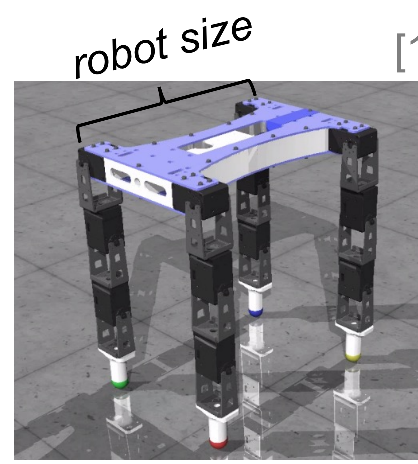
Ye Yuan [*1], Can (Sam) Chen [*†1,2],
Zixuan Liu [3], Willie Neiswanger [4], Xue Liu [1]

[1] McGill University [2] Mila [3] University of Washington [4] Stanford University

[*]Equal contribution with random order; [†]Corresponding to can.chen@mila.quebec

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.



[1]

[1] Trabucco, Brandon, Xinyang Geng, Aviral Kumar, and Sergey Levine. "Design-Bench: Benchmarks for Data-Driven Offline Model-Based Optimization."

Problem Formulation

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

$\boldsymbol{x_i}$: some design (robot size);

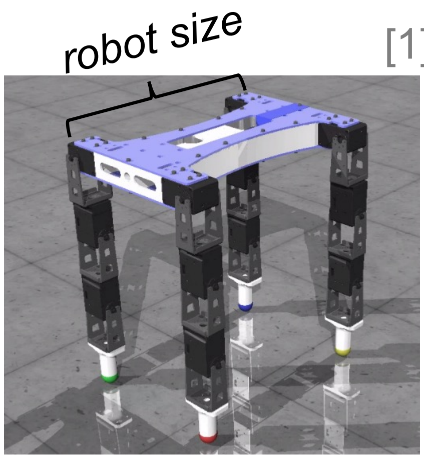
$y_i = f(\boldsymbol{x}_i)$: some property (robot speed);

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

$\mathcal{A}$: some algorithm outputs K candidates

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

[1]



[1] Trabucco, Brandon, Xinyang Geng, Aviral Kumar, and Sergey Levine. “Design-Bench: Benchmarks for Data-Driven Offline Model-Based Optimization.”

Related Work

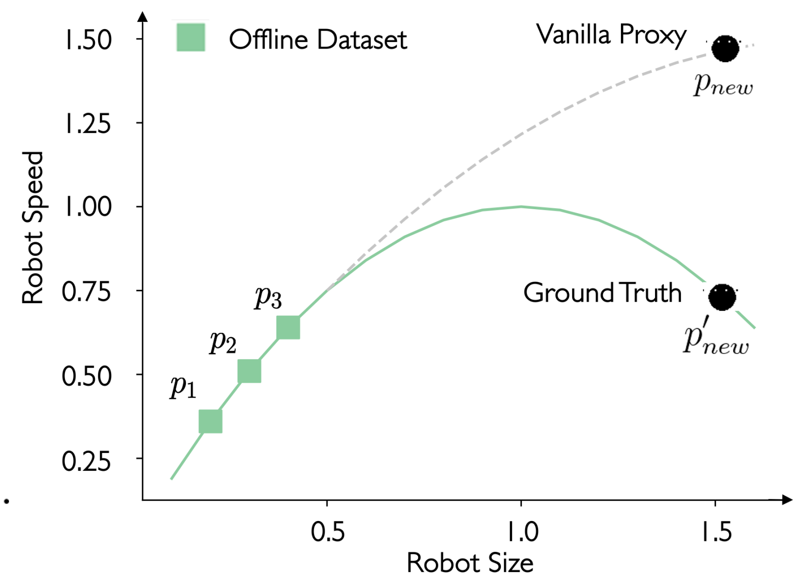
A common approach consists of 2 steps:

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

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

2) Perform gradient ascent:

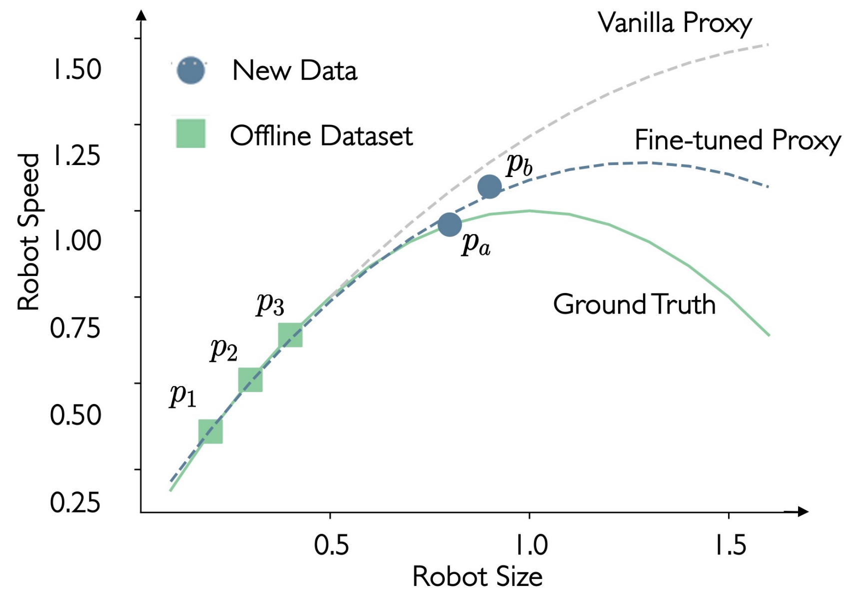
$$\boldsymbol{x}_t = \boldsymbol{x}_{t-1} + \eta \nabla_{\boldsymbol{x}} f_{\boldsymbol{\theta}}(\boldsymbol{x})\Big|_{\boldsymbol{x}=\boldsymbol{x}_t}, \quad \text{for } t \in [1, T].$$



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.

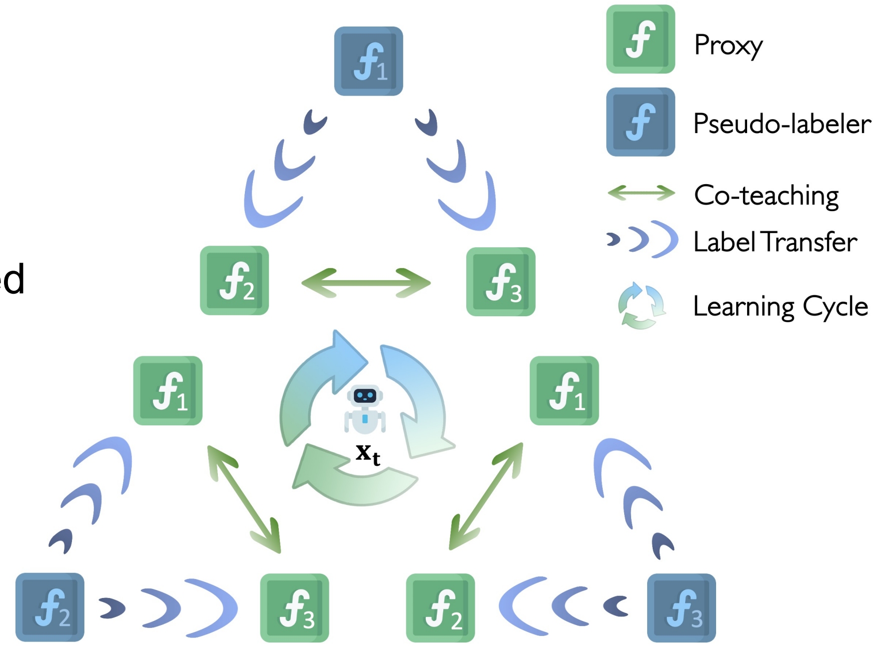
Motivation

- 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.



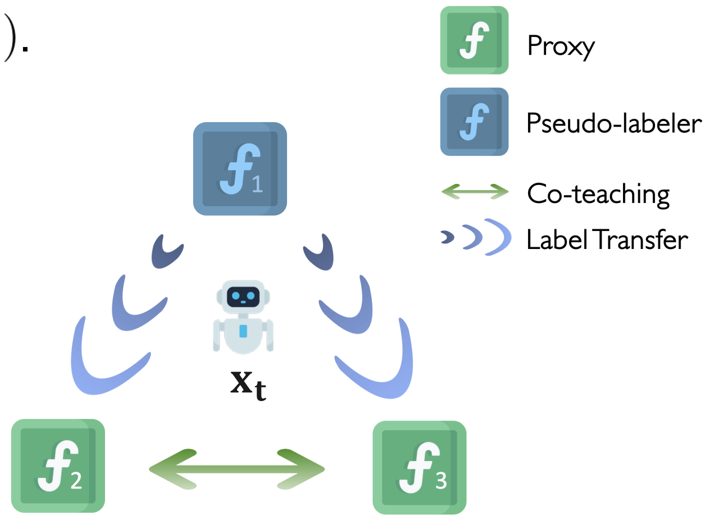
Methodology: Pseudo-label-driven Co-teaching

- 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, $f_{\boldsymbol{\theta_1}}(\cdot)$, $f_{\boldsymbol{\theta_2}}(\cdot)$, and $f_{\boldsymbol{\theta_3}}(\cdot)$.
- Generate a set of points near the current point $\boldsymbol{x}_t$ 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.



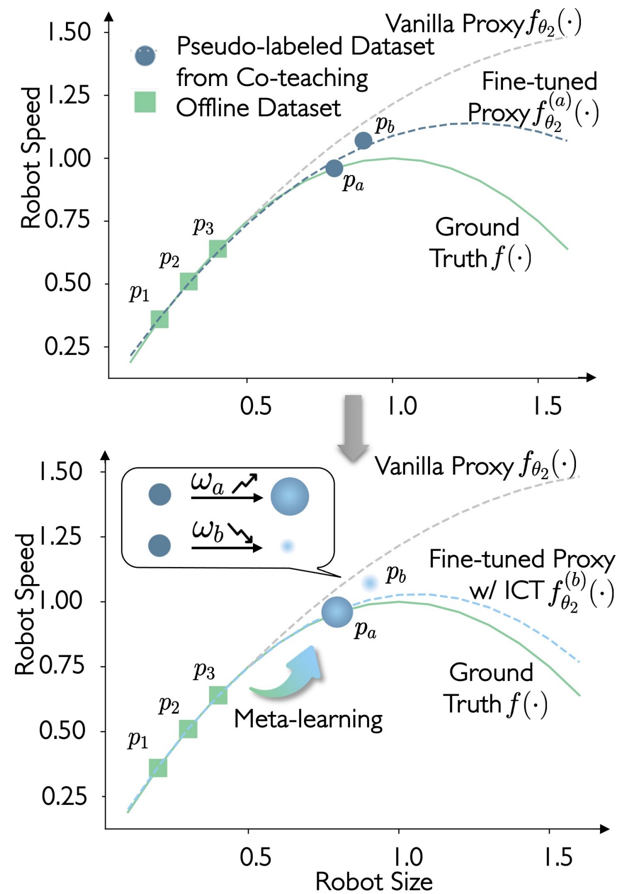
Methodology: Meta-learning-based Sample Reweighting

- Assign an importance weight $\boldsymbol{\omega}_i$ to the i^{th} selected sample.

- Leverage the supervision signals from the offline dataset to update the weight:

$$\boldsymbol{\omega}_i' = \boldsymbol{\omega}_i - \beta \frac{\partial \mathcal{L}(\boldsymbol{\theta}^*(\boldsymbol{\omega}))}{\partial \boldsymbol{\theta}} \frac{\partial \boldsymbol{\theta}^*(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}_i}$$
$$= \boldsymbol{\omega}_i + \frac{\alpha\beta}{K} \frac{\partial \mathcal{L}(\boldsymbol{\theta}^*(\boldsymbol{\omega}))}{\partial \boldsymbol{\theta}} \frac{\partial (f_{\boldsymbol{\theta}}(\boldsymbol{x}_i^s) - \bar{y}_i^s)^2}{\partial \boldsymbol{\theta}^\top},$$

where $\mathcal{L}(\boldsymbol{\theta}^*(\boldsymbol{\omega})) = \arg\min_{\boldsymbol{\omega}} \frac{1}{N} \sum_{i=1}^{N} (f_{\boldsymbol{\theta}^*(\boldsymbol{\omega})}(\boldsymbol{x}_i) - y_i)^2$ 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
$\mathcal{D}$(**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	$\mathbf{1.214 \pm 0.732}$	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	$\mathbf{0.503 \pm 0.069}$	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	$\mathbf{0.954 \pm 0.010}$	$\mathbf{1.791 \pm 0.182}$
Mean	0.497 ± 0.011	0.943 ± 0.012	$\mathbf{0.961 \pm 0.012}$	$\mathbf{1.815 \pm 0.111}$
Min	$\mathbf{0.505 \pm 0.017}$	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	$\mathbf{0.510 \pm 0.015}$	0.917 ± 0.030	0.927 ± 0.013	1.740 ± 0.188
NEMO	0.502 ± 0.002	0.952 ± 0.002	$\mathbf{0.950 \pm 0.001}$	0.483 ± 0.005
BDI	$\mathbf{0.513 \pm 0.000}$	0.906 ± 0.000	0.919 ± 0.000	$\mathbf{1.993 \pm 0.000}$
IOM	$\mathbf{0.520 \pm 0.018}$	0.918 ± 0.031	0.945 ± 0.012	1.176 ± 0.452
ICT$_{\text{(ours)}}$	$\mathbf{0.503 \pm 0.017}$	$\mathbf{0.961 \pm 0.007}$	$\mathbf{0.968 \pm 0.020}$	$\mathbf{2.104 \pm 0.357}$

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
$\mathcal{D}$(**best**)	0.439	0.467	0.436		
BO-qEI	0.798 ± 0.083	0.652 ± 0.038	$\mathbf{1.079 \pm 0.059}$	9.9/15	11/15
CMA-ES	$\mathbf{0.953 \pm 0.022}$	0.670 ± 0.023	0.985 ± 0.079	6.1/15	3/15
REINFORCE	$\mathbf{0.948 \pm 0.028}$	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	$\mathbf{0.705 \pm 0.000}$	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$_{\text{(ours)}}$	$\mathbf{0.958 \pm 0.008}$	$\mathbf{0.691 \pm 0.023}$	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