ICLR 2023

A System for Morphology-Task Generalization via Unified Representation and Behavior Distillation

Hiroki Furuta[1,2], Yusuke Iwasawa[1], Yutaka Matsuo[1], Shixiang Shane Gu[2,1]

[1]The University of Tokyo, [2]Google Research, Brain Team

Contact: furuta@weblab.t.u-tokyo.ac.jp

Recent Success of Large-Scale Foundation Models

- **GPT-3, PaLM, ChatGPT, GPT-4, LLaMA, …** in NLP
- **CLIP, DALL-E, Imagen, Stable Diffusion, …** in CV

and more!

When did International Conference on Learning Representations (ICLR) start?

The International Conference on Learning Representations (ICLR) was first held in 2013. The inaugural conference was organized by Yoshua Bengio, Yann LeCun, and Corinna Cortes and took place in Scottsdale, Arizona, USA. Since then, ICLR has become one of the top-tier conferences in the field of machine learning, specifically focused on the study of deep learning and representation learning.

ChatGPT [2022]

Explaining a joke

Prompt

```
Explain this joke:
Joke: Did you see that Google just hired an eloquent whale for their TPU team? It showed them how to communicate between two different pods!
```

Model Response

```
Prediction: TPUs are a type of computer chip that Google uses for deep learning. A "pod" is a group of TPUs. A "pod" is also a group of whales. The joke is that the whale is able to communicate between two groups of whales, but the speaker is pretending that the whale is able to communicate between two groups of TPUs.
```

PaLM [2022]

CLIP [2021]

DALL-E [2021]

What is the key ingredient of large-scale models?

Key Ingredients for “Generalist” Model

Considering the recent success of the large-scale models, they must have …

- curating **massive diverse dataset**

Key Ingredients for “Generalist” Model

Considering the recent success of the large-scale models, they must have …

- curating **massive diverse dataset**
- defining **unified IO representation**

Key Ingredients for “Generalist” Model

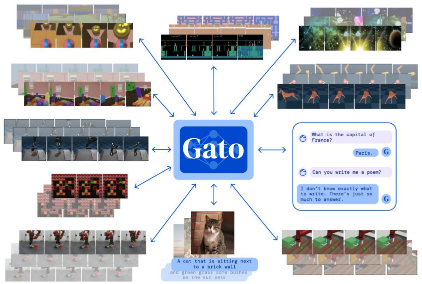
Considering the recent success of the large-scale models, they must have ...

- curating **massive diverse dataset**
- defining **unified IO representation**
- performing **efficient representation and architecture selection**

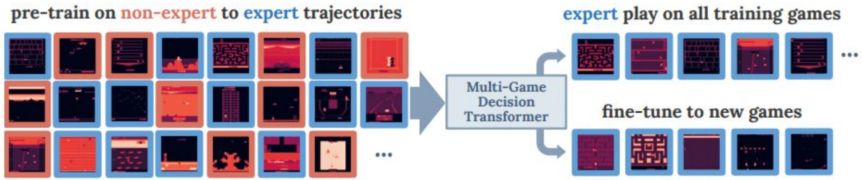
Key Ingredients for “Generalist” Model

Considering the recent success of the large-scale models, they must have …

- curating **massive diverse dataset**
- defining **unified IO representation**
- performing **efficient representation and architecture selection**



GATO [2022]

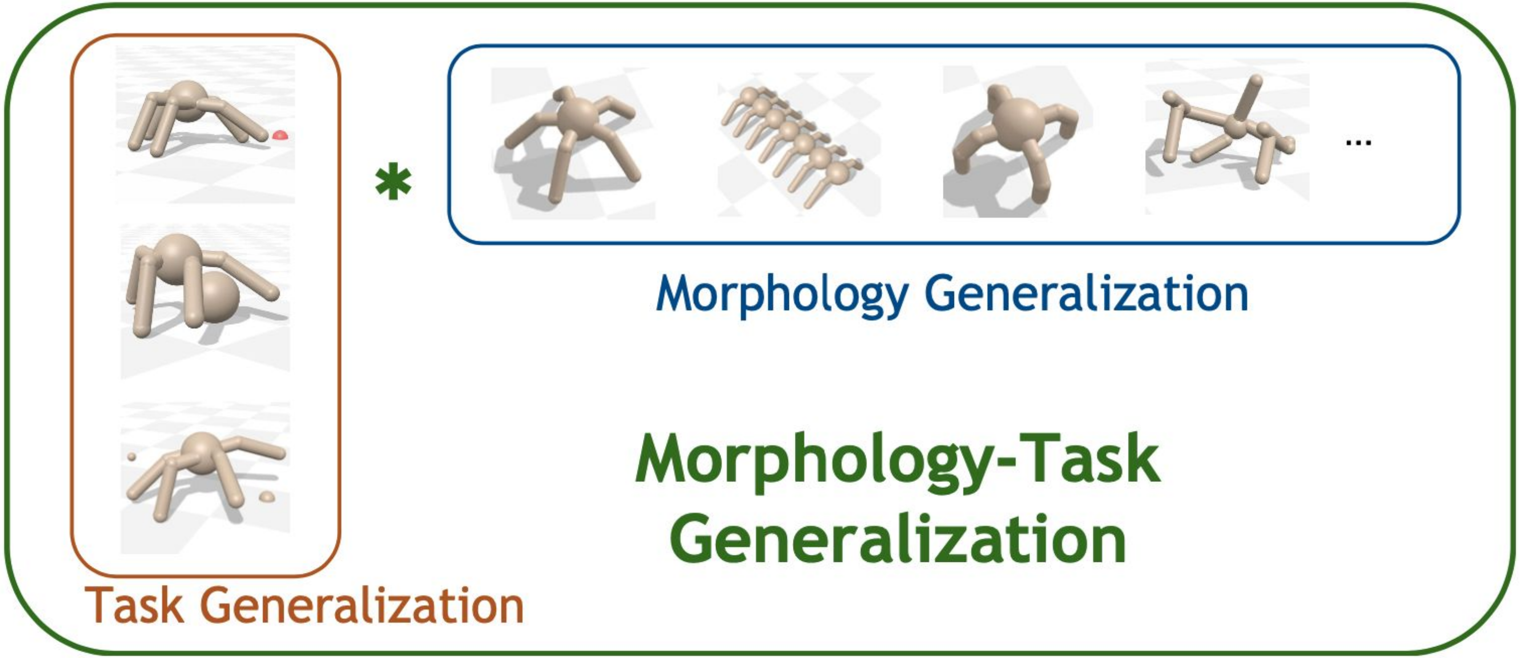


MGDT [2022]

We try to build “multitask learners” over **task** and **morphology** for RL.

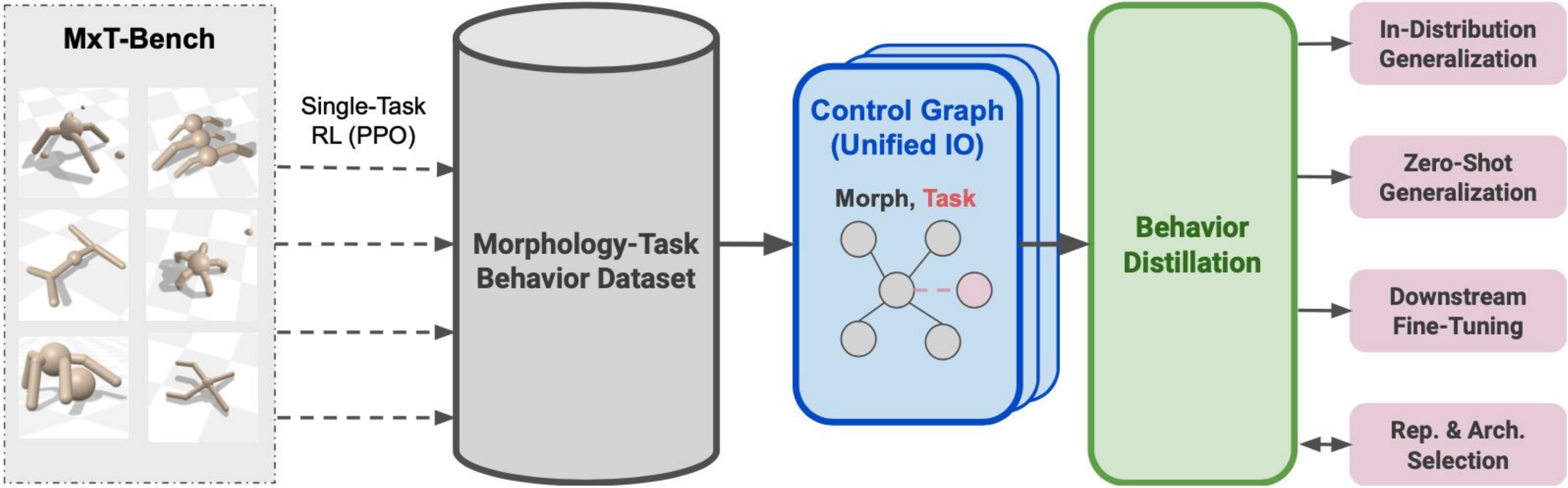
Morphology-Task Generalization

Compared to prior works, we consider the “**cartesian product**” of morphology generalization and task generalization.



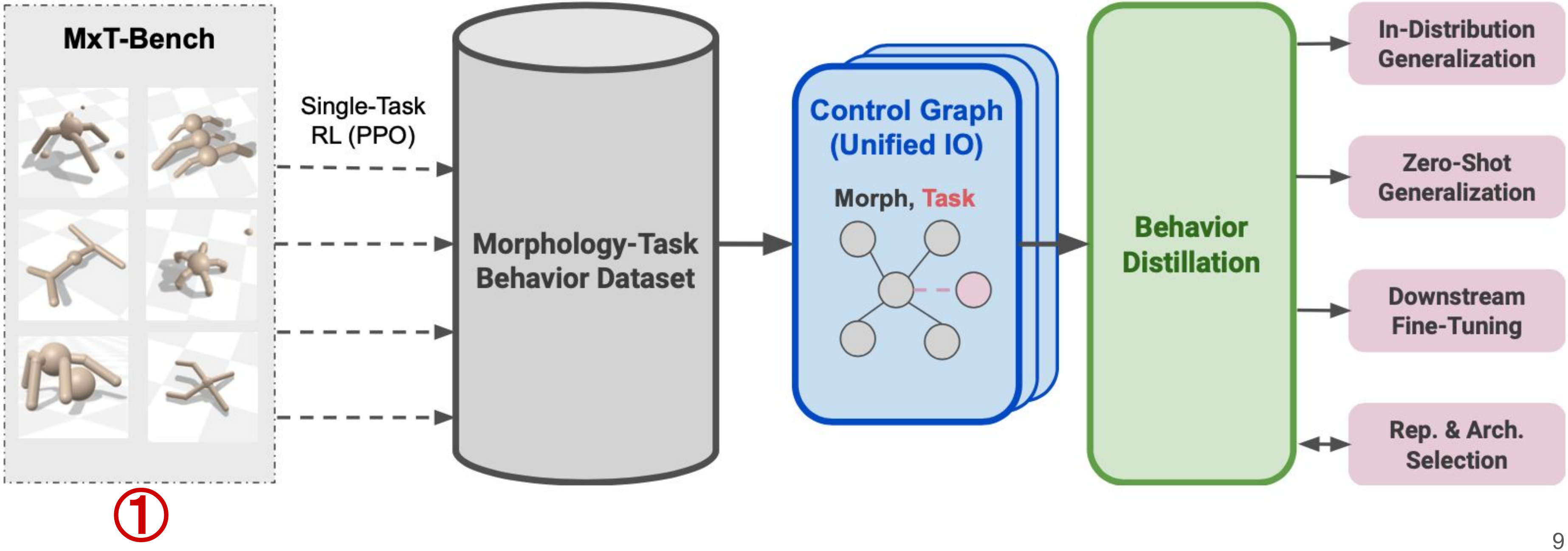
Overview: Behavior Distillation Pipeline

We present a pipeline to enable a single policy to learn **multiple tasks and morphologies** via offline **behavior distillation** with **unified IO** (control graph).



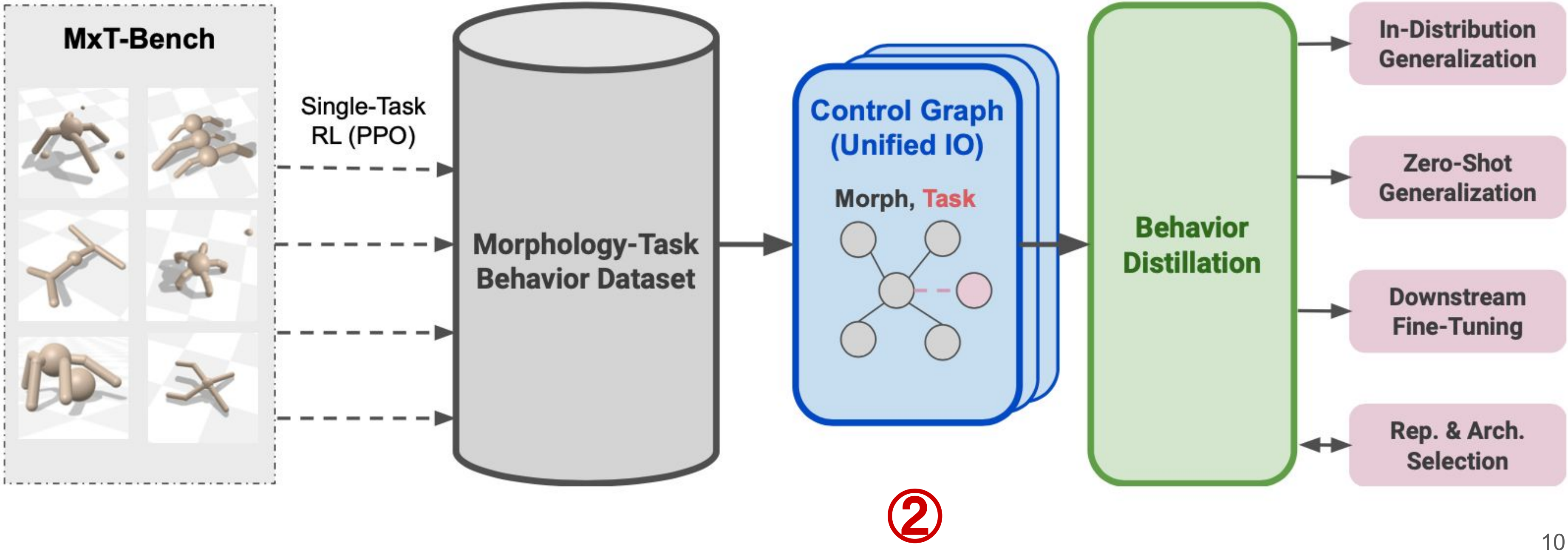
Overview: Behavior Distillation Pipeline

We present a pipeline to enable a single policy to learn **multiple tasks and morphologies** via offline **behavior distillation** with **unified IO** (control graph).



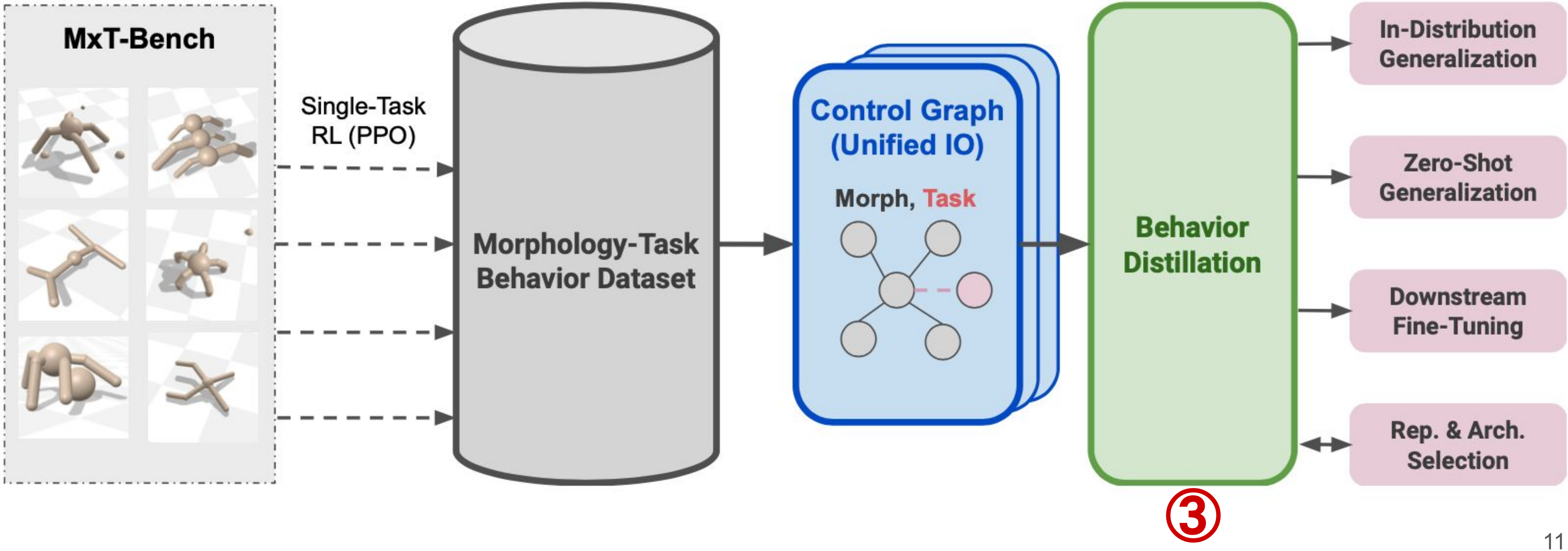
Overview: Behavior Distillation Pipeline

We present a pipeline to enable a single policy to learn **multiple tasks and morphologies** via offline **behavior distillation** with **unified IO** (control graph).



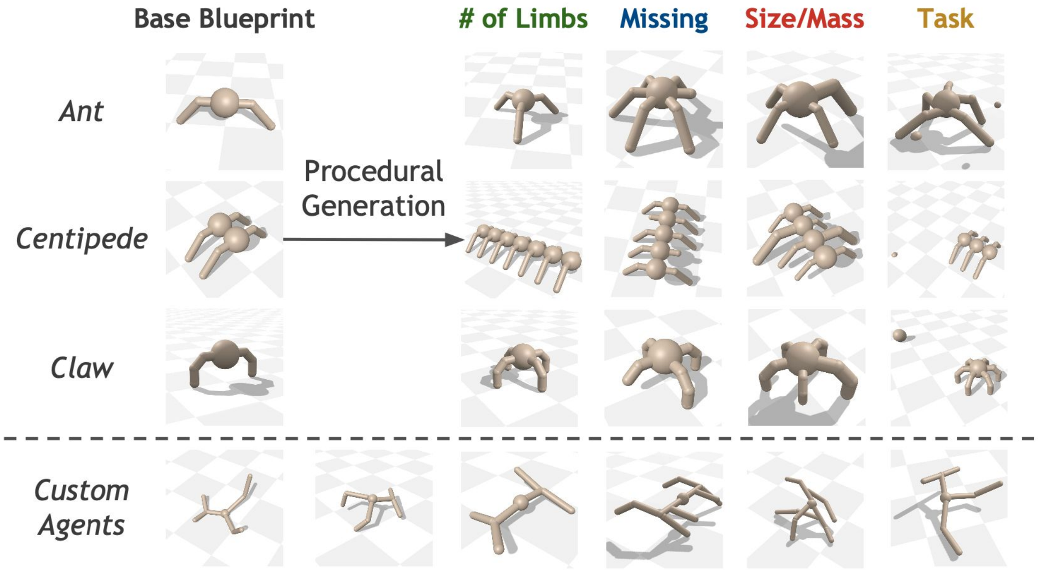
Overview: Behavior Distillation Pipeline

We present a pipeline to enable a single policy to learn **multiple tasks and morphologies** via offline **behavior distillation** with **unified IO** (control graph).



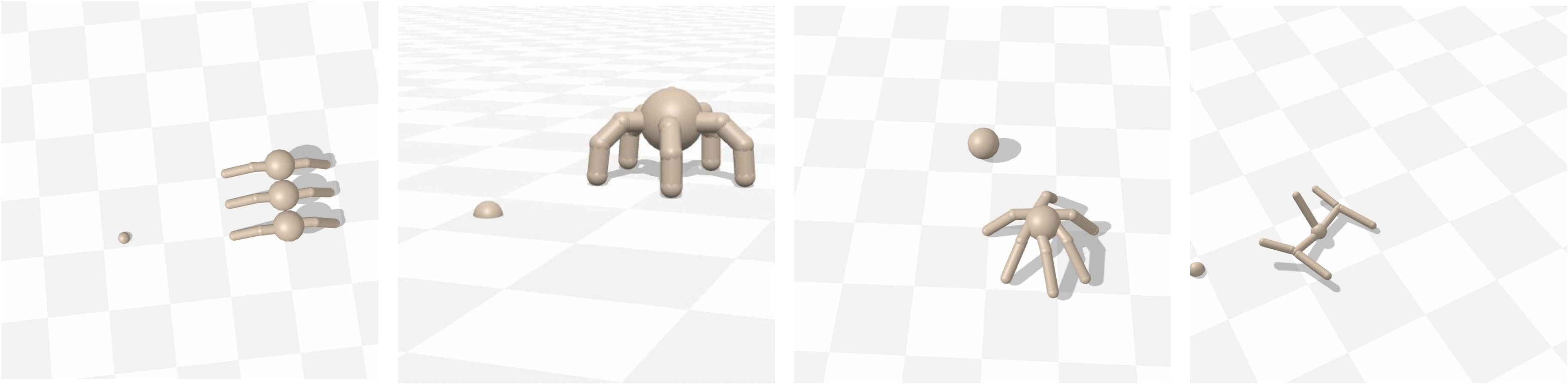
MxT-Bench

- A testbed of morphology-task generalization
- Fast behavior generation with hardware-accelerated simulator Brax [Freeman et al. 2021]
- Procedural generation of morphology and task with minimal blueprints



Example Morphology and Task

In MxT-Bench, we prepare several base morphology (ant, centipede, worm, claw, unimal [Gupta et al. 2022]) and task (reach, touch, twisters).



Task Performance Metric

Task = Goal Reaching

- *Reach* ➔ Distance between Leg and Goal
- *Touch* ➔ Distance between Body and Ball
- *Twisters* ➔ Sum of Distance between Legs and Goals

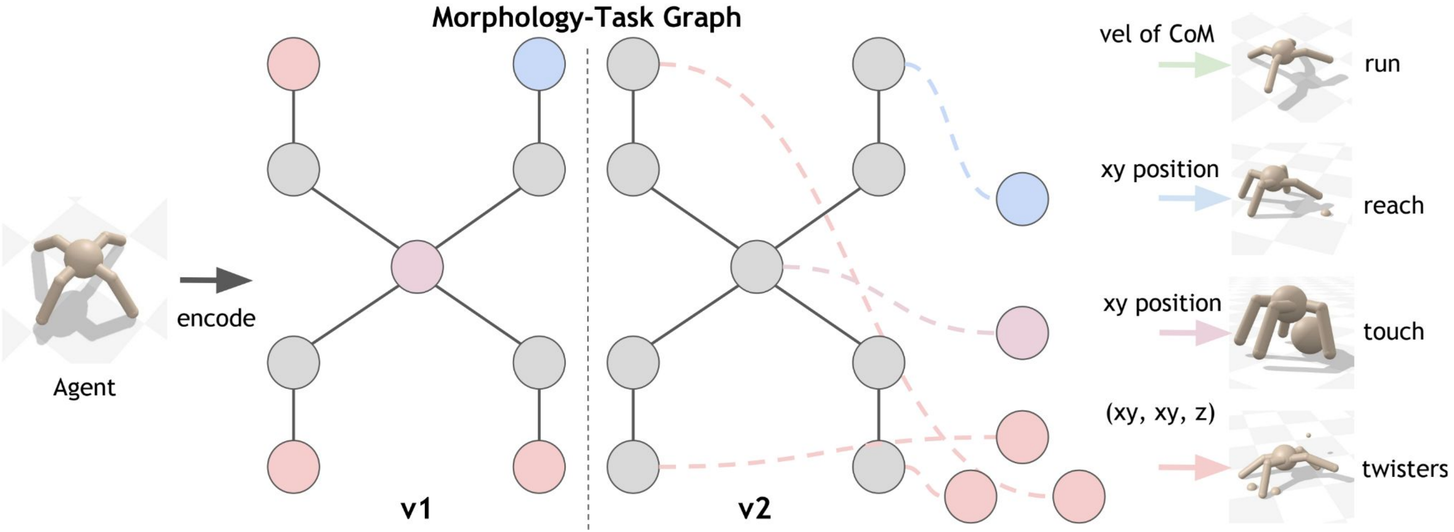
Performance of policy *π* =

Average normalized final distance over Morphology *M* and Task *Ψ*

$$\bar{d}(\mathcal{M},\Psi;\pi) := \frac{1}{|\mathcal{M}||\Psi|} \underbrace{\sum_{m}^{\mathcal{M}}}_{\text{Morph.}} \underbrace{\sum_{\psi}^{\Psi}}_{\text{Task}} \underbrace{\mathbb{E}_{s_g \sim p_\psi}}_{\text{Goal}} \left[\sum_{i=1}^{N_\psi} \frac{\overbrace{d_\psi(\boldsymbol{s}_T^m, s_g^i)}^{\text{Distance}} - d_{\min}^{i,m,\psi}}{d_{\max}^{i,m,\psi} - d_{\min}^{i,m,\psi}} \right]$$

Morphology-Task Graph (as Unified IO representation)

- *Scene Graph*: 3D relational information of a scene in CV
- *Morphology Graph*: agent's geometry and actions in RL
- **Morphology-Task Graph** (ours): unified interface to encode observations, actions, and goals (i.e. tasks) as nodes in the shared graph representation



Architecture & Representation Selection

Behavior Distillation: *supervised learning* over morphology-task expert demo

$$\mathcal{L}_\pi = -\mathbb{E}_{m,\psi\sim\mathcal{M},\Psi}\left[\mathbb{E}_{\boldsymbol{s}^m,\boldsymbol{a}^m,s_g\sim\mathcal{D}_{m,\psi}}\left[\log\pi_\theta(\boldsymbol{a}^m|\{\boldsymbol{s}^m,s_g\})\right]\right]$$

Offline training allows us **efficient representation and architecture selection** (cf. multi-task online RL).

Architecture Candidates

- **MLP** → Baseline w/o geometric bias
- **GNN** → Straightforward approach
- **Transformer** → Consider each node = token and assume fully-connected graph

Architecture & Representation Selection

Considering prior works (SMP, Amorpheus, MetaMorph), we define **base_set representation** as:

- Position, Velocity, Quaternion, Angular velocity, Joint Angle, Joint Range

We test the combination to other observations:

- **jv** = joint vel, **id** = limb id, **rp** = relative pos, **rr** = relative rot, **m** = morph information

	Node features	+jv	+id	+rp	+rr	+m	Average Dist.
	base_set						0.4330 ± 0.02
Prior Work (1)	base_set-id		✔				0.4090 ± 0.02
	base_set-rp			✔			0.3820 ± 0.01
	base_set-rr				✔		0.4543 ± 0.01
Best Feature Set	base_set-m					✔	$\mathbf{0.3128} \pm \mathbf{0.02}$
	base_set-rp-rr			✔	✔		0.3869 ± 0.01
	base_set-jv-rp-rr	✔		✔	✔		0.4000 ± 0.01
Prior Work (2)	base_set-jv-rp-rr-m	✔		✔	✔	✔	0.3323 ± 0.01

Morphology-Task Generalization Results

We test 3 different types of generalization:

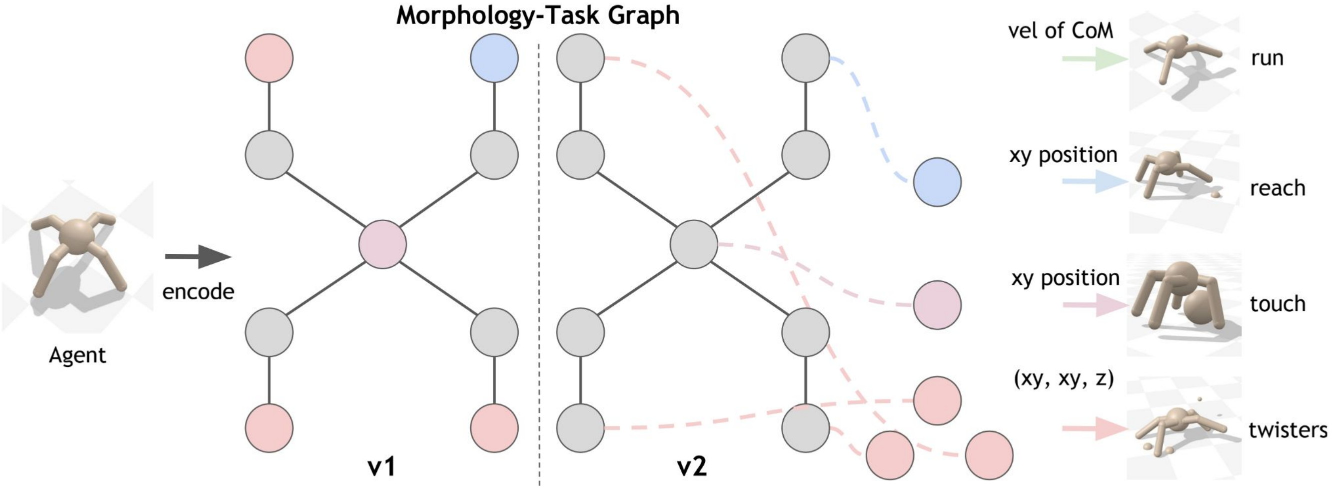
- **In-distribution** generalization
- **Compositional** generalization
- **OOD** generalization

		Task	
		Seen	Unseen
Morphology	Seen	In-distribution	Compositional (Task)
	Unseen	Compositional (Morphology)	OOD

Morphology-Task Generalization Results

Morphology-Task Graph v2 + Transformer shows better generalization

	Random	MLP	GNN (MTGv1)	Transformer (MTGv1)	Transformer (MTGv2)	Token-MTGv2
In-Distribution	1.2019 ± 0.41	0.5150 ± 0.01	0.4776 ± 0.01	0.4069 ± 0.02	$\mathbf{0.3128 \pm 0.02}$	0.3402 ± 0.01
In-Distribution (unimal)	0.9090 ± 0.03	0.6703 ± 0.01	–	0.4839 ± 0.02	$\mathbf{0.4178 \pm 0.01}$	–
Compositional (Morphology)	1.1419 ± 0.41	0.7216 ± 0.01	–	0.4940 ± 0.01	$\mathbf{0.4066 \pm 0.01}$	–
Compositional (Task)	0.8932 ± 0.01	0.6849 ± 0.01	–	0.5395 ± 0.04	$\mathbf{0.4461 \pm 0.05}$	–
Out-of-Distribution	0.8979 ± 0.01	0.7821 ± 0.02	–	0.6144 ± 0.04	$\mathbf{0.5266 \pm 0.04}$	–

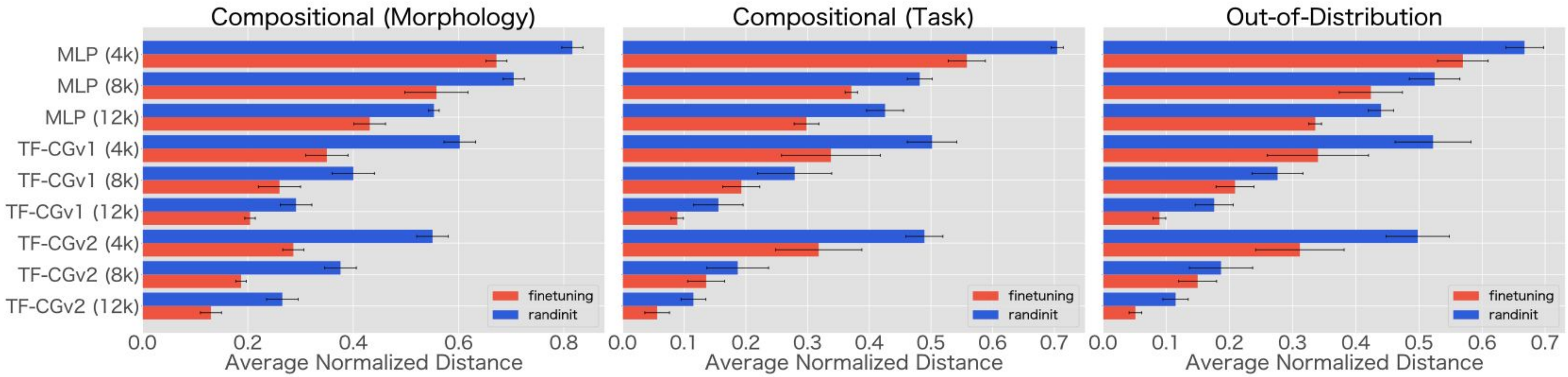


Downstream Fine-Tuning Results

Setting: **Multi-task Imitation Learning** (for unseen morphology/task)

Comparing: ①pretraining v.s. from scratch, ②architecture & representation

Control Graph v2 + Transformer also provides better control prior

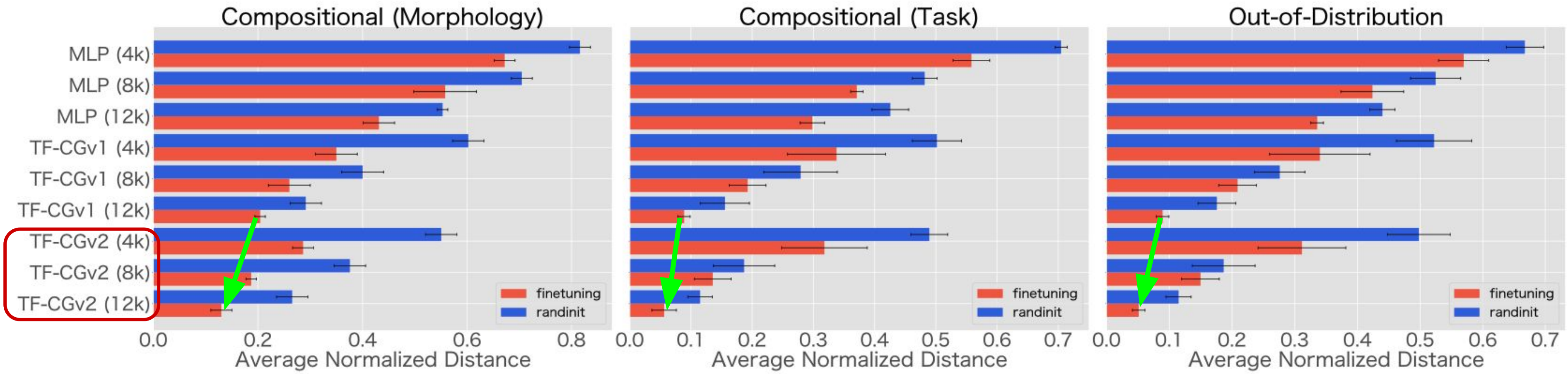


Downstream Fine-Tuning Results

Setting: **Multi-task Imitation Learning** (for unseen morphology/task)

Comparing: ①pretraining v.s. from scratch, ②architecture & representation

Morphology-Task Graph v2 + Transformer also provides better control prior!



Summary

Learning a generalist model seems to requires the same key ingredients:

- Massive diverse dataset → **MxT-Bench**

Summary

Learning a generalist model seems to requires the same key ingredients:

- Massive diverse dataset → **MxT-Bench**
- Unified IO representation → **Morphology-Task Graph**

Summary

Learning a generalist model seems to requires the same key ingredients:

- Massive diverse dataset ➜ **MxT-Bench**
- Unified IO representation ➜ **Morphology-Task Graph**
- Efficient rep. & arch. selection ➜ **Behavior Distillation**

Summary

Learning a generalist model seems to requires the same key ingredients:

- Massive diverse dataset → **MxT-Bench**
- Unified IO representation → **Morphology-Task Graph**
- Efficient rep. & arch. selection → **Behavior Distillation**

We show morphology-task graph v2 + transformer generalize both morphology and task in **in-distribution**, **compositional**, and **out-of-distribution** settings.

Morphology-task graph v2 + transformer also has a good inductive bias for downstream tasks (multi-task imitation learning).

Summary

Learning a generalist model seems to requires the same key ingredients:

- Massive diverse dataset ➜ **MxT-Bench**
- Unified IO representation ➜ **Morphology-Task Graph**
- Efficient rep. & arch. selection ➜ **Behavior Distillation**

We hope this work inspires the community to more focus on scalability & generalization in RL.