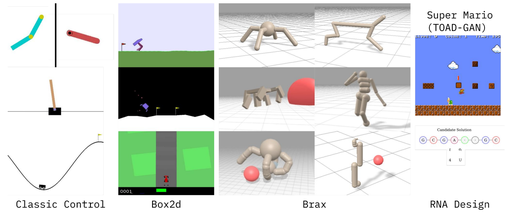
CARL: A Benchmark for Contextual and Adaptive Reinforcement Learning

Carolin Benjamins[1], Theresa Eimer[1], Frederik Schubert[1], André Biedenkapp[2], Bodo Rosenhahn[1], Frank Hutter[2,3], Marius Lindauer[1]

[1]Leibniz University Hannover, [2]Albert-Ludwigs University Freiburg, [3]Bosch Center for Artificial Intelligence

1 TL;DR

- CARL = benchmark for contextual RL
- Goal: generalization over different contexts
- Varying the context → more difficult
- Making context explicit → helps learning



Included Environments

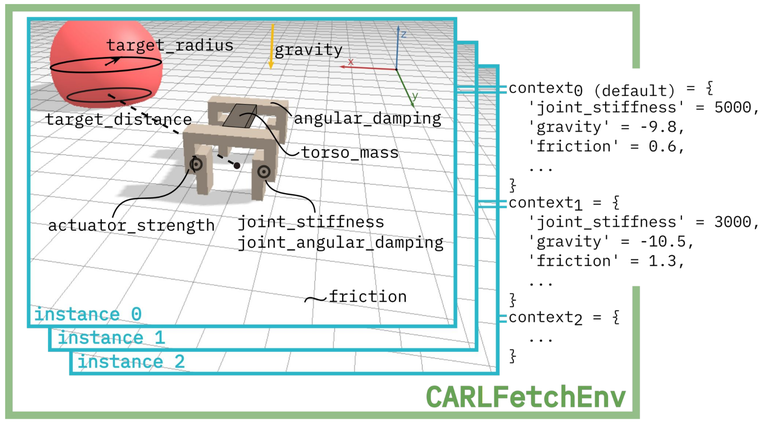
2 Motivation & The Benchmark

Goal

- Train for **generalization** over different instances (contexts) of the same environment (env)
- Ultimately: create general agents

CARL

- Extends existing RL environments
- Makes the **context** defining the behavior of the env **explicit, visible** & **configurable**

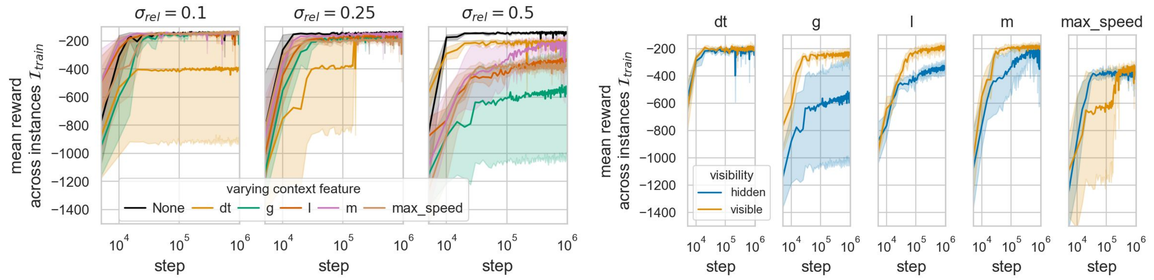


3 What is contextual RL?

- A cMDP $\mathbf{M_I := \{M_i\}_{i \in I}}$ consists of an MDP $\mathbf{M_i}$ for each instance **i** of an instance set **I**
- Instances are defined by their context **c,** e.g. gravity
- Between different $\mathbf{M_i}$ actions **A** and state space **S** stay the same
- Transition dynamics **T** and reward function **R** can vary depending on the instance context
- Requires generalization: success is usually measured across a set of contexts from the training distribution

4 Key Insights

- Varying the context during training increases difficulty even on simple environments
- The more variation of the context, the more difficult
- Showing context instead of hiding can help learning
- In-distribution generalization: agent generalizes to unseen contexts from the train distribution



(a) different context distributions

(b) the effect of visible/explicit context

Training performance on CARLPendulumEnv (DDPG)

5 Future Works

- Separate **representation learning** from policy learning by comparing learned representation to ground truth **Uncertain dynamics** through arbitrarily perturbed context features
- Testbed for **explainability and interpretability** through ground truth
- New challenges for **AutoRL and hyperparameter optimization** (see our poster on Hyperparameters in cRL)
- Explicit context settings for fields like **Safe RL and Continual Learning**

Leibniz Universität Hannover

UNI FREIBURG

AutoML.org

@AutoML_org

github