OFF-THE-GRID MULTI-AGENT REINFORCEMENT LEARNING

A REPOSITORY OF DATASETS FOR TRAINING RL AGENTS FOR MULTI-AGENT TASKS WITHOUT ACCESS TO THE ENVIRONMENT

Author: *Claude Formanek*

Supervisors: *Associate Professor Jonathan Shock*

Dr. Arnu Pretorius

INTRO

- For real-world tasks we might not have access to an environment simulator to train RL agents.
- Offline RL is about using datasets of previously collected experience to train RL agents.
- But many real-world tasks depend on cooperating agents. Offline multi-agent RL is relativly under-researched.

BACKGROUND

- Offline RL is challenging due to out-of-distribution actions, i.e. the deployed agent may choose actions that weren't in the dataset.
- Such actions lead to extrapolation-error and hurt performance.
- In offline multi-agent RL, extrapolation error accumulates exponentially in the number of agents.

AIMS & OBJECTIVES

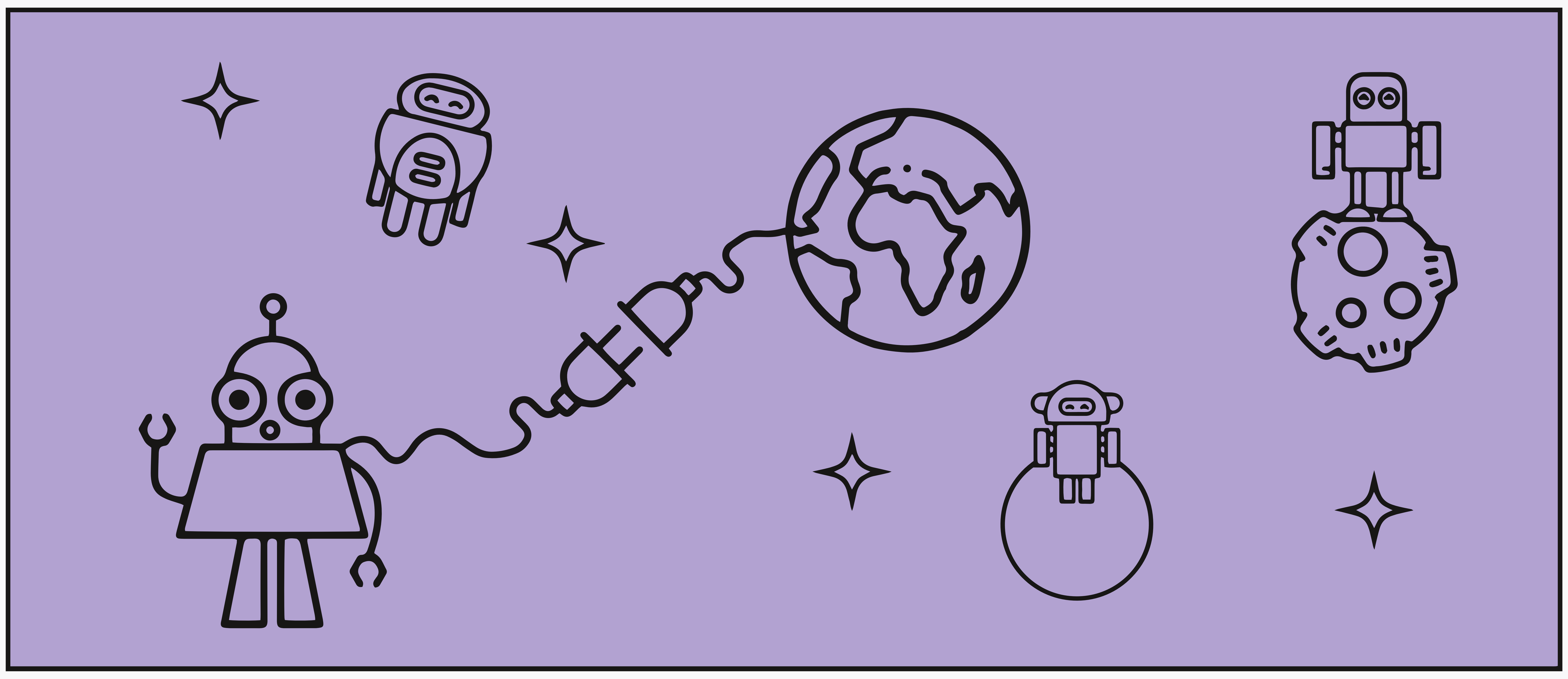
- Offline multi-agent RL is an important stepping stone towards applying RL to real-world problems.
- This work aims to lower the barrier for entry for future researchers in offline multi-agent RL by providing:
  - >Pre-generated datasets
  - >Strong baseline algorithms
  - >Useful utilities such as an experience logger to generate new datasets.

METHOD

- Generate datasets of experience on a range of multi-agent tasks with varying levels of agent proficiency (e.g. expert, medium).
- Implement several multi-agent RL algorithms and demonstrate that they can learn from offline datasets.

RESULTS

- Successfully generated a repository of datasets on 12 popular multi-agent tasks.
- Implemented 6 offline multi-agent RL algorithms and demonstrated their effectiveness on offline multi-agent tasks.



ABSTRACT

Being able to harness the power of large static datasets for developing intelligent multi-agent systems could unlock enormous value for real-world applications.

Many important industrial systems are multi-agent in nature and are difficult to model using bespoke simulators. However, in industry, distributed system processes can often be recorded during operation and large quantities of static data can be collected. Offline multi-agent reinforcement learning (MARL) provides a promising paradigm for building effective controllers from static data.

But offline MARL is still in its infancy, and unlike more mature subfields of RL, lacks standardised benchmarks, baselines and evaluation protocols. This deficiency makes it difficult for the community to sensibly measure progress.

In this work, we aim to fill this gap by releasing off-the-grid MARL (OG-MARL), a set of diverse datasets and baselines for offline MARL together with a standardised evaluation protocol.

Our datasets provide settings that are characteristic of real-world systems including complex dynamics, non-stationarity, partial observability, suboptimality and sparse rewards and are generated from popular online MARL benchmarks. Our recorded behaviour span different levels of proficiency useful for comparing more advanced offline approaches to simpler forms of behaviour cloning.

Furthermore, we provide implementations and benchmarking results for state-of-the-art offline MARL algorithms using our proposed evaluation protocol. We hope OG-MARL will serve the community and help steer progress in offline MARL while also providing an easy entry point for researchers new to the field.