ChemGymRL

An Interactive Framework for Reinforcement Learning for Digital Chemistry

Chris Beeler, Sriram Ganapathi Subramanian, Kyle Sprague, Nouha Chatti, Mitchell Shahen, Nicholas Paquin, Mark Baula, Amanuel Dawit, Zihan Yang, Xinkai Li, **Mark Crowley, Isaac Tamblyn, Colin Bellinger**

Join us! - Building Up Through Collaboration
Contribute your own models! Make new benches, try out existing or new RL algorithms!

chemgymrl.com

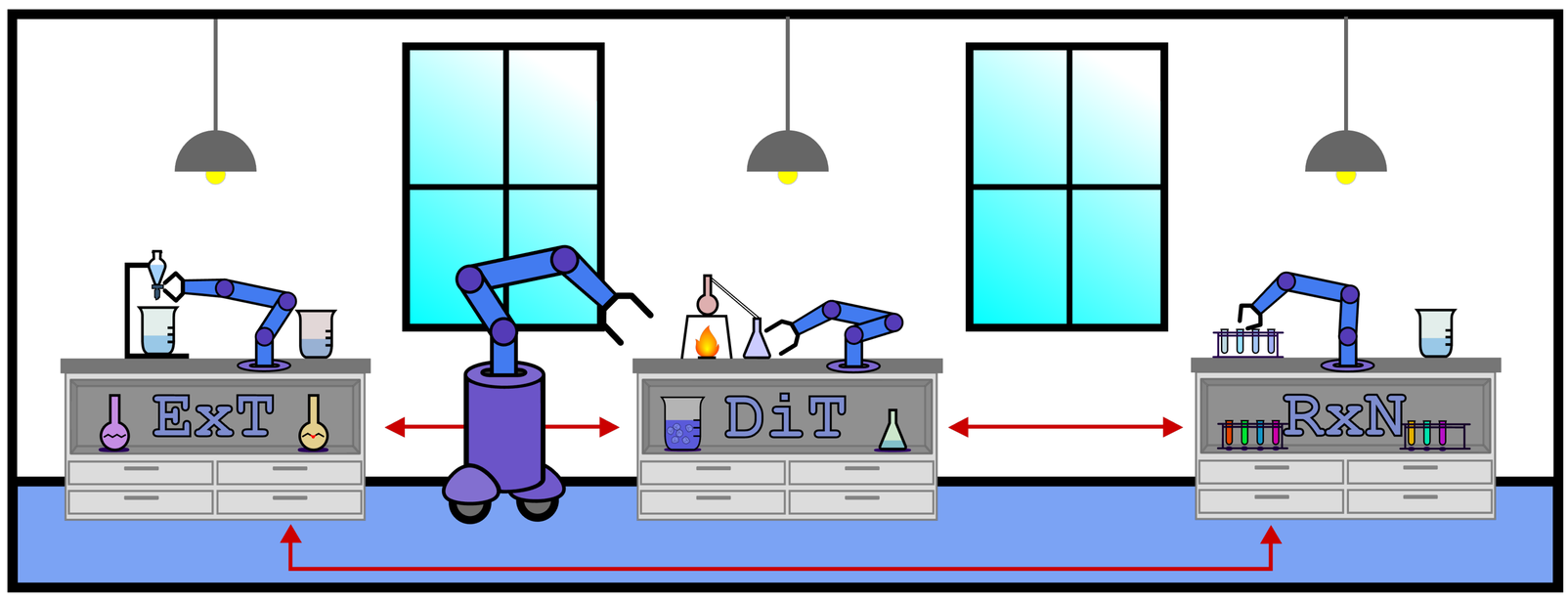
WATERLOO.AI
WATERLOO ARTIFICIAL INTELLIGENCE INSTITUTE

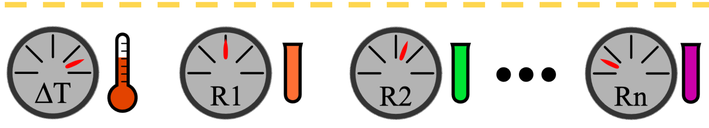
What is ChemGymRL?

Problem: RL is *very* data intensive, training robotic chemistry agents by taking actions in the real world is *infeasible* and ~~possibly~~ *dangerous*.

Benefits for RL: ML researches always need new challenges, and Digital Chemistry involves challenges which are not commonly found in RL benchmarks and therefore offer a rich space to work in.

Our Solution: We introduce a set of highly customizable and *open-source* RL environments, implementing the standard *Gymnasium API*.





Why RL for Chemistry?

Reinforcement Learning(RL) is class of Machine Learning algorithms that learn by taking actions, making observations, viewing the results, and updating its model (or hypothesis is you like)/policy/beliefs.

In other words...RL is a perfect analogy for the experimental scientist!

THE LABORATORY

The laboratory can be thought of as a virtual chemistry laboratory consisting of different benches where a variety of tasks can be completed.

The laboratory is comprised of 3 basic elements: **Vessels, Shelves, Benches**
Vessels contain materials, in pure or mixed form
tracking the hidden internal state of its contents
shelves hold any vessels not currently in use or inputs/outputs of experiments.

All Benches Have...

Input: target material given as one-hot vector
State: vessels and contained materials

REACTION BENCH

The reaction bench (RxN) allows the agent to **transform available reactants** into various products via a chemical reaction.
The agent has the ability to control

- the **temperature of the vessel** and
- the **amounts of reactants added**.

The goal of the agent operating this bench is to modify the reaction parameters, in order to increase and/or decrease the yield of certain desired/undesired materials.

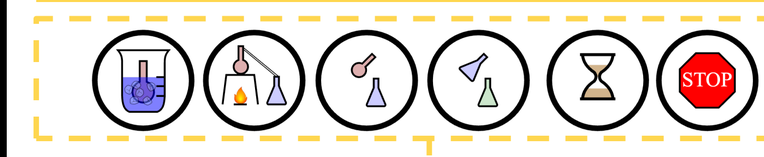
Actions: The key to the agent's success in this bench is *learning how best to allow certain reactions to occur* such that

- the yield of the desired material is **maximized**
- While the yield of the undesired material is **minimized**.

Rewards: After the 20 steps have elapsed, the agent receives a reward equal to the molar amount of the target material produced.

Experiments

- To test and demonstrate the framework, we trained RL agents were trained for 100K time steps across 10 environments in parallel (for a total of 1M time steps).
- Samples: 256 time steps of experience (in each environment) to update policies/Value-functions.
- Replay buffer: 1M experiences for **DQN**, **SAC**, and **TD3**.
- Exploration: first 30K steps of DQN used linear schedule from 1.0 down to 0.01, then fixed.
- Implementations: Stable Baselines 3



DISTILLATION BENCH

The distillation bench (DiT) aims to isolate certain materials from an inputted vessel containing multiple materials (albeit with a different process).

Actions: *Transferring* materials between a number of vessels and heating/cooling the vessel to separate materials from each other.

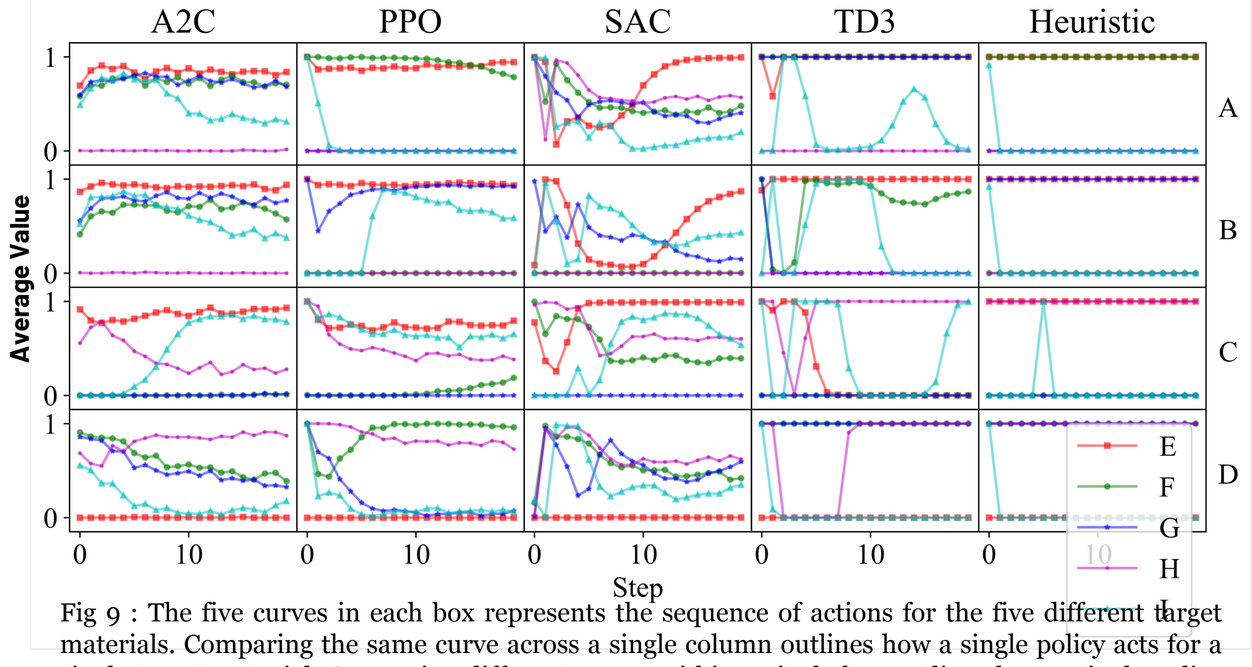


Fig 9 : The five curves in each box represents the sequence of actions for the five different target materials. Comparing the same curve across a single column outlines how a single policy acts for a single target material. Comparing different curves within a single box outlines how a single policy acts differently between different target materials. Comparing the same curve across a single row outlines how different policies act for the same target material.

EXTRACTION BENCH

Extraction is a method to **separate out undesired products** from the outputs of chemical reactions.

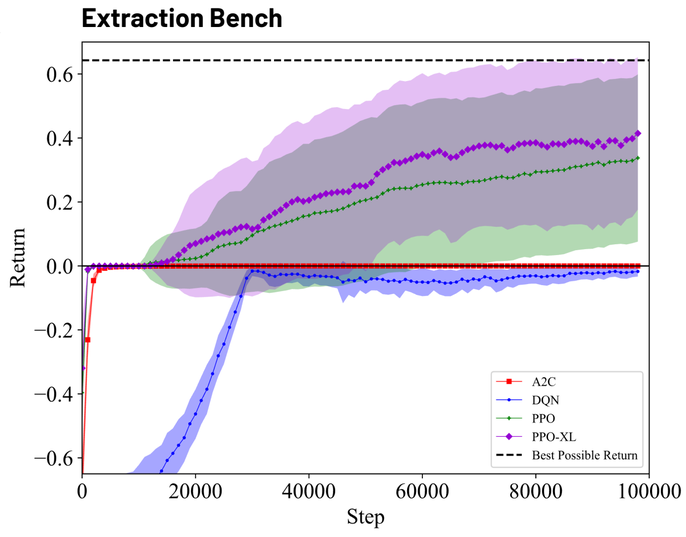
The extraction bench (ExT) aims to *isolate* and *extract* certain dissolved materials from the input vessels.

Actions: *Transferring* materials between different vessels and utilizing specifically selected solvents to separate materials from each other.

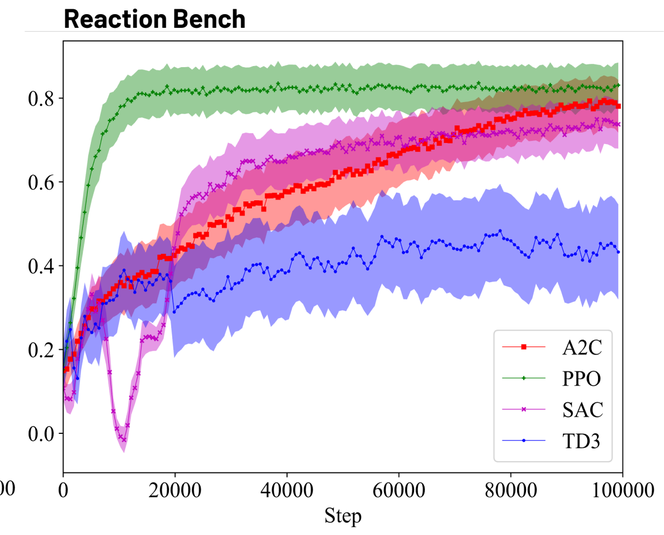
CHARACTERIZATION BENCH

The characterization bench is the primary method to obtain insight as to what the vessel contains. The purpose of the characterization bench is not to manipulate the input vessel, but to subject it to analysis techniques that observe the state of the vessel, possibly including the materials inside it and their relative quantities. This does not mean that the contents of the input vessel cannot be modified by the characterization bench. This allows an agent or user to observe vessels, determine their contents, and allocate the vessel to the necessary bench for further experimentation.

Extraction Bench



Reaction Bench



AFFILIATIONS
1) Department of Mathematics and Statistics, University of Ottawa
2) Department of Electrical and Computer Engineering, University of Waterloo
3) Digital Technologies, National Research Council of Canada,
4) Vector Institute for Artificial Intelligence, Toronto
5) Department of Physics, University of Ottawa