Hyperparameters in Reinforcement Learning and How to Tune Them

[Eimer et al. ICML'23]



LUH|AI LUH|AI Meta

Our Key Findings

1. Tuning RL hyperparameters is easy

Our Key Findings

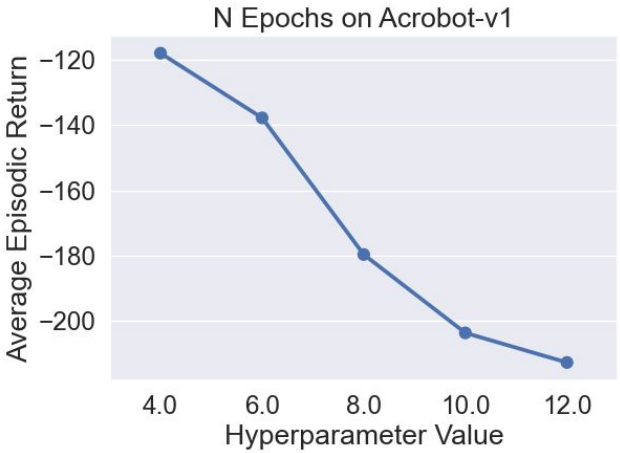
1. Tuning RL hyperparameters is easy
2. Automatically tuning RL is cheaper than manual hyperparameter search

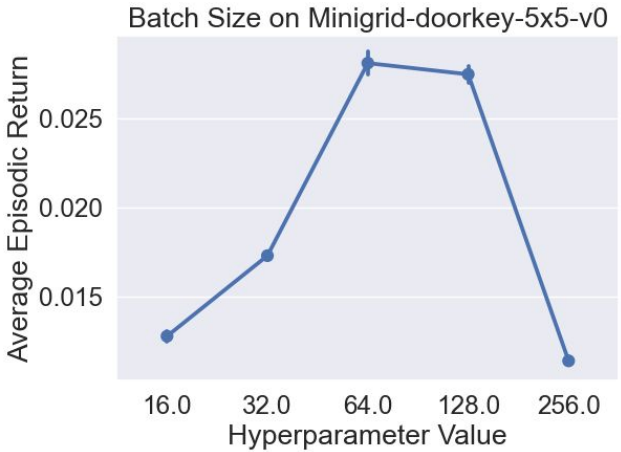
Our Key Findings

1. Tuning RL hyperparameters is easy
2. Automatically tuning RL is cheaper than manual hyperparameter search
3. Reporting hyperparameter optimization is crucial for reproducibility

Key Point 1: RL Hyperparameters Are Easy

Across 128 algorithm/environment/HP combinations, **only 8** show a hyperparameter not being relevant for the success of the agent



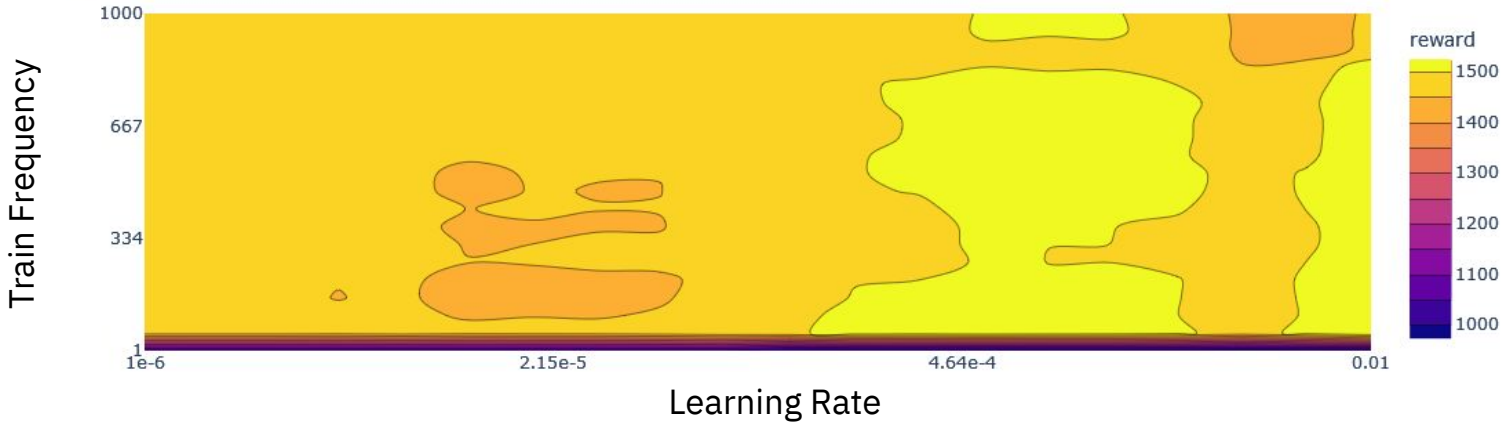


Left: PPO's number of epochs on Acrobot.

Right: DQN's batch size on MiniGrid DoorKey 5x5.

Key Point 1: RL Hyperparameters Are Easy

- Partial Dependency Plots [Moosbauer et al. ‘21] of selected environments show few interaction effects between hyperparameters
- The hyperparameter ranges where agents perform well is fairly wide

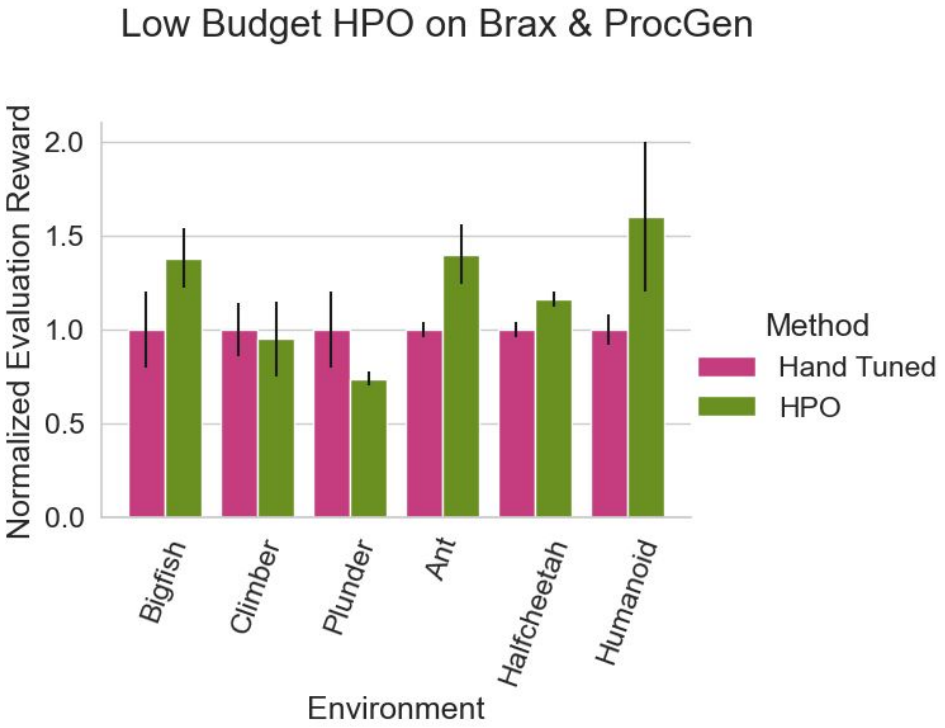


PDP: Train Frequency and Learning Rate of SAC on Pendulum

Key Point 2: Tuning RL is Cheap

With only 64 runs:

- IDAAC on Procgen & PPO on Brax
- Large search spaces (14/11 HPs)
- Overall improvement over hand tuned default settings
- Up to 10x cheaper than manual baselines

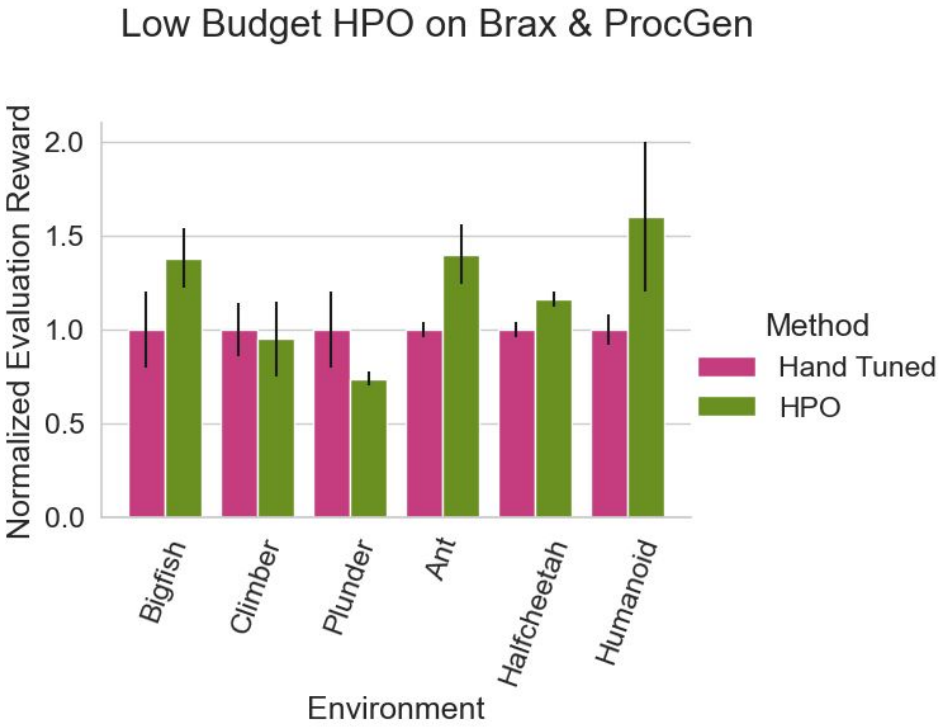


Key Point 2: Tuning RL is Cheap

With only 64 runs:

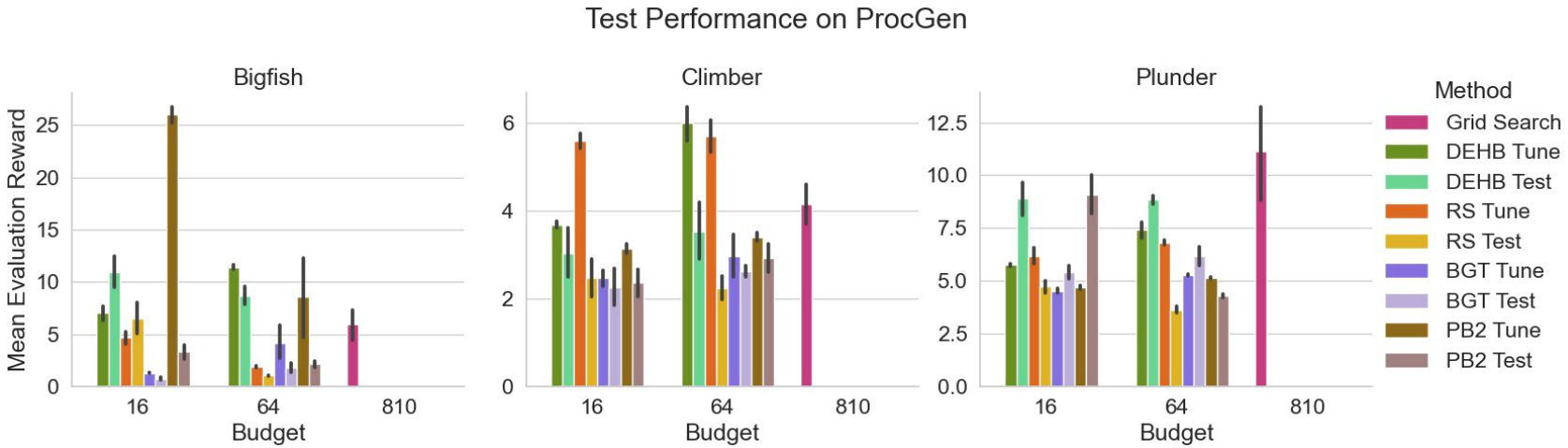
- IDAAC on Procgen & PPO on Brax
- Large search spaces (14/11 HPs)
- Overall improvement over hand tuned default settings
- Up to 10x cheaper than manual baselines

But...



Key Point 3: Reporting is crucial

Up to 8x worse performance on test seed, even when tuning across multiple seeds - **independent of the tuning method**



Key Point 3: Reporting is crucial

Best practices for reporting:

- Report tuning setting, including tuning and test seeds
- Report tuning method, including the settings of the tuner
- Report budget used for all methods and baselines

Key Point 3: Reporting is crucial

Best practices for reporting:

- Report tuning setting, including tuning and test seeds
- Report tuning method, including the settings of the tuner
- Report budget used for all methods and baselines

Full checklist for your paper here.

Tuning RL Agents

- Tuning hyperparameters is important in RL but also achievable
- Ideally: tune all hyperparameters for all investigated methods
- Modern HPO methods work better than grid search even on small budgets
- Reporting HPO details is important for future comparability of results

Get In Touch!

LUH|AI

@The_Eimer

LUH-AI

@luh-ai

For more information, check out our paper website, blog post and GitHub repository:

Website

Blog

GitHub

Funded by:

European Research Council
Established by the European Commission

Hyperparameter Optimization in RL Currently

- Existing RL algorithms require HPO to work on new environments

Hyperparameter Optimization in RL Currently

- Existing RL algorithms require HPO to work on new environments
  But: Inefficient grid searches are common [e.g. Hambro et al. ‘22]

Hyperparameter Optimization in RL Currently

- Existing RL algorithms require HPO to work on new environments
  But: Inefficient grid searches are common [e.g. Hambro et al. ‘22]
- RL-specific tuning methods exist [Jaderberg et al. ‘17, Wan et al. ‘22]

Hyperparameter Optimization in RL Currently

- Existing RL algorithms require HPO to work on new environments
  - But: Inefficient grid searches are common [e.g. Hambro et al. ‘22]
- RL-specific tuning methods exist [Jaderberg et al. ‘17, Wan et al. ‘22]
  - But: Hard to use research code, incomplete state of the art

Hyperparameter Optimization in RL Currently

- Existing RL algorithms require HPO to work on new environments
  - But: Inefficient grid searches are common [e.g. Hambro et al. ‘22]
- RL-specific tuning methods exist [Jaderberg et al. ‘17, Wan et al. ‘22]
  - But: Hard to use research code, incomplete state of the art
- RL papers by now tend to include hyperparameter settings

Hyperparameter Optimization in RL Currently

- Existing RL algorithms require HPO to work on new environments
  - But: Inefficient grid searches are common [e.g. Hambro et al. ‘22]
- RL-specific tuning methods exist [Jaderberg et al. ‘17, Wan et al. ‘22]
  - But: Hard to use research code, incomplete state of the art
- RL papers by now tend to include hyperparameter settings
  - But: Details like the time spent tuning each method are still unclear