Advanced Topics in Reinforcement Learning

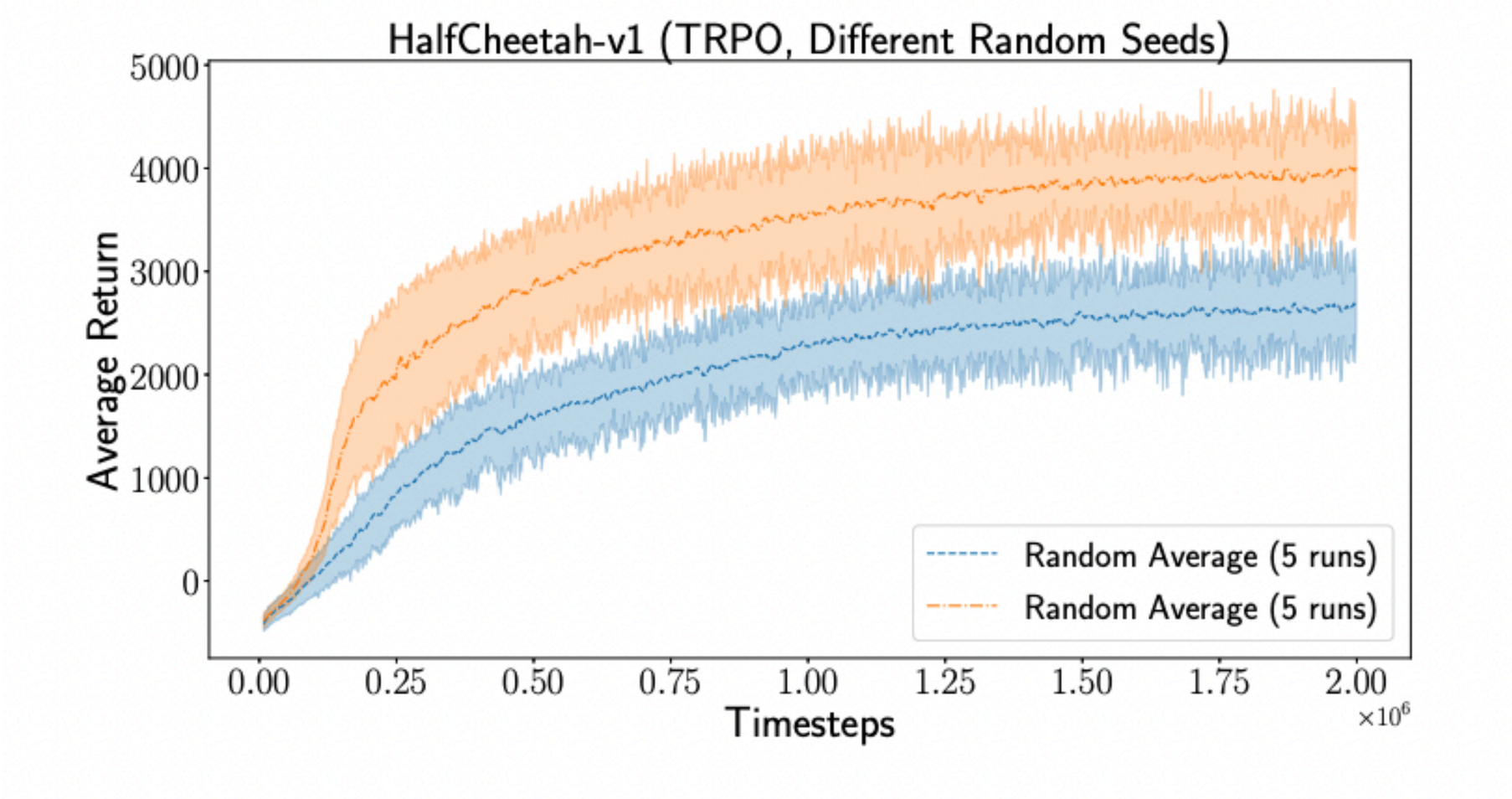
Lecture 22: Evaluation and Reproducibility

Josiah Hanna
University of Wisconsin — Madison

Announcements

- Next week: Multi-agent RL
- No class Thursday! 🎉

Variation in RL



Variation in RL

- Any single run of an RL algorithm may mislead you about expected performance.
- High stochasticity in deep RL algorithms:
  - MDP is stochastic.
  - Exploration policy is usually stochastic.
  - Mini-batch gradient descent with a replay buffer is stochastic.
  - Network weight initialization is stochastic.
  - GPUs have some non-determinism.

Experimental Protocol

- For each algorithm considered, run multiple trials.
- Report aggregate statistics for each algorithm such as mean, median, and mode.
- Report confidence measures such as standard error, %-confidence interval, or interquartile range.
- How many trials?
  - Depends…
  - If computational requirements of experiments are light, can just re-run with a larger number of trials.
  - Otherwise, use power analysis to decide if sample size is sufficient.

Common Random Seeds

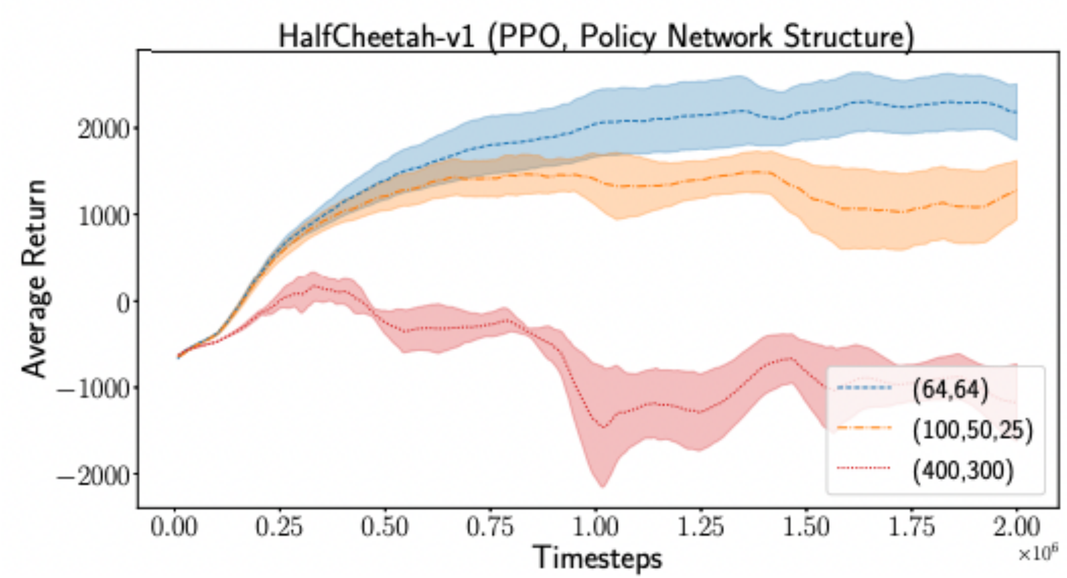
- You plan to compare algorithms A and B in simulated environment X.
- Since data in RL is stochastic, what if algorithm A gets unlucky with observed data?
- To prevent this, fix a set of common random seeds. Ensure that each algorithm is ran once with each seed.

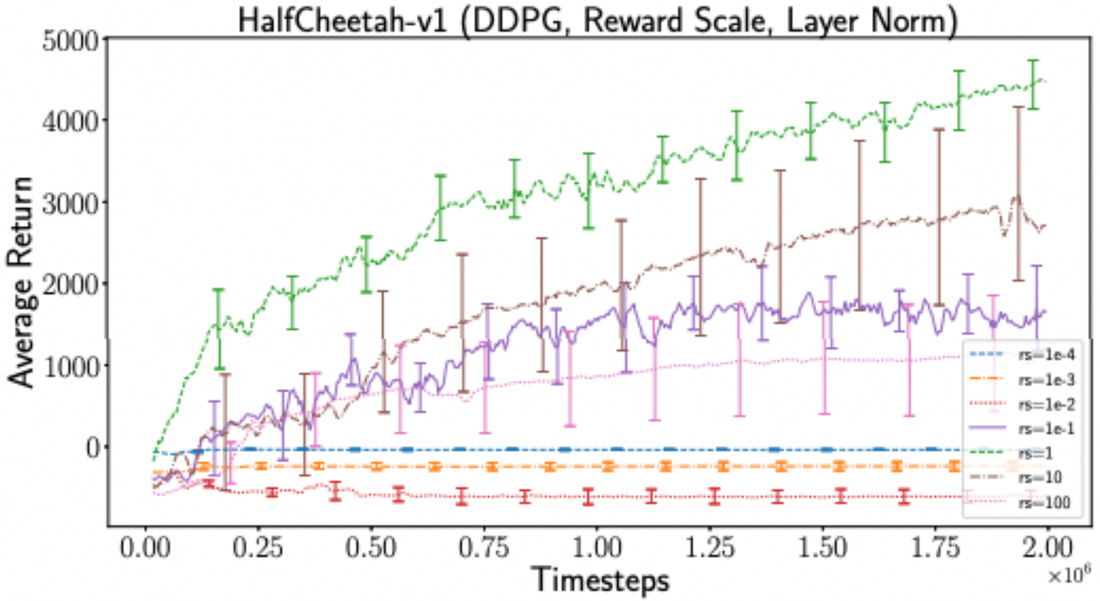
```python
import random

seeds = [random.randint(1e6) for _ in range(n_trials)]

for seed in seeds:
    # Run algorithm A with seed
    ...

    # Run algorithm B with seed
    ...
```

Yunfu’s Presentation

- Slides

Hyper-parameters Matter





Hyper-parameter Selection

- Report how hyper-parameters selected. Why?
  - Best hyper-parameters are problem dependent and so the process is more important than the final hyper-parameter values.
- How to select:
  - Graduate student descent
  - Random search
  - Grid search
  - More advanced: Bayesian optimization, population-based training, neural architecture search.

Community Standards

- What should conference and journal standards for reproducibility be?
  - Reproducibility challenges
  - Reproducibility check-lists
- Culture challenges:
  - Emphasis on positive results and novelty.
  - Low benefit for reproducing work of others and considering incremental questions.

Sharing Code

- Is code sharing sufficient?
  - Not necessarily, may require missing compute and data resources.
- Why do different code-bases give different performances?
  - Deep RL implementations often have hidden tricks that go beyond the base algorithm.
  - Example: observation normalization, i.e., divide observed state variables by a running average of their standard deviation.

Summary

- Deep RL experiments are very stochastic. Makes evaluation of deep RL algorithms a challenge.
- Deep RL algorithms can be sensitive to their hyper-parameters. Reporting these values is necessary for replicability.
- Community standards and values matter for what type of science is done.

Action Items

- Multi-agent RL reading for next week.
- Continue making progress on your final project.