Reinforcement Learning for Goal Recognition

Gabriel Waengertner Henrique

Introduction

- Reinforcement Learning

    - Capable of learning robust policies for environment interaction

    - Policies can be made to follow any desired behavior

- Goal Recognition is an important ability for interaction

    - Intelligent agents implicitly recognize objectives given a behavior

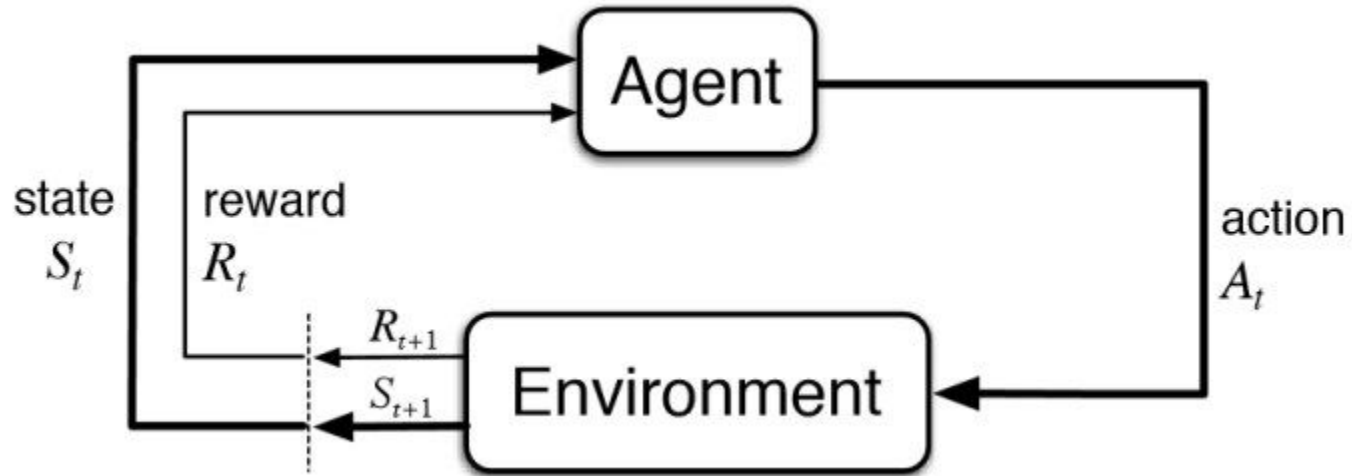   - Observations can be noisy

Introduction

Can we learn policies using reinforcement learning

in order to recognize goals?
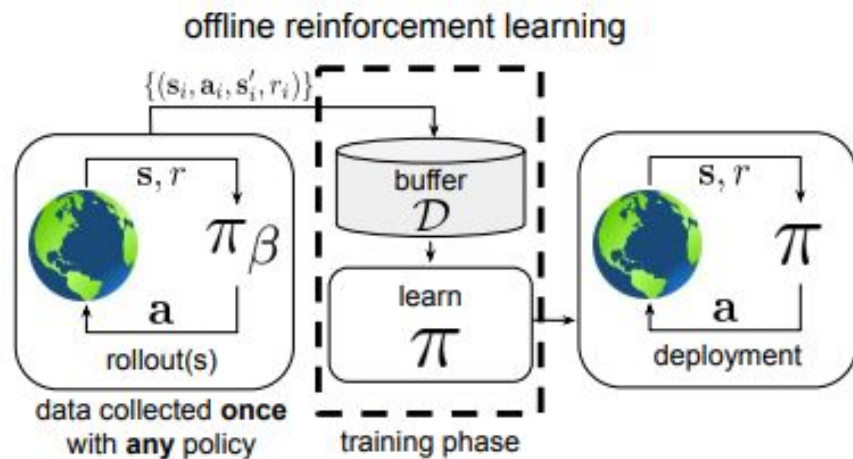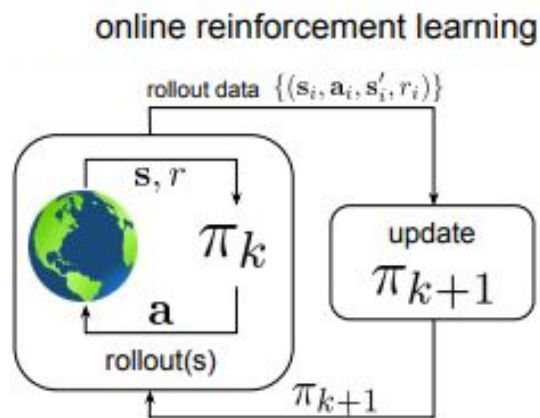
Reinforcement Learning

- A Reinforcement Learning (RL) problem consists of an MDP $\mathcal{M} = \langle \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma \rangle$
  - $\mathcal{S}$ is a finite set of states
  - $\mathcal{A}$ is a finite set of actions
  - $\mathcal{P}$ is a transition function
  - $\mathcal{R}$ is a reward function
  - $\gamma \in [0, 1]$ is a discounting factor
- Learn a policy π that maximizes the expected discounted reward

Reinforcement Learning

Reinforcement Learning

online reinforcement learning

rollout data $\{(\mathbf{s}_i, \mathbf{a}_i, \mathbf{s}'_i, r_i)\}$

$\mathbf{s}, r$

π_k

$\mathbf{a}$

rollout(s)

update π_{k+1}

π_{k+1}

offline reinforcement learning

$\{(\mathbf{s}_i, \mathbf{a}_i, \mathbf{s}'_i, r_i)\}$

$\mathbf{s}, r$

π_β

$\mathbf{a}$

rollout(s)

data collected **once** with **any** policy

buffer $\mathcal{D}$

learn π

training phase

$\mathbf{s}, r$

π

$\mathbf{a}$

deployment

Goal Recognition

- A Goal Recognition (GR) problem is a tuple $P_G = \langle \Xi,\ \mathcal{I},\ \mathcal{G},\ O \rangle$, where
  - Ξ is the domain definition
  - $\mathcal{I}$ is an initial state
  - $\mathcal{G}$ is a finite set of candidate goals
  - O is a sequence of observations $\langle o_1,\ o_2,\ \ldots,\ o_n \rangle$
- The solution to a GR problem is a goal $g \in \mathcal{G}$ that is most consistent with observations O

Goal Recognition

- Observations
  - Noisy
  - Missing
  - Different representations
- Datasets available[1]

[1] https://github.com/pucrs-automated-planning/goal-plan-recognition-dataset

Proposal

- Use Reinforcement Learning as a solution to a Goal Recognition Problem

- Initial steps on a larger study
  - Learn domain definitions using Reinforcement Learning

First Steps

- Train a tabular RL method to learn a different policy for each goal
- Find the most likely goal
    - Extract a trajectory using an off-the-shelf planner
    - Measure distance between policies when following a trajectory
    - KL Divergence
- Toy domains
    - Blocksworld

Schedule

Week 1 - Develop tabular Reinforcement Learning methods to learn policies in small state space domains

Week 2 - Start testing Inverse Reinforcement Learning as a solution to Goal Recognition

Week 3 - Use functions approximators, starting with linear ones

Week 4 - Expand to larger state space domains

Week 5 - Compare with other works and write final paper with results

Thank you!

Reinforcement Learning for Goal Recognition
Gabriel Waengertner Henrique