Robot grocery shopping in partially observable settings

Rodrigo Gomes, Xiaomin Wang, Dustin Tran

May 13, 2015

MIT, 6.834j Cognitive Robotics

Outline

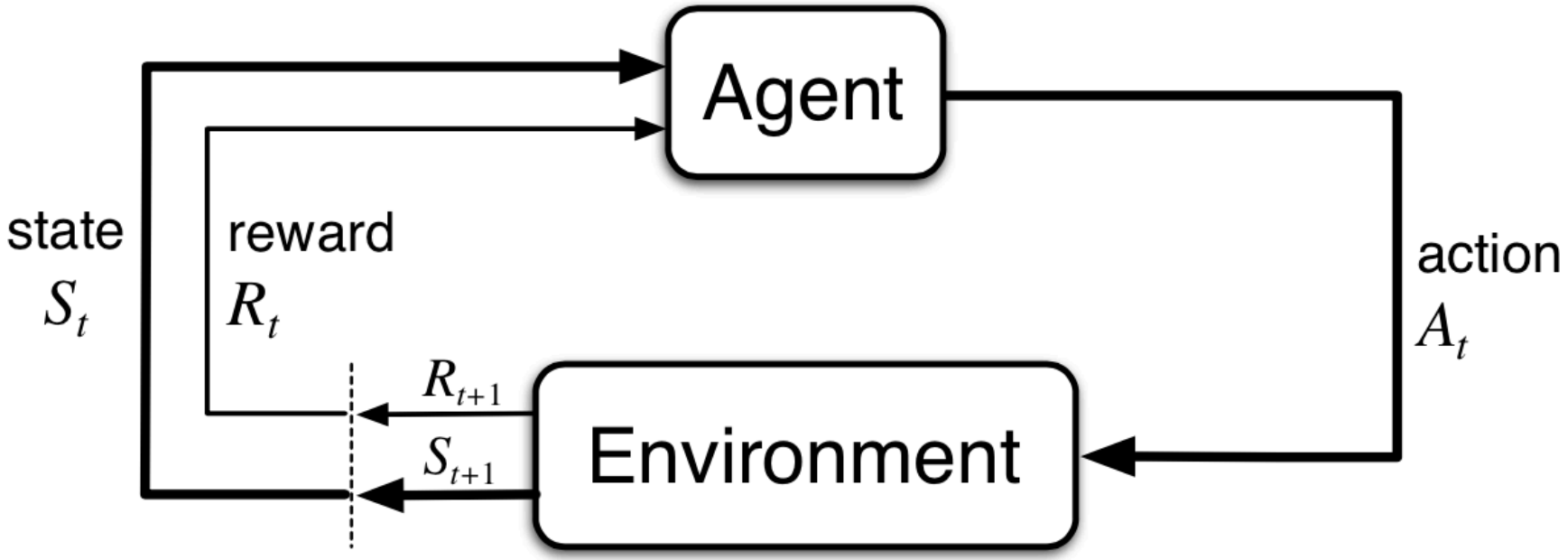
1. (2 min) Background on POMDPs, belief-state MDP, MDP solvers we have
2. (2 min) Setup: Grocery shopping as planning in a POMDP
3. (4 min) Demo
4. (2 min) The solver actually used (value iteration)
5. (1 min) Things that failed (Thompson sampling)
6. (1 min) Q&A

Partially observable Markov decision process (POMDPs)

Collection of objects (S, A, Ω, R, T, O)

- S: state space
- A: action space
- Ω: observation space
- T: transition operator. $T(s' \mid s, a)$ is probability of next state s' given state s and action a
- O: observable operator. $O(o \mid s)$ is probability of observing o given at state s
- $R : S \times A \to \mathbb{R}$ reward function

Partially observable Markov decision process (POMDPs)



re 3.1: The agent–environment interaction in reinforcement learn

Grocery shopping

the task in the POMDP framework

demo

Demo

link from here to a bunch of videos and an interactive demo

Results

the task in the POMDP framework

Play with it!

github.com/dustinvtran/bayesrl