

Robot grocery shopping in partially observable settings

Rodrigo Gomes, Xiaomin Wang, Dustin Tran

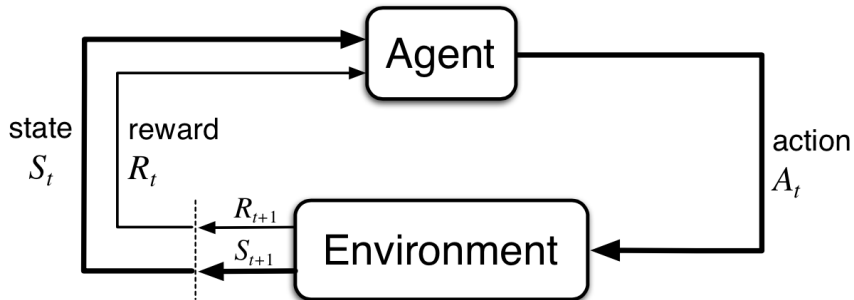
May 13, 2015

MIT, 6.834j Cognitive Robotics

1. Background on POMDPs
2. Grocery shopping as planning in a POMDP
3. Demo!
4. What worked
5. What failed

A *partially observable Markov decision process* (POMDP) is a collection of objects (S, A, Ω, R, T, O)

- S : state space
- A : action space
- Ω : observation space
- $R : S \times A \rightarrow \mathbb{R}$ reward function
- 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



Belief-state MDP

Implemented MDP solvers:

- ☐ Q-learning
- ☐ SARSA
- ☐ R-MAX
- ☐ Thompson sampling

There are a lot!

- ☐ Function approximations with adaptive basis functions
- ☐ BOSS
- ☐ Spectral methods
- ☐ Skill chaining
- ☐ ...

the task in the POMDP framework

how the software works, etc.

demo

how to do value iteration

- ☐ Value iteration as a belief-state MDP
- ☐ Thompson sampling
- ☐ ...

Play with it!



github.com/dustinvtran/bayesrl