POMDP Pacman

Shengliang Xu

Problem Statement

- Implement POMDP algorithms in the Pacman code base

- MDP environment setting
  - A uniform random ghost
  - Pacman can observe its own position and the food status, but can only observe Manhattan distance to the Ghost rather than the position of the ghost

Main Obstacles

- Exact computation of POMDP is impossible if the number of states is more than dozens

- In Pacman setting, a small layout can result in thousands of states
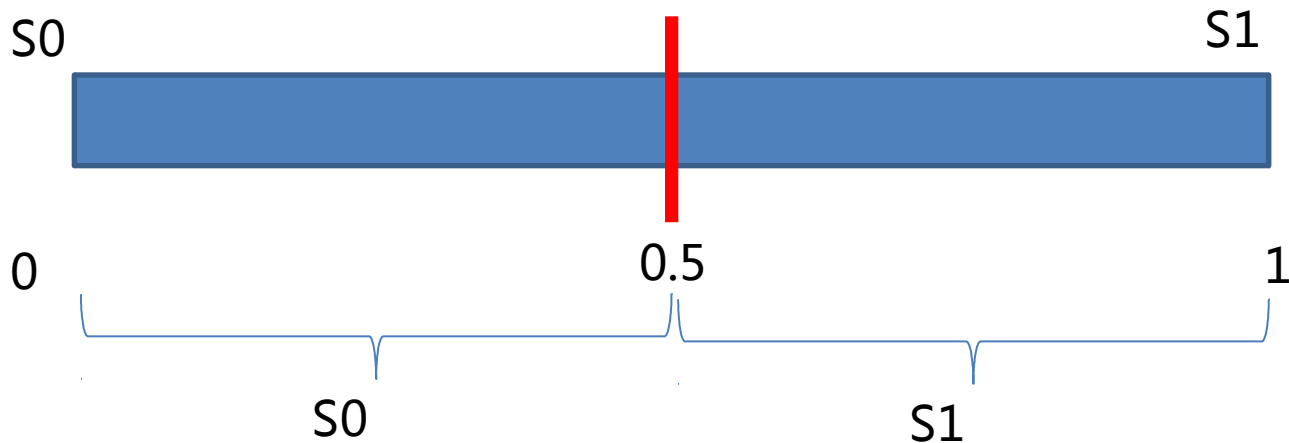
  - ~2600 states

-

-

Proposed Method

PBVI: Point based Value Iteration for POMDP (Pineau, Gordon and Thrun, ijcai03)

- It is able to handle problems with millions of states

Experiment: Baseline

- Adapt MDP directly

  – Linearly and equally discretize the belief simplex, i.e. update belief using m*aximum* posterior *probability*

Experiment: Baseline

- Adapt MDP directly
  - Precompute optimum policy for infinite horizon of the underlying MDP
  - At runtime, keep a belief of the states, update the belief using observations
  - At each step, compute the most likely position of ghost using updated belief, assume the ghost is actually there and get the action from underlying MDP

Experiment: DEMO

- MDP of Pacman

- AdaptivePOMDP of Pacman

-

- PBVI is still in development