Markov Decision Processes

Alice Gao

Lecture 14

Based on work by K. Leyton-Brown, K. Larson, and P. van Beek

Outline

Learning Goals

By the end of the lecture, you should be able to

- Describe/trace value iteration for a Markov Decision Process.

CQ: A stochastic environment

CQ: The robot is in s_{14} and tries to move to the right, what is the probability that the robot stays in s_{14}?

(A) 0.1

(B) 0.2

(C) 0.8

(D) 0.9

(E) 1.0

CQ: A stochastic environment

CQ: True or False: The optimal solution to this problem is the fixed action sequence: down, down, right, right, and right.

(A) True

(B) False

(C) I don't know

CQ: A stochastic environment

CQ: True or False: The fixed action sequence "down, down, right, right, and right" could take us to any square in the environment with positive probability.

(A) True

(B) False

(C) I don't know

CQ: A stochastic environment

CQ: True or False: The solution to this problem should be a fixed sequence of actions. For example, a fixed sequence of actions is down, down, right, right, right.

(A) True

(B) False

(C) I don't know

CQ: The optimal policy

CQ: Take a guess. What do you think is the optimal action in state s_{13}?

(A) Up

(B) Down

(C) Left

(D) Right

Revisiting the Learning Goals

By the end of the lecture, you should be able to

- Describe/trace value iteration for a Markov Decision Process.