Programming Assignment

Report

Student: Zhiwei Han
Matrikel Nr. 03672554

Lehrstuhl für Datenverarbeitung
Technische Universtät München
November 30, 2016

1. Solution with Dynamic Programming

1.1 Find a mapping to linearize the state space into a vector

The approach is just vectorize the 2D maze array and kick out the wall grides.

1.2 Visualize the optimal policy with a 2d plot, where colored boxes indicate the direction of the optimal action

[images] Optimal policy by policy iteration:

images

1.3 Do the two methods generate the same policy?

Yes, from the graphs above we know the optimal policies from value iteration and policy iteration are the same.

2. A Study of Algorithm Properties

2.1 Vary the discounting factor α. Does this have a influence on the final policy?

Yes, the discount factor α decides how further the algorithm looks into future state value. If it is too small, the algorithms cares only about the step cost but ignore future state value.

2.2 Run policy evaluation, policy iteration and value iteration until each have converged. Regard the solution of those runs as the ground truth value function and policy.

[images]

2.3 Now run the three algorithms again and plot the error to the ground truth with respect to the iterations.

[images]

2.4 Run this error plot again for three sensible values of α.

[images]