Artificial Intelligence - Project 2 part 2

Universitat Pompeu Fabra
TA: Nicolás Gastón Rozado

November 16, 2018

After doing Q1-3, answer the following:

1. How many iterations of Value iteration (approximately), would you say we need to run before the Policy converges?
Answer qualitatively, by running "python gridworld.py -a value -i x" where x is the iteration number. Assume it converges when there are no changes after 3 iterations.

2. Look at the output in Q1. Following that policy and settings (noise, gamma, living reward).

>   Will an agent starting in state(0,0) (bottom-left corner) ever fall into the fire pit (-1 red state)?
>   What if it starts in (2,0) (top-left corner)? Explain why.

3. Use "python gridworld.py -a value -i 100 -g DiscountGrid –discount 0.9 –noise 0 –livingReward 0" to test options and answer, following Q3 Grid setting. You can modify the discount, the noise and the living reward.

i. Modify the parameters to get a "suicidal" policy that makes our agent jump off the cliff (at least for cliffs 1, 2 and 4). Explain why does your setting cause that.

ii. Give (if possible) positive values to the parameters noise and discount so that states other than terminal have a negative value. Don't change the Living Reward, fix it to be 0. Explain your answer.

iii. Why do all states except terminal states get a value of 0 with discount = 0 and transition rewards = 0. Can we solve it by increasing the noise? Why?

iv. With noise = 0.9, discount = 0.9 and living reward = 0 we get a pretty weird policy. Why does it choose the action that hits the walls in so many states? Why doesn't always choose the action that leads to terminal states from neighboring ones?