Amit Nijjar

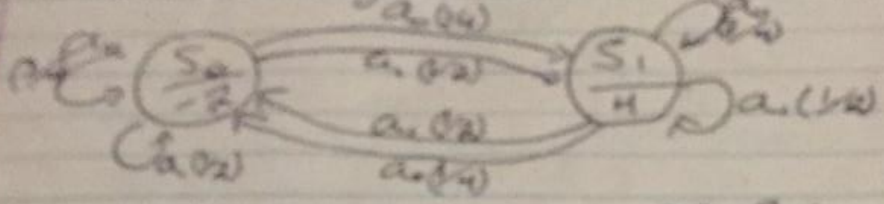
A11489111

CSE 150 Assignment 5

Part 1

CSE 150 Assignment 5

a)



$\gamma = 2/3$

$V^\pi(s) = R(s) + \gamma \sum_{s'} P(s'|s, \pi(s)) V^\pi(s')$

$V^\pi(0) = R(0) + 2/3 \left(V(1) 1/4 + V(0) 3/4 \right)$

$V^\pi(0) = -2 + 1/6 V^\pi(1) + 1/2 (V^\pi(0))$

$3 V^\pi(0) = -12 + V^\pi(1)$

$\boxed{V^\pi(0) = -3/2}$

$V^\pi(1) = R(1) + 2/3 \left(V(1) 3/4 + V(0) 1/4 \right)$

$= 4 + 1/2 V^\pi(1) + 1/6 V^\pi(0)$

$3 V(1) = 24 + V^\pi(0)$

$\boxed{V^\pi(1) = 15/2}$

b) $P(0|0, a=1) V(0) + P(1|0, a=1) V(1) = 3$

s=0 $(V^\pi(0) + 2) 3/2 = 3/4$

a=1 a=0

$\pi'(0) = 1$

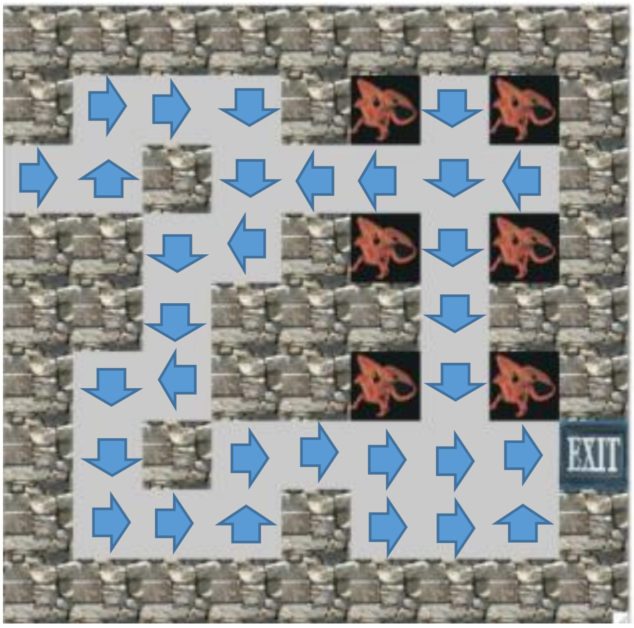
$P(0|1, a=1) V(0) + P(1|1, a=1) V(1) = 3$

$(V^\pi(1) - 4) 3/2 = 21/4$

$\pi'(1) = 0$

$\pi'(0) = 1; \quad \pi'(1) = 0$

Part 2



I made a mdp object to make things simpler. The first part of both files are similar, with different imports.

a) Values:

   a. (3, 71.39423924450648, 'EAST')
   b. (11, 72.98307548141419, 'EAST')
   c. (12, 72.17386511668235, 'NORTH')
   d. (15, 79.8281034704119, 'SOUTH')
   e. (16, 80.72376573333952, 'SOUTH')
   f. (17, 81.62947722483922, 'EAST')
   g. (20, 73.80192201304537, 'EAST')

h. (22, 77.20028225357903, 'SOUTH')

i. (23, 78.06648130841413, 'SOUTH')

j. (24, 78.94237892205466, 'WEST')

k. (26, 82.54535069648871, 'EAST')

l. (29, 74.62997079468701, 'SOUTH')

m. (30, 75.46731058752621, 'SOUTH')

n. (31, 76.34295045844954, 'WEST')

o. (34, 84.40804092457351, 'EAST')

p. (35, 83.4715001645621, 'NORTH')

q. (39, 74.39781781782916, 'WEST')

r. (43, 85.35508956563932, 'EAST')

s. (48, 64.8863097098779, 'WEST')

t. (52, 86.3127639849926, 'EAST')

u. (53, 90.51903456910568, 'EAST')

v. (56, 59.66756995481123, 'SOUTH')

w. (57, 68.94981992730506, 'SOUTH')

x. (58, 70.31432055607144, 'SOUTH')

y. (59, 80.32521158926627, 'SOUTH')

z. (60, 81.47292839949202, 'SOUTH')

aa. (61, 92.20298052160265, 'EAST')

bb. (62, 91.62118640382897, 'EAST')

cc. (66, 59.66756995481123, 'WEST')

dd. (70, 93.67475869133705, 'EAST')

ee. (71, 92.63575448194284, 'NORTH')

ff. (79, 99.9999999999992, 'EAST')

b) Policy:

a. (3, 'EAST')

b. (11, 'EAST')

c. (12, 'NORTH')

d. (15, 'SOUTH')

e. (16, 'SOUTH')

f. (17, 'EAST')

g. (20, 'EAST')

h. (22, 'SOUTH')

i. (23, 'SOUTH')

j. (24, 'WEST')

k. (26, 'EAST')

l. (29, 'SOUTH')

m. (30, 'SOUTH')

n. (31, 'WEST')

o. (34, 'EAST')

p. (35, 'NORTH')

q. (39, 'WEST')

r. (43, 'EAST')

s. (48, 'WEST')

t. (52, 'EAST')

u. (53, 'EAST')

v. (56, 'SOUTH')

w. (57, 'SOUTH')

x. (58, 'SOUTH')

y. (59, 'SOUTH')

z. (60, 'SOUTH')

aa. (61, 'EAST')

bb. (62, 'EAST')

cc. (66, 'WEST')

dd. (70, 'EAST')

ee. (71, 'NORTH')

ff. (79, 'NORTH')

c) The way I computed value is I made two lists and got all the probabilities of all of the states. I copied them to a list and recalculated the probabilities with discount factors to get neighboring states. Then I updated the list so that in the end it would contain the maximum utility per state

d) Implementing policy evaluation I used the policy parameter to get the evaluated utility for each of the individual states

I did this assignment on my own.