CIS 471: Homework 3

Zach Domke

3/14/19

1. * 1. V\*(A) = 85, because we need to add the reward of our next choice (C) and the discounted value of that choice (75)

V\*(C) = 100, because we need to take the reward of our next choice (D) and the discounted value of that choice, which is nothing as it’s a terminal node

π\*(A) = 85, because the policy should be the same as the value in this case. We try to direct Pacman toward the D node as Pacman tries moving to the next highest valued node that it can.

* + 1. V(A) = 63.75, as that is the value achieved from going to the next highest valued node available (node C)

V(C) = 71.67, because it is the sum of the expected values of the options made for Pacman on node C (three best options)

* + 1. V(A) = 65, as Pacman will decide to visit node B instead of node C. Node B has a lower reward but is more guaranteed than node C.
  1. Q\*(s,a) = .5(R(s,a,s1’) + γV(s1’)) + .3(R(s,a,s2’) + γV(s2’)) +.2(R(s,a,s3’) + γV(s3’))

V\*(s) = maxa (.5(R(s,a,s1’) + γV(s1’)) + .3(R(s,a,s2’) + γV(s2’)) +.2(R(s,a,s3’) + γV(s3’)))

(Note that V(s1’) ≥ V(s2’) ≥ V(s3’))



|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| t | st | at | st+1 | rt | (a) Q Values | (b) T(s, a, s’) | (b) R(s, a, s’) |
| 0 | A | Down | B | 2 | Q(A, Down) = 2 | 1 | 2 |
| 1 | A | Down | B | 2 | Q(A, Down) = 2 | 1 | 2 |
| 2 | A | Down | C | -3 | Q(A, Down) = -.5 | .5 | -3 |
| 3 | B | Down | C | 3 | Q(B, Down) = 1.5 | .75 | 3 |
| 4 | B | Down | B | 0 | Q(B, Down) = 1.125 | .625 | 0 |
| 5 | B | Up | A | 1 | Q(B, Up) = .5 | .6875 | 1 |
| 6 | C | Down | B | 2 | Q(C, Down) = 1.282 | .65 | 2 |
| 7 | C | Up | B | -2 | Q(C, Up) = -.719 | .657 | -2 |
| 8 | A | Up | C | 1 | Q(A, Up) = .821 | .67 | 1 |
| 9 | C | Down | A | 2 | Q(C, Down) = 1.846 | .68 | 2 |
| 10 | A | Up | B | 3 | Q(A, Up) = 2.192 | .675 | 3 |

|  |  |  |
| --- | --- | --- |
| State | Action | (a) Q Value |
| A | Up | 2.192 |
| A | Down | -.5 |
| B | Up | .5 |
| B | Down | 1.125 |
| C | Up | -.719 |
| C | Down | 1.846 |

* 1. 1. V(s) =

π(s) =

1. 1. Probability that you chose coin 2 is

P(+sequence|+coin2) = .008789 or 9/1024

P(+coin2|+sequence) = (P(+seq|+c2) \* P(+c2)) / (P(+seq)) ->

= (P(+seq|+c2) \* P(+c2)) / (P(+seq|+c2) \* P(+c2) + P(+seq|-c2) \* P(-c2)) ->

P(+coin2|+sequence) = 0.002838 = .28% = (96/33829)



10. P(X = 1, Y = 1) = 63/80, because the sum of all other P(Y = 1) = 17/80

11. P(X = 1, Y = 2) = 3/40, because the sum of all other P(Y = 2) = 1/4

12. P(X = 1, Y = 3) = 7/80, because the sum of all other P(Y = 3) = 1/8

13. P(X = 1, Y = 4) = 1/20, because the sum of all other P(Y = 4) = 1/2

14. P(X = 2, Y = 1) = 13/20, because the sum of all other P(X = 2) = 7/20

15. P(X = 3, Y = 1) = 23/30, because the sum of all other P(X = 3) = 7/30

16. P(X = 4, Y = 1) = 17/24, because the sum of all other P(X = 4) = 7/24

1. 1. 1. No, it does not hold
      2. No, it does not hold
      3. Yes, it does hold
      4. No, it does not hold
      5. Yes, it does hold
      6. Yes, it does hold
      7. No, it does not hold
   2. 1. P(X6) P(x4|X6) P(X5|X6) P(x3) P(x2|x3, x4, X5) P(X1|x2, x3)

Choose: X6

Multiply: P(X6) P(x4|X6) P(X5|X6) -> P(x4, X5|X6)

Sum out: X6 -> P(x4, X5)

P(x4, X5) P(x3) P(x2|x3, x4, X5) P(X1|x2, x3)

Choose: X5

Multiply: P(x4, X5) P(x2|x3, x4, X5) -> P(x2|x3, x4, X5)

Sum out: X5 -> P(x2|x3, x4)

P(x3) P(x2|x3, x4) P(X1|x2, x3)

Finish with: X1

Multiply: P(x3) P(x2|x3, x4) P(X1|x2, x3) -> P(x2, x3, x4, X1)

Normalize: P(x2, x3, x4, X1)-> P(X1|x2, x3, x4)

* + 1. P(X6) P(x4|X6) P(x5|X6) P(x3) P(X2|x3, x4, x5) P(X1|X2, x3)

Choose: X6

Multiply: P(X6) P(x4|X6) P(x5|X6) -> P(x4, x5|X6)

Sum out: X6 -> P(x4, x5)

P(x4, x5) P(x3) P(X2|x3, x4, x5) P(X1|X2, x3)

Finish with: X1, X2

Multiply: P(x4, x5) P(x3) P(X2|x3, x4, x5) P(X1|X2, x3) -> P(x3, x4, x5, X1, X2)

Normalize: P(x3, x4, x5, X1, X2) -> (X1, X2|x3, x4, x5)

* + 1. P(x6) P(X4|x6) P(X5|x6) P(X3) P(X2|X3, X4, X5) P(X1|X2, X3)

Choose: X5

Multiply: P(X5|x6) P(X2|X3, X4, X5) -> P(X2, X5|x6, X3, X4)

Sum out: X5 -> P(X2|x6, X3, X4)

P(x6) P(X4|x6) P(X3) P(X2|x6, X3, X4) P(X1|X2, X3)

Choose: X4

Multiply: P(X4|x6) P(X2|x6, X3, X4) -> P(X2, X4| x6, X3)

Sum out: X4 -> P(X2|x6, X3)

P(x6) P(X2|x6, X3) P(X1|X2, X3)

Choose: X3

Multiply: P(X2|x6, X3) P(X1|X2, X3) -> P(X1, X2|x6, X3)

Sum out: X3 -> P(X1, X2|x6)

P(x6) P(X1, X2|x6)

Finish with: X1, X2

Multiply: P(x6) P(X1, X2|x6) -> P(X1, X2, x6)

Normalize: P(X1, X2, x6) -> P(X1, X2|x6)