Assign 2 Part 1 2019114009

(col, row) notation

①

from state	action	(1,1)	(2,1)	(1,2)	(2,2)
(1,1)	right	0.2	0.8	0	0
	up	0.2	0	0.8	0
(1,2)	right	0	0	0.75	0.25
	down	0.8	0	0.2	0
(2,1)	left	0.8	0.2	0	0
	up	0	0.2	0	0.8

② ACR — Probab → 0.8 × 0.25 = 0.2
└ cost → −4

ABR → Probab → 0.8 × 0.8 = 0.64
└ cost → −5

here we can see probability wise the optimal solution is ABR while cost wise the optimal solⁿ is ACR. Also this probabilities are given to the respective paths.

if we choose ACR instead of ABR then is 20% drop in cost. but probability goes down from 64% to 20%.
So, i think ABR is the best path.

(3) $U_0(A) = 0 \quad U_0(B) = 0 \quad U_0(C) = 0$

$U(R) = 16.5$

Roll no. = 2019114009

0	16.5
0	0

$$U_1(A) = \max_a \left[R(A,a) + \gamma \sum_j U_0(j) P(j \mid A,a) \right)$$

$a = \{ up, right \}$

$$R(A,a) = R(A,a,j) * \sum P(j \mid A,a) \qquad j = A, B, C, R$$

$$R(A, up) = 0.8 \times -1 + 0.2 \times -1 \qquad = -1$$

$$R(A, right) = 0.2 \times -1 + 0.8 \times -1 \qquad = -1$$

$$U_1(A) = \max_a \left[\cancel{B}(-1+0), \; \cancel{0+1} \; (-1+0) \right]$$

$$= -1$$

Page No.
Date:
www.mirejmulticolour.com

$U_1(B) = \max_a \left[R(B,a) + \gamma \sum_j U_0(j) P(j \mid B,a) \right]$

$a = \{left, up\}$

~~R(B,A) =~~

$R(B, left) = 0.8 \times -1 + 0.2 \times -1 = -1$

$R(B, up) = 0.8 \times -4 + 0.2 \times -1 =$
$\cancel{0.32} - 0.2$
-3.2
$= -3.4$

$U_1(B) = \max_a \left[(-3.4 + 16.5 \times 0.8 \times 0.2), (-1 + 16.5 \times 0 \times 0.2) \right]$

$\cancel{-3.4+}$ 0, $\cancel{-3.4}$.
$(-0.76, -1)$

$U_1(B) = -0.76$

~~$U_1(C) = \max_a [-1.16, -1.14]$.~~

$U_1(C) = \max_a \left[R(C,a) + \gamma \sum_j U_0(j) P(j \mid C,a) \right]$

$a = \{Right, down\}$

$R(C, right) = 0.25 \times -3 + 0.75 \times -1$
$= -1.5$

$R(C, down) = 0.8 \times -1 + 0.2 \times -2$
$= -1$

Page No.
Date:

$U_1(C) = \cancel{\max[(-1.16), (-1.14)]}$

$\max[(-1.0), (-0.67)]$

$= -0.67$

iter-1 →

-0.67	16.5
-1	-0.76

iter-2

$U_2(A)$

$U_{t+1}(i) = \max_a \left[R(i,a) + \sum_j U_t(j) P(j \mid i,a) \right]$

$U_2(A) = \max_a \left[R(A,a) + \sum_j U_1(j) P(j \mid A,a) \right]$

$a = \{ up, right \}$

$U_2(A) = \max [(-1.16), (-1.14)]$

$\underline{U_2(A) = -1.14}$

$U_2(B) = \max_a \left[R(B,a) + \sum_j U_1(j) P(j \mid B,a) \right]$

$a = \{ up, left \}$

$U_2(B) = \max[-1.19, -0.79]$

$U_2(B) = -0.79$

$U_2(C) = \max_a \left[R(C,a) + \sum_j U(j) P(j \mid C,a) \right]$

$a = \{ \text{right}, \text{down} \}$

$= \max [-1.18, -0.77]$

$U_2(C) = -0.77$

-0.77	16.5
-1.14	-0.79

iter 3

$U_3(A) = \max_a \left[R(A,a) + \sum_j U(j) P(j \mid A,a) \right]$

$= \max [-1.172, -1.1701]$

$U_3(A) = -1.1701$

$U_3(B) = \max [-1.21, -0.79]$

$= -0.79$

$U_3(C) = \max [-1.21, -0.791]$

$= -0.791$

-0.79	16.5
-1.17	-0.79

(Bellman error) delta of A and C is greater than 0.01 so continue

Date:

iter 4

$U_4(A) = \max[\cancel{-1.47}, -1.1734348, -1.17346]$

$= -1.173434$

$U_4(B) = \max[-1.21, -0.79]$

$= -0.79$

$U_4(C) = \max[-1.21, -0.79]$

$= -0.79$

-0.793715625

-0.79	16.5
-1.173 ~~-1.173~~	-0.79

→ -0.791664639

now Bellman error is less than 0.01 in all states. So ~~to~~ convergence.

Ans 4

If you are at A, next highest utility is B, so.

A → B → R would be best acc. to VI

Yes, our privious guess was correct. :)

Page No.
Date:

Ans 5 ~~B~~. As we calculated the values of $U_t(B)$, $U_t(A)$ and $U_t(C)$, $U_t(R)$ in ques - 3.

they are dependent on the R-value (Reward)

So what i observe is that when the value of R is more then $A \to B \to R$ is more likely whereas when $A \to C \to R$ is more likely in lesser values of R.

Also if you calculate for some more values of R then you will see for smaller values the best path would be $A \to C \to R$ ~~and~~ whereas in larger values the path will be $A \to B \to R$.

So, obviously there ~~should~~ will be some value of R for which the cost of the paths $A \to C \to R$ and $A \to B \to R$ will be same.