HW2

2023年3月3日 16:02

Problem 1

1.

$Q^1(\mathrm{A},1)$	$Q^1(\mathrm{A},2)$	$Q^1(\mathrm{A},3)$	$Q^1(\mathrm{B},1)$	$Q^1(\mathrm{B},2)$	$Q^1(\mathrm{B},3)$
0	2	0	6	0	3
$Q^2(\mathrm{A},1)$	$Q^2(\mathrm{A},2)$	$Q^2(\mathrm{A},3)$	$Q^2(\mathrm{B},1)$	$Q^2(\mathrm{B},2)$	$Q^2(\mathrm{B},3)$
6	4	4	10	2	7

2.

$i=1$, $Q^i=Q^1$, $\pi^{i+1}=2$

For $\pi^2(A)$

$\forall s \in S$, $\pi^{i+1}(s) = \arg\max_a Q^i(s,a)$

$= \arg\max(Q^1(A,1), Q^1(A,2), Q^1(A,3))$

$= \arg\max(0,2,0)$

$= Q^1(A,2) \Rightarrow$ optimal action is 2

For $\pi^3(B)$

$\forall s \in S$, $\pi^{i+1}(s) = \arg\max_a Q^i(s,a)$

$= \arg\max(Q^1(B,1), Q^1(B,2), Q^1(B,3))$

$= \arg\max(6,0,3)$

$= Q^1(B,1) \Rightarrow$ optimal action is 1

For $\pi^3(A)$

$\forall s \in S$, $\pi^{i+1}(s) = \arg\max_a Q^i(s,a)$

$= \arg\max(Q^2(A,1), Q^2(A,2), Q^2(A,3))$

$= \arg\max(6,4,4)$

$= Q^1(A,1) \Rightarrow$ optimal action is 1

For $\pi^3(B)$

For $\pi^3(B)$

$\forall s \in S,\ \pi^{i+1}(s) = \text{argmax}\ Q^i(s,a)$

$= \text{argmax}(Q^2(B,1), Q^2(B,2), Q^2(B,3))$

$= \text{argmax}(10, 2, 7)$

$= Q^1(B,1) \Rightarrow$ optimal action is 1

Problem 2.

$Q(s_1,\uparrow) = (1-\alpha) * Q(s,\uparrow) + \alpha(r + \gamma * \max(s_4, a))$

$= 0.2 \times 4 + 0.8(0 + 0.3 \times 8)$

$= 2.72$

$Q(s_4,\rightarrow) = (1-\alpha) * Q(s,\rightarrow) + \alpha(r + \gamma * \max(s_5, a))$

$= 0.2 \times 8 + 0.8(0 + 0.3 \times 16)$

$= 5.44$

$Q(s_5,\downarrow) = (1-\alpha) * Q(s,\downarrow) + \alpha(r + \gamma * \max(s_2, a))$

$= 0.2 \times 4 + 0.8(0 + 10 \times 0.3)$

$= 3.2$

$Q(s_2,\rightarrow) = (1-\alpha) * Q(s,\rightarrow) + \alpha(r + \gamma * \max(s_3, a))$

$= 0.2 \times 10 + 0.8(0 + 20 \times 0.3)$

$= 6.8$

$Q(s_3,\uparrow) = (1-\alpha) * Q(s,\uparrow) + \alpha(r + \gamma * \max(s_6, a))$

$= 0.2 \times 20 + 0.8 \times (20 + 0)$

$= 20$