HW1

Thursday, 2 February 2023 4:19 AM

Problem 1

For $p(s', r / s, a)$ the colums are s, a, s', r

Hence, from the table given, we can deduce the following:

Since the rewards are deterministic, probability distribution will be,

s	a	s'	r	$p(s', r \mid s, a)$
high	search	high	r_{search}	α
high	search	low	r_{search}	$1-\alpha$
low	search	high	-3	$1-\beta$
low	search	low	r_{search}	β
high	wait	high	r_{wait}	1
high	wait	low	–	0
low	wait	high	–	0
low	wait	low	r_{wait}	1
low	recharge	high	0	1
low	recharge	low	–	0

For $p(s', r / s, a) > 0$, we will have

s	a	s'	r	$p(s', r \mid s, a)$
high	search	high	r_{search}	α
high	search	low	r_{search}	$1-\alpha$
low	search	high	-3	$1-\beta$
low	search	low	r_{search}	β
high	wait	high	r_{wait}	1
low	wait	low	r_{wait}	1
low	recharge	high	0	1

Ans

Problem 2

$$p(s' \mid s, a) = P_r\{S_t = s' \mid S_{t-1} = s, A_{t-1} = a\} = \sum_{r \in R} p(s', r \mid s, a)$$

& $$r(s, a) = E[R_t \mid S_{t-1} = s, A_{t-1} = a] = \sum_{r \in R} r \sum_{s' \in S} p(s', r \mid s, a)$$

We can write the 4 Bellman equations for 4 value functions (v_π, v_*, q_π, q_*) as follows:

$$v_\pi(s) = E_\pi[G_t \mid S_t = s]$$

$$\Rightarrow \quad v_\pi(s) = \sum_a \pi(a|s)\left(R_s^a + \gamma \sum_{s'} P_{ss'}^{a'} v_\pi(s')\right)$$

$$\Rightarrow \quad \boxed{v_\pi(s) = \sum_{a \in A} \pi(a|s)\left[r(s,a) + \gamma \sum_{s' \in S} p(s'|s,a)\, v_\pi(s')\right]}$$

$$\begin{aligned} v_*(s) &= \max_a q_*(s,a) \\ &= \max_a \left(R_s^a + \gamma \sum_{s' \in S} P_{ss'}^a v_*(s')\right) \end{aligned}$$

$$\boxed{v_*(s) = \max_a \left[r(s,a) + \gamma \sum_{s'} p(s'|s,a)\, v_*(s')\right]}$$

$$q_\pi(s,a) = E_\pi[G_t \mid S_t = s, A_t = a]$$

$$\Rightarrow \quad = R_s^a + \gamma \sum_{s' \in S} P_{ss'}^a \sum_{a' \in A} \pi(a'|s')\, q_\pi(s',a')$$

$$\Rightarrow \quad \boxed{q_\pi(s,a) = r(s,a) + \gamma \sum_{s'} p(s'|s,a) \sum_{a'} \pi(a'|s')\, q_\pi(s',a')}$$

$$\begin{aligned} q_*(s,a) &= R_s^a + \gamma \sum_{s' \in S} P_{ss'}^a v_*(s') \\ &= R_s^a + \gamma \sum_{s'} P_{ss'}^a \max_{a'} q_*(s',a') \end{aligned}$$

$$\Rightarrow \quad \boxed{q_*(s,a) = r(s,a) + \gamma \sum_{s'} p(s'|s,a) \max_{a'} q_*(s',a')}$$