

9.1

$$\sum_{k \geq t} r^k r_k \leq \sum_{k \geq t} r^k$$

$$\begin{aligned} \sum_{k \geq t} r^k &= \frac{r^t}{1-r} \\ &= \frac{e^{t \log r}}{1-r} \end{aligned}$$

$$\therefore \log r < r-1$$

$$\frac{e^{t \log r}}{1-r} \leq \frac{e^{t(r-1)}}{1-r}$$

$$\frac{e^{t(r-1)}}{1-r} = \frac{e^{-t(1-r)}}{1-r} = h e^{-t/h}$$

$$\therefore \sum_{k \geq t} r^k r_k \leq h e^{-t/h}$$

9.2

$$\begin{aligned} (a) \textcircled{1} V^\pi(s=2) &= R(2) + \gamma P(s'=1 | s=2, a=\uparrow) V^\pi(1) \\ &\quad + \gamma P(s'=2 | s=2, a=\uparrow) V^\pi(2) \\ &\quad + \gamma P(s'=3 | s=2, a=\uparrow) V^\pi(3) \\ &= 30 + \frac{2}{3} \cdot \frac{1}{2} \cdot (-18) + \frac{2}{3} \cdot \frac{1}{2} V^\pi(s=2) \\ &= 24 + \frac{1}{3} V^\pi(s=2) \end{aligned}$$

$$\boxed{V^\pi(s=2) = 36}$$

$$\begin{aligned} \textcircled{2} V^\pi(s=3) &= R(3) + \gamma P(s'=1 | s=3, a=\downarrow) V^\pi(1) \\ &\quad + \gamma P(s'=2 | s=3, a=\downarrow) V^\pi(2) \\ &\quad + \gamma P(s'=3 | s=3, a=\downarrow) V^\pi(3) \\ &= -25 + \frac{2}{3} \cdot \frac{1}{4} \cdot (30) + \frac{2}{3} \cdot \frac{3}{4} V^\pi(s=3) \\ &= -20 + \frac{1}{2} V^\pi(s=3) \end{aligned}$$

$$\boxed{V^\pi(s=3) = -38}$$

(b)

from a , we know $V^\pi(1), V^\pi(2), V^\pi(3)$

$$\Rightarrow Q^\pi(s=1, a=\uparrow) = V^\pi(1) = -18$$

$$Q^\pi(s=2, a=\uparrow) = V^\pi(2) = 36$$

$$Q^\pi(s=3, a=\downarrow) = V^\pi(3) = -38$$

$$\begin{aligned} \textcircled{1} Q^\pi(s=3, a=\uparrow) &= \gamma P(s'=1 | s=3, a=\uparrow) V^\pi(1) \\ &\quad + \gamma P(s'=2 | s=3, a=\uparrow) V^\pi(2) \\ &\quad + \gamma P(s'=3 | s=3, a=\uparrow) V^\pi(3) \\ &= -25 + \frac{2}{3} \cdot \frac{3}{4} \cdot (36) + \frac{2}{3} \cdot \frac{1}{4} \cdot (-38) \\ &= -\frac{40}{3} \end{aligned}$$

$$\begin{aligned} \textcircled{2} Q^\pi(s=1, a=\downarrow) &= \gamma P(s'=1 | s=1, a=\downarrow) V^\pi(1) \\ &\quad + \gamma P(s'=2 | s=1, a=\downarrow) V^\pi(2) \\ &\quad + \gamma P(s'=3 | s=1, a=\downarrow) V^\pi(3) \\ &= -15 + \frac{2}{3} \cdot \frac{1}{4} \cdot (-18) + \frac{2}{3} \cdot \frac{3}{4} \cdot (36) \\ &= 0 \end{aligned}$$

③

$$\begin{aligned} Q^{\pi}(s=2, a=\downarrow) &= R(2) + \gamma P(s'=1 | s=2, a=\downarrow) V^{\pi}(1) \\ &\quad + \gamma P(s'=2 | s=2, a=\downarrow) V^{\pi}(2) \\ &\quad + \gamma P(s'=3 | s=2, a=\downarrow) V^{\pi}(3) \\ &= 30 + \frac{2}{3} \cdot \frac{1}{2} \cdot (36) + \frac{2}{3} \cdot \frac{1}{2} \cdot (38) \\ &= \frac{88}{3} \end{aligned}$$

$$\pi'(1) = \operatorname{argmax} Q^{\pi}(s=1, a) = \downarrow$$

$$\pi'(2) = \operatorname{argmax} Q^{\pi}(s=2, a) = \uparrow$$

$$\pi'(3) = \operatorname{argmax} Q^{\pi}(s=3, a) = \uparrow$$

9.3

$$(a) V^\pi(s) = R(s) + \gamma \sum_{s'=0} P(s' | \pi(s)) V^\pi(s')$$

$$V^\pi(s) = s + \gamma \left(\frac{2}{3} V^\pi(s) + \frac{1}{3} V^\pi(s+1) \right)$$

$$(b) as+b = s + \gamma \left(\frac{2}{3} (as+b) + \frac{1}{3} (a(s+1)+b) \right)$$

$$\Rightarrow as+b = \left(1 + \frac{2}{3} a\gamma + \frac{1}{3} a\gamma \right) s + \gamma \left(b + \frac{a}{3} \right)$$

$$\Rightarrow a = 1 + a\gamma \Rightarrow a = \frac{1}{1-\gamma}$$

$$b = \gamma \left(b + \frac{a}{3} \right) \Rightarrow b = \frac{1}{(1-\gamma)} \times \frac{1}{3} \frac{\gamma}{(1-\gamma)}$$

$$= \frac{\gamma}{3(1-\gamma)^2}$$