9.1

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

$$\sum_{k\geq t}^{\infty} r^k = \frac{r^t}{1-r}$$

$$= \frac{e^{t\log r}}{1-r}$$

$\because \log r < r-1$

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

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

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

9.2

(a) ① $V^{\pi}(S=2) = R(2) + \gamma P(s'=1 \mid s=2, a=\uparrow) V^{\pi}(1)$

$+ \gamma P(s'=2 \mid s=2, a=\uparrow) V^{\pi}(2)$

$+ \gamma P(s'=3 \mid s=2, a=\uparrow) V^{\pi}(3)$

$= 30 + \frac{2}{3} \cdot \frac{1}{2} \cdot (-18) + \frac{2}{3} \cdot \frac{1}{2} V^{\pi}(S=2)$

$= 24 + \frac{1}{3} V^{\pi}(S=2)$

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

② $V^{\pi}(S=3) = R(3) + \gamma P(s'=1 \mid s=3, a=\downarrow) V^{\pi}(1)$

$+ \gamma P(s'=2 \mid s=3, a=\downarrow) V^{\pi}(2)$

$+ \gamma P(s'=3 \mid s=3, a=\downarrow) V^{\pi}(3)$

$= -25 + \frac{2}{3} \cdot \frac{1}{4} \cdot (30) + \frac{2}{3} \cdot \frac{3}{4} V^{\pi}(S=3)$

$= -20 + \frac{1}{2} V^{\pi}(S=3)$

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

(b)

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

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

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

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

① $Q^{\pi}(s=3, a=\uparrow) = \gamma P(s'=1 \mid s=3, a=\uparrow) V^{\pi}(1)$

$+ \gamma P(s'=2 \mid s=3, a=\uparrow) V^{\pi}(2)$

$+ \gamma P(s'=3 \mid s=3, a=\uparrow) V^{\pi}(3)$

$= -25 + \frac{2}{3} \cdot \frac{3}{4} \cdot (36) + \frac{2}{3} \cdot \frac{1}{4} (-38)$

$= -\frac{40}{3}$

② $Q^{\pi}(s=1, a=\downarrow) = \gamma P(s'=1 \mid s=1, a=\downarrow) V^{\pi}(1)$

$+ \gamma P(s'=2 \mid s=1, a=\downarrow) V^{\pi}(2)$

$+ \gamma P(s'=3 \mid s=1, a=\downarrow) V^{\pi}(3)$

$= -15 + \frac{2}{3} \cdot \frac{1}{4} \cdot (-18) + \frac{2}{3} \cdot \frac{3}{4} (36)$

$= 0$

(3)

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

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

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

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

9.3

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

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

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

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

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

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

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