

$$\begin{aligned}
 Q3 \\
 a) \quad V_R(s) &= \mathbb{E} \left[\sum_{\pi \text{ aint } \Pi(\cdot|s)} \gamma^i R(s_{t+i}, a_{t+i}, s_{t+i+1}) \mid s_t = s \right] \\
 &\quad \left. \longleftrightarrow \text{this sums from } i=0 \text{ to } \infty \right]
 \end{aligned}$$

$$\begin{aligned}
 &= \mathbb{E} \left[R(s_t, a_t, s_{t+1}) + \gamma \mathbb{E} \left[\sum_{i=2}^{\infty} \gamma^i R(s_{t+i}, a_{t+i}, s_{t+i+1}) \mid s_t = s \right] \right]
 \end{aligned}$$

$$\mathbb{E} \left[\gamma^2 \sum_{i=0}^{\infty} \gamma^i R(s_{t+2+i}, a_{t+2+i}, a_{t+2+i+1}) \mid s_{t+2} = s_{t+2} \right]$$

$$= \gamma^2 \mathbb{E} \left[\sum_{i=0}^{\infty} \gamma^i R(s_{t+2+i}, a_{t+2+i}, a_{t+2+(i+1)}) \mid s_{t+2} \right]$$

$$\Rightarrow \gamma^2 V_R(s_{t+2})$$

$$\begin{aligned}
 \therefore V_R(s) &= \mathbb{E} [R(s_t, a_t, s_{t+1}) + \gamma R(s_{t+1}, a_{t+1}, s_{t+2}) \\
 &\quad + \gamma^2 V_R(s_{t+2}) \mid s_{t+2} = s]
 \end{aligned}$$