

$$\pi(a^c | S) = 1, 0 \text{ if } A \in \{a^A, a^B\}$$

$$b(a^x | s^x) = 0.5, \text{ else } 0.25$$

e_1

$$S_0 = S^A, A_0 = a^B, R_1 = 10$$

$$S_1 = S^B, A_1 = a^C, R_2 = 10$$

$$S_2 = S^C, A_2 = a^A, R_3 = 90$$

$$S_3 = S^A, A_3 = a^C, R_4 = 30$$

e_2

$$S_0 = S^C, A_0 = a^A, R_1 = 50$$

$$S_1 = S^C, A_1 = a^C, R_2 = 70$$

$$S_2 = S^C, A_2 = a^B, R_3 = 10$$

$$S_3 = S^B, A_3 = a^B, R_4 = 20$$

$$P'_0 = \frac{\pi(a^B | S^A)}{b(a^B | S^A)} = \frac{0}{0.25} = 0$$

$$P'_1 = \frac{\pi(a^C | S^B)}{b(a^C | S^B)} = \frac{1}{.25} = 4$$

$$P'_2 = \frac{\pi(a^A | S^C)}{b(a^A | S^C)} = \frac{0}{.25} = 0$$

$$P'_3 = \frac{\pi(a^C | S^A)}{b(a^C | S^A)} = \frac{1}{.25} = 4$$

$$P^2_0 = \frac{\pi(a^A | S^C)}{b(a^A | S^C)} = \frac{0}{.25} = 0$$

$$P^2_1 = \frac{\pi(a^C | S^C)}{b(a^C | S^C)} = \frac{1}{.5} = 2$$

$$P^2_2 = \frac{\pi(a^B | S^C)}{b(a^B | S^C)} = 0$$

$$P^2_3 = \frac{\pi(a^B | S^B)}{b(a^B | S^B)} = 0$$

Note: I am using $V_\pi(s) = \sum P G$

I was ~~penalized~~ Penalized for using the average / weighted importance

Sampling formula on assignment 5: $(\frac{1}{\sum P} \cdot \sum P \cdot G)$

$$\cancel{P_{t:T-1} = \prod_{k=t}^{T-1} \frac{\pi(A_k | S_k)}{b(A_k | S_k)}}$$

$$P_{t:T-1} = \prod_{k=t}^{T-1} \frac{\pi(A_k | S_k)}{b(A_k | S_k)}$$

Andrew Cade
M. Stern

$$V_{\pi}(s) = \sum p \cdot G$$

$$V_{\pi}(s^A) = \sum p_{0:T}^1 \cdot G_0^1 + p_{1:T}^1 \cdot G_1^1 = (0 \cdot 4 \cdot 0 \cdot 4) [(20 + 8 \cdot 10 + 8^2 \cdot 40 + 8^3 \cdot 30) + 30] \cdot \frac{1}{2} \\ + 4 \cdot 30$$

$$\boxed{V_{\pi}(s^A) = 120}$$

$$V_{\pi}(s^B) = p_{1:T}^1 \cdot G_1^1 + p_{3:T}^2 \cdot G_3^2 = p_{1:T}^1 \cdot G_1^1 \\ \text{46 } p_3^2 = 0, \quad = [4 \cdot 0 \cdot 4] \cdot G_1^1$$

$$\boxed{V_{\pi}(s^B) = 0}$$

$$V_{\pi}(s^C) = \int_{2:T}^1 \cdot G_2^1 + \int_{0:T}^2 G_0^2 + \int_{1:T}^2 \cdot G_1^2 + \int_{2:T}^2 \cdot G_2^2$$

$$p_{2:T}^1 = [0 \cdot 4] = 0$$

$$p_{0:T}^2 = [0 \cdot 2 \cdot 0 \cdot 0] = 0$$

$$p_{1:T}^2 = [2 \cdot 0 \cdot 0] = 0$$

$$p_{2:T}^2 = [0 \cdot 0] = 0$$

$$\boxed{V_{\pi}(s^C) = 0}$$