

Practice final

1) Backprop
gradient descent } What's the difference?

2) Tabular = max disc, min gen
so C

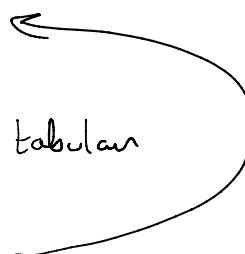
TC = Tilings :improve disc

Big tiles have big gen d is best answer
Small tiles have big disc maybe could argue e
with good explanation

3) a) yes

b) yes because we assumed tabular

c) FA breaks ↑ but not



4) This Q is way harder than anything on the final

a) probably not. It could if $v_\pi(s)$ is representable.

b) definitely, yes

c) yes (with appropriate α)

d) yes because of (b) above

5) a) we know very little about π and π^* .
we know $q_{\pi}(s,a) \geq q_{\pi^*}(s,a) \quad \forall s,a \in S \times A$

b) Nope. Could be oscillating between two optimal policies

Practice Final 2

a) priority \times server statuses

b) accept + reject

c) continuing

d) whatever

e) give def'n markov. make argument after

technically not markov, rejecting a job tells us something about the dist of future jobs we might observe.

does this mean we can't use Q-learning?

f) Can't use DP. Continuous state space makes sweeping hard
will never know true P and R.

2) a) BOE

b) BE

c) yup

3) We didn't teach you this. Not on final

$$a) \pi(\text{switch} | s_1) = 1$$

$$\pi(\text{switch} | s_2) = \frac{1}{2}$$

$$\mu(s_1) = 0\mu(s_1) + \frac{1}{2}\mu(s_2)$$

$$\mu(s_2) = 1\mu(s_1) + \frac{1}{2}\mu(s_2)$$

$$\mu(s_1) + \mu(s_2) = 1$$

$$\begin{aligned} \mu(s_1) &= \frac{1}{3} \\ \mu(s_2) &= \frac{2}{3} \end{aligned}$$

$$P = \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

$$\mu = \mu^P$$

skip until
end

b) focus FA power on states we actually visit

alternatives and why?
does it matter with tabular?

Extra Q

$$a) Q(z, \text{forward}) = \frac{1}{1-\gamma} 2 \\ = 4$$

$$Q(z, \text{stay}) = \frac{1}{1-\gamma} 1$$

$$= 2$$

$\pi^*(z) = \text{forward}, v^*(z) = 4$

$$Q(y, \text{forward}) = -1 + \gamma_2 v^*(z)$$

$$= -1 + 2$$

$$= 1$$

$$Q(y, \text{stay}) = \frac{1}{1-\gamma} 2$$

$$= 4$$

$\pi^*(y) = \text{Stay}, v^*(y) = 4$

$$Q(x, \text{forward}) = -1 + \gamma v^*(y)$$

$$\leq -1 + 2$$

$$= 1$$

$$Q(x, \text{stay}) = \frac{1}{1-\gamma} 1$$

$$= 2$$

$\pi^*(x) = \text{Stay}, v^*(x) = 2$

b) $Q(z, \text{forward}) = \frac{1}{1-\gamma} 2$

$$b) Q(z, \text{forward}) = \frac{1}{1-\gamma} 2$$

$$= 4 \cdot 2$$

$$= 8$$

$$Q(z, \text{stay}) = \frac{1}{1-\gamma} 1 = 4$$

$$\pi_*(z) = \text{forward}, v_*(z) = 8$$

$$Q(y, \text{forward}) = -1 + \frac{3}{4} v_*(z)$$

$$= -1 + 6$$

$$= 5$$

$$Q(y, \text{stay}) = \frac{1}{1-\gamma} 2 = 8$$

$$\pi_*(y) = \text{Stay}, v_*(y) = 8$$

$$Q(x, \text{forward}) = -1 + \frac{3}{4} v_*(y)$$

$$= -1 + 6$$

$$= 5$$

$$Q(x, \text{stay}) = \frac{1}{1-\gamma} 1 = 4$$

$$\pi_*(x) = \text{forward}, \quad v_*(x) = 5$$