

# CS 747, Autumn 2020: Week 4, Q&A

Shivaram Kalyanakrishnan

Department of Computer Science and Engineering  
Indian Institute of Technology Bombay

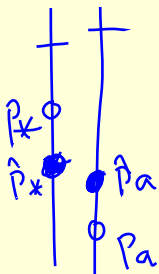
Autumn 2020

# Question 1

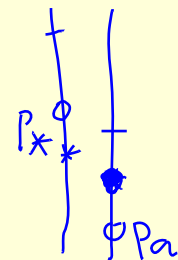
- What is the role of  $\bar{u}_a^T$  in the UCB regret proof?
-

# Question 1

- What is the role of  $\bar{u}_a^T$  in the UCB regret proof?



early on,  
few pulls



After many  
pulls

$$P_* - P_a = \Delta_a.$$

for arm  $*$ , we expect ucb to always exceed  $P_*$  (w.h.p.).  
for arm  $a$ , we only expect/show  $ucb < P_*$  after  $\bar{u}_a^T$  pulls.

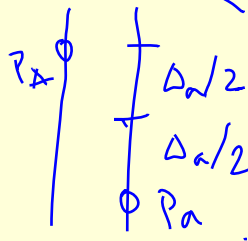
# Question 1

- What is the role of  $\bar{u}_a^T$  in the UCB regret proof?

$$\bar{u}_a^T = \sqrt{\frac{8}{\Delta_a^2} \ln T}$$

Hence

$$u_a^t \geq \bar{u}_a^T \Rightarrow \text{width} = \sqrt{\frac{2}{u_a} \ln t}$$

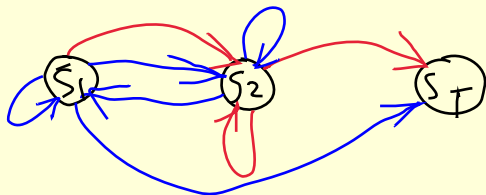
$$\leq \sqrt{\frac{2}{\sqrt{\frac{8}{\Delta_a^2} \ln T}}} \ln t \geq \frac{\Delta_a}{2}.$$


## Question 2

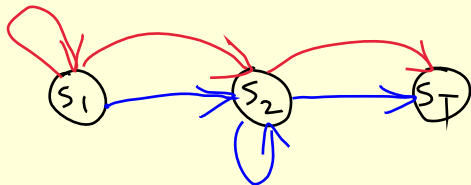
- Can you explain the constraint we impose on episodic MDPs for using  $\gamma = 1$ ?
-

## Question 2

- Can you explain the constraint we impose on episodic MDPs for using  $\gamma = 1$ ?



X



✓