Sept 8th Announcements

- Make sure you are on **SLACK** and the **PRIVATE** session on coursera
    - ….don’t know what I am talking about? Email me!
- Discussion submissions due **Sunday** nights. Please do use upvoting feature
    - This week’s discussion sub is due Thursday (this week only)
- Graded items (e.g., notebooks) due **Thursday** nights
- Thursday lecture is a Coursera review, but we will have more time for help with your notebooks …
- Office hours today! Every Tuesday after class

Discussion submissions

- First week is a bit like a practice run
- Starting next week:
  - poor grammar and formatting; zero.
    - If the TAs don’t understand your submission; zero
- **The goal of discussion questions is to show you read the textbook and thought about the material**
  - If I can say “the answer is on page 25” that’s not good
- Don’t just ask questions. Provide answers!

Your discussion submissions C1M1

- Given the limitations of UCB & optimistic initial methods, can we say that epsilon-greedy is the best option to deal with nonstationary and non-bandit problems?

- In applying ε-greedy (with linear annealing) to non-stationary problems, may we attempt to spot distribution shifts in the rewards and reset ε?

- Goal is to maximize expectedreward.But how to deal with high expected rewards with high variance?We might prefer a slightly low expectedreward with low variance

- In practice, how often are hyperparameter tuning methods (grid search, Bayesian optimization) applied to RL models like UCB and ε-greedy to find c or ε?

More C1M1 …

- Are there methods other than UCB that have been studied under bandit settings but haven't been generalized to the full RL problem (and that we'd like to)?

- Can we explore the world first and then exploit what we learned for an evaluation? Why does an agent have to explore and exploit simultaneously in one episode?

- Comparing two algorithms, how do we make sure that our results are statistically significant? When can we safely use the SE, and when is it not appropriate?