Cpt_S 580

(Reinforcement Learning)
• Actions: Wave, Stand, Clap
• Observations: Colors, Rewards
• Goal: find an optimal *policy*
• What is your policy?
• What does the world look like?
Formalize the Problem

• Knowns
• Unknowns
- (UP) Wave
- (RIGHT) Stand
- (DOWN) Clap
• Agent can change policy
  – Representation and Reward as given

• Pros:
  – Just specify goals
  – can be much less work than programming
  – unexpected situations

• Cons:
  – Can be slow
  – need to pick representation & rewards
Goals for Our Class

- Solid foundation for RL
- More focus on empirical than theoretical
RL is “Hot” right now

• Where have you heard about it?

• https://www.youtube.com/watch?v=1qa7oRhZvbM
• Introductions
• Web page, Syllabus, Web tools
• Makeup classes
• Burlap
• OpenAI Gym
• Grading
• Tas
• Room, coffee
Reduced Formalism

• Knowns
• Unknowns
• What is Reinforcement Learning?

• A problem, not an approach

• Vs supervised and unsupervised?
• Observations vs. State
• Value-based vs. Policy Search
• Dynamic Programming vs. Model Free vs. Model-based (Model-learning)
• Explore vs. Exploit