CS394R
Reinforcement Learning: Theory and Practice

Scott Niekum and Peter Stone

Department of Computer Science
The University of Texas at Austin
Good Morning Colleagues
Good Morning Colleagues

- Are there any questions?
Logistics
Logistics

- Registering for the course
Logistics

- Registering for the course
- Nice responses!
Logistics

- Registering for the course
- Nice responses!
  - Length and content mostly good
Logistics

- Registering for the course
- Nice responses!
  - Length and content mostly good
  - Be clear and specific
Logistics

- Registering for the course
- Nice responses!
  - Length and content mostly good
  - Be clear and specific
  - Short and focussed is fine
Logistics

- Registering for the course
- Nice responses!
  - Length and content mostly good
  - Be clear and specific
  - Short and focussed is fine
  - Help us help you
Logistics

- Registering for the course
- Nice responses!
  - Length and content mostly good
  - Be clear and specific
  - Short and focused is fine
  - Help us help you
  - Also ask in class or on discussion board
Logistics

- Registering for the course
- Nice responses!
  - Length and content mostly good
  - Be clear and specific
  - Short and focussed is fine
  - Help us help you
  - Also ask in class or on discussion board
  - default: 9/10
Logistics

- Registering for the course
- Nice responses!
  - Length and content mostly good
  - Be clear and specific
  - Short and focussed is fine
  - Help us help you
  - Also ask in class or on discussion board
  - default: 9/10
- Do the first exercises and programming assignment
Logistics

- Registering for the course
- Nice responses!
  - Length and content mostly good
  - Be clear and specific
  - Short and focussed is fine
  - Help us help you
  - Also ask in class or on discussion board
  - default: 9/10
- Do the first exercises and programming assignment
- If you missed last Thursday . . .
Logistics

• Registering for the course
• Nice responses!
  – Length and content mostly good
  – Be clear and specific
  – Short and focussed is fine
  – Help us help you
  – Also ask in class or on discussion board
  – default: 9/10
• Do the first exercises and programming assignment
• If you missed last Thursday . . .
  – Watch intro lecture video
  – Read webpage carefully
Our Role
Our Role

- Our role isn’t to teach RL
Our Role

• Our role isn’t to teach RL

• It’s to help you learn RL
Our Role

- Our role isn’t to teach RL
- It’s to help you learn RL
  - provide context
  - guide your learning (assign readings, exercises, activities)
  - clarify misconceptions
Our Role

- Our role isn’t to teach RL
- It’s to help you learn RL
  - provide context
  - guide your learning (assign readings, exercises, activities)
  - clarify misconceptions
- You have to do the learning
Our Role

• Our role isn’t to teach RL

• It’s to help you learn RL
  – provide context
  – guide your learning (assign readings, exercises, activities)
  – clarify misconceptions

• You have to do the learning

• Read, write, ask, answer, program (investigate)
Logistics

- Week numberings
Logistics

- Week numberings
- Email both instructors and TAs
Logistics

- Week numberings
- Email both instructors and TAs
- Next readings posted
Logistics

- Week numberings
- Email both instructors and TAs
- Next readings posted
  - 2nd edition!!
Logistics

- Week numberings
- Email both instructors and TAs
- Next readings posted
  - 2nd edition!!
  - Dynamic programming and Monte Carlo
Logistics

• Week numberings

• Email both instructors and TAs

• Next readings posted
  - 2nd edition!!
  - Dynamic programming and Monte Carlo
  - Mostly chapter 4 Tuesday, then chapter 5
Logistics

- Week numberings
- Email both instructors and TAs
- Next readings posted
  - 2nd edition!!
  - Dynamic programming and Monte Carlo
  - Mostly chapter 4 Tuesday, then chapter 5
- Look at resources page
Let’s Play!
Let’s Play!

• I’m a 2-armed bandit
Let’s Play!

- I’m a 2-armed bandit
- As a class, you choose which arm
Let’s Play!

- I’m a 2-armed bandit
- As a class, you choose which arm
- Maximize your payoff.
Let’s Play!

• I’m a 2-armed bandit
• As a class, you choose which arm
• Maximize your payoff.
• The answer:
Let’s Play!

- I’m a 2-armed bandit
- As a class, you choose which arm
- Maximize your payoff.
- The answer:

```lisp
(defun l () (+ 5 (random 7)))
(defun r ()
  (let ((x (random 3)))
    (case x
      (0 20)
      (1 0)
      (2 (+ 7 (random 11)))
    )))
```

- What about minimizing risk?
N-armed bandit in practice?
N-armed bandit in practice?

- Choosing mechanics
- Choosing a barber/hairdresser
Assignments

• Monitor and contribute to discussion forums!
Assignments

- Monitor and contribute to discussion forums!
- 1st exercises and programming assignment
Assignments

- Monitor and contribute to discussion forums!
- 1st exercises and programming assignment
- Read Chapters 4 and 5
Assignments

• Monitor and contribute to discussion forums!
• 1st exercises and programming assignment
• Read Chapters 4 and 5
• Submit a reading response by 5pm Monday