Good Afternoon Colleagues
Good Afternoon Colleagues

- Are there any questions?
Logistics
Logistics

- Responses
Logistics

- Responses
  - Work for clarity and substance
Logistics

- Responses
  - Work for clarity and substance
  - Look for programming assignment opportunities
Logistics

• Responses
  – Work for clarity and substance
  – Look for programming assignment opportunities
  – Share responses?
Logistics

- Responses
  - Work for clarity and substance
  - Look for programming assignment opportunities
  - Share responses?
    * You? Me? Just selected ones?
Logistics

- Responses
  - Work for clarity and substance
  - Look for programming assignment opportunities
  - Share responses?
    * You? Me? Just selected ones?

- Stars on the sections?
Logistics

- Responses
  - Work for clarity and substance
  - Look for programming assignment opportunities
  - Share responses?
    * You? Me? Just selected ones?

- Stars on the sections?

- Class on Thursday
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: 2 times around.
Let’s Play!

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

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

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

  (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?

Peter Stone
N-armed bandit in practice?
N-armed bandit in practice?

- Choosing mechanics
- Choosing a barber/hairdresser
Evaluative Feedback

- Understanding “binary” bandit.
Evaluative Feedback

- Understanding “binary” bandit.
- Why $L_{R-I}$ better on bandit A than B?
Evaluative Feedback

- Understanding “binary” bandit.
- Why $L_{R-I}$ better on bandit A than B?
  - Why better than action values?
Evaluative Feedback

• Understanding “binary” bandit.
• Why $L_{R-I}$ better on bandit A than B?
  – Why better than action values?
• Why supervised smooth for Bandit B?
Evaluative Feedback

- Understanding “binary” bandit.
- Why $L_{R-I}$ better on bandit A than B?
  - Why better than action values?
- Why supervised smooth for Bandit B?
- Ex. 2.4
• Which exploration methods work in practice?
Questions

• Which exploration methods work in practice?

• What’s the intuition behind pursuit and reinforcement comparison?
  – Why better than $\epsilon$-greedy?
Questions

- Which exploration methods work in practice?
- What’s the intuition behind pursuit and reinforcement comparison?
  - Why better than $\epsilon$-greedy?
- Ex. 2.1
- Ex. 2.8