CS394R
Reinforcement Learning: Theory and Practice

Peter Stone

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

• Are there any questions?
Logistics

- Make progress on final projects!
Logistics

• Make progress on final projects!

• This week’s readings: game playing
Logistics

- Make progress on final projects!
- This week’s readings: game playing
  - How to represent domains?
Logistics

- Make progress on final projects!
- This week’s readings: game playing
  - How to represent domains?
- Next week’s readings: financial domains
  - Moody and Saffell: portfolio optimization
Logistics

- Make progress on final projects!

- This week’s readings: game playing
  - How to represent domains?

- Next week’s readings: financial domains
  - **Moody and Saffell**: portfolio optimization
  - learning policies vs. learning values
  - Find Q-learning isn’t the way to go
Logistics

- Make progress on final projects!

- This week’s readings: game playing
  - How to represent domains?

- Next week’s readings: financial domains
  - **Moody and Saffell**: portfolio optimization
  - learning policies vs. learning values
  - Find Q-learning isn’t the way to go
  - **Nevmyvaka, Feng, and Kearns**: trade execution
Logistics

• Make progress on final projects!

• This week’s readings: game playing
  – How to represent domains?

• Next week’s readings: financial domains
  – Moody and Saffell: portfolio optimization
  – learning policies vs. learning values
  – Find Q-learning isn’t the way to go
  – Nevmyvaka, Feng, and Kearns: trade execution
  – Interesting problem choice, custom algorithm, careful choice of representation
Surveys

- Thanks!
Surveys

- Thanks!
- Talk about readings for next class
Surveys

- Thanks!
- Talk about readings for next class
- Examples of assignments and projects
Surveys

- Thanks!
- Talk about readings for next class
- Examples of assignments and projects
- Healthier snacks
Surveys

- Thanks!
- Talk about readings for next class
- Examples of assignments and projects
- Healthier snacks
- Student-led discussions aren’t the best use of class time
Game Playing

- Is it a worthwhile pursuit for AI?
Background

- Backgammon
- NNs
How does it work?

• What’s the role of the NN?
How does it work?

- What’s the role of the NN?
- What’s the input representation?
How does it work?

• What’s the role of the NN?

• What’s the input representation?
  – Truncated unary encoding (units for 1,2,3,4+ pieces on each space)
  – eventually: strength of blockade, probability of being hit
How does it work?

• What’s the role of the NN?

• What’s the input representation?
  – Truncated unary encoding (units for 1, 2, 3, 4+ pieces on each space)
  – eventually: strength of blockade, probability of being hit

• What’s the output representation?
How does it work?

- What’s the role of the NN?
- What’s the input representation?
  - Truncated unary encoding (units for 1, 2, 3, 4+ pieces on each space)
  - eventually: strength of blockade, probability of being hit
- What’s the output representation?
- Is it Sarsa or Q-learning?
How does it work?

• What’s the role of the NN?

• What’s the input representation?
  – Truncated unary encoding (units for 1, 2, 3, 4+ pieces on each space)
  – eventually: strength of blockade, probability of being hit

• What’s the output representation?

• Is it Sarsa or Q-learning?

• What’s lambda?
What does the NN learn?

An examination of the input-to-hidden weights in this network revealed interesting spatially organized patterns of positive and negative weights, roughly corresponding to what a knowledge engineer might call useful features for game play.
What does the NN learn?

An examination of the input-to-hidden weights in this network revealed interesting spatially organized patterns of positive and negative weights, roughly corresponding to what a knowledge engineer might call useful features for game play.

- Learn non-linear concepts?
Class Discussion - Yuchen

- TD vs. coevolution?
Who won the dispute?
Background

- Go
- GGP
Class Discussion - Shun

- Minimax vs. RL
Making UCT more practical

- Gelly slides