Good Afternoon, Colleagues

Are there any questions?
Good Afternoon, Colleagues

Are there any questions?

- Other LL domains?
Good Afternoon, Colleagues

Are there any questions?

- Other LL domains?
- Other hierarchical learning approaches?
Logistics

- Surveys from Poland
Logistics

- Surveys from Poland
- Final reports due to Mazda tomorrow by 8pm
The Tournament

1. **Soccer Fascists**
   - Sura and Hwang
2. **G-Cipher**
   - Barksdale and Morris
3. **Kablip.FC**
   - Kane, Issen, and Parkeh
4. **Ottomans**
   - Deligonul and Ciftici
5. **CG United**
   - Su and Bradley
6. **MISC**
   - Lewis
7. **PG-11**
   - Li and Fayyaz
8. **The Big O’z**
   - Shao and Jones
9. **Serendipity**
   - Trimble and Hatfield
10. **Node Warrior**
    - Fakhreddine and Clark
11. **Team Quarks**
    - Chuah and Dasler
12. **Team Stamina**
    - High and Ulrich
Machine Learning

Hypothesis space: set of possible functions

Training examples: the data

Learning method: training examples $\mapsto$ hypothesis
Machine Learning

Hypothesis space: set of possible functions

Training examples: the data

Learning method: training examples $\mapsto$ hypothesis

Agent Learning

Policy: how to act (generate training examples)

neural network training, Q-learning, decision tree training, clustering, genetic algorithms, genetic programming, ...
Genetic algorithms

• Keep a population of individuals

• Each generation
  – Evaluate their fitness
  – Throw out the bad ones
  – Change the good ones randomly
  – Repeat
Genetic algorithms

- Keep a population of individuals

- Each generation
  - Evaluate their fitness
  - Throw out the bad ones
  - Change the good ones randomly
  - Repeat

The fitness function matters

- Playing against top-notch competition → no info
- Playing against a single foe → too brittle
Class Discussion

Brian Jones on competitive coevolution
Collaborative Co-Evolution

• Learn **collaborative** behaviors simultaneously
Collaborative Co-Evolution

- Learn **collaborative** behaviors simultaneously
- Applied in pursuit domain among others
Collaborative Co-Evolution

- Learn **collaborative** behaviors simultaneously
- Applied in pursuit domain among others
- Could be used in context of layered learning
  - Research here with Shimon Whiteson
Collaborative Co-Evolution

- Learn **collaborative** behaviors simultaneously
- Applied in pursuit domain among others
- Could be used in context of layered learning
  - Research here with Shimon Whiteson
- Simultaneous learning by teammates could be thought of in this way as well.
3 vs. 2 Keepaway  (joint with Rich Sutton)

- Play in a **small area** (20m × 20m)
- **Keepers** try to keep the ball
- **Takers** try to get the ball

**Episode:**
- Players and ball reset randomly
- Ball starts near a keeper
- Ends when taker gets the ball or ball goes out

- Performance measure: **average possession duration**

- Use **CMUnited-99 skills**:
  - HoldBall, PassBall($k$), GoToBall, GetOpen
Available Skills (from CMUnited-99)

**HoldBall()**: Remain stationary while keeping possession of the ball.

**PassBall($k$)**: Kick the ball directly to keeper $k$.

**GoToBall()**: Intercept a moving ball or move directly towards a stationary ball.

**GetOpen()**: Move to a position that is free from opponents and open for a pass from the ball’s current position (using SPAR (Veloso et al., 1999))

**BlockPass($k$)**: Get in between the ball and keeper $k$.
The Keepers’ Policy Space

- GetOpen
- Ballnot kickable
- Ball kickable
- Teammate with ball or can get there faster

\{HoldBall, PassBall(k)\}
(k is another keeper)
The Keepers’ Policy Space

Example Policies

**Random:** HoldBall or PassBall\((k)\) randomly

**Hold:** Always HoldBall

**Hand-coded:**

- If no taker within 10m: HoldBall
- Else If there’s a good pass: PassBall\((k)\)
- Else HoldBall
Mapping Keepaway to RL

Discrete-time, episodic, distributed RL

- Simulator operates in discrete time steps, $t = 0, 1, 2, \ldots$, each representing 100 msec

- Episode: $s_0, a_0, r_1, s_1, \ldots, s_t, a_t, r_{t+1}, s_{t+1}, \ldots, r_T, s_T$

- $a_t \in \{\text{HoldBall}, \text{PassBall}(k), \text{GoToBall}, \text{GetOpen}\}$

- $r_t = 1$

- $V^\pi(s) = E\{T \mid s_0 = s\}$

- Goal: Find $\pi^*$ that maximizes $V$ for all $s$
Representation

- Full soccer state
- Few continuous state variables (13)
- Huge binary feature vector (about 400 1’s and 40,000 0’s)
- Sparse, coarse, tile coding
- Linear map
- Action values
13 Continuous State Variables

- 11 distances among players, ball, and center
- 2 angles to takers along passing lanes
Function Approximation: Tile Coding

- Form of sparse, coarse coding based on CMACS (Albus, 1981)

- Tiled state variables individually (13)
Policy Learning

- Learn $Q^\pi(s, a)$: Expected possession time
Policy Learning

- Learn $Q^\pi(s, a)$: Expected possession time
- Linear Sarsa($\lambda$) — each agent learns independently
  - On-policy method: advantages over e.g. Q-learning
  - Not known to converge, but works (e.g. (Sutton, 1996))
Main Result

1 hour = 720 5-second episodes
## Varied Field Size

<table>
<thead>
<tr>
<th>Keepers</th>
<th>Testing Field Size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>15x15</td>
</tr>
<tr>
<td>Trained on field of size</td>
<td></td>
</tr>
<tr>
<td>15x15</td>
<td>11.0</td>
</tr>
<tr>
<td>20x20</td>
<td>10.7</td>
</tr>
<tr>
<td>25x25</td>
<td>6.3</td>
</tr>
<tr>
<td>Benchmarks</td>
<td></td>
</tr>
<tr>
<td>Hand</td>
<td>4.3</td>
</tr>
<tr>
<td>Hold</td>
<td>3.9</td>
</tr>
<tr>
<td>Random</td>
<td>4.2</td>
</tr>
</tbody>
</table>

- Single runs
- Learning specific to fields
  - mechanism generalizes better than policies
4 vs. 3 Keeper Learning

- Preliminary: taker learning successful as well
Course recap

- You’ve read.
Course recap

- You’ve read.
- You’ve reacted and formed opinions.
Course recap

- You’ve read.
- You’ve reacted and formed opinions.
- You’ve spoken.
Course recap

- You’ve read.
- You’ve reacted and formed opinions.
- You’ve spoken.
- You’ve written.
Course recap

- You’ve read.
- You’ve reacted and formed opinions.
- You’ve spoken.
- You’ve written.
- You’ve coded for a task with no right answer and no way of knowing that you’re done.
Course recap

- You’ve read.
- You’ve reacted and formed opinions.
- You’ve spoken.
- You’ve written.
- You’ve coded for a task with no right answer and no way of knowing that you’re done.

Do you like CS research?
<table>
<thead>
<tr>
<th>What have we covered?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Autonomous agents:</td>
</tr>
</tbody>
</table>

Peter Stone
What have we covered?

1. Autonomous agents: What is an agent?
2. Agent architectures: Subsumption, TCA
What have we covered?

1. Autonomous agents: What is an agent?
2. Agent architectures: Subsumption, TCA
3. Multiagent Systems: Overview, subsumption
What have we covered?

1. Autonomous agents: What is an agent?
2. Agent architectures: Subsumption, TCA
3. Multiagent Systems: Overview, subsumption
4. Communication and Teamwork: KQML, Joint Intentions
What have we covered?

<table>
<thead>
<tr>
<th>1. Autonomous agents:</th>
<th>What is an agent?</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Agent architectures:</td>
<td>Subsumption, TCA</td>
</tr>
<tr>
<td>3. Multiagent Systems:</td>
<td>Overview, subsumption</td>
</tr>
<tr>
<td>4. Communication and Teamwork:</td>
<td>KQML, Joint Intentions</td>
</tr>
<tr>
<td>5. RoboCup case studies</td>
<td></td>
</tr>
</tbody>
</table>

Peter Stone
What have we covered?

1. Autonomous agents: What is an agent?
2. Agent architectures: Subsumption, TCA
3. Multiagent Systems: Overview, subsumption
4. Communication and Teamwork: KQML, Joint Intentions
5. RoboCup case studies
6. Swarms and ant-based approaches: “Go to the Ant”
What have we covered?

1. Autonomous agents: What is an agent?
2. Agent architectures: Subsumption, TCA
3. Multiagent Systems: Overview, subsumption
4. Communication and Teamwork: KQML, Joint Intentions
5. RoboCup case studies
6. Swarms and ant-based approaches: “Go to the Ant”
7. Applications: Air traffic, electric elves
<table>
<thead>
<tr>
<th>What have we covered?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Autonomous agents:</td>
</tr>
<tr>
<td>2. Agent architectures:</td>
</tr>
<tr>
<td>3. Multiagent Systems:</td>
</tr>
<tr>
<td>4. Communication and Teamwork:</td>
</tr>
<tr>
<td>5. RoboCup case studies</td>
</tr>
<tr>
<td>6. Swarms and ant-based approaches:</td>
</tr>
<tr>
<td>7. Applications:</td>
</tr>
<tr>
<td>8. Game theory:</td>
</tr>
</tbody>
</table>
What have we covered?

1. **Autonomous agents:**
   - What is an agent?

2. **Agent architectures:**
   - Subsumption, TCA

3. **Multiagent Systems:**
   - Overview, subsumption

4. **Communication and Teamwork:**
   - KQML, Joint Intentions

5. **RoboCup case studies**

6. **Swarms and ant-based approaches:**
   - “Go to the Ant”

7. **Applications:**
   - Air traffic, electric elves

8. **Game theory:**
   - Nash equilibrium

9. **RoboCup rescue:**
What have we covered?

1. **Autonomous agents:**
   - What is an agent?

2. **Agent architectures:**
   - Subsumption, TCA

3. **Multiagent Systems:**
   - Overview, subsumption

4. **Communication and Teamwork:**
   - KQML, Joint Intentions

5. **RoboCup case studies**

6. **Swarms and ant-based approaches:**
   - “Go to the Ant”

7. **Applications:**
   - Air traffic, electric elves

8. **Game theory:**
   - Nash equilibrium

9. **RoboCup rescue:**

10. **Agent modeling:**
    - coaching, RMM, tracking teams
### What have we covered?

1. **Autonomous agents:** What is an agent?
2. **Agent architectures:** Subsumption, TCA
3. **Multiagent Systems:** Overview, subsumption
4. **Communication and Teamwork:** KQML, Joint Intentions
5. **RoboCup case studies**
6. **Swarms and ant-based approaches:** “Go to the Ant”
7. **Applications:** Air traffic, electric elves
8. **Game theory:** Nash equilibrium
9. **RoboCup rescue**
10. **Agent modeling:** coaching, RMM, tracking teams
11. **Distributed rational decision making:** voting, ...
## What have we covered?

1. Autonomous agents: What is an agent?
2. Agent architectures: Subsumption, TCA
3. Multiagent Systems: Overview, subsumption
4. Communication and Teamwork: KQML, Joint Intentions
5. RoboCup case studies
6. Swarms and ant-based approaches: “Go to the Ant”
7. Applications: Air traffic, electric elves
8. Game theory: Nash equilibrium
9. RoboCup rescue:
10. Agent modeling: coaching, RMM, tracking teams
11. Distributed rational decision making: voting, ...
12. Auctions: FCC spectrum auctions, TAC
What have we covered?

1. Autonomous agents: What is an agent?
2. Agent architectures: Subsumption, TCA
3. Multiagent Systems: Overview, subsumption
4. Communication and Teamwork: KQML, Joint Intentions
5. RoboCup case studies
6. Swarms and ant-based approaches: “Go to the Ant”
7. Applications: Air traffic, electric elves
8. Game theory: Nash equilibrium
9. RoboCup rescue:
10. Agent modeling: coaching, RMM, tracking teams
11. Distributed rational decision making: voting, ...
12. Auctions: FCC spectrum auctions, TAC
13. Entertainment agents cobot, chatbots
What have we covered?

1. **Autonomous agents:** What is an agent?
2. **Agent architectures:** Subsumption, TCA
3. **Multiagent Systems:** Overview, subsumption
4. **Communication and Teamwork:** KQML, Joint Intentions
5. **RoboCup case studies**
6. **Swarms and ant-based approaches:** “Go to the Ant”
7. **Applications:** Air traffic, electric elves
8. **Game theory:** Nash equilibrium
9. **RoboCup rescue**
10. **Agent modeling:** coaching, RMM, tracking teams
11. **Distributed rational decision making:** voting, . . .
12. **Auctions:** FCC spectrum auctions, TAC
13. **Entertainment agents** cobot, chatbots
14. **Multiagent learning:** layered learning, co-evolution
The original question

- What is an agent?
Course recap

- I’ve been impressed by the levels of discussions we’ve had in class
- I’m happy with the progress in writing and speaking that many of you have made
- I’m proud of all of you for sticking with it through such a demanding course
Course recap

• I’ve been impressed by the levels of discussions we’ve had in class

• I’m happy with the progress in writing and speaking that many of you have made

• I’m proud of all of you for sticking with it through such a demanding course

THANKS!!!
Surveys

- Mazda’s and my surveys
- Positive and negative feedback useful
Surveys

- Mazda’s and my surveys
- Positive and negative feedback useful
- Invitation to do more on-line surveys
Surveys

- Mazda’s and my surveys
- Positive and negative feedback useful
- Invitation to do more on-line surveys
  - Still anonymous
  - Fill it out only what you feel like
  - Should the course be run again?
  - How should it change?
Next Meeting

• The tournament!
Next Meeting

- The tournament!
- Thursday, May 13th
- ACES 6.304
- 10:30am–12:30pm
Next Meeting

- The tournament!
- Thursday, May 13th
- ACES 6.304
- 10:30am–12:30pm
- Come prepared to talk (informally) about your team