

CS395T
Reinforcement Learning:
Theory and Practice
Fall 2004

Peter Stone

Department of Computer Sciences
The University of Texas at Austin

Week1a: Tuesday, August 31st

Good Afternoon Colleagues



Good Afternoon Colleagues

- Are there any questions?

Logistics



Logistics

- Nice responses!

Logistics

- Nice responses!
 - Length and content good

Logistics

- Nice responses!
 - Length and content good
 - Be clear and specific

Logistics

- Nice responses!
 - Length and content good
 - Be clear and specific
 - Look for programming assignment opportunities

Logistics

- Nice responses!
 - Length and content good
 - Be clear and specific
 - Look for programming assignment opportunities
- Some of you got replies

Logistics

- Nice responses!
 - Length and content good
 - Be clear and specific
 - Look for programming assignment opportunities
- Some of you got replies
- Today: self-introductions

Logistics

- Nice responses!
 - Length and content good
 - Be clear and specific
 - Look for programming assignment opportunities
- Some of you got replies
- Today: self-introductions
- Thursday: Discussion leader assignments

Some Questions

- What's a model?

Some Questions

- What's a model?
- Reward function vs. value function

Some Questions

- What's a model?
- Reward function vs. value function
 - Tic-tac-toe example

Some Questions

- What's a model?
- Reward function vs. value function
 - Tic-tac-toe example
 - Phil making breakfast example

Some Questions

- What's a model?
- Reward function vs. value function
 - Tic-tac-toe example
 - Phil making breakfast example
- Distinction with evolutionary methods?
 - Tic-tac-toe example

Some Questions

- What's a model?
- Reward function vs. value function
 - Tic-tac-toe example
 - Phil making breakfast example
- Distinction with evolutionary methods?
 - Tic-tac-toe example
 - Phil making breakfast example

Reduced Formalism

Knowns:

- $\mathcal{S} = \{\text{Blue, Red, Green, Black, } \dots\}$
- Rewards in \mathbb{R}
- $\mathcal{A} = \{\text{Wave, Clap, Stand}\}$

$s_0, a_0, r_0, s_1, a_1, r_1, s_2, \dots$

Unknowns:

- $\mathcal{R} : \mathcal{S} \times \mathcal{A} \mapsto \mathbb{R}$
- $\mathcal{T} : \mathcal{S} \times \mathcal{A} \mapsto \mathcal{S}$

$$r_i = \mathcal{R}(s_i, a_i)$$

$$s_{i+1} = \mathcal{T}(s_i, a_i)$$