

**CS395T**  
**Reinforcement Learning:**  
**Theory and Practice**  
**Fall 2004**

**Peter Stone**

Department of Computer Sciences  
The University of Texas at Austin

# BE a reinforcement learner

---

# BE a reinforcement learner

---

- You, as a class, act as a learning agent

# BE a reinforcement learner

---

- You, as a class, act as a learning agent
- **Actions:** Wave, Stand, Clap

# BE a reinforcement learner

---

- You, as a class, act as a learning agent
- **Actions:** Wave, Stand, Clap
- **Observations:** colors, reward

# BE a reinforcement learner

---

- You, as a class, act as a learning agent
- **Actions:** Wave, Stand, Clap
- **Observations:** colors, reward
- **Goal:** Find an optimal *policy*

# BE a reinforcement learner

---

- You, as a class, act as a learning agent
- **Actions:** Wave, Stand, Clap
- **Observations:** colors, reward
- **Goal:** Find an optimal *policy*
  - Way of selecting actions that gets you the most reward

# How did you do it?

---



# How did you do it?

---

- What is your policy?
- What does the world look like?

# Formalizing What Just Happened

---

Knowns:

# Formalizing What Just Happened

---

## Knowns:

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

$o_0, a_0, r_0, o_1, a_1, r_1, o_2, \dots$

# Formalizing What Just Happened

---

## Knowns:

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

$o_0, a_0, r_0, o_1, a_1, r_1, o_2, \dots$

## Unknowns:

# Formalizing What Just Happened

---

## Knowns:

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

$o_0, a_0, r_0, o_1, a_1, r_1, o_2, \dots$

## Unknowns:

- $\mathcal{S} = 4 \times 3$  grid
- $\mathcal{R} : \mathcal{S} \times \mathcal{A} \mapsto \mathbb{R}$
- $\mathcal{P} = \mathcal{S} \mapsto \mathcal{O}$
- $\mathcal{T} : \mathcal{S} \times \mathcal{A} \mapsto \mathcal{S}$

# Formalizing What Just Happened

---

## Knowns:

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

$o_0, a_0, r_0, o_1, a_1, r_1, o_2, \dots$

## Unknowns:

- $\mathcal{S} = 4 \times 3$  grid
- $\mathcal{R} : \mathcal{S} \times \mathcal{A} \mapsto \mathbb{R}$
- $\mathcal{P} = \mathcal{S} \mapsto \mathcal{O}$
- $\mathcal{T} : \mathcal{S} \times \mathcal{A} \mapsto \mathcal{S}$

$$o_i = \mathcal{P}(s_i)$$

# Formalizing What Just Happened

---

## Knowns:

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

$o_0, a_0, r_0, o_1, a_1, r_1, o_2, \dots$

## Unknowns:

- $\mathcal{S} = 4 \times 3$  grid
- $\mathcal{R} : \mathcal{S} \times \mathcal{A} \mapsto \mathbb{R}$
- $\mathcal{P} = \mathcal{S} \mapsto \mathcal{O}$
- $\mathcal{T} : \mathcal{S} \times \mathcal{A} \mapsto \mathcal{S}$

$$o_i = \mathcal{P}(s_i)$$

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

# Formalizing What Just Happened

---

## Knowns:

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

$o_0, a_0, r_0, o_1, a_1, r_1, o_2, \dots$

## Unknowns:

- $\mathcal{S} = 4 \times 3$  grid
- $\mathcal{R} : \mathcal{S} \times \mathcal{A} \mapsto \mathbb{R}$
- $\mathcal{P} = \mathcal{S} \mapsto \mathcal{O}$
- $\mathcal{T} : \mathcal{S} \times \mathcal{A} \mapsto \mathcal{S}$

$$o_i = \mathcal{P}(s_i)$$

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

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

# This Course

---

- Reinforcement Learning theory (start)

# This Course

---

- Reinforcement Learning theory (start)
- Reinforcement Learning in practice (end)

# The Big Picture

---

- AI

# The Big Picture

---

- AI  $\longrightarrow$  ML

# The Big Picture

---

- AI  $\longrightarrow$  ML  $\longrightarrow$  RL

# The Big Picture

---

- AI  $\longrightarrow$  ML  $\longrightarrow$  RL
- Types of Machine Learning

# The Big Picture

---

- AI  $\longrightarrow$  ML  $\longrightarrow$  RL
- Types of Machine Learning
  - Supervised learning:** learn from labeled examples

# The Big Picture

---

- AI  $\longrightarrow$  ML  $\longrightarrow$  RL

- Types of Machine Learning

**Supervised learning:** learn from labeled examples

**Unsupervised learning:** cluster unlabeled examples

# The Big Picture

---

- AI  $\longrightarrow$  ML  $\longrightarrow$  RL

- Types of Machine Learning

**Supervised learning:** learn from labeled examples

**Unsupervised learning:** cluster unlabeled examples

**Reinforcement learning:** learn from interaction

# The Big Picture

---

- AI  $\longrightarrow$  ML  $\longrightarrow$  RL

- Types of Machine Learning

**Supervised learning:** learn from labeled examples

**Unsupervised learning:** cluster unlabeled examples

**Reinforcement learning:** learn from interaction

- Defined by the problem

# The Big Picture

---

- AI  $\longrightarrow$  ML  $\longrightarrow$  RL

- Types of Machine Learning

**Supervised learning:** learn from labeled examples

**Unsupervised learning:** cluster unlabeled examples

**Reinforcement learning:** learn from interaction

- Defined by the problem

- Many approaches possible (including evolutionary)

# The Big Picture

---

- AI  $\longrightarrow$  ML  $\longrightarrow$  RL

- Types of Machine Learning

**Supervised learning:** learn from labeled examples

**Unsupervised learning:** cluster unlabeled examples

**Reinforcement learning:** learn from interaction

- Defined by the problem
- Many approaches possible (including evolutionary)
- Book focusses on a particular class of approaches

# Syllabus

---

- Available on-line

# Assignments for Tuesday

---

- Join the mailing list!

# Assignments for Tuesday

---

- Join the mailing list!
- Read Chapter 1

# Assignments for Tuesday

---

- Join the mailing list!
- Read Chapter 1
- Send a question or comment by 10pm Monday