Good Afternoon, Colleagues
Good Afternoon, Colleagues

Are there any questions?
Logistics

- Problems with the assignment?
Logistics

• Problems with the assignment?
  – 3.1: Goal formulation vs. problem formulation
Logistics

- Problems with the assignment?
  - 3.1: Goal formulation vs. problem formulation

- Mailing list
Logistics

- Problems with the assignment?
  - 3.1: Goal formulation vs. problem formulation

- Mailing list
  - CC Daniel (urieli@cs), and me on everything
Logistics

• Problems with the assignment?
  – 3.1: Goal formulation vs. problem formulation

• Mailing list
  – CC Daniel (urieli@cs), and me on everything
  – Use correct subject line
Logistics

- Problems with the assignment?
  - 3.1: Goal formulation vs. problem formulation

- Mailing list
  - CC Daniel (urieli@cs), and me on everything
  - Use correct subject line

- Next week’s assignments up
Logistics

- Problems with the assignment?
  - 3.1: Goal formulation vs. problem formulation

- Mailing list
  - CC Daniel (urieli@cs), and me on everything
  - Use correct subject line

- Next week’s assignments up

- Comments on Forum for AI talk last Friday?
Logistics

• Problems with the assignment?
  – 3.1: Goal formulation vs. problem formulation

• Mailing list
  – CC Daniel (urieli@cs), and me on everything
  – Use correct subject line

• Next week’s assignments up

• Comments on Forum for AI talk last Friday?

• Reading responses good overall
An Example
An Example

- 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 the Activity

Knowns:
Formalizing the Activity

Knowns:

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

\[o_0, a_0, r_0, o_1, a_1, r_1, o_2, \ldots\]
Formalizing the Activity

Knowns:

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

Unknowns:

- \( o_0, a_0, r_0, o_1, a_1, r_1, o_2, \ldots \)
Formalizing the Activity

Knowns:
- \( \mathcal{O} = \{ \text{Blue, Red, Green, Black, \ldots} \} \)
- Rewards in \( \mathbb{R} \)
- \( \mathcal{A} = \{ \text{Wave, Clap, Stand} \} \)

Unknowns:
- \( \mathcal{S} = 4 \times 3 \) grid
- \( \mathcal{R} : \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R} \)
- \( \mathcal{P} = \mathcal{S} \rightarrow \mathcal{O} \)
- \( \mathcal{T} : \mathcal{S} \times \mathcal{A} \rightarrow \mathcal{S} \)

\( o_0, a_0, r_0, o_1, a_1, r_1, o_2, \ldots \)
Formalizing the Activity

Knowns:

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

```
\begin{align*}
o_0, a_0, r_0, o_1, a_1, r_1, o_2, \ldots
\end{align*}
```

Unknowns:

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

\[ o_i = P(s_i) \]
Formalizing the Activity

Knowns:
- \( O = \{ \text{Blue, Red, Green, Black, \ldots} \} \)
- Rewards in \( \mathbb{R} \)
- \( A = \{ \text{Wave, Clap, Stand} \} \)
- \( o_0, a_0, r_0, o_1, a_1, r_1, o_2, \ldots \)

Unknowns:
- \( S = 4 \times 3 \) grid
- \( \mathcal{R} : S \times A \rightarrow \mathbb{R} \)
- \( P = S \rightarrow O \)
- \( T : S \times A \rightarrow S \)
- \( o_i = P(s_i) \quad r_i = \mathcal{R}(s_i, a_i) \)
Formalizing the Activity

**Knowns:**
- $O = \{\text{Blue, Red, Green, Black, \ldots}\}$
- Rewards in $\mathbb{R}$
- $A = \{\text{Wave, Clap, Stand}\}$

\[
o_0, a_0, r_0, o_1, a_1, r_1, o_2, \ldots
\]

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

\[
o_i = P(s_i) \quad r_i = R(s_i, a_i) \quad s_{i+1} = T(s_i, a_i)
\]
Formalizing the Activity

Knocks:
- \( \mathcal{O} = \{ \text{Blue, Red, Green, Black, \ldots} \} \)
- Rewards in \( \mathbb{R} \)
- \( \mathcal{A} = \{ \text{Wave, Clap, Stand} \} \)

\[ o_0, a_0, r_0, o_1, a_1, r_1, o_2, \ldots \]

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

\[ o_i = \mathcal{P}(s_i) \quad r_i = \mathcal{R}(s_i, a_i) \quad s_{i+1} = \mathcal{T}(s_i, a_i) \]
PEAS description

- Performance measure:
- Environment:
- Actuators:
- Sensors:
PEAS description

- Performance measure:
- Environment:
- Actuators:
- Sensors:

- fully observable vs. partially observable (accessible)
PEAS description

- Performance measure:
- Environment:
- Actuators:
- Sensors:

- fully observable vs. partially observable (accessible)
- single-agent vs. multiagent
PEAS description

- Performance measure:
- Environment:
- Actuators:
- Sensors:

- fully observable vs. partially observable (accessible)
- single-agent vs. multiagent
- deterministic vs. non-deterministic
PEAS description

- Performance measure:
- Environment:
- Actuators:
- Sensors:

- fully observable vs. partially observable (accessible)
- single-agent vs. multiagent
- deterministic vs. non-deterministic
- episodic vs. sequential
PEAS description

- Performance measure:
- Environment:
- Actuators:
- Sensors:
  - fully observable vs. partially observable (accessible)
  - single-agent vs. multiagent
  - deterministic vs. non-deterministic
  - episodic vs. sequential
  - static vs. dynamic
PEAS description

- Performance measure:
- Environment:
- Actuators:
- Sensors:

- fully observable vs. partially observable (accessible)
- single-agent vs. multiagent
- deterministic vs. non-deterministic
- episodic vs. sequential
- static vs. dynamic
- discrete vs. continuous
PEAS description

- Performance measure:
- Environment:
- Actuators:
- Sensors:

- fully observable vs. partially observable (accessible)
- single-agent vs. multiagent
- deterministic vs. non-deterministic
- episodic vs. sequential
- static vs. dynamic
- discrete vs. continuous
- known vs. unknown
Next week: Relaxing the Assumptions

- Textbook readings: continuous spaces, nondeterminism, partial observations, unknown environments, ...
- Responses both Monday and Wednesday
- Search programming assignment