

CS394R/ECE381V
Reinforcement Learning:
Theory and Practice

Amy Zhang and Peter Stone

Departments of ECE and CS
The University of Texas at Austin

Good Morning Colleagues

- Are there any (logistics) questions?

Logistics

- Next step: literature surveys

Logistics

- Next step: literature surveys
 - Build on proposal

Logistics

- Next step: literature surveys
 - Build on proposal
 - Have some progress: working prototype!

Logistics

- Next step: literature surveys
 - Build on proposal
 - Have some progress: working prototype!
- Next week's readings

Logistics

- Next step: literature surveys
 - Build on proposal
 - Have some progress: working prototype!
- Next week's readings
 - Exploration and intrinsic motivation

Abstraction

- What are the two types?

Abstraction

- What are the two types?
 - State abstraction
 - Temporal abstraction

Options

- Extension of RL to temporal abstraction

Options

- Extension of RL to temporal abstraction
- They don't address **what** temporal abstraction to use — they just show how it can fit into the RL formalism

Options

- Extension of RL to temporal abstraction
- They don't address **what** temporal abstraction to use — they just show how it can fit into the RL formalism
 - Why couldn't it before?

Options

- Extension of RL to temporal abstraction
- They don't address **what** temporal abstraction to use — they just show how it can fit into the RL formalism
 - Why couldn't it before?
- Markov vs. Semi-markov:
 - states, actions
 - mapping from (s, a) to expected discounted reward
 - well-defined distribution of next state, transit time

Thinking about Options

- I'm a robot learning to get from here to the tower
 - I can turn 45 degrees in either direction or move 100cm forward
 - Starting with a random policy, how long will it take to learn?

Thinking about Options

- I'm a robot learning to get from here to the tower
 - I can turn 45 degrees in either direction or move 100cm forward
 - Starting with a random policy, how long will it take to learn?
 - Assume we add 1 option to the action space with a correct policy for getting from any location to the nearest door of this room
 - In what ways will it speed up/slow down learning?
 - What could you do to improve the situation?

Thinking about Options

- I'm a robot learning to get from here to the tower
 - I can turn 45 degrees in either direction or move 100cm forward
 - Starting with a random policy, how long will it take to learn?
 - Assume we add 1 option to the action space with a correct policy for getting from any location to the nearest door of this room
 - In what ways will it speed up/slow down learning?
 - What could you do to improve the situation?
- Options can be detrimental without good state abstractions (slides)

Questions about Options

- Will a good set of learned options with a good state abstraction always lead to an optimal policy?

Questions about Options

- Will a good set of learned options with a good state abstraction always lead to an optimal policy?
- Different forms of optimality (slides)

Policy-Irrelevance Abstractions

- Reminder from Tuesday

Policy-Irrelevance Abstractions

- Reminder from Tuesday
- I'm that same robot navigating to the tower
 - I have trouble navigating through crowds
 - Included in my state representation is how many people are in every 10m x 10m zone on campus

Policy-Irrelevance Abstractions

- Reminder from Tuesday
- I'm that same robot navigating to the tower
 - I have trouble navigating through crowds
 - Included in my state representation is how many people are in every 10m x 10m zone on campus
 - What would be a good policy-irrelevance abstraction?

Policy-Irrelevance Abstractions

- Reminder from Tuesday
- I'm that same robot navigating to the tower
 - I have trouble navigating through crowds
 - Included in my state representation is how many people are in every 10m x 10m zone on campus
 - What would be a good policy-irrelevance abstraction?
- Taxi domain - same question

Policy-Irrelevance Abstractions

- Reminder from Tuesday
- I'm that same robot navigating to the tower
 - I have trouble navigating through crowds
 - Included in my state representation is how many people are in every 10m x 10m zone on campus
 - What would be a good policy-irrelevance abstraction?
- Taxi domain - same question
- How could you learn them?
- How might they support transfer learning? (slides)

Policy-Irrelevance Abstractions

- Reminder from Tuesday
- I'm that same robot navigating to the tower
 - I have trouble navigating through crowds
 - Included in my state representation is how many people are in every 10m x 10m zone on campus
 - What would be a good policy-irrelevance abstraction?
- Taxi domain - same question
- How could you learn them?
- How might they support transfer learning? (slides)

Successor Features

- What would you do if you could explore before you get a reward function?