Good Afternoon Colleagues

- Are there any questions?
Good Afternoon Colleagues

• Are there any questions?

• Pending questions:
  – Weaknesses of subgoal discovery as presented.
  – Why can state abstraction prevent achieving planned values?
Logistics

- Tom Dietterich visiting next Friday:
  “Three Challenges for Machine Learning Research”
  3pm, ACES 2.302
MAXQ

• Defines how to learn given a task hierarchically
MAXQ

- Defines how to learn given a task hierarchically
- Does not address how to construct the hierarchy
MAXQ

- Defines how to learn given a task hierarchically
- Does not address how to construct the hierarchy
- Strives for recursive optimality
MAXQ

- Defines how to learn given a task hierarchically
- Does not address how to construct the hierarchy
- Strives for **recursive optimality**— local optimality for each subtask
MAXQ

- Defines how to learn given a task hierarchically
- Does not address how to construct the hierarchy
- Strives for recursive optimality—local optimality for each subtask
- Enables reuse of subtasks
MAXQ

- Defines how to learn given a task hierarchically
- Does not address how to construct the hierarchy
- Strives for **recursive optimality**— local optimality for each subtask
- Enables reuse of subtasks
- Enables useful state abstraction
Some details

- a means both primitive actions and subtasks (options)
Some details

- $a$ means both primitive actions and subtasks (options)
- Higher-level subtasks are essentially policies over options
  - But subtasks are learned too
  - And the values propagate correctly
Some details

- $a$ means both primitive actions and subtasks (options)
- Higher-level subtasks are essentially policies over options
  - But subtasks are learned too
  - And the values propagate correctly
- What does $C_i^\pi(s, a)$ mean?
Some details

- $a$ means both primitive actions and subtasks (options)
- Higher-level subtasks are essentially policies over options
  - But subtasks are learned too
  - And the values propagate correctly
- What does $C_i^\pi(s, a)$ mean? (Dietterich slides)
Some details

- $a$ means both primitive actions and subtasks (options)
- Higher-level subtasks are essentially policies over options
  - But subtasks are learned too
  - And the values propagate correctly
- What does $C_{i}^{\pi}(s, a)$ mean? (Dietterich slides)
- How does equation (2) relate to flat Q?
Some details

- $a$ means both primitive actions and subtasks (options)
- Higher-level subtasks are essentially policies over options
  - But subtasks are learned too
  - And the values propagate correctly
- What does $C_i^\pi(s, a)$ mean? (Dietterich slides)
- How does equation (2) relate to flat Q?
- The parameterization is deceptive, but there IS reuse.
Student-led discussion

• Jon on safe abstraction
Discussion Points

• What does MAXQ-Q buy you over flat?
Discussion Points

- What does MAXQ-Q buy you over flat?
- What does polling buy you over flat?
Discussion Points

• What does MAXQ-Q buy you over flat?

• What does polling buy you over flat?

• Would learning the subtasks from the bottom up help?
Discussion Points

• What does MAXQ-Q buy you over flat?

• What does polling buy you over flat?

• Would learning the subtasks from the bottom up help?