CS394R
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

- Project proposals!

Logistics

- Project proposals!

- Practice midterm available

Logistics

- Project proposals!

- Practice midterm available

- Midterm after spring break

Logistics

- Project proposals!

- Practice midterm available

- Midterm after spring break

  - 1 page of notes - turn it in
  - No internet searches or prior code
  - Calculators or code written on the fly is OK

Logistics

- Project proposals!

- Practice midterm available

- Midterm after spring break

    – 1 page of notes - turn it in
    – No internet searches or prior code
    – Calculators or code written on the fly is OK

- Next readings

Logistics

- Project proposals!

- Practice midterm available

- Midterm after spring break

  – 1 page of notes - turn it in
  – No internet searches or prior code
  – Calculators or code written on the fly is OK

- Next readings

  – Applications

Learning a Quadruped Walk

- Introduction to the Aibo

Learning a Quadruped Walk

- Introduction to the Aibo

- Motor commands sent at 10Hz
  - PID control - choose a set point

Learning a Quadruped Walk

- Introduction to the Aibo

- Motor commands sent at 10Hz
  - PID control - choose a set point

- Goal: get it to learn how to walk fast
  - Learn only on the real robot

Learning a Quadruped Walk

- Introduction to the Aibo

- Motor commands sent at 10Hz
  - PID control - choose a set point

- Goal: get it to learn how to walk fast
  - Learn only on the real robot

- Formulate it as a value function learning problem

Learning a Quadruped Walk

- Introduction to the Aibo

- Motor commands sent at 10Hz
  - PID control - choose a set point

- Goal: get it to learn how to walk fast
  - Learn only on the real robot

- Formulate it as a value function learning problem

- Formulate it as a policy gradient problem

Learning a Quadruped Walk

- Introduction to the Aibo

- Motor commands sent at 10Hz
  - PID control - choose a set point

- Goal: get it to learn how to walk fast
  - Learn only on the real robot

- Formulate it as a value function learning problem

- Formulate it as a policy gradient problem
  - Start with a parameterized walk