

16-899C ACRL:

Adaptive Control and Reinforcement Learning

Machine Learning Techniques for Decision Making, Planning and Control

Time and Day: Spring, 2010, 4:30-5:50, Tuesday and Thursday, NSH ????

Instructors: Drew Bagnell (dbagnell@ri.cmu.edu)

Office Hours: Drew Bagnell, Tuesday and Thursday AM, by appointment



Why?

Machine learning has escaped from the cage of perception. A growing number of state-of-the-art systems from field robotics, acrobatic autonomous helicopters, to the leading computer Go player and walking robots rely upon learning techniques to make decisions. This change represents a truly fundamental departure from traditional classification and regression methods as such learning systems must cope with a) their own effects on the world, b) sequential decision making and long control horizons, and c) the

exploration and exploitation trade-off.

In the last 5 years, techniques and understanding of these have developed dramatically. One key to the advance of learning methods has been a tight integration with optimization techniques, and as such our case studies will focus on this.

What? (Things we will cover)

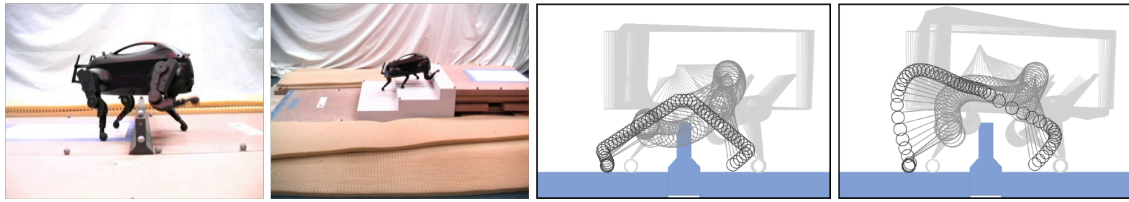
Planning and Optimal Control Techniques

- Differential Dynamic Programming
- Elastic Bands and Functional Optimization over the Space of Trajectories
- Iterative Learning Control
- Approximately optimal search and fast replanning: D*, ARA*
- Linear and neighboring optimal control

Imitation Learning and Inverse Optimal Control

- Imitation Learning as Structured Prediction
- Inverse Optimal Control
- LEARning to searCH and Maximum Margin Planning
- Maximum Entropy Inverse Optimal Control
- Personally customized routing navigation





Reinforcement Learning and Adaptive Control

- Temporal Differences
- Approximate Policy Iteration
- Function approximation for values and policies

Exploration

- Bandit algorithms for limited feedback learning
- Contextual bandits and optimal decision making
 - o “Sliding Autonomy” by contextual bandit methods
- Dual Control
 - o “Bayesian” Reinforcement learning and optimal control for uncertain models
- “Unscented” linear quadratic regulation

Policy Search Methods

- Direct Policy Search Methods and Stochastic Optimization
 - o Optimization of walking gaits and stabilizing controllers
- Conservative Policy Iteration
- Policy Search by Dynamic Programming
- REINFORCE and Policy Gradient Methods

Motion Planning

- Motion Planning that learns from experience
 - o Trajectory libraries
 - o Learning heuristics to speed planning

Design for Learnability

- Identifying feedback sources
- Modular learning design and structured problem
- Engineering insight as features and priors

Planning/Decision making under Uncertainty

Value-functions and stochastic planning

Partially Observed Markov Decision

Processes and Information Space Planning

Belief Compression

Value of information and active learning



Who?

This course is directed to students—primarily graduate although talented undergraduates are welcome as well—interested in developing adaptive *software that makes decisions that affect the world*. Although much of the material will be driven by applications within mobile robotics, anyone interested in applications of learning to planning and control techniques or an interest in building complex adaptive systems is welcome.

Prerequisites

As an advanced course, familiarity with basic ideas from probability, machine learning, and control/decision making are strongly recommended. Useful courses to have taken in advance include Machine Learning, Statistical Techniques in Robotics, Artificial Intelligence, and Kinematics, Dynamics, and Control. As the course will be project driven, prototyping skills including C, C++, and/or Matlab will also be important. Creative thought and enthusiasm are required.

How?

The course will include a mix of homework assignments that exercise the techniques we study, quizzes to demonstrate proficiency with the theoretical tools, and a strong emphasis on a significant research project.

Grading

Final grades will be based on the homeworks (30%), midterm (20%), final project (40%), and class participation and attendance (10%)

Late homework policy:

You will be allowed 2 total late days without penalty for the entire semester. Once those days are used, you will be penalized according to the following policy:

- Homework is worth full credit at the beginning of class on the due date
- It is worth half credit for the next 48 hours.
- It is worth zero credit after that.

You must turn in all homework, even if for zero credit.

Collaboration on homeworks:

Unless otherwise specified, homeworks will be done individually and each student must hand in their own assignment. It is acceptable, however, for students to collaborate in figuring out answers and helping each other understand the underlying concepts. You must write on each homework the names of the students you collaborated with.

Project

Projects may be done in groups of up to 3 students. The project is an opportunity to make a significant exploration into the application of ideas from the course to a robotics problem. More information to follow.

Exams

There will be a midterm exam but no final. It will be open book and open notes (no computers allowed).

Scribed Notes

We use a scribing system in lectures that worked well last year in the course. Since the lectures are very open ended and mostly done on the board, every member of class will take turns taking detailed notes on the lectures which they will type up with any necessary figures (preferably using LaTeX) to be posted on the website. This will help maintain detailed course notes that everyone can look back at and study from later. This will mean that each person in the class (including those auditing) will scribe about 2 lectures. Please be thorough since people will be using these to review for assignments and exams. **Your writeups will be graded for a portion of the homework grade.**

Scribing will be done in alphabetical order

A LaTeX template for you to use has been posted to:

www.cs.cmu.edu/~dbagnell/ACRL/scribe.tar.gz

Lecture notes will be due within 3 days of the lecture.

Auditing

If you do not wish to take the class for credit, you must register to audit the class. To satisfy the auditing requirement, you must either:

- Do two homework assignments, at least one of which must be one of the homeworks requiring the implementation of algorithms discussed in class
- Work with a team on a class project

Textbooks (all optional) that will benefit discussion

Optional Textbook: [Probabilistic Robotics](#), Sebastian Thrun, Wolfram Burgard, Dieter Fox

Optional Textbook: [Pattern Recognition and Machine Learning](#), Chris Bishop

Optional Textbook: [Optimal Control and Estimation](#), R. Stengle

Optional Textbook: *Model-Based Control of a Robot Manipulator*, C. H. An, C. G. Atkeson, and J. M. Hollerbach,

Optional Textbook: [Adaptive Control](#), K. J. Astrom

Optional Textbook: [Convex Optimization](#), Stephen Boyd and Lieven Vandenberghe

Optional Textbook: [Reinforcement Learning: An Introduction](#), R. Sutton and A. Barto

Additional readings will be posted on the course website.



Quality Of Life Technology Center Class

Many examples throughout the class, homeworks, as well as available class projects will focus on the research needs of the Quality of Life Technology Center ERC. QoLT students are particularly encouraged to use research projects as part of their in class exercises, and to work with the instructors to encourage QoLT relevant homework's and problem sets.