CSE 6369.001 - Reinforcement Learning

Spring 2014 - TTh 2:00 - 3:20

Instructor: Manfred Huber (huber@cse.uta.edu)

Course Description

Contents and Objectives:

Machine learning techniques are increasingly employed in a wide range of areas to model and analyze data as well as to facilitate decision support and autonomous decision making by computer systems. Reinforcement learning is an important machine learning paradigm in particular in the context of decision support and decision making, but also in the context of modeling when only limited feedback is available. This course will introduce the Reinforcement Learning paradigm and its underlying formalisms, and will cover a wide range of basic and advanced Reinforcement Learning algorithms as well as aspects of model learning, hierarchy and abstraction, and reward modeling. Throughout, this course will study these techniques in the context of a wide range of application areas, including robotics, computer vision, security, control, scheduling, and data analysis. Students completing this course will gain an understanding of the field and be able to apply modern, state-of-the-art Reinforcement Learning techniques to a wide range of problems and applications.

Prerequisites:

Many of the techniques covered in this course are based on probabilities and random processes and a basic background in statistics is required for the course. Prerequisites for the course are an advanced statistics and random processes course (*CSE 5301* or similar), or consent of instructor. In addition, experience with programming will be useful for assignments and projects.

Course Materials:

The course will mainly use the following textbook:
R. Sutton and A. Barto, *Reinforcement Learning: An Introduction*, MIT Press, 1998.

In addition the course will use readings from other books as well as papers from technical conferences and journals. These materials will be made available through the engineering library or the course site.

E-mail and WWW page:

There is a course web page at *http://www-cse.uta.edu/∼huber/cse6369_rl* . All changes and supplementary course materials will be available from this site. In addition, necessary changes or important announcements will also be distributed by e-mail.

Tentative Office Hours:

Office hours for the course will be held by the instructor in ERB 128 or in ERB 522, TTh 3:30 - 4:30, and TTh 6:30 - 7:00. Times are subject to change and will be posted. If you can not make it to any of these office hours, please inform the instructor.
Course related emails should be sent to the instructor at huber@cse.uta.edu and should list the course number in the subject line.

1 Course Work and Grading

In-class presentation of a technical paper:
Each student will be given a technical topic/paper to present in class and to lead the following discussion of the material.

Homework Assignments:
Two hands-on homework assignments will be given where learned techniques are applied to practical problems.

Projects:
Two small projects will be assigned where students implement and test some of the techniques.

In addition, every student will perform a final project that will be presented at the end of the course.

Grading Policy:
The final grade will be calculated using the following policy:

Presentation & Class Participation	15 %
Assignments	30 %
Projects	30 %
Final Project	25 %

2 Course Topics

Topics covered in this course include:

- **Reinforcement Learning methodology and underlying models**
- **Basic Reinforcement Learning Approaches:** Value Iteration, Actor-Critic, Policy iteration
- **Exploration / Exploitation Tradeoff:** Exploration strategies
- **Model-based Reinforcement Learning:** Algorithms and model learning techniques
- **Reinforcement Learning in Partially Observable systems**
- **Hierarchical Reinforcement Learning**
- **Inverse Reinforcement Learning**
- **Multiagent Reinforcement Learning**

Exact topics are subject to change and may depend on available time.

3 Tentative Class Schedule

CSE 6369
Reinforcement Learning
Tentative Lecture and Assignment Schedule
Spring Semester 2014 - TTh 2:00 - 3:20

Class	Date	Readings	Lecture Topics	Assignments
1	01/14	1.1 - 1.6	Course Overview and Introduction	
2	01/16	2.1 - 2.4	Background - Evaluative Feedback	
3	01/21	2.5 - 2.10	Background - Utility and Decision Theory	
4	01/23	3.1 - 3.10	Background - Probabilistic models and MDPs	
5	01/28	4.4, 4.7	Dynamic Programming Methods - Value Iteration	
6	01/30	4.1 - 4.6	Dynamic Programming Methods - Policy Iteration	
7	02/04	5.1 - 5.8	Monte Carlo Methods	
8	02/06	6.1 - 6.5	Temporal Difference Learning	
9	02/11	6.6 - 6.9	Actor-Critic Models	Homework 1 due
10	02/13		Actor-Critic Models	
11	02/18		On-Policy vs Off-Policy Learning	
12	02/20		Function Approximation in Reinforcement Learning	
13	02/25		Function Approximation in Reinforcement Learning	
14	02/27		Exploration vs. Exploitation Tradeoff	
15	03/04		Model-Based Learning	Project 1 due
16	03/06		Learning Models for Reinforcement Learning	
	03/11		*Spring Break - No Class*	
	03/13		*Spring Break - No Class*	
17	03/18		Learning Models for Reinforcement Learning	
18	03/20		Efficient Model-Based Learning	
19	03/25		Efficient Model-Based Learning	
20	03/27		Learning in Partially Observable Systems	
21	04/01		Learning in Partially Observable Systems	Homework 2 due
22	04/03		Learning in Partially Observable Systems	
23	04/08		Hierarchical Reinforcement Learning	
24	04/10		Hierarchical Reinforcement Learning	
25	04/15		Hierarchical Reinforcement Learning	
26	04/17		Inverse Reinforcement Learning	Project 2 due
27	04/22		Reinforcement Learning in Multiagent Domains	
28	04/24		Student Presentations	
29	04/29		Student Presentations	
30	05/01		Student Presentations & Current Challenges	
31	05/06		**Final Project Presentations (2:00pm-4:30pm)**	

[1] All information is tentative and subject to change.

4 University Policies and Services

Grade Grievances:

Any appeal of a grade in this course must follow the procedures and deadlines for grade-related grievances as published in the current undergraduate catalog.

Drop Policy:

The standard UTA drop policy applies to this course. Students may drop or swap (adding and dropping a class concurrently) classes through self-service in MyMav from the beginning of the registration period through the late registration period. After the late registration period, students must see their academic advisor to drop a class or withdraw. Undeclared students must see an advisor in the University Advising Center. Drops can continue through a point two-thirds of the way through the term or session. It is the student's responsibility to officially withdraw if they do not plan to attend after registering. Students will not be automatically dropped for non-attendance. Repayment of certain types of financial aid administered through the University may be required as the result of dropping classes or withdrawing. For more information, contact the Office of Financial Aid and Scholarships (http://wweb.uta.edu/ses/fao).

Americans With Disabilities Act:

The University of Texas at Arlington is on record as being committed to both the spirit and letter of all federal equal opportunity legislation, including the Americans with Disabilities Act (ADA). All instructors at UT Arlington are required by law to provide "reasonable accommodations" to students with disabilities, so as not to discriminate on the basis of that disability. Any student requiring an accommodation for this course must provide the instructor with official documentation in the form of a letter certified by the staff in the Office for Students with Disabilities, University Hall 102. Only those students who have officially documented a need for an accommodation will have their request honored. Information regarding diagnostic criteria and policies for obtaining disability-based academic accommodations can be found at www.uta.edu/disability or by calling the Office for Students with Disabilities at (817) 272-3364.

Academic Integrity:

All students enrolled in this course are expected to adhere to the UT Arlington Honor Code: I pledge, on my honor, to uphold UT Arlington's tradition of academic integrity, a tradition that values hard work and honest effort in the pursuit of academic excellence. I promise that I will submit only work that I personally create or contribute to group collaborations, and I will appropriately reference any work from other sources. I will follow the highest standards of integrity and uphold the spirit of the Honor Code.
Instructors may employ the Honor Code as they see fit in their courses, including (but not limited to) having students acknowledge the honor code as part of an examination

or requiring students to incorporate the honor code into any work submitted. Per UT System Regents' Rule 50101, paragraph 2.2, suspected violations of university's standards for academic integrity (including the Honor Code) will be referred to the Office of Student Conduct. Violators will be disciplined in accordance with University policy, which may result in the student's suspension or expulsion from the University.

Student Support Services:

UT Arlington provides a variety of resources and programs designed to help students develop academic skills, deal with personal situations, and better understand concepts and information related to their courses. Resources include tutoring, major-based learning centers, developmental education, advising and mentoring, personal counseling, and federally funded programs. For individualized referrals, students may visit the reception desk at University College (Ransom Hall), call the Maverick Resource Hotline at 817-272-6107, send a message to resources@uta.edu, or view the information at www.uta.edu/resources.

Electronic Communication:

UT Arlington has adopted MavMail as its official means to communicate with students about important deadlines and events, as well as to transact university-related business regarding financial aid, tuition, grades, graduation, etc. All students are assigned a MavMail account and are responsible for checking the inbox regularly. There is no additional charge to students for using this account, which remains active even after graduation. Information about activating and using MavMail is available at http://www.uta.edu/oit/cs/email/mavmail.php.

Student Feedback Survey:

At the end of each term, students enrolled in classes categorized as lecture, seminar, or laboratory shall be directed to complete a Student Feedback Survey (SFS). Instructions on how to access the SFS for this course will be sent directly to each student through MavMail approximately 10 days before the end of the term. Each student's feedback enters the SFS database anonymously and is aggregated with that of other students enrolled in the course. UT Arlington's effort to solicit, gather, tabulate, and publish student feedback is required by state law; students are strongly urged to participate. For more information, visit http://www.uta.edu/sfs.

Final Review Week:

A period of five class days prior to the first day of final examinations in the long sessions shall be designated as Final Review Week. The purpose of this week is to allow students sufficient time to prepare for final examinations. During this week, there shall be no scheduled activities such as required field trips or performances; and no instructor shall assign any themes, research problems or exercises of similar scope that have a completion date during or following this week unless specified in the class syllabus. During Final Review Week, an instructor shall not give any examinations constituting 10% or more of the final grade,

except makeup tests and laboratory examinations. In addition, no instructor shall give any portion of the final examination during Final Review Week. During this week, classes are held as scheduled. In addition, instructors are not required to limit content to topics that have been previously covered; they may introduce new concepts as appropriate.

Emergency Exit Procedures:

Should we experience an emergency event that requires us to vacate the building, students should exit the room and move toward the nearest exit, which is located to the right when exiting the room. When exiting the building during an emergency, one should never take an elevator but should use the stairwells. Faculty members and instructional staff will assist students in selecting the safest route for evacuation and will make arrangements to assist handicapped individuals.