Reinforcement learning
Spring 2008
Wednesdays 2-4pm*
815 Meyer Hall

Instructor: Nathaniel Daw
Office: 873 Meyer Hall
Office hours: By appointment
email: daw@gatsby.ucl.ac.uk

*** Important note on scheduling:**

The course as currently scheduled conflicts with the Psychology core course in Cognition, which starts at 3:30. If this is a problem for you, don't worry: it's a problem for me also, since I'm supposed to teach it sporadically. If you want to take both classes for credit, go ahead and register for both. (You may need to visit the Psych Academic Affairs office to approve it – they have both instructors' permission to do so.) Since efforts to fix this problem officially have not met with any success, we'll unofficially negotiate among ourselves what to do once we see who's in the class. In the worst case, we'll let those people who need to go (including, sometimes, myself) leave early.

Overview:

This course will be structured around an overview of methods and concepts from reinforcement learning (RL), the field of artificial intelligence concerned with learned trial-and-error decision making. Our goal is not to develop detailed expertise in the computer science per se, but rather to draw on the frameworks that computer scientists have developed, the problems they have elucidated and the solutions they have suggested, as jumping off points for understanding the neuroscience and psychology of learning. The thesis we will explore throughout the semester is that the careful and concrete formalizations of issues in learning developed by engineers may be a useful conceptual resource for designing and interpreting experiments.

To that end, we will jump between many areas in RL, pairing readings from the computer science literature, where possible, with readings from the neural and behavioral sciences. In some areas, the analogies are already well developed and explicit; in others, they are as yet nonexistent. In these cases, the intent of our discussion will be to imagine how these connections might be drawn.

Format:

We will not assume a quantitative or technical background (and we will not have problem sets or programming assignments or the like), though of course one of our goals will be to develop at least a big-picture familiarity and facility with formal approaches to learning. To ease the transition, for the first two meetings (and sporadically as nececessary), I will deliver background lectures to introduce some basic formal concepts; after this we will switch to a more traditional journal-club-like discussion-oriented seminar format.

For these sessions, *everyone will be expected to have read the readings and come prepared to discuss* *them.* Students will switch off responsibility for leading the discussion; this will involve, first, a structured (roughly half-hour) presentation of the material and, more importantly, asking questions and facilitating discussion for the remaining time. Students auditing the class will also be expected to both participate and present; how many presentations per student (and students per presentation) will depend on the size of the class.

Course requirements:

33% for presentations and leading discussions.

33% for participation in discussions.

33% for the final paper.

About that: Students enrolled in the class will be expected to write a final paper, due on the last day of class. The paper can be either in the format of an NRSA application (Specific Aims, Background Literature, Preliminary Data if applicable, Research Plan) proposing some experiment designed to address one of the issues considered during the term, or a primary research article. Students should check with the instructor as to the topic, but the intent is to construe the possibilities quite broadly so as to enable you both to pursue a topic and approach of interest, and if possible to enable you to devote your efforts to an application or article that will be of additional use to you in your ongoing graduate studies. Students will briefly present their final papers on the last day of class.

Readings:

Background textbook: Sutton & Barto, *Reinforcement Learning: An Introduction*. We won't stick particularly close to this (except in the first couple of weeks) but it is a good resource and available online at:

http://cognet.mit.edu/library/books/view?isbn=0262193981
http://www.cs.ualberta.ca/%7Esutton/book/ebook/the-book.html

Readings will be from the primary literature, and will be provided at least a week before class.

Preliminary schedule of topics:

1/23: Preliminaries & details, Markov decision processes (lecture)
1/30: Dynamic programming, on-line learning (lecture)
2/6: Prediction errors & dopamine
2/13: Model-based learning & planning
2/20: Function approximation & generalization
2/27: Gradient vs value function techniques
3/5: Non-Markov situations, observation noise & partial observability
3/12: Partial observability 2: sensory decision making
3/19: No class (spring break)
3/26: Will reschedule (honeymoon): Temporal abstraction, action sequences & options
4/2: Multi-effector control

4/9: Learning in multiplayer games
4/16: The explore/exploit dilemma
4/23: Metalearning and learning to learn
4/30: Inverse reinforcement learning