

RL and Planning Under Uncertainty (ANU, Sem2, 2008)
Lab1: Learning to Drive using Value Iteration & RTDP
Lab Instructor: Scott Sanner

READ THE SUTTON AND BARTO BOOK, CHAPTERS 1-4!!!

It's online if you're too cheap to buy it... but it is a great book, and I highly recommend buying it (in stock in bookstore).

Agenda:

1. Discuss asynchronous dynamic programming.
2. Discuss real-time dynamic programming (RTDP) algorithm and why it converges without touching all states.
3. Discuss code.

Student lab work:

1. Implement value iteration
2. Implement RTDP (simple: update states encountered during trial)
3. (Optional for 10% assignment Extra Credit) Improve RTDP substantially
 - a. Backward sweep after completing trajectory?
 - b. Use dual admissible bounds, prioritize exploration.
 - c. Use lower bound policy during actual trials.
 - d. Other ideas? See paper below.

For further reading:

Learning to Act Using Real-time Dynamic Programming: Barto, Bradtke, Singh (1995), discusses connections to LRTA*

Labelled RTDP: Bonet and Geffner (2003), discusses connections to AO*

Bounded RTDP: McMahan and Gordon (2005)

Focused RTDP: Smith and Simmons (2006)