On The Virtues of Linear Learning and Trajectory Distributions

Richard S. Sutton
University of Massachusetts at Amherst
rich@gte.com

In contrast to recent work by Boyan and Moore, I have obtained excellent results applying conventional VFA methods to control tasks such as "Mountain Car" and "Acrobot". The difference in our results is due to the form of function approximator and the nature of the training. I used a local function approximator known as a CMAC—essentially a linear learner using a sparse, coarse-coded representation of the state. Although not a complete solution, I argue that this approach solves the "curse of dimensionality" as well as one can expect, and enables VFA systems to generalize as well as other artificial learning systems. Also, I used TD(λ) and trained on state transitions from actual, experienced trajectories. A theorem by Dayan assures stability of TD(λ) when trained using such *trajectory distributions*. Gordon and Tsitsiklis and Van Roy have recently shown that TD(λ) can be unstable when trained with other distributions. Finally, I present a small compendium of results all showing much faster learning when λ is slightly less than 1.

Much of this work is described in detail in a paper submitted to NIPS, "Generalization in Reinforcement Learning: Successful Examples Using Sparse Coarse Coding". That paper is available electronically via `http://envy.cs.umass.edu/People/sutton/archive.html` or `ftp://ftp.cs.umass.edu/pub/anw/pub/sutton/sutton-inprep.ps`.