Hierarchical adaptive critics for neural aircraft control

Thaddeus T. Shannon and George G. Lendaris
Systems Science Program, Portland State University,
P.O. Box 751, Portland, OR 97207,
E-mail: tads@sysc.pdx.edu, lendaris@sysc.pdx.edu

Purpose
Many control tasks that intelligent control systems are designed for contain multiple objectives together with dynamics on multiple time scales. Adaptive critic based approximate dynamic programming techniques offer a principled, generally applicable design methodology for such problem contexts. This paper demonstrates an approach for using adaptive critics that takes advantage of multiple objectives and time scales to simplify the design process. We show how to use multiple critics in a hierarchical manner to decompose the learning process for a neurocontroller. This approach can lead to simpler controllers, more parsimonious critic structures, faster learning and increased transparency of the design process.

Method
Our demonstration project is the design of a neurocontroller for a simulated experimental aircraft. The aircraft is naturally unstable, so our design objectives include stabilization of the aircraft within a nominal attitude and velocity envelope as well as the tracking of a predefined flight path. Our basic design methodology uses the Dual Heuristic Programming (DHP) method of approximate dynamic programming to train a neurocontroller. We structure this controller using two separate networks, one for the stabilization task and one for the maneuvering task. We then train the controllers using a hierarchy of critics, a fast-dynamic stability critic, a slow-dynamic maneuvering critic and a slow-dynamic stability critic.

Results
We demonstrate that our hierarchy of critics can successfully train a neurocontroller that satisfactorily achieves both the stability and maneuvering objectives for our simulated aircraft.

New aspect of work
This is the first demonstration of the DHP methodology that focuses on the use of multiple critics to enable a hierarchical decomposition of objectives and plant dynamics. In previous work we have advocated incorporating as much problem specific information as possible into both the critic and controller structures used for DHP. This current proposal extends that idea by enabling the use of both expert and common sense knowledge to structure the actual learning process.

Conclusions
This paper demonstrates a generally applicable approach for implementing adaptive critic methods for control problems involving multiple objectives and/or time scales. Our approach structures controller training in a hierarchical manner, with the goal of simplifying the learning process.