Chapter 5

Adaptive Dynamic Programming Applied to a 6DoF Quadrotor

Petru Emanuel Stingu
University of Texas at Arlington, USA

Frank L. Lewis
University of Texas at Arlington, USA

ABSTRACT

This chapter discusses how the principles of Adaptive Dynamic Programming (ADP) can be applied to the control of a quadrotor helicopter platform flying in an uncontrolled environment and subjected to various disturbances and model uncertainties. ADP is based on reinforcement learning. The controller (actor) changes its control policy (action) based on stimuli received in response to its actions by the critic (cost function, reward). There is a cause and effect relationship between action and reward. Reward acts as a reinforcement signal that leads to learning of what actions are likely to generate it. After a number of iterations, the overall actor-critic structure stores information (knowledge) about the system dynamics and the optimal controller that can accomplish the explicit or implicit goal specified in the cost function.

INTRODUCTION

There is currently a dichotomy between optimal control and adaptive control. Adaptive Control algorithms learn online and give controllers with guaranteed performance for unknown systems. On the other hand, optimal control design is performed off line and requires full knowledge of the system dynamics. In this research we designed Optimal Adaptive Controllers, which learn online in real-time and converge to optimal control solutions. For linear time-invariant systems, these controllers solve the Riccati equation online in real-time by using data measured along the system trajectories. These results show how to approximately solve the optimal control problem for nonlinear systems online in real-time, while

DOI: 10.4018/978-1-60960-551-3.ch005
simultaneously guaranteeing that the closed-loop system is stable, i.e. that the state remains bounded. This solution requires knowledge of the plant dynamics, but in future work it is possible to implement algorithms that only know the structure of the system and not the exact dynamics.

The main focus of this chapter is to present different mechanisms for efficient learning by using as much information as possible about the system and the environment. Learning speed is crucial for a real-time, real-life application that has to accomplish a useful task. The control algorithm isn’t usually allowed to generate the best commands suitable for exploration and for learning, because this would defeat the purpose of having the controller in the first place, which is to follow a designated trajectory. The information gathered along the trajectory has to be used efficiently to improve the control policy. There is a big amount of data that has to be stored for such a task. The system is complex and has a large number of continuous state variables. The value function and the policy that corresponds to the infinite number of combinations of state variable values and possible commands have to be stored using a finite number of parameters. The coding of these two functions is made using function approximation with a modified version of Radial Basis Function (RBF) neurons. Due to their local effect on the approximation, the RBF neurons are best suited to hold information that corresponds to training data generated only around the current operating point, which is what one can obtain by following a normal trajectory without exploration. The usual approach of using multilayer perceptrons that have a global effect suffers from having to do a compromise between learning speed and the dispersion of the training samples. For samples that are concentrated around the operating point, learning has to be very slow to avoid deteriorating the approximation precision for states that are far away.

Two very important characteristics of learning are generalization and classification. The amount of information gathered by the system corresponds only to particular state trajectories and particular commands. Still, the value of being in a certain state and of using a certain command has to be estimated over an infinite continuous space. The RBF neurons are able to interpolate between the specific points where data samples are stored. They don’t provide a global solution, but they certainly cover the space around the states likely to be visited in normal conditions.

The neural network structure is adaptive. Neurons are added or removed as needed. If for a specific operating point the existing neurons can’t provide enough accuracy to store a new sample, then a new neuron is added in that point. The modified RBF neurons are created initially with a global effect in all dimensions. It is only on the dimensions where there is a need to discern between different values of the state variable that the effect is local. This mechanism allows neurons to partition the state space very efficiently. If some state variables do not affect the value function or the control policy corresponding to a certain region of the state space, then the neurons in the vicinity of that region are global on those dimensions. This organization of the RBF network falls in line with the idea that if the function to be approximated is not very complicated, then a reasonably small number of parameters should be sufficient to achieve a small error even if the number of dimensions of the input space is large. This applies to smooth and nice behaving functions. In the worst case, the number of parameters needed grows exponentially with the number of inputs. For the current implementation, the total number of neurons is kept at a reasonable value by pruning the ones in regions that have been visited in the distant past and thus diluting the approximation precision in those regions.
Related Content

Social Impact and Challenges of Virtual Reality Communities
www.igi-global.com/chapter/social-impact-challenges-virtual-reality/43410?camid=4v1a

Artificial Neural Networks and Learning Techniques
www.igi-global.com/chapter/artificial-neural-networks-and-learning-techniques/140392?camid=4v1a

Prediction of Temperature in Buildings using Machine Learning Techniques
Juan José Carrasco, Juan Caravaca, Mónica Millán-Giraldo, Gonzalo Vergara, José M. Martínez-Martínez, Javier Sanchis and Emilio Soria-Olivas (2016). Handbook of Research on Advanced Computational Techniques for Simulation-Based Engineering (pp. 314-333).
www.igi-global.com/chapter/prediction-of-temperature-in-buildings-using-machine-learning-techniques/140395?camid=4v1a

The Future of Augmented Reality Gaming
www.igi-global.com/chapter/future-augmented-reality-gaming/10173?camid=4v1a