Project Proposal

Reinforcement Learning for Control of Communications Networks

Lead: *Martha Steenstrup*
Sponsor: *Richard Sutton*
Amount Requested: *$25,400*
Previous Proposal: *None*
Related Projects: *None*

Scientific Rationale

Unchanged? no

We will explore the application of reinforcement learning to control problems in the domain of communications networks. Our research, if successful, will produce network control algorithms that substantially reduce the costs of network planning, operation, and management and that promote more efficient use of network resources, thereby enabling a network to support a larger number of users and a wider range and better quality of service for those users. Moreover, lessons learned through exercising reinforcement learning methods for control in the complicated domain of communications networks may provide insight into new approaches for attacking some of the fundamental and as yet unsolved problems in the domain of reinforcement learning.

The quintessential communications network is a large, heterogeneous, dynamic, distributed system whose behavior depends on the properties of its component devices, the characteristics of the environment in which it operates, the behavior of its users, and the set of control algorithms that determine where, when, and how data is transported among communicating devices. Controllable aspects of a network vary widely in temporal and spatial scale, ranging from individual bits to the entire network infrastructure. Moreover, in most communications networks, the devices that collect observations of state, make decisions - i.e., take actions - based on these observations, and implement these actions are usually distributed throughout the network and often behave autonomously. Therefore, interactions between the control applied to different portions and aspects of a network can be complex and not always easily understood.

Traditionally, algorithms for network control have been designed by humans according to sound engineering principles, but their performance is inherently limited by human biases, in particular a propensity for adopting optimistic and simplified models of behavior for the network, users, and interactions among them that do not necessarily hold in practice. We expect that in most cases, network control algorithms that can autonomously learn effective control policies without innate knowledge of detailed behavioral models or supervisory input, make appropriate decisions under uncertainty, and extract useful features from observations of state - i.e., network control algorithms based on reinforcement learning - will be superior to human-engineered solutions in terms of accuracy, agility, robustness, and efficiency.

Reinforcement learning principles have been applied to network control as early as the 1970s when learning automata were proposed to increase the number of successful call completions in telephone networks. Since that time, reinforcement learning has been applied sporadically to network control problems - principally but not exclusively routing - but there has yet to be a concerted effort to systematically explore the use of reinforcement learning in network control and the potential performance gains of doing so. The long-term goal of our research is to conduct such an investigation.

Network control poses several interesting challenges that must be addressed for the successful application of reinforcement learning in this domain. Cast in the context of Markov decision processes, much of network control exhibits the following characteristics: both the state and action spaces are large and may be continuous; observations of state and reward are often delayed, partial, and noisy; and accurate and detailed models of the network, its users, and its operating environment are not necessarily available. Interactions among multiple control agents acting independently in a network may be synergistic or destructive. The degree of coordination necessary for constructive interaction among agents has yet to be fully characterized in the context of purely adaptive agents let alone those that learn control policies. Feature selection, critical for situation-appropriate decision making, tends to be idiosyncratic when performed by humans. Algorithms that learn useful features through exploration are not subject to these human biases, but efficient search for key features in large state spaces remains an illusive goal. We will explore each of these challenges from the outset of the project.

Proposal

The research project outlined above is large in scope and not something that can be completed quickly. To enable us to obtain some concrete results during the first year, we will carve out a small portion of the project on which to concentrate our efforts. In particular, we will choose a single well-defined network control problem, likely one related to transmission scheduling, and we will investigate the efficacy of reinforcement learning methods in solving this problem. This will be our initial case study. Our experiments will be conducted in small networks so that we will be better able to understand the underlying reasons for the observed behavior of our algorithms. For the first year, the proposed work comprises four tasks, each marked by a milestone as described below, and the first two of which can be performed in parallel.

Project Milestones

We plan to reach each of the following milestones over the next year.

A. *Characterization of the Problem*

We will select a network control problem expected to benefit from a reinforcement-learning-based solution, analyze the essential properties of the selected problem, and ultimately formulate the problem as a Markov decision process, defining the state and action spaces and the characteristics of the reward and state observation signals. Based on the properties of the problem thus formulated, we will determine which reinforcement learning methods are likely to yield the best performance on our chosen network control problem, and we will design solutions using these methods. We will also choose one or more existing state-of-the-art adaptive-only

solutions to the network control problem, or design our own if necessary, as points of comparison for the solutions that incorporate reinforcement learning.
Contributors: *Two Masters students to be determined and Martha Steenstrup*

B. *Design of a Network Simulation Testbed*

We will design and build a simple simulation testbed that will permit us to rapidly implement networks, control algorithms, and test scenarios and to easily modify these as needed. The testbed is not intended as an all-encompassing tool, but rather one that contains only those components required for our experiments, thus enabling us to spend the majority of our time designing and experimenting with our algorithms instead of developing software.
Contributors: *Two Masters students to be determined and Martha Steenstrup*

C. Design and Execution of Simulation Experiments

We will design a suite of experiments intended to expose both the strengths and weaknesses of each of our candidate algorithms - both those that employ reinforcement learning and those that are adaptive only - for solving our chosen network control problem. Our experiments will be instrumented to provide performance information including per-flow and network-wide throughput and end-to-end delay as well as volume of control traffic that must be transmitted in the network and amount of computation required in order for agents to make their decisions, thus enabling us to determine the relative performance of the candidate algorithms in terms of quality of service provided to user data traffic and the costs of providing that quality of service.
Contributors: *Two Masters students to be determined and Martha Steenstrup*

C. Analysis of Results Obtained from Simulation Experiments

Our analysis of the results obtained from experimentation will provide not only numerical assessment of the comparative performance of the selected algorithms with respect to the chosen network control problem but also an in-depth understanding of the reasons for the observed behavior of each algorithm, which in turn may lead us to modifications of the algorithms to improve their performance.
Contributors: *Two Masters students to be determined and Martha Steenstrup*

Success Criteria / Outcomes

We *will judge the success of this project in the coming year primarily on the results of our experiments and whether they enable us to draw meaningful conclusions about the behavior of our algorithms. We anticipate that lessons learned from this particular case study can be generalized to other problems in network control.*

By working on this project and by sharing knowledge between them, the two Masters students involved will develop expertise in both communications networks and reinforcement learning. They will also be expected to publish the results of their research so that others may benefit from what they have discovered.