DOAS 2006 PROJECT: REINFORCEMENT LEARNING OF TRAFFIC LIGHT CONTROLLERS ADAPTING TO ACCIDENTS

Bram Bakker   Leon Kester

Intelligent Autonomous Systems group, Informatics Institute
University of Amsterdam, Kruislaan 403, 1098 SJ Amsterdam

Abstract

Last year we started a project concerned with intelligent traffic control. Using a simulator that models urban road traffic, we developed an improved traffic light controller based on measuring traffic congestion on the roads and reinforcement learning. This year an important focus will be on dealing with traffic accidents. In this student project we want to investigate a learning traffic control system that will detect accidents and handle them intelligently. It will learn to direct traffic around accident spots such that traffic delay is minimized.

1 Introduction

Traffic is an integral and important part of modern society. However, with an increasing population and increasing mobility, traffic jams are becoming a more and more common sight, especially in and around large urban areas. This leads to significant economical damage and environmental pollution.

Traffic in urban areas is, in large part, controlled by means of traffic lights. Inefficient configuration of traffic lights can lead to unnecessarily long waiting times for cars and even to traffic jams. Such inefficient configuration is unfortunately still the case in a lot of urban areas; many traffic lights are based on a "fixed cycle" protocol, which basically means that the lights are set to green for a fixed amount of time and consecutively to red for a fixed amount of time. Usually this is not optimal, since such a policy does not take into account the current traffic situation at a particular traffic light, meaning that situations can occur in which a long queue of waiting cars is being held up by a shorter queue (or in the worst case, by no queue at all). What one would want to happen in such a situation is that the policy decides to let longer queues have an "advantage" over the shorter ones in order to improve the total traffic flow; but this should not lead to situation where cars at a quieter part of a junction will never be able to continue their journey.

The question that arises from these considerations is how to determine a policy that optimally takes into account the particular traffic situation at each traffic light junction. "Hand-coding" all possible situations around a traffic junction and determining the traffic light configurations according to those situations is infeasible due to the large number of possible traffic situations and the complexity of the control problem.

A promising approach is to make use of machine learning techniques. Such methods allow the control system to automatically learn a good, or even optimal policy. Techniques that have been used for this problem include genetic algorithms [5] and fuzzy logic [6]. We have used reinforcement learning (RL) to optimize the traffic light controllers in a traffic network. We used a model-based RL algorithm, alternatively viewed as approximate real-time dynamic programming. The methods we proposed optimize an individual traffic light junction in the traffic network locally, but the optimization takes into account traffic congestion at neighboring junctions in the traffic network, such that there is implicit cooperation between the traffic lights [3, 1].
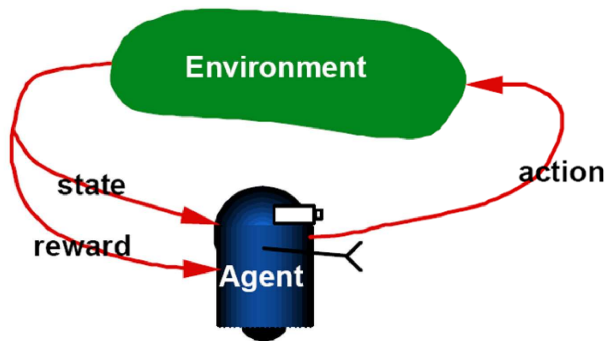
Figure 1: An agent interacting with an environment. The agent executes actions based on perceived states, and adapts it behavior based on rewards.

2 Reinforcement learning and traffic control

2.1 Reinforcement Learning

RL has been applied to various domains with considerable success. One of the most impressive results so far is in backgammon [7], where a program based on this method is able to play at world-class expert level.

RL is a technique for learning control strategies for autonomous agents from trial and error [4, 2]. The agents interact with the environment by trying out actions, and use resulting feedback (reward and the consecutive state) to reinforce behavior that leads to desired outcomes (see figure 1 for an illustration).

2.2 The traffic simulator

In our work we deal with detailed traffic patterns and traffic light control in a (simulated) city environment. We use a freely available traffic simulator, the *Green Light District* (GLD) traffic simulator [8, 9].[1]

The main entities in the simulations are cars driving through a traffic network and traffic light controllers controlling individual traffic lights at traffic light junctions (see figure 2). The cars drive on roads consisting of different driving lanes going in different directions at a junction. Every lane consists of a discrete number of positions, each of which can be occupied by a car. As in the real world, particular driving lanes and corresponding traffic lights constrain the direction in which a car can go. For example, the left lane may be for cars that go straight or turn left, the right lane may be for cars that turn right.

2.3 Traffic light control with RL

The goal in our traffic light control problem is to minimize the total travel time of all cars. The reward function is, therefore, based on waiting time of cars.

In our RL method for traffic light control, traffic light decisions are mainly based on local state information around a single junction. However, traffic light junctions also take into account, as extra information, the amount of traffic at neighboring junctions, a "congestion factor". This means that next to local state information, more global (but not completely global) information is also used for traffic light optimization, and there is a form of implicit cooperation between neighboring traffic light junctions.

This RL method has been shown to outperform a variety of alternative, non-learning, handcoded traffic light controllers as well as alternative learning controllers [3, 1].

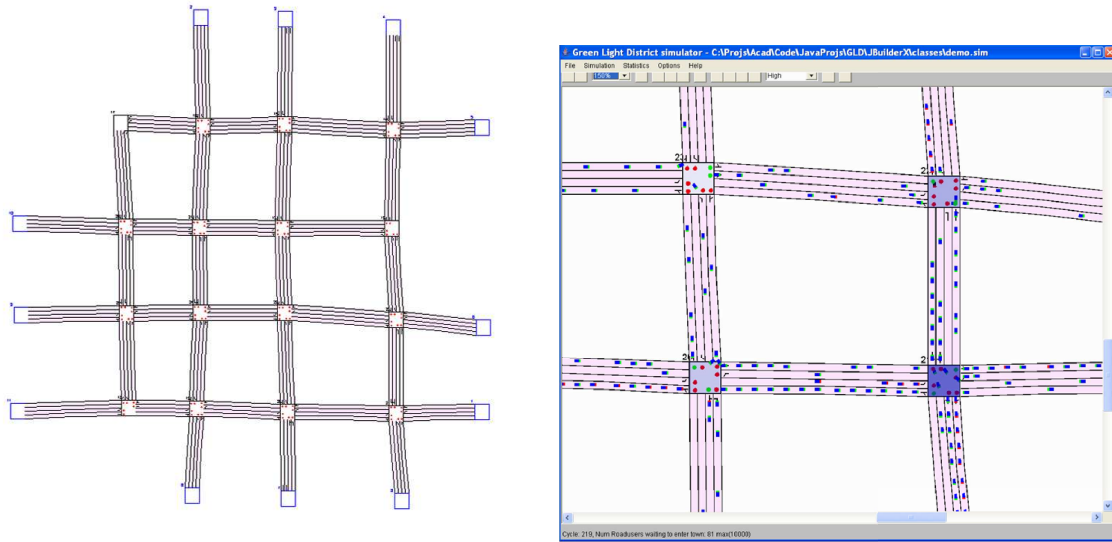[1]Available from http://sourceforge.net/projects/stoplicht

Figure 2: Fig. a (left). A GLD traffic network used in our work, with 12 edge nodes, 15 traffic sign junctions and 1 junction that does not have a traffic light. Fig. b (right). Snapshot of traffic in part of the traffic network. Darker nodes indicate more congested traffic.

3 DOAS 2006 project

3.1 Dealing with accidents

This year an important focus will be on dealing with traffic accidents. In this student project we want extend the RL traffic control system such that it will detect accidents and handle them intelligently. It will learn to direct traffic around accident spots such that traffic delay is minimized.

3.2 Possible mechanisms for dealing with accidents

We already built into the simulator the possibility of accidents happening, effectively shutting down certain traffic lanes temporarily. The idea is that the traffic control system may be extended in two ways.

First, based on the information of accidents and associated lanes becoming unavailable, the traffic light controller will have to learn to set traffic lights such that traffic directly toward the unavailable lane(s) temporarily becomes impossible, and instead give preference to traffic lights for alternative routes around the accident spot.

The second extension is based on the individual cars having route planning systems. The traffic control system, which knows the position of accidents and congestion, will share this information with the cars' route planning systems, such that they can continuously and adaptively replan their routes given accidents and congestion.

3.3 Concrete steps

The students in this DOAS project will perform the following tasks:

- Study basic RL papers and the RL for traffic light control papers
- Familiarize yourself with the GLD traffic simulator and the existing RL algorithms for it
- Experiment with and test the simulation of accidents
- Implement the traffic control system extension(s) for dealing with accidents
- Perform experiments to compare the new algorithms with naive and alternative methods
- Write a report

3.4 Opportunities for students

This project offers students the opportunity to learn about and get hands-on experience with:

- state-of-the-art reinforcement learning algorithms;
- simulation and control of a model of an interesting and important real-world phenomenon: traffic;
- scientific experimentation to compare various algorithms.

Acknowledgments

References

[1] B. Bakker, M. Steingröver, R. Schouten, E. Nijhuis, and L. Kester. Cooperative multi-agent reinforcement learning of traffic lights. In *Proceedings of the Workshop on Cooperative Multi-Agent Learning, European Conference on Machine Learning, ECML'05*, 2005.

[2] L. P. Kaelbling, M. L. Littman, and A. W. Moore. Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, 4:237–285, 1996.

[3] M. Steingröver, R. Schouten, S. Peelen, E. Nijhuis, and B. Bakker. Reinforcement learning of traffic light controllers adapting to traffic congestion. In *Proceedings of the Belgium-Netherlands Artificial Intelligence Conference, BNAIC'05*, 2005.

[4] R. S. Sutton and A. G. Barto. *Reinforcement learning: An introduction*. MIT Press, Cambridge, MA, 1998.

[5] H. Taale, T. Bäck, M. Preuss, A. E. Eiben, J. M. de Graaf, and C. A. Schippers. Optimizing traffic light controllers by means of evolutionary algorithms. In *Proceedings of EUFIT'98*, 1998.

[6] K. K. Tan, M. Khalid, and R. Yusof. Intelligent traffic lights control by fuzzy logic. *Malaysian Journal of Computer Science*, 9-2, 1995.

[7] G. Tesauro. TD-gammon, a self-teaching backgammon program, achieves master-level play. *Neural Computation*, 6(2):215–219, 1994.

[8] M. Wiering, J. van Veenen, J. Vreeken, and A. Koopman. Intelligent traffic light control. Technical report, Dept. of Information and Computing Sciences, Universiteit Utrecht, 2004.

[9] M. Wiering, J. Vreeken, J. van Veenen, and A. Koopman. Simulation and optimization of traffic in a city. In *IEEE Intelligent Vehicles symposium (IV'04)*, 2004.