A Junction Tree based Reinforcement Learning Algorithm for Coordinated Multi Agent Systems to Solve Network level Signal Control Problems

Feng Zhu (zhu214@purdue.edu), H. M. Abdul Aziz (aziz@purdue.edu), Xinwu Qian (xinwu.qian.1@purdue.edu); Purdue University
Satish V. Ukkusuri, Associate Professor, Purdue University (sukkusuri@purdue.edu)

Introduction

Optimized traffic control systems:
- Travel time reduction
- Saving in fuel consumption
- Cut down emissions

Traditional control:
SCOOT (Hunt et al., 1982), SCATS (Lowrie, 1982), PRODYN (Fargas et al., 1983), OPAC (Gartner, 1983), RHODES (Mirchandani and Head, 2001), UTOPIA (Mauro and Taranto, 1989)

Limitation:
- Isolated learning, no network wide coordination
- No consideration of dynamic feedback

Research objectives
1) To propose multi-agent RL based signal control algorithm where agents coordinate their decisions
2) Compare the coordinated learning and isolated learning algorithms
3) To demonstrate the coordinated control algorithmas a potential application in the CVenvironment
4) To assess the environmental impacts of the proposed controller using a dynamic emissions simulator (MOVES2010)

Problem description
- Major assumption: The intersections coordinate with each other to benefit the entire system.
- Signal control: To find the optimal policy (mapping between the phase activations and traffic states)
- Develop a coordinated Reinforcement Learning algorithm: Traffic lights in the network act as agents and take actions (activating the phase) according to RL algorithm.

Methodology

- Modeling Framework
  - VISSIM: 1. provides state information (e.g. traffic charc.) to C++
  - C++: 2. receives states information (e.g. traffic charc.)
  - C++: 3. implement RL algorithm to provide policy to VISSIM
  - VISSIM: 4. receives policy from C++ to take actions

Junction Tree Algorithm

Formulation:
\[ \max_{\pi} \sum_{t=0}^{\infty} \gamma^t Q(s_t, a_t) \]

Five steps:

a. Moralization:
   - converts a directed graph into an undirected graph
b. Introduction of potential:
   - a new term to describe the characterization of the cluster
c. Triangulation:
   - ensure that every loop of length 4 or more has a chord
d. Construction of junction tree:
   - find out the maximal spanning tree
e. Propagation of messages:
   - forward and backward message passing

Test case study

Experiment design

Demand scenarios:
- Low (500vph~600vph)
- Medium (600vph~900vph)
- High (900vph~1200vph)

Comparison of algorithms:
- JTA: junction tree algorithm
- Max-plus: coordinated RL algorithm (Kuyer et al., 2008)
- Q-learning: independent RL algorithm (Abdulhai et al., 2003)
- LQF: longest queue first algorithm (Wunderlich et al., 2008)
- Fixed-control: fixed time signal control

Result analysis

Conclusions

- Development of a coordinated RL based algorithm for signal control
- Junction tree algorithm to obtain best joint actions
- Explore the environmental benefits of the algorithm
- Test results show significant advantages of coordinated learning over independent learning of agents.