Smart Traffic Lights that Learn!

Multi-Agent Reinforcement Learning Integrated Network of Adaptive Traffic Signal Controllers

MARLIN

Samah El-Tantawy, Ph.D.
Baher Abdulhai, Ph.D., P.Eng.
Hossam Abdelgawad, Ph.D., P.Eng.

Post Doctoral Fellow, Dept of Civil Engineering
Director, ITS Centre and Testbed, Dept of Civil Engineering
Manager of ITS Centre and Testbed

ACGM 2013 - Intelligent Transport for Smart Cities
1. In a Nutshell

2. Theory in Brief
   - Reinforcement Learning and Game Theory

3. Applications
   - City of Toronto Testbed

4. Hardware in the Loop Testing
   - Approach
   - Integration with PEEK ATC-1000

☐ Next Steps
☐ Q&A
In a Nutshell

- **Grand objective**
  - Intersections "talk to each other",
  - Each is affected by what is happening upstream
  - Each affects what is happening downstream –
  - Whole network control in one shot from a grand brain is the dream

- **Issue**
  - Intractable theoretically,
  - Too complex practically,
  - Requires massive and very expensive communication

- **Solution**
  - Decentralized,
  - Self learning: *agents learn to control* their local intersection, and
  - Game theory based: *agents learn to collaborate*
What is MARLIN?

- Artificial-intelligence-based control software
- Enables traffic lights to self-learn and self-collaborate with neighbouring traffic lights
- Cuts down motorists’ delay, fuel consumption and the negative environmental effects of congestion
- Easier to operate (self learning)
- Less expensive communication if even necessary (less costly)
Evolution of “Adaptive” Signal Control

MARLIN-ATSC: Level 4

Level 0
- Fixed-Time and Actuated Control
- SCATS
- 1969, UK
- >50 installations worldwide

Level 1
- Centralized Control, Off-line Optimization
- SCATS
- 1979, Australia
- >50 installations worldwide

Level 2
- Centralized Control, On-line Optimization
- SCOOT
- 1981, UK
- >170 installations worldwide

Level 3
- Distributed Control, Model-Based
  - OPAC, RHODES
  - 1992, USA
  - 5 installations in USA

Level 4
- Distributed Self-Learning Control
  - MARLIN-ATSC
  - 2011, Canada
## Issues with Leading ATSC Technologies?

<table>
<thead>
<tr>
<th>Centralized</th>
<th>Model-Based</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expensive</td>
<td>Relying on an accurate traffic modelling framework</td>
</tr>
<tr>
<td>Not scalable</td>
<td>the accuracy of which is questionable</td>
</tr>
<tr>
<td>Not robust</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Curse of Dimensionality</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Increasing the complexity of the system exponentially with the increase in the number of intersections/controllers</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Human Intervention Requirements</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Requiring highly skilled labour to operate due to their complexity.</td>
</tr>
</tbody>
</table>
Why is MARLIN Different?

- Self-Learning
- Decentralized
- Model-Free
- Scalable
- Coordinated
- Pattern Sensitive
- Generic
- Specific Design
- Human Intervention Requirements
- Centralized
- Model-Based
- Prediction Requirement
- Inefficient Coordination
- Curse of Dimensionality
- Centralized
- Inefficient Coordination
- Prediction Requirement
- Generic
Learning the Control Law: Reinforcement Learning Architecture

**Goal:** Optimal Control law = mapping between states and actions

\[ Q^{k+1}(s^k, a^k) = Q^k(s^k, a^k) + \alpha[r^{k+1} + \gamma \max_a Q^k(s^{k+1}, a) - Q^k(s^k, a^k)] \]

\[ a^{k+1} = \arg \max_a Q^k(s^{k+1}, a) \]  
Balancing exploration and exploitation

### Q Table

<table>
<thead>
<tr>
<th>Q</th>
<th>a₁</th>
<th>a₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>s₁</td>
<td>-10</td>
<td>-5</td>
</tr>
<tr>
<td>s₂</td>
<td>-3</td>
<td>-15</td>
</tr>
</tbody>
</table>
RL-based ATSC Architecture

Traffic Simulation Environment

State (Queue Lengths)

Reward (Delay Savings)

Action (Extend/Switch)

RL Software Agent
Each agent plays a game with each adjacent intersection in its neighborhood.
MARLIN-ATSC Available Modes

- MARLIN-ATSC: (a) Independent Mode, (b) Integrated Mode
## Large-Scale Application
### Network-Wide MOE in the Normal Scenario

<table>
<thead>
<tr>
<th>System MOE</th>
<th>BC</th>
<th>% Improvements MARL-TI Vs. BC</th>
<th>% Improvements MARLIN-IC Vs. BC</th>
<th>% Improvements MARLIN-IC Vs. MARL-TI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Intersection Delay (sec/veh)</td>
<td>35.27</td>
<td>27%</td>
<td>38%</td>
<td>14%</td>
</tr>
<tr>
<td>Throughput (veh)</td>
<td>23084</td>
<td>3%</td>
<td>6%</td>
<td>3%</td>
</tr>
<tr>
<td>Avg Queue Length (veh)</td>
<td>8.66</td>
<td>24%</td>
<td>32%</td>
<td>11%</td>
</tr>
<tr>
<td>Std. Avg. Queue Length (veh)</td>
<td>2.12</td>
<td>23%</td>
<td>31%</td>
<td>10%</td>
</tr>
</tbody>
</table>
Large-Scale Application

% Improvement in Average Delay

MARLIN-IC vs BC

Area 1

Area 2

Area 3

% Improvement

- 0 - 10
- 10 - 20
- 20 - 40
- 40 - 60
- 60 - 100
Large-Scale Application
Average Route Travel Time for Selected Routes

Gardiner EB

Average Travel Time (min)
Time Interval (5 min)
Gardiner EB
BC MARL-TI MARLIN-IC

Freeway

Limited downstream intersection capacity
Alternative Route
Large-Scale Application
Average Route Travel Time for Selected Routes

![Graph showing average travel time for LakeShore EB to Spadina NB]

- **BC**: Blue line
- **MARL-TI**: Red line
- **MARLIN-IC**: Green line

**Time Interval (5 min)**

- **Average Travel Time (min)**

**Major Arterial**

- Spadina St.
- Blue Jays St.
Controller Interface Device (CID)  
RS485 to USB

Paramics Modeller

Traffic Signal Controller

Industrial Computer

RS485 - SDLC protocol

USB - SDLC protocol

Ethernet - NTCIP protocol

MARLIN-HILS Architecture
HILS Setup: Demo
Conclusion

- MARLIN state of the art gen4+
- Thoroughly developed and tested
- Patent Pending Status
- On going:
  - HILS & PEEK ATC-1000 Integration
  - Potential Field Operation Test
  - Productization
- From TSP to People Priority (PSP)
Samah El-Tantawy
samah.el.tantawy@utoronto.ca

Baher Abdulhai
baher.abdulhai@utoronto.ca

Hossam Abdelgawad
h.abdel.gawad@utoronto.ca

ACGM 2013 - Intelligent Transport for Smart Cities
Smart Traffic Lights that Learn!

Multi-Agent Reinforcement Learning Integrated Network of Adaptive Traffic Signal Controllers

M A R L I N

Samah ElTantawy, Ph.D. 
Baheer Abdulhai, Ph.D., P.Eng. 
Hossam Abdelgawad, Ph.D., P.Eng.

Post Doctoral Fellow, Dept of Civil Engineering
Director, ITS Centre and Testbed, Dept of Civil Engineering
Manager of ITS Centre and Testbed

ACGM 2013 - Intelligent Transport for Smart Cities