The Intersection of Reinforcement Learning and Traffic Light Scheduling

Drew Kristensen
University of Puget Sound

April 28, 2018

What is the issue

We can all relate to being at a traffic light, with no one else around, and having to wait for a red to turn green. Wouldn't it be great if we could fix this?

What is the issue

We can all relate to being at a traffic light, with no one else around, and having to wait for a red to turn green. Wouldn't it be great if we could fix this?

More Seriously:
Traffic congestion is becoming more and more of a problem for both humans and the environment.
In 2013:

- 65 hours
- 3.1 megatons

My project seeks to solve this problem of traffic congestion by controlling the traffic lights.

Definitions

Traffic control agents (TCA)	The agent that decides which signal to show for a traffic light

Definitions

Traffic control agents (TCA)	The agent that decides which signal to show for a traffic light
State Space	The variables and characteristics we can observe to use as input

Definitions

Traffic control agents (TCA)	The agent that decides which signal to show for a traffic light
State Space	The variables and characteristics we can observe to use as input
Action Space	The variables that we will control with our output

Definitions

Traffic control agents (TCA)	The agent that decides which signal to show for a traffic light
State Space	The variables and characteristics we can observe to use as input
Action Space	The variables that we will control with our output
Reward	Some value dependent on the state that tells us how we are doing

Definitions

Traffic control agents (TCA)	The agent that decides which signal to show for a traffic light
State Space	The variables and characteristics we can observe to use as input
Action Space	The variables that we will control with our output
Reward	Some value dependent on the state that tells us how we are doing
Neural Networks (NN)	Function approximators made from collections of "neurons"

Definitions

Traffic control agents (TCA)	The agent that decides which signal to show for a traffic light
State Space	The variables and characteristics we can observe to use as input
Action Space	The variables that we will control with our output
Reward	Some value dependent on the state that tells us how we are doing
Neural Networks (NN)	Function approximators made from collections of "neurons"
Reinforcement Learning (RL)	An algorithm which uses positive and negative rewards to "teach" a NN to produce the behavior we want

Q Learning

Difficult to know what action to pick at each time.
Train a neural network to approximate the reward of choosing a particular action at each time

Q Learning

Difficult to know what action to pick at each time.
Train a neural network to approximate the reward of choosing a particular action at each time

$$Q(s_t, a_t, \pi) = \mathbb{E}\left\{\sum_{k=0}^{\infty} \gamma^k r_{t+k} \middle| s_t, a_t, \pi\right\}$$

For some policy π.
To choose an action, we take

$$a_t = \arg\max_a Q(s_t, a, \pi)$$

.

Q Learning

Difficult to know what action to pick at each time.
Train a neural network to approximate the reward of choosing a particular action at each time

$$Q(s_t, a_t, \theta) = \mathbb{E}\left\{ \sum_{k=0}^{\infty} \gamma^k r_{t+k} \middle| s_t, a_t, \theta \right\}$$

where the weights to our NN are θ.
To choose an action, we take

$$a_t = \arg\max_a Q(s_t, a, \theta)$$

.

Q Learning: Simply put

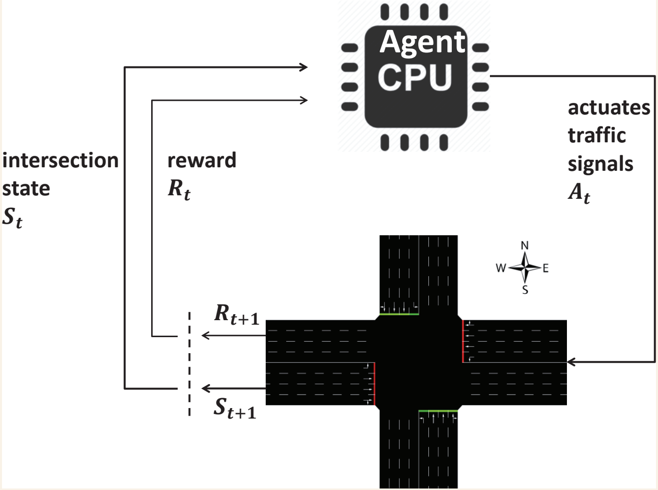


Figure: Gao et al.

Current Approaches

Current methods that cities implement to limit congestion include

- Fixed timing
- Adaptive control
- Coordinated control

Approach

Action Space:

12 values

- 3 possible pairs of inbound edges
- 4 total non-red light settings for each pair

Approach

Action Space:
12 values

- 3 possible pairs of inbound edges
- 4 total non-red light settings for each pair

State Space:
72 values

- 6 possible incoming roads
- 5 lane types
- 2 features per lane type
- "One hot" vector holding index of current light signal

Approach

Action Space:
12 values

- 3 possible pairs of inbound edges
- 4 total non-red light settings for each pair

State Space:
72 values

- 6 possible incoming roads
- 5 lane types
- 2 features per lane type
- "One hot" vector holding index of current light signal

Reward Function:
1 value
Sum the squares of all lane waiting times
Take the difference across consecutive time steps

Differences from Previous Research: State Space

Previous Research

- Omnipotent state space
- Positions of cars in lanes
- Speeds of cars in lanes
- Fixed sizes

Current Method

- Still somewhat Big Brother
- Number of cars in each lane
- Number of cars under speed threshold
- Can be extended to any number/type of lanes

State Space

The variables and characteristics we can observe to use as input

Differences from Previous Research: Action Space

Previous Research

- Fixed timing
- Fixed path (ie yellow follows green)
- Only works for 4 way intersections

Current Method

- No fixed timing
- No fixed path
- Can be generalized for any intersection*

Action Space

The variables that we will control with our output

Differences from Previous Research: Reward

Previous Research

- Raw difference in wait time
- Linear wait

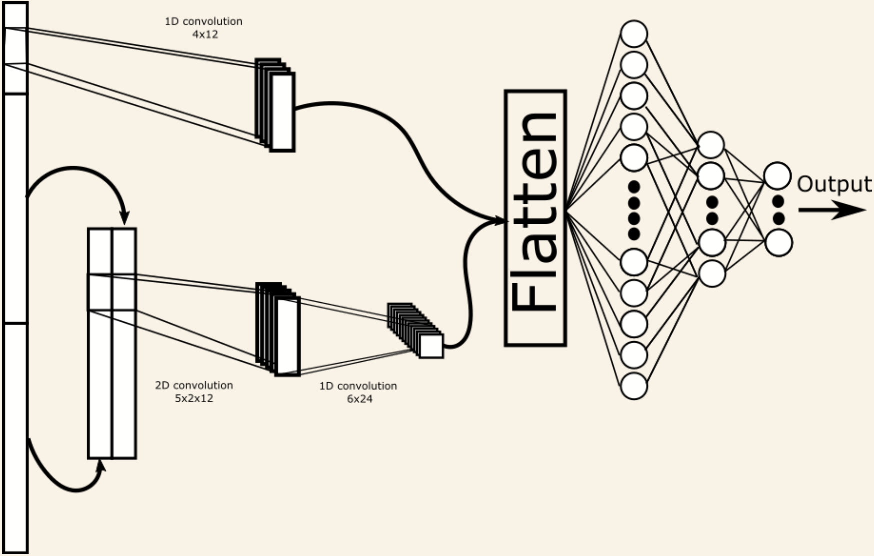
Current Method

- Uses proportion of change in wait
- Squares wait times per vehicle

Reward

Some value dependent on the state that tells us how we are doing

Model Architecture



Demo

Results

4 way, single lane intersection		
Traffic Control Agent	Throughput	Avg Delay
Fixed timing	141.0	19
Q-learning	115.00	119.8

Results

4 way, single lane intersection

Traffic Control Agent	Throughput	Avg Delay
Fixed timing	141.0	19
Q-learning	115.00	119.8

4 way, four lane intersection

Traffic Control Agent	Throughput	Avg Delay
Fixed timing	0	0
Q-learning	0	0

Results

4 way, single lane intersection

Traffic Control Agent	Throughput	Avg Delay
Fixed timing	141.0	19
Q-learning	115.00	119.8

4 way, four lane intersection

Traffic Control Agent	Throughput	Avg Delay
Fixed timing	0	0
Q-learning	0	0

6 way, two lane intersection

Traffic Control Agent	Throughput	Avg Delay
Fixed timing	0	0
Q-learning	0	0

Roadblocks

Obviously did not get as far as I had hoped
What were some of the problems?

Roadblocks

Obviously did not get as far as I had hoped
What were some of the problems?

Understanding theory $\neq$ Easy implementation

Sometimes code doesn't work how you would expect

Conclusion

Lots to explore now

Compare to existing RL algorithms

Evaluate for 3, 5, and 6 way intersections