Adaptive Traffic Signal Control based on Reinforcement Learning

Mengyu Guo, Pin Wang, Ching-Yao Chan

Acknowledging generous support of HUAWEI

BDD Spring Retreat
March 27, 2019

Berkeley DeepDrive

Project Overview

- **Objective:**
  - **To develop a Traffic Signal Control System for demonstrating the benefits of advanced machine learning techniques in optimizing traffic flow in road networks.**
- **Focus of Current Phase:**
  - **Implementation of simulation platform for case studies**
  - **Development of RL algorithms for traffic signal control for single and multiple intersections**

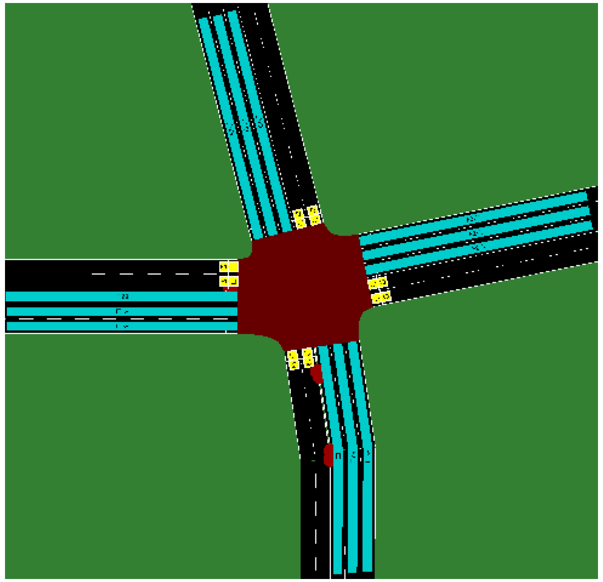
Building Road Network

- **Road Network: Construct a Real-world Intersection in SUMO**

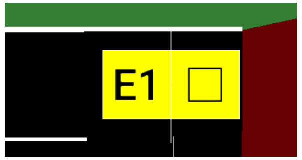
Test site description

- Total 12 approaching lanes
- 3 approaching lanes of each directions
- 1 left-turn lane, 2 through lanes
- Right-most through lane can be used for turning right

SUMO platform



Inductive loop detector

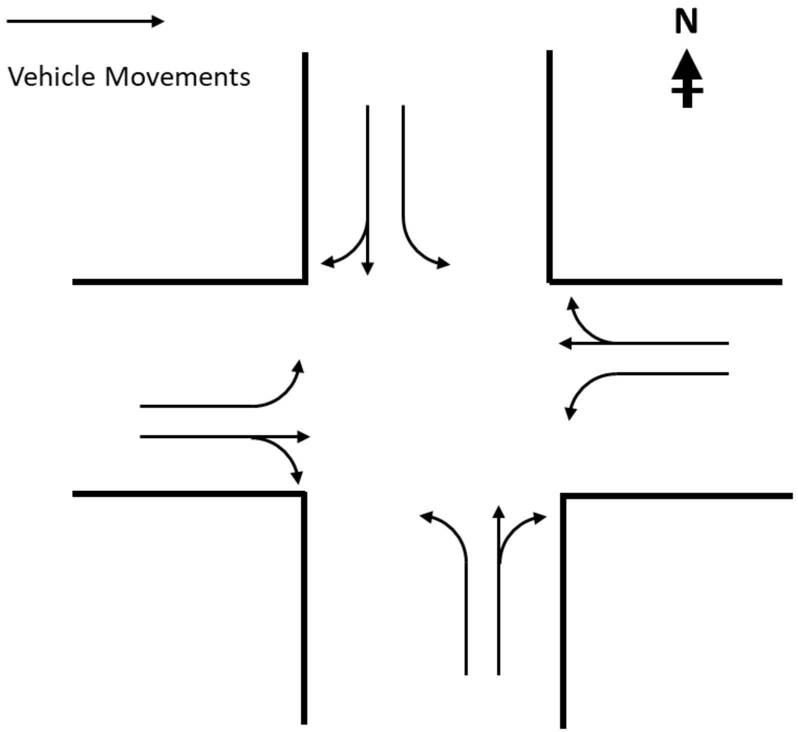


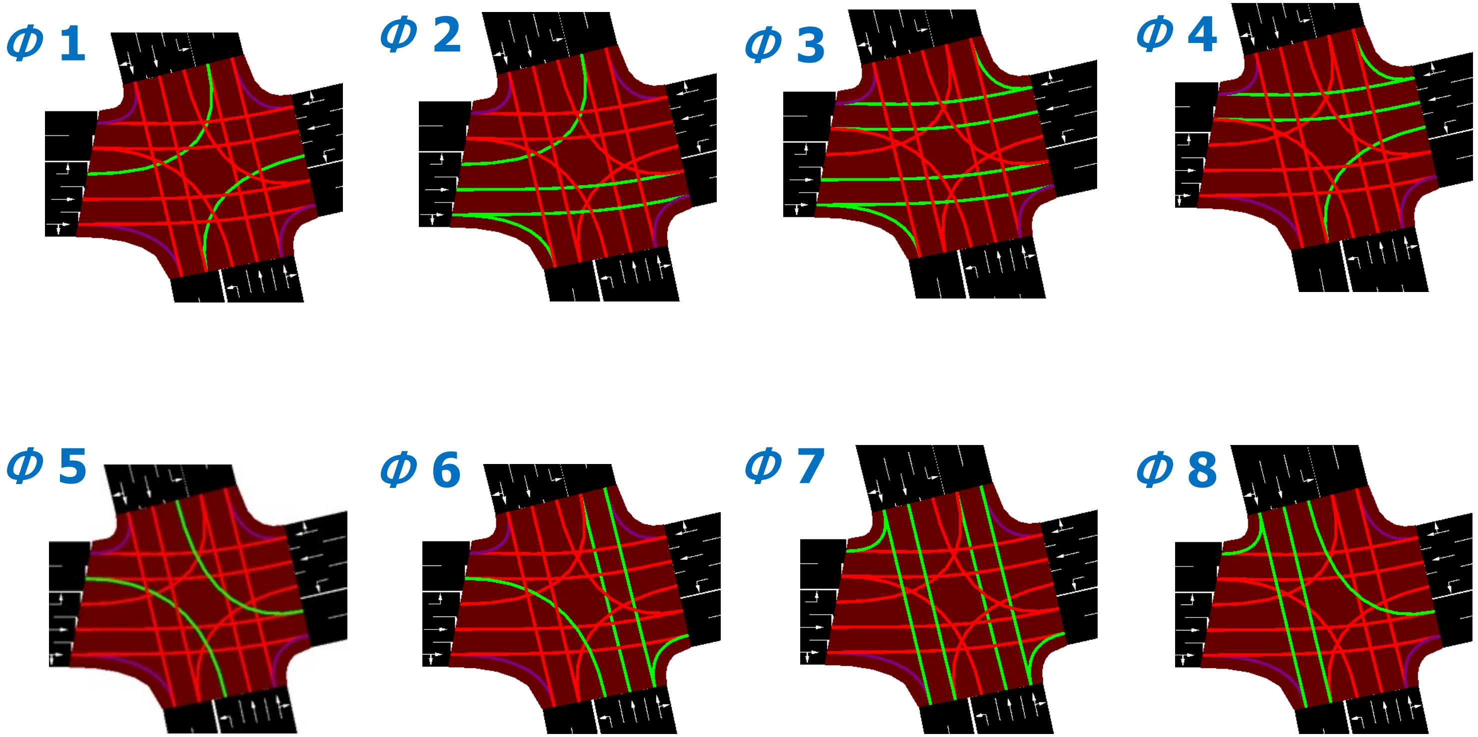
Lane area detector

Berkeley DeepDrive

Traffic Signal Plan

- **Signal Phases**
  - **Movements in one phase can operate concurrently**





Berkeley DeepDrive

Traffic Signal Control Scheme

- Fixed vs. Flexible Phase Sequence
  - Fixed sequence
    - Vastly exist in practice
    - More interpretable
    - Pre-programmed
  - Flexible sequence (fully-actuated)
    - More efficient in dealing with real-time traffic
    - More robust in accidental scenarios
    - More like policeman directing traffic at intersection

Reinforcement Learning Algorithm - State

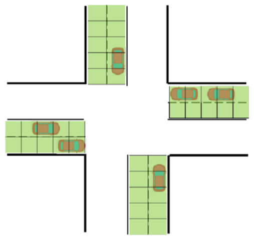
- **State**
  - **In current phase simulated traffic is assumed to be data source**
  - **In future phase we will go deeper with proper real-world dataset**

Sensor Inputs

Traffic video

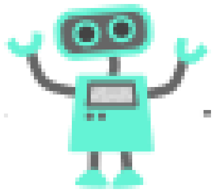
Simulated environment

Individual vehicle location & speed



1	0	3	…	0	6	0	2	0

Halting vehicle number of 12 approaching lanes

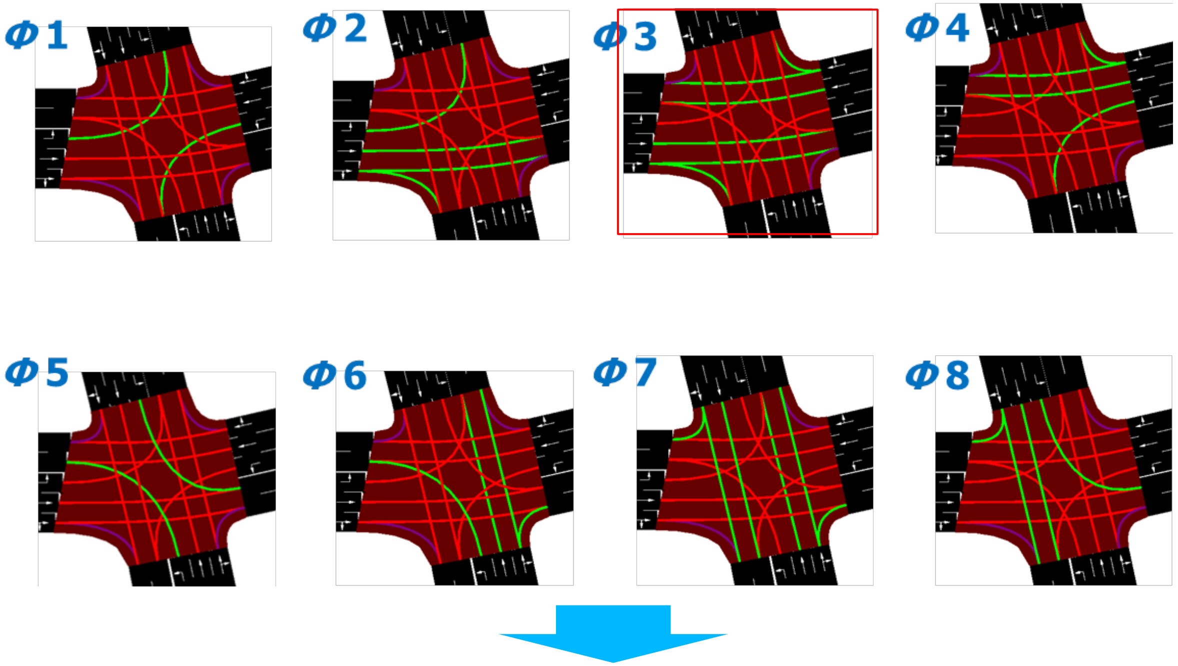


RL agent

Berkeley DeepDrive

Reinforcement Learning Algorithm - Action

- **Action**
  - **Two strategies: fixed phase pair sequence VS. flexible phase pair sequence**
  - **Flexible sequence strategy:**



Action:
- 8-dimensional vector
- Give phase X green light for Δ

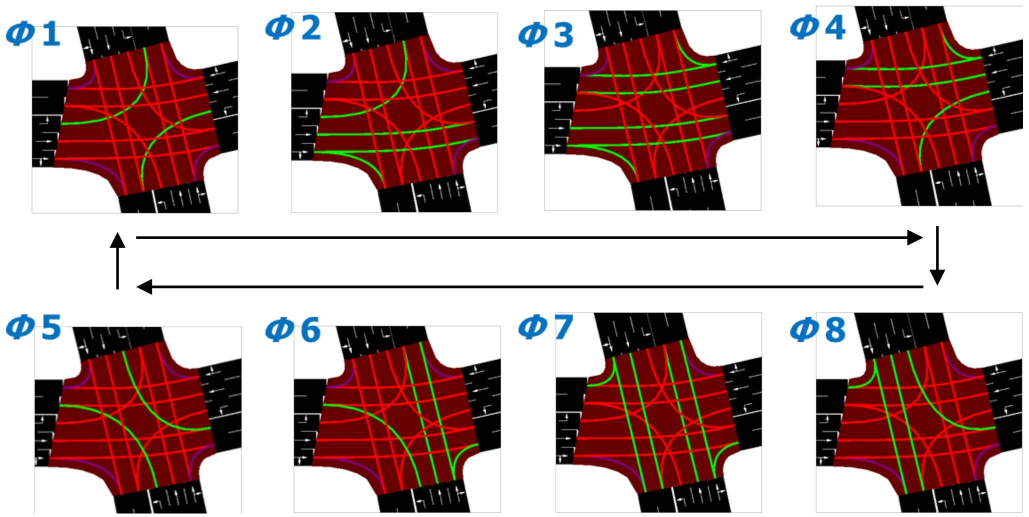
1	2	3	4	5	6	7	8

Berkeley DeepDrive

Reinforcement Learning Algorithm - Action

- **Action**
  - **Two strategies: fixed phase pair sequence VS. flexible phase pair sequence**
  - **Fixed sequence strategy:**

The phase sequence:



Parameters:

- T_{min} - minimum green time
- Δ - additive time span

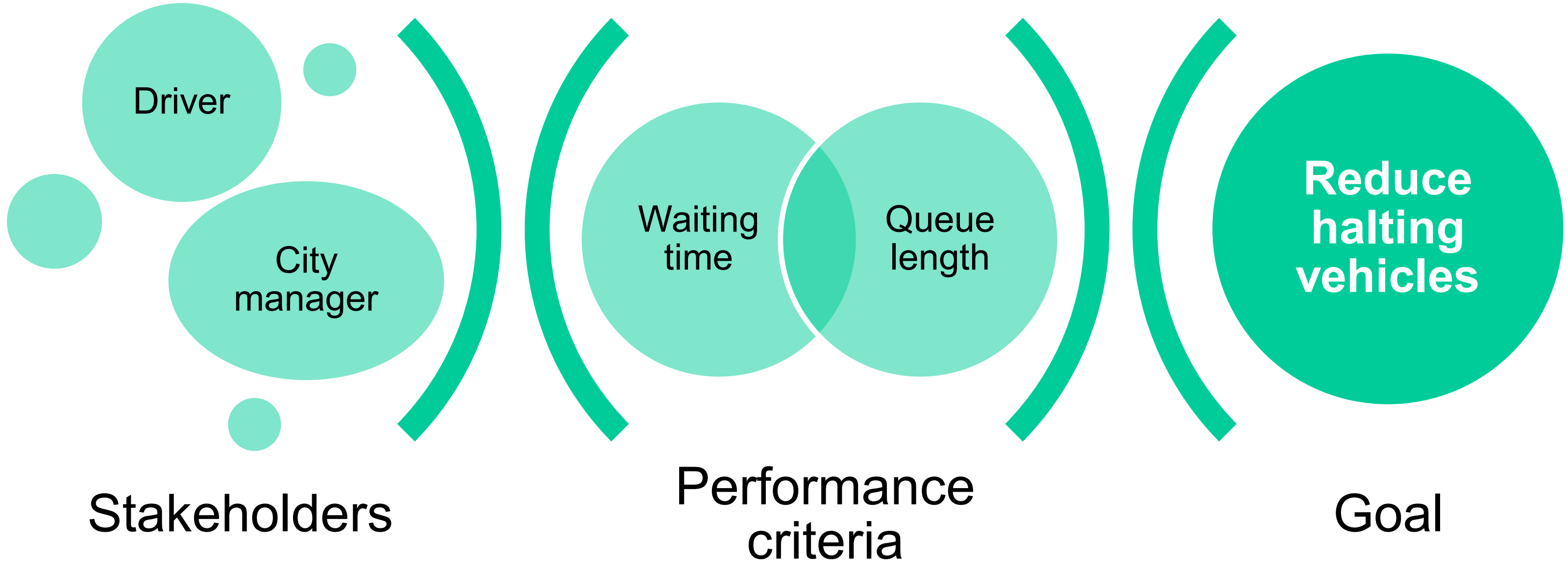
Action:

- {0, 1}
- 0 - switch to next phase pair then run T_{min}
- 1 - extend current green phase pair by Δ then switch to next phase pair

Berkeley DeepDrive

Reinforcement Learning Algorithm - Reward

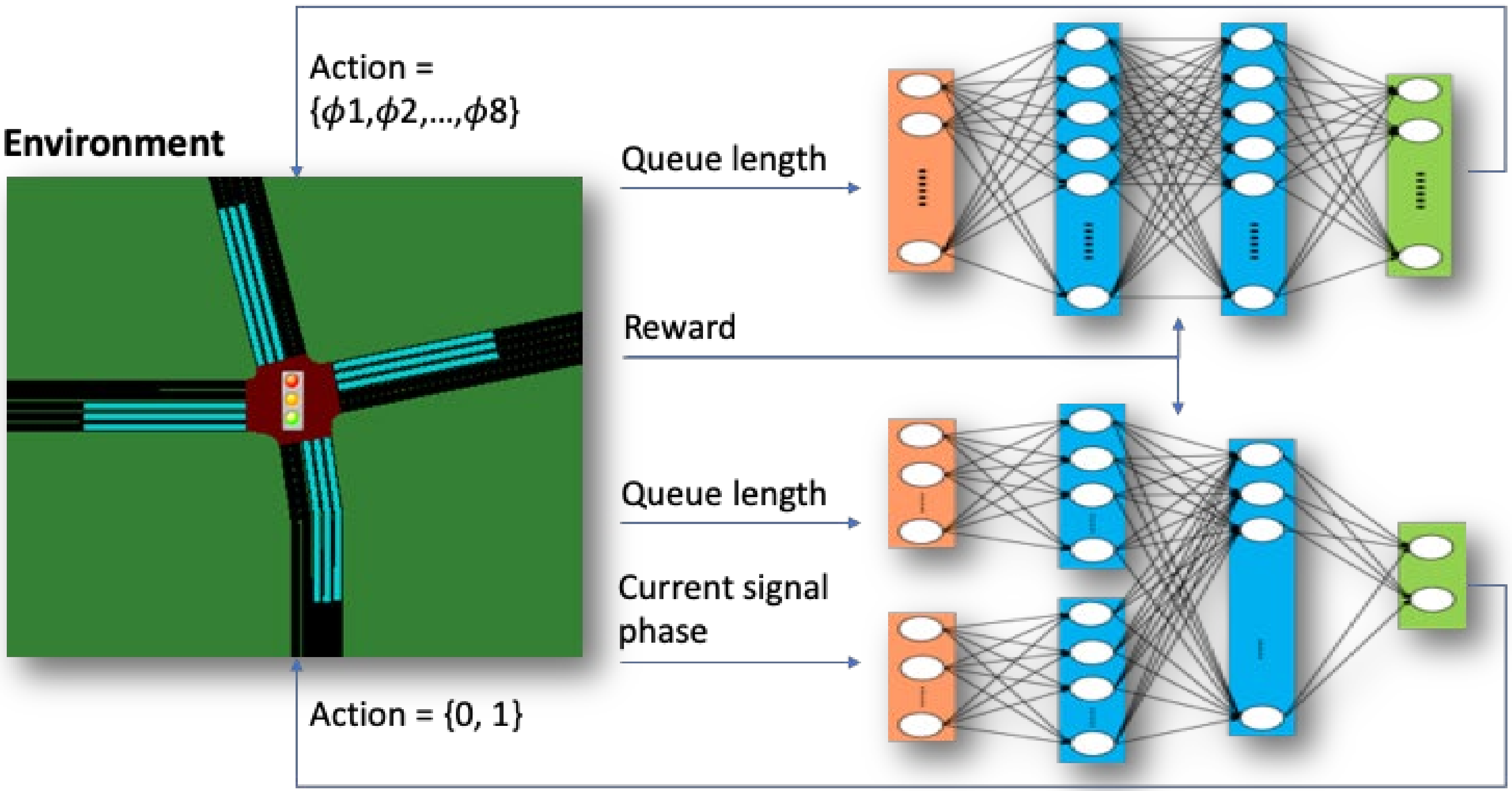
- **Reward**
  - **Reward = percentage of halting vehicles reduced by an action (i.e. released during passed green phase pair)**



Halting vehicle: speed < 0.1 m/s

Reinforcement Learning Algorithm - Network

- **Single Agent RL Architecture**



RL Agent (flexible sequence of signal phases)

- Input layer (1×12)
- First hidden layer (1×64)
- Second hidden layer (1×64)
- Output layer (1×8)

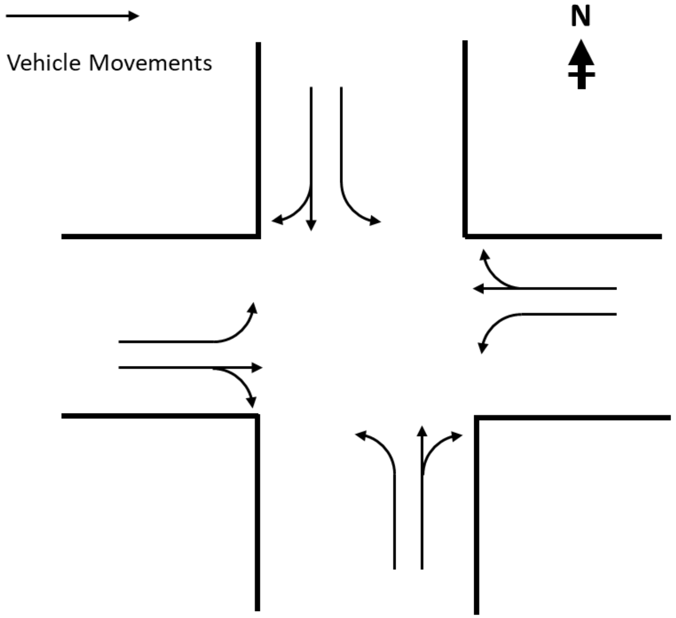
RL Agent (fixed sequence of signal phases)

- Input layers (1×12 & 1×8)
- First hidden layers (1×64 & 1×64)
- Second hidden layer (1×128)
- Output layer (1×2)

Berkeley DeepDrive

Case Study – Different Traffic Scenarios

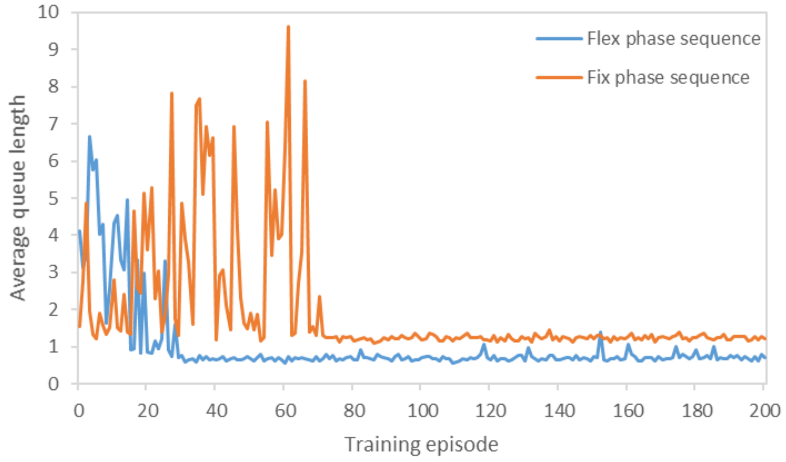
- **Traffic Scenarios**
  - **Scenario 1 – major-minor roads**
    - west-east higher than north-south
  - **Scenario 2 – rush hour**
    - north and east higher than the other two directions
  - **Scenario 3 – heavy left turn**
    - left-turn in west-east directions higher than through lanes of the same directions
  - **Scenario 4 – varying traffic flow**
    - significant variations alternating in west-east and north-south directions



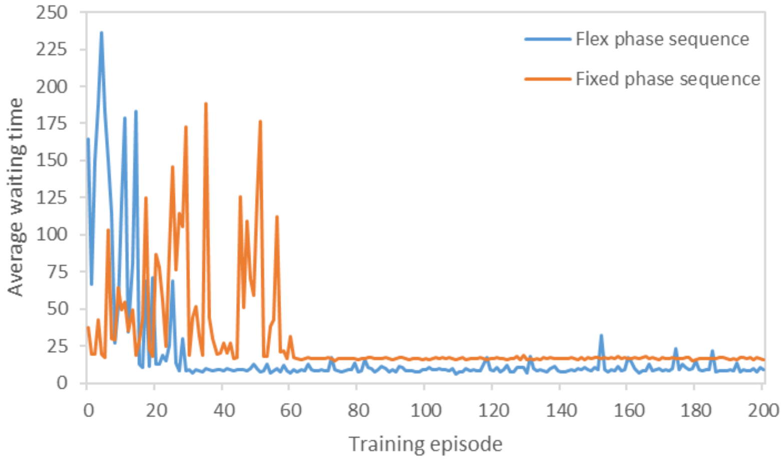
Berkeley DeepDrive

Reinforcement Learning Algorithm - Convergence

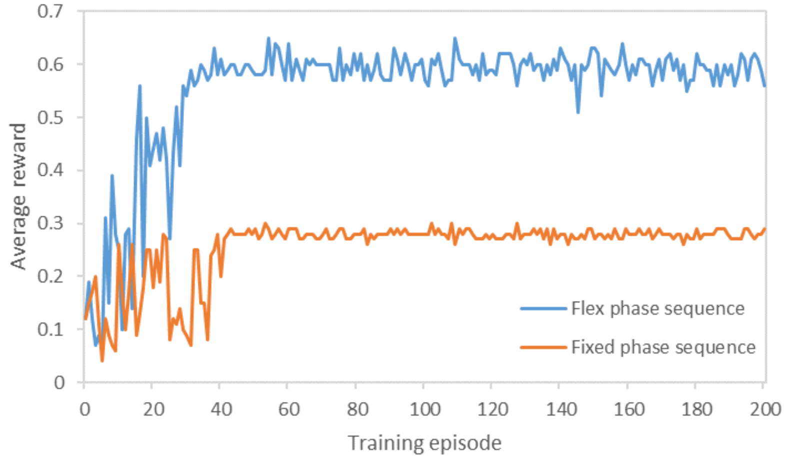
- **Performance of RL Algorithm**
  - **Training: improvement and convergence (under scenario 1)**



Queue Length
(vehicle per approaching lane)



Waiting Time
(second per approaching vehicle)



Average Reward

Berkeley DeepDrive

Reinforcement Learning Algorithm - Benchmarking

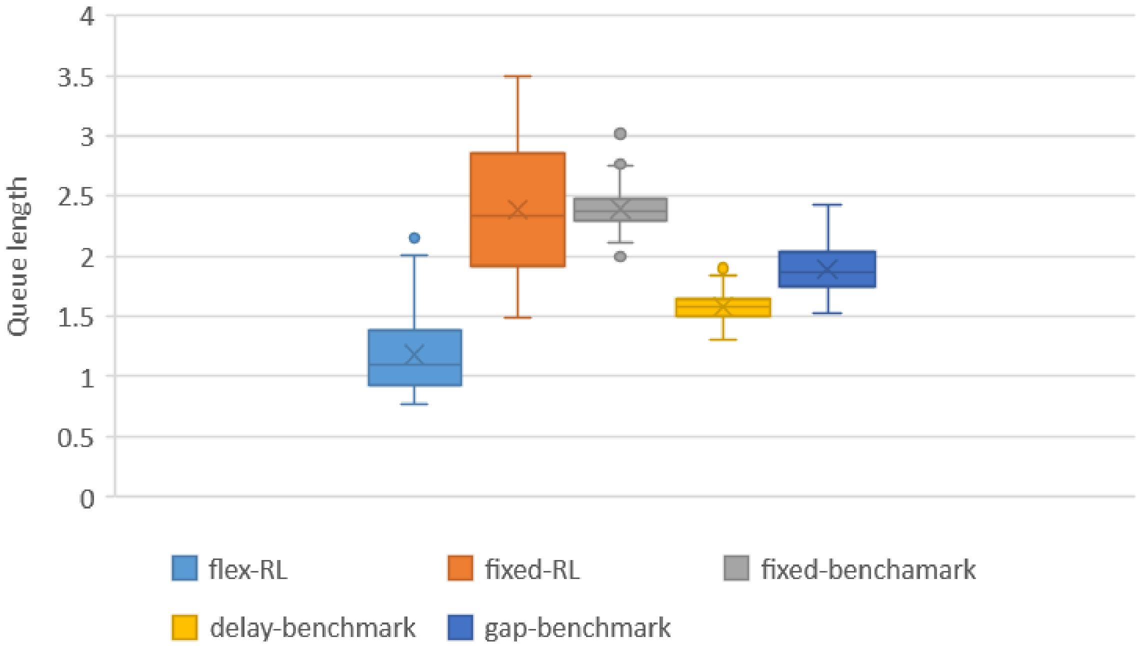
- **Benchmarks:**
  - **Fixed time control:** phase duration is fixed during operation
  - **Gap-based adaptive control:** prolong traffic phases whenever a continuous (i.e. maximum time gap between successive vehicle < 5) stream of traffic is detected
  - **Time loss based adaptive control:** prolong traffic phases whenever there exists vehicle with accumulated time loss (i.e. $1 - v/v_{max}$) exceeds 1s

https://sumo.dlr.de/wiki/Simulation/Traffic_Lights#Improving_Generated_programs_with_knowledge_about_traffic_demand

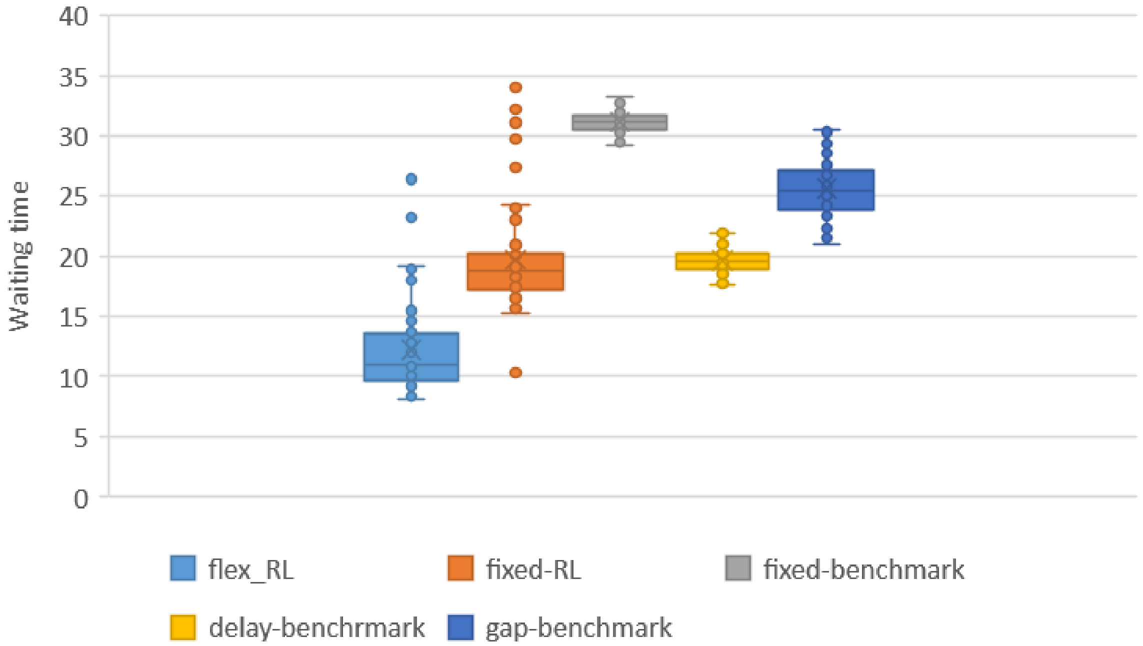
Berkeley DeepDrive

Reinforcement Learning Algorithm - Benchmarking

- **Comparison with Benchmarks (under Scenario 2)**



Queue length



Waiting time

Berkeley DeepDrive

Reinforcement Learning Algorithm - Generalization

- **Generalization across scenarios**

		Scenario 1 (Trained)	Scenario 2 (Test)	Scenario 3 (Test)	Scenario 4 (Test)
Flex-RL	Queue length	0.72 (0.11)	1.11 (0.26)	1.28 (0.32)	0.55 (0.28)
	Waiting time	9.84 (3.37)	12.62 (8.57)	12.73 (5.67)	10.95 (4.14)
Fixed-RL	Queue length	1.24 (0.06)	1.67 (0.32)	2.75 (0.60)	1.03 (0.42)
	Waiting time	16.66 (0.58)	17.09 (1.17)	22.29 (4.18)	10.81 (6.52)

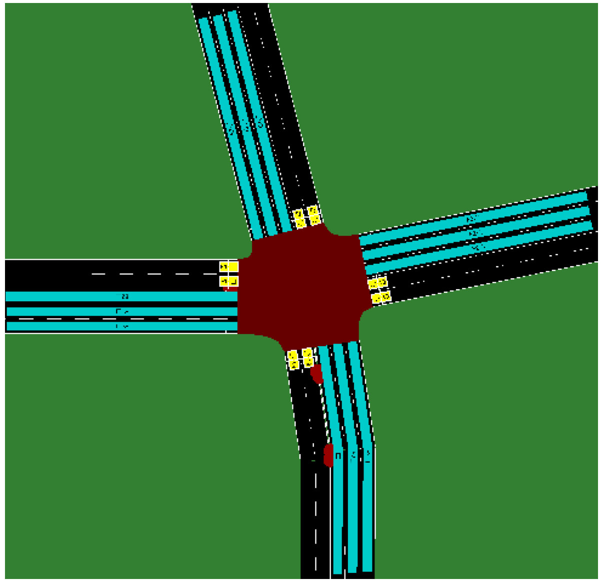
Statistics were obtained after 100 runs

Berkeley DeepDrive

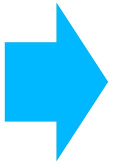
Future Work

- **Multi-Agent RL for Multi-Intersection Traffic Signal Control**

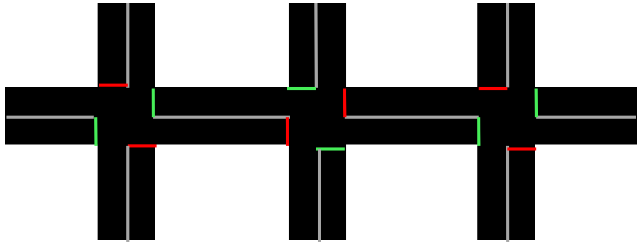
Single intersection



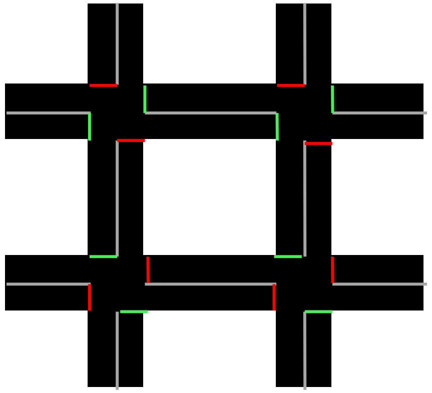
Topology expansion



Multiple intersections



Arterial street



Downtown network

Berkeley DeepDrive

Thank You!

Questions?

Berkeley DeepDrive