Title: Traffic Signal Control Optimization with SUMO

Group 24 Members:

Victor Wang 101265938 victorwang5@cmail.carleton.ca
Gator Guo 101267370 gatorguo@cmail.carleton.ca
Fatih Ozer 101240725 fatihozer@cmail.carleton.ca
Lewis He 101259286 lewishe@cmail.carleton.ca
Julien Larivière-Chartier 101173643 julienlarivierechartier@cunet.carleton.ca

Problem Statement: Traffic Signal Control (TSC) to optimize traffic flow for a single 4-way intersection with pedestrian crossings. Experiment with different reinforcement learning algorithms and compare results.

Feasibility: One or many custom traffic scenario(s) can be modelled as MDPs where the rewards are obtained by letting objects (cars and pedestrians) cross in the least amount of time. The actions would be setting a number of seconds for each direction's crossing. The traffic controller agent would learn to optimize adjusting crossing lights' timing such that each object has crossed in the least amount of time possible. The reward structure could be a small reward for each quick crossing (i.e. dense rewards) and a larger reward when the scenario is completed (i.e., all objects have crossed). The size of the end reward would be proportional to how quickly the scenario is completed.

We will be basing our implementation on an existing framework for traffic and intersection modelling called SUMO (https://github.com/smtg-bham/sumo) and we will use an already existing gymnasium-compatible environment for this: SUMO-RL (https://github.com/LucasAlegre/sumo-rl).

Papers:

A Comparative Study of Algorithms for Intelligent Traffic Signal Control [1]
- Single intersection
- Different scenarios
- Compared a few algorithms such as Deep Q-Network (DQN) and Advantage Actor-Critic (A2C)
- Methods were tested on a simulation of a real-world intersection in Bengaluru, India

From Local to Global: A Curriculum Learning Approach for Reinforcement Learning-based Traffic Signal Control [2]
- Discusses the approach of applying a curriculum to RL to solve TSC.
- A curriculum means designing a training sequence of tasks or environments that gradually increase in difficulty, instead of exposing the agent to the hardest version of the problem from the start.
- Uses an existing Q-learning based TSC where each agent controls one traffic light.

- Takes information such as the status of traffic light and numbers of vehicles queuing on the incoming lanes as the observation

Recent Advances in Reinforcement Learning for Traffic Signal Control: A Survey of Models and Evaluation [3]

- Surveys on recent RL approaches and methods for traffic signal control, such as value based (DQN) and policy-based (actor-critic, DDPG) methods
- Discusses key design choices: reward functions (queue length, waiting time, throughput), state representations, and action schemes
- Highlights open challenges such as benchmarking datasets, standard baselines, and improving learning efficiency
- Summarizes efficiency and usefulness using simulation tools like SUMO and CityFlow with metrics such as travel time, queue length, and throughput

Reinforcement Learning Benchmarks for Traffic Signal Control [4]

- Introduces RESCO, a standardized RL testbed for traffic signal control built on top of SUMO and SUMO-RL.
- Provides benchmark single- and multi-agent tasks based on realistic real-world traffic scenarios (Cologne and Ingolstadt), enabling fair comparisons across algorithms.
- Implements state-of-the-art RL controllers (IDQN, IPPO, MPLight, FMA2C) and baseline controllers (fixed-time, max-pressure, greedy) for evaluation.
- Reports that many algorithms which claim state-of-the-art performance fail in realistic scenarios, while decentralized DQN-based approaches are often more robust.

More relevant papers can be found here:
https://lucasalegre.github.io/sumo-rl/examples/publications/

Milestones:

Sept.28 (Proposal) - Submit proposal and finalize team roles.

Oct.15 - Environment setup, install SUMO, run SUMO-RL and add pedestrian signals.

Oct.30 (Environment Demo) - Define state/action/reward, and show interaction with the agent.

Nov.15 - Implement and test algorithms (e.g. Q-learning, etc).

Nov.30 (Result Demo) - Compare algorithms on traffic and pedestrian scenarios.

Dec.7 (Final Report) - Analyze results and submit the final report.

References:

[1] H. Chaudhuri, V. Masti, V. Veerendranath, and S. Natarajan, “A comparative study of algorithms for intelligent traffic signal control,” in *Machine Learning and Autonomous Systems*, Singapore: Springer Nature Singapore, 2022, pp. 271–287. doi: 10.1007/978-981-16-7996-4_19.

[2] N. Zheng, J. Li, Z. Mao and K. Tei, "From Local to Global: A Curriculum Learning Approach for Reinforcement Learning-based Traffic Signal Control," *2022 IEEE 2nd International Conference on Software Engineering and Artificial Intelligence (SEAI)*, Xiamen, China, 2022, pp. 253-258, doi: 10.1109/SEAI55746.2022.9832372.

[3] H. Wei, G. Zheng, V. Gayah and Z. Li, “Recent advances in reinforcement learning for traffic signal control: A survey of models and evaluation,” *ACM SIGKDD Explorations Newsletter*, vol. 21, no. 1, pp. 12–18, 2019, doi: 10.1145/3343484.3343488.

[4] Ault J. and Sharon G., “Reinforcement learning benchmarks for traffic signal control,” (2021), *Proceedings of the 35th NeurIPS conference on Neural Information Processing Systems Track on Datasets and Benchmarks,* pp. 1-10, https://openreview.net/forum?id=LqRSh6V0vR