TRAFFIC OPTIMIZATION

MURAT AKİF DUMLU
FURKAN ILISU
KAZIM SANLAV

11.05.2019

OUTLINE

- Introduction and What is RL?
- Problem Definition and Literature Review
- Potential Applications
- Our Journey

Introduction: What is RL

- Reinforcement Learning
  - RL is a machine learning paradigm which aims to train a model via maximizing cumulative rewards of agents resulting from the agents' actions.
- Multi Agent Reinforcement Learning
  - In Multi Agent Reinforcement Learning (MARL), several individual agents aims to maximize their cumulative rewards via co-operating and/or competing each other.

Problem Definition and Literature Review

- Problem: Aim to decrease the delay and congestion on a traffic network utilizing reinforcement learning methods.
- General Approaches: Model the traffic optimization problem as a Markov Decision Process (MDP), and apply Q-learning algorithm to find an optimal policy.

Problem Definition and Literature Review

- An Efficient Deep Reinforcement Learning Model for Urban Traffic Control [1]
- Traffic Flow Optimization using Reinforcement Learning [2]
- Deep Reinforcement Learning for

  Traffic Light Control in Vehicular Networks [3]

Potential Applications

- Simply many of the real world problems that requires cooperation
- Multiplayer online games
- Cooperative robots in the production factories
- Military surveillance systems
- Self-Driving Cars

Self Driving Cars and DRL

- Advantages of DRL on the problem:

  - Can utilize the simulation environment, hence unlimited resource for self play.

  - Whereas, it takes much more time to train a model in physical world.

  - Example: Boston Dynamics, refined a robot model using DRL, with the exact same hardware.

Self Driving Cars and DRL

- So Why not just a single agent DRL?
- In single agen RL, each player tries to maximize its own reward
- A zero sum game
- Which poses an already well known problem:
  - Selfish Routing.

Selfish Routing and Price of Anarchy

- In game theory, there is a concept which describes the problematic nature of zero sum games: Even though each player is rational, there might be situations such that the equilibrium (e.g. Nash Eq, or some other Eq) leads to a socially disadvantageous state. (e.g. Prisoners' Dilemma)

  - Price of anarchy introduced by Koutsoupias and Papadimitriou [4]

    - Aims to measure the degradation in a system's effectiveness due to selfish behaviour.

Prisoner's dilemma

Consider the 2x2 game called prisoner's dilemma, given by the following cost matrix:

	Cooperate	Defect
Cooperate	1, 1	7, 0
Defect	0, 7	5, 5

Selfish Routing and Price of Anarchy

- Braess's paradox

  - Postulated in 1968 by German mathematician Dietrich Braess, who noticed that adding a road to a congested road traffic network could increase overall journey time due to selfish behaviour

  - And it has been used to explain instances of improved traffic flow when existing major roads are closed.

- Hence multi-agent RL have the advantage of cooperation of agents.

Challenges

- MARL is not scalable, heavily rely on hyperparameters
- Continuous action/state space
- Computationally demanding
- Uncertain Environment
- Hard to implement a real world alike physical world simulation

Process

- Literature review went ok.
- Suffered to develop a base-line implementation.
- Could not end up with a working model.

What Did Go Wrong?

- External:
  - Non-trivial Problem
  - Lack of library for multi-agent RL
- Internal:
  - Lack of experience on the domain
  - Communication Problems.

THANK YOU FOR YOUR ATTENTION