CITY-BRAIN CHALLENGE

Phase-traffickers (GroupFour):
Charlotte
Faheem
Max
Build a multi-agent system that minimizes travel time of vehicles
City Brain Challenge - KDD Cup 2021

Total Served Vehicles = max. 1047 vehicles
Delay Index = 1 (best)
<table>
<thead>
<tr>
<th>Data Set</th>
<th>Intersection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warm Up</td>
<td>36</td>
</tr>
<tr>
<td>Round 2</td>
<td>2048</td>
</tr>
<tr>
<td>Round 3</td>
<td>92344</td>
</tr>
</tbody>
</table>
Visualization
Tasks and Goals in our Project

- Infrastructure Set Up - MLFlow & Slurm

- Baselines
  - CoLight
  - Presslight
  - FRAP

- Extensions to Baselines
  - Ape-X
  - QMix
Infrastructure Setup

Set up LRZ CC VM and Slurm

- **Dashboard**
- **mlflow server**

---

**SLURM:**

- Image + Container for Environment Simulation → Installation of Cuda
- Training

---

<table>
<thead>
<tr>
<th>Start Time</th>
<th>Parameters</th>
<th>Metrics</th>
<th>Tags</th>
</tr>
</thead>
<tbody>
<tr>
<td>2021-07-15 12:24:16</td>
<td>env</td>
<td>agent_timesteps</td>
<td>DQN_MultiFlowC...</td>
</tr>
<tr>
<td></td>
<td>env_config</td>
<td>episode_len_mean</td>
<td></td>
</tr>
<tr>
<td></td>
<td>multiagent</td>
<td>episode_reward</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>trial_name</td>
<td></td>
</tr>
<tr>
<td>2021-07-15 12:16:11</td>
<td>{'simulator cfg': 'policies': {'defa...}</td>
<td>120480</td>
<td>-810344</td>
</tr>
<tr>
<td></td>
<td></td>
<td>120</td>
<td>DQN_MultiFlowC...</td>
</tr>
</tbody>
</table>

- log artifacts, hyperparameters, metrics
Baseline: Presslight

- Reward function based on pressure

PressLight: Learning Max Pressure Control to Coordinate Traffic Signals in Arterial Network

Hua Wei†, Chacha Chen‡, Guanjie Zheng†, Kan Wu†, Vikash Gayah†, Kai Xu§, Zhenhui Li†
Baseline: Colight

CoLight: Learning Network-level Cooperation for Traffic Signal Control

Hua Wei†, Nan Xu‡, Huichu Zhang‡, Guanjie Zheng†, Xinshi Zang†, Chacha Chen‡, Weinan Zhang‡, Yanmin Zhu‡, Kai Xu§, Zhenhui Li†
RLLib

https://docs.ray.io/en/master/rllib.html
Extension: Ape-X
Extension: QMix

QMix: Monotonic Value Function Factorisation for Deep Multi-Agent Reinforcement Learning

Tabish Rashid, Mikayel Samvelyan, Christian Schroeder de Witt, Gregory Farquhar, Jakob Foerster, Shimon Whiteson
Training Results

Mean Episode Reward vs Iterations

CoLight

Presslight

QMix with LSTM
## Evaluation Results

Total vehicles served, total delay

<table>
<thead>
<tr>
<th>Model</th>
<th>Dataset</th>
<th>Total vehicles served</th>
<th>Total delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max Pressure (RLLib)</td>
<td>Warm Up</td>
<td>1047</td>
<td>1.3185</td>
</tr>
<tr>
<td>Presslight (RLLib)</td>
<td>Warm Up</td>
<td>1047</td>
<td>1.2609</td>
</tr>
<tr>
<td>CoLight (RLLib)</td>
<td>Warm Up</td>
<td>1047</td>
<td>16.307</td>
</tr>
<tr>
<td>QMix (RLLib)</td>
<td>Warm Up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presslight (Original, single hidden)</td>
<td>Round2</td>
<td>22962</td>
<td>1.6034</td>
</tr>
<tr>
<td>Presslight (Original, extra hidden)</td>
<td>Round2</td>
<td>29128</td>
<td>1.7507</td>
</tr>
<tr>
<td>Starter-kit DQN</td>
<td>Round2</td>
<td>70026</td>
<td>1.6566</td>
</tr>
</tbody>
</table>
Challenges

- Familiarisation with RL concepts under a very tight deadline
- Task redistribution when 2 members left our team
- Switching to new starter-kit with RLlib
- Poor documentation of starter-kit & RLlib
- RLlib’s low GPU utilization on PyTorch
- Limited resources on LRZ AI System
What have we learned?

- Fundamentals of multi-agent reinforcement learning
- Coding RL models with PyTorch, later using RLLib
- Setting up ML environment
- Team collaboration

What would we have done differently in retroperspective

- Stick to TensorFlow for (GPU) facilitation
- Clear prioritization of tasks around submission deadline
Questions?

- Phase-traffickers