A Study on Predictive Deep Reinforcement Learning for Isolated Intersection Signal Control

予測深層強化学習の単独交差点信号制御への適用性に関する一考察

東京大学 生産技術研究所 大口研究室（交通制御工学）
http://www.transport.iis.u-tokyo.ac.jp/
Tianyang HAN, Masaki ITO, Ken SHIRAHATA, Takashi OGUCHI

1.Traffic Signal Control (TSC)

- Basic idea
  - Separate vehicle from different direction temporally to ensure safety. Bring extra waiting time to road users.
- Approach
  - Produce phases and corresponding duration for response arriving vehicle (demand).

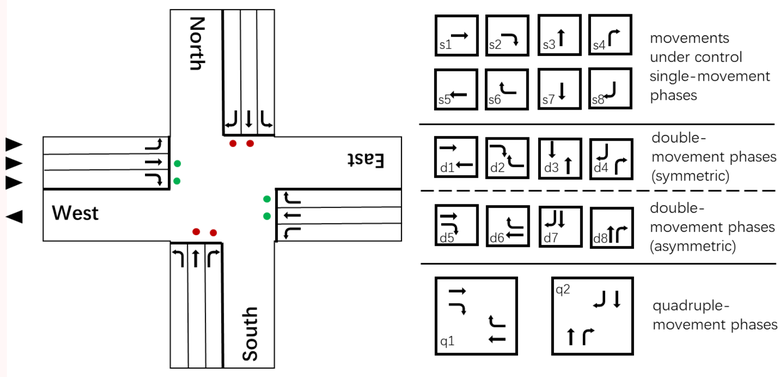


Fig.1 the isolated intersection and its phase combinations

2.Reinforcement Learning(RL) Modelling

- The RL is an autonomous control method to take responsive actions (a) with observed (s) states. Supposing the effect of actions could be measured as the reward r, the RL only needs to learn action under various states to achieve higher reward.
- TSC problem could be modelled as:
  - States: real time traffic demand (how many vehicle incoming?)
  - Action: control logic (which phase to choose?)
  - Reward: measure of efficiency (how effective can the vehicles be served?)

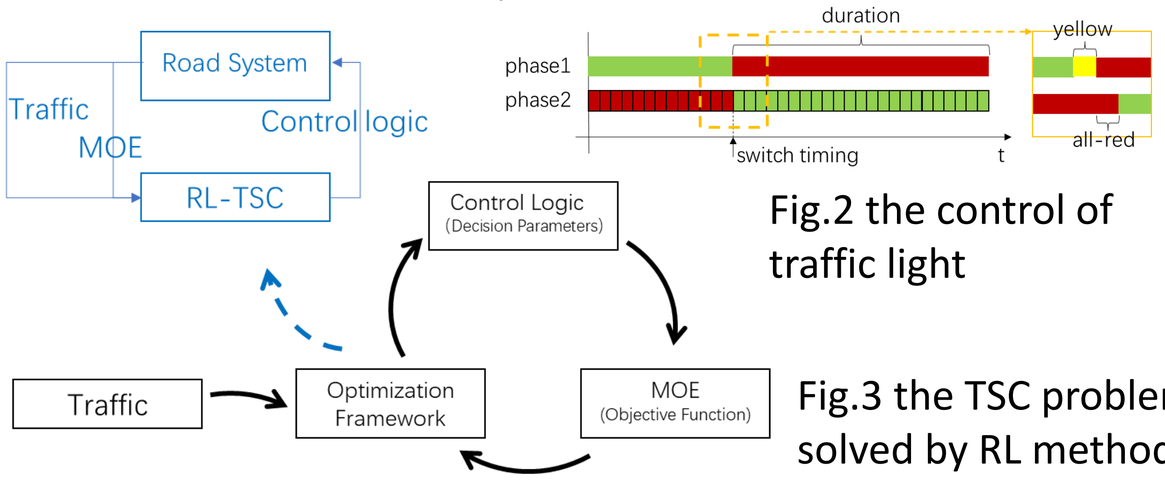


Fig.2 the control of traffic light

Fig.3 the TSC problem solved by RL methods

3.Predictive RL with Queue Estimation

- A model-free RL has difficulty to understand how traffic state change through the time due to the naïve Markov assumption.
- A queue estimation using input-output (IO) model or shock-wave (SW) model can predict future traffic state by considering real-time demand and the control behavior.

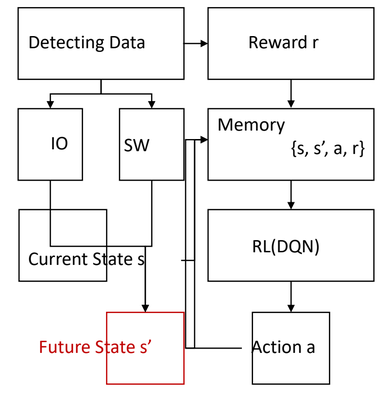


Fig.4 predictive framework

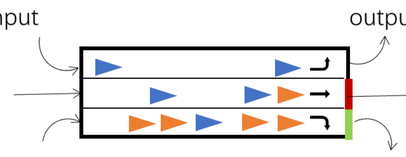


Fig.5 Input-output idea

Queue Estimation — Markovian Process — Markovian Process

$$s(t) \longrightarrow s(t+1) \longrightarrow s(t+2) \longrightarrow \cdots$$

4.Simulation Experiment & Results

- The validation of the proposed method (QueueLearner) is conducted with SUMO (Simulation of Urban MObility).
- Comparison with both traffic engineering methods (FIX, LQF, RHS) and traditional reinforcement (RL, DRL) learning methods.
- Better robustness for demand increasing.

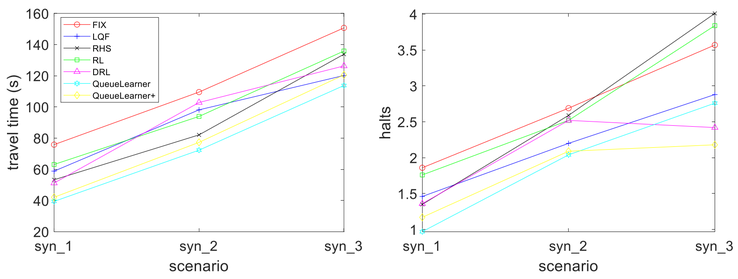
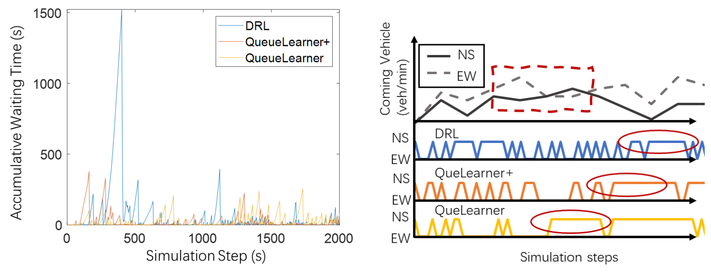


Fig.6 average travel time and stop comparison (demand increasing)

- Fast adaption of traffic change than traditional RL

Fig.7 accumulated waiting time and demand-supply comparison



5.More Related Information

- Coordinated intersections along one artery by communication among decentralized multi-agent RL system.
  - Share-brain for computational economic
  - Green-wave for less stoppage
- Toward a more general urban road network
- Subscribe my google scholar for future outcomes [https://scholar.google.com/citations?user=jDMroEYAAAAJ&hl=zh-CN].

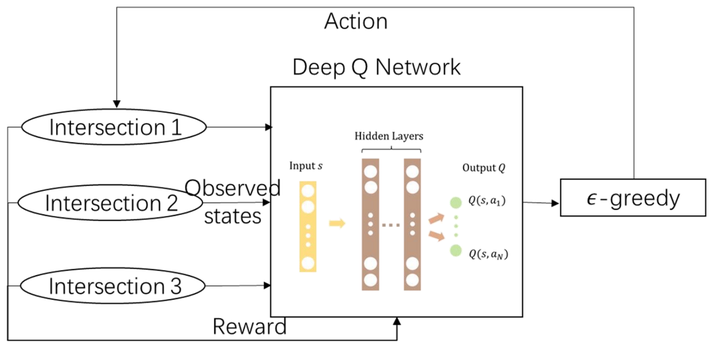


Fig.8 Share-brain multi-agent learning

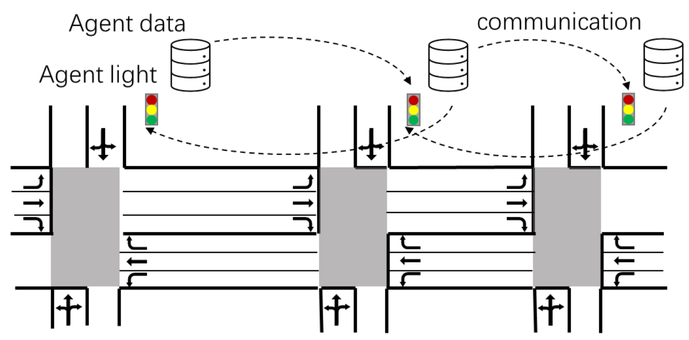


Fig.9 Arterial control