

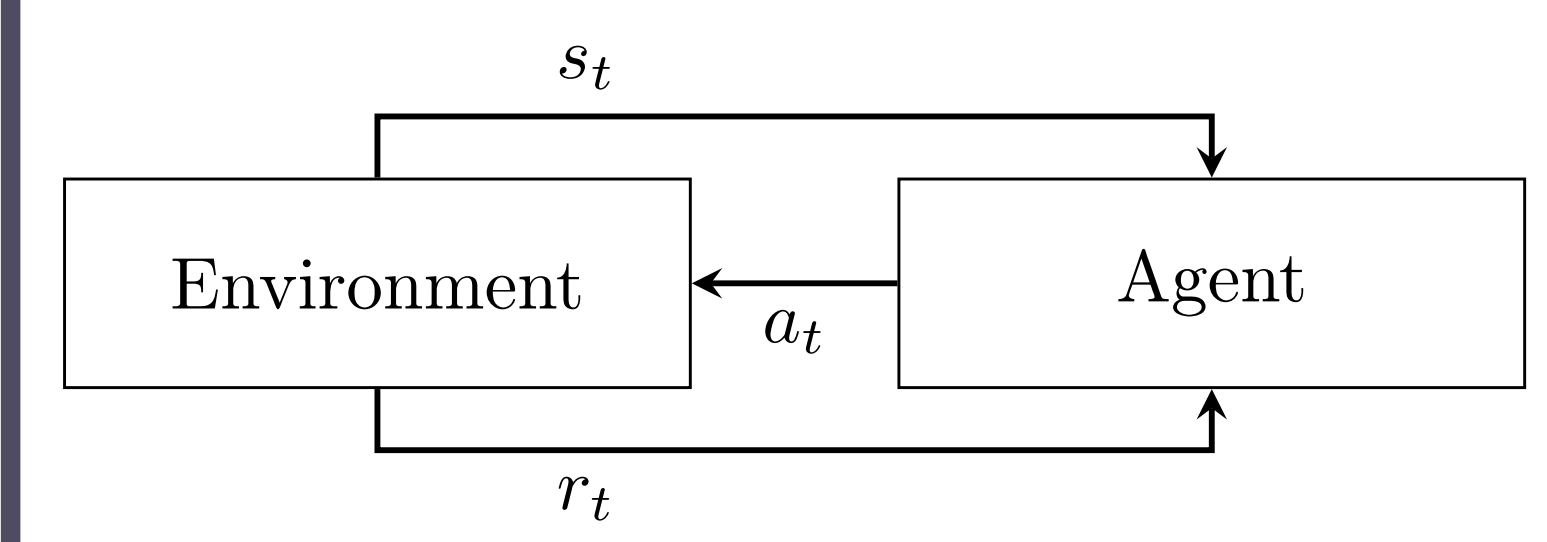
Guidelines for Action Space Definition in Traffic Signal Control Systems

Maxime Tréca, Julian Garbiso, Dominique Barth
{maxime.treca, julian.garbiso}@vedecom.fr, dominique.barth@uvsq.fr

Introduction

- Reinforcement Learning algorithms have been applied to Traffic Signal Control from more than 20 years.
- The controller observes the road network **state**, chooses an **action** to change its traffic lights, and obtains a **reward**.
- Controllers can use phase-based or step-based action schemes, but they are not compared in the literature.

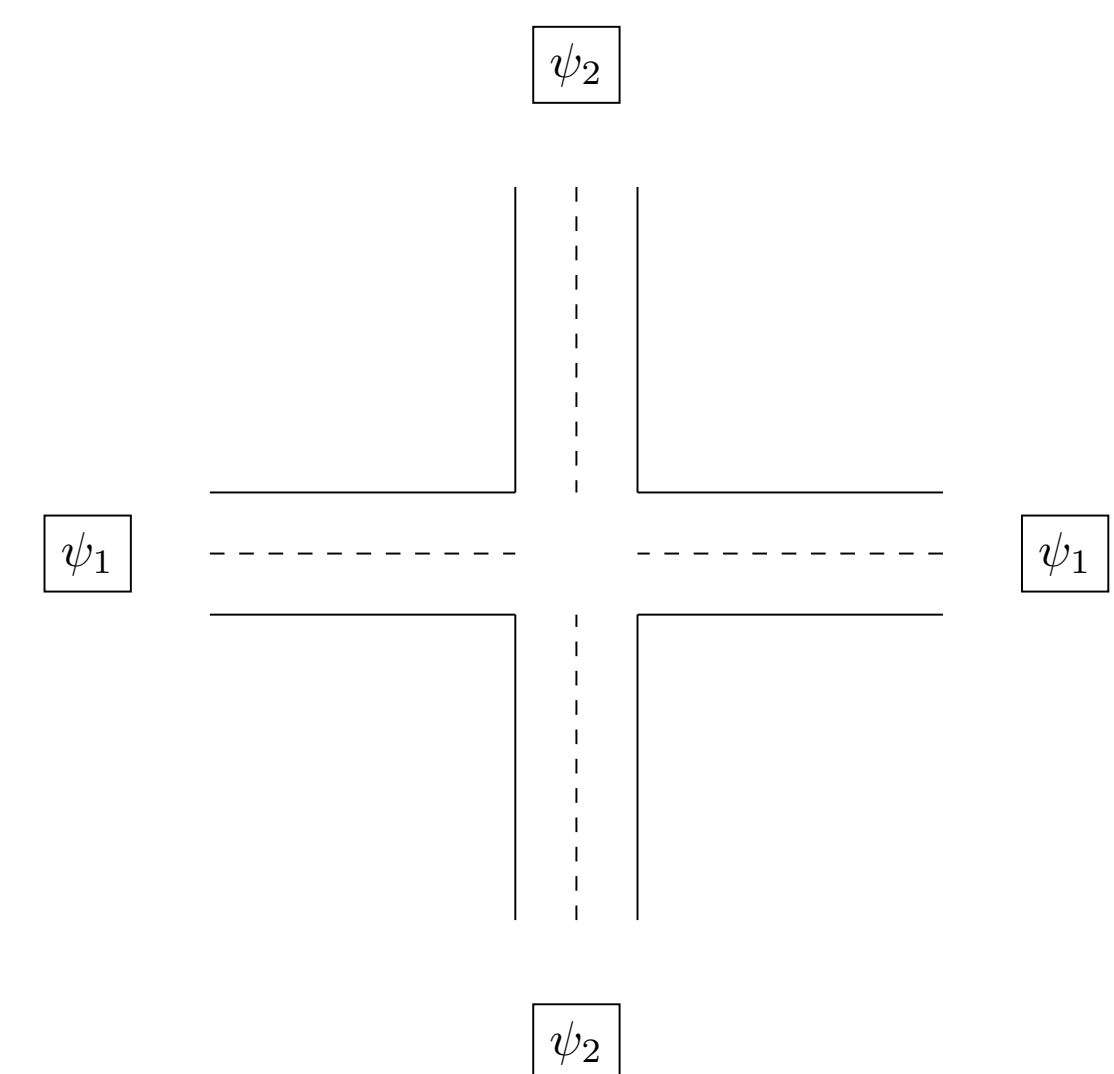
RL Loop



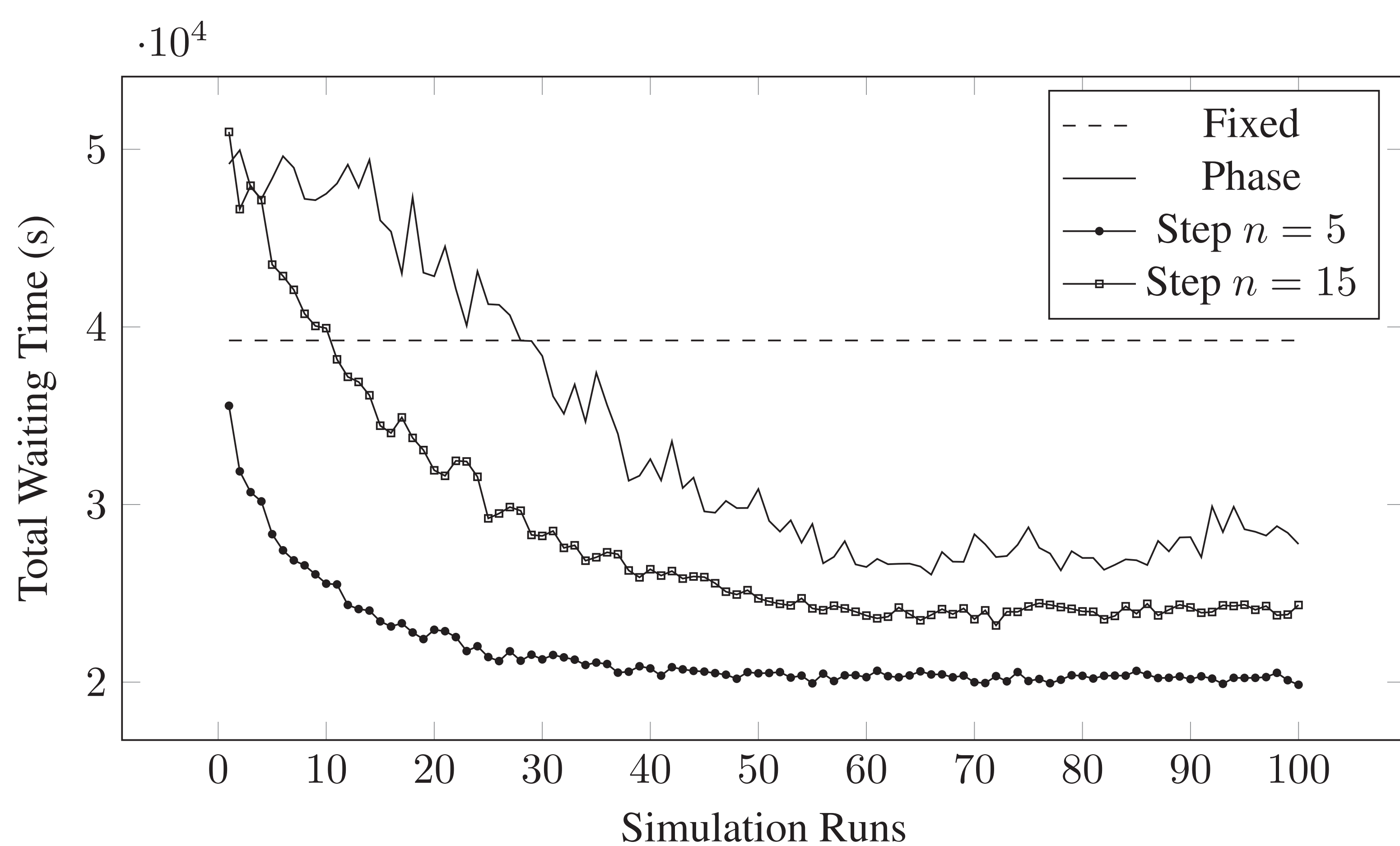
Model

- Single intersection with a single Q-learning agent.
- State : current phase and duration ψ_i, d_i , vehicles per phase.
- Action (phase-based) : The entire length of phase ψ_i .
- Action (step-based) : Whether to extend ψ_i by k steps or not.
- Reward: Difference in accumulated waiting time.

Network



Results



Simulation

- Measure total waiting time of vehicles over 100 simulation iterations on the SUMO simulator.
- We measure action space performance for increasingly skewed traffic scenarios.

Guidelines

- 1 Step-based is always preferable to phase-based actions.
- 2 Shorter action interval is preferable for uniform traffic demand.
- 3 Optimal and realistic step interval between 5 and 10 seconds.

References

- [1] Daniel Krajzewicz, Jakob Erdmann, Michael Behrisch, and Laura Bieker. Recent development and applications of SUMO - Simulation of Urban MObility. *International Journal On Advances in Systems and Measurements*, 5(3&4):128–138, December 2012.
- [2] Christopher JCH Watkins and Peter Dayan. Q-learning. *Machine learning*, 8(3-4):279–292, 1992.
- [3] MA Wiering. Multi-agent reinforcement learning for traffic light control. In *Machine Learning: Proceedings of the Seventeenth International Conference (ICML'2000)*, pages 1151–1158, 2000.