Q-value reuse between state abstractions for traffic light control

Author: Emanuel Kuhn
Supervisors: Jinke He, Rolf Starre, Frans Oliehoek

Background

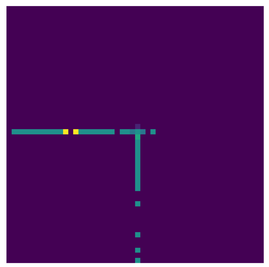
- Agent learns Q-value for each state, action: $Q(s, a)$
- Q-value reuse[1] of source task's learned Q-values to learn target task Q-function:
  $Q(s, a) = Q_{source}(s', a') + Q_{target}(s, a)$

Questions

- Can a simplified state space still learn traffic control?
- Does using a simplified state space lead to learning a (suboptimal) policy faster?
- Can training with the complex state space be sped up using the simplified state space as heuristic?

Method

- Evaluated state abstractions:



Car position matrix
with traffic light values

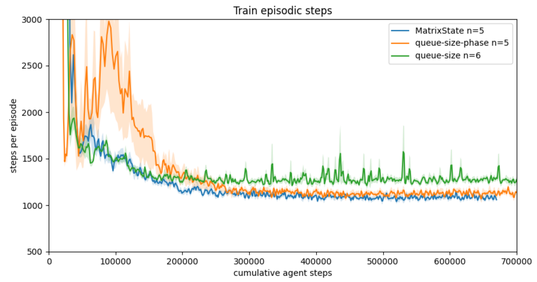
queue sizes

queue sizes and
traffic phase

- Performed Q-value reuse from **queue sizes** to **position matrix**

Results

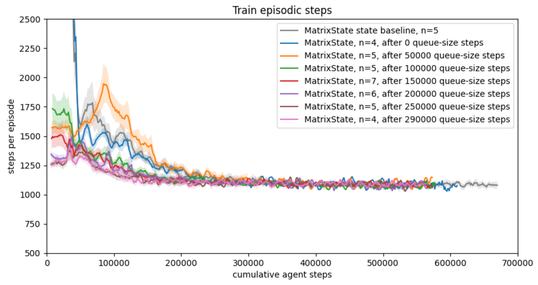
Comparing state abstractions:



- Queue size **without** phase diverges
- Queue size **with** phase converges to **similar performance** as position matrix

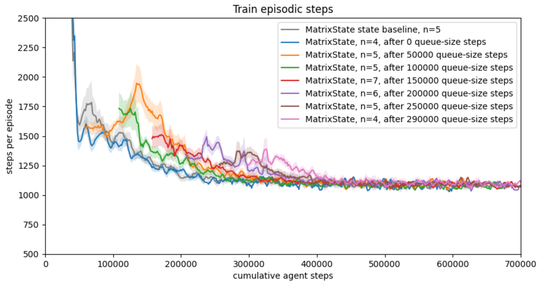
Q-value reuse: queue sizes ➡️ position matrix:

Without source cost:



- More training with source abstraction leads to **fewer** training steps in target abstraction for same performance

With source cost:



- More training with source abstraction leads to **more total** training steps for same performance

Conclusions

- A traffic policy can be learned using **queue sizes** and **current phase** as state space
- A simplified state space **does not** lead to quicker convergence
- Suboptimal Q-values with only **queue sizes** as state space can improve training performance with more **matrix states** agent

Future work

- Compare state abstractions on more complex scenarios
- Try Q-value reuse to learn global Q-function for small multi-intersection scenario

References

1. Matthew E Taylor, Peter Stone, and Yaxin Liu. Transfer learning via inter-task mappings for temporal difference learning. Journal of Machine Learning Research, 8(Sep):2125–2167, 2007.
2. Elise Van der Pol and Frans A Oliehoek. Coordinated deep reinforcement learners for traffic light control. Proceedings of Learning, Inference and Control of Multi-Agent Systems (at NIPS 2016), 2016.

Full paper available at:
http://resolver.tudelft.nl/uuid:aea7f1d8-cd87-4d12-ba43-9bf8ca7c4479

TU Delft