Time Intelligence in Reinforcement Learning

Ashlesha Akella, Chin-Teng Lin; CIBCI lab, Centre for Artificial Intelligence, the School of Computer Science, University of Technology Sydney

INTRODUCTION

Being able to measure and subjectively experience the passage of time is one of the striking aspects of intelligent behaviour. It enables humans and primates to anticipate the future and execute actions with precise timing. The ability to measure time is also vital in artificial agents. Most of the existing reinforcement learning algorithms do not directly address the notion of timing. One way to achieve this is to let an agent develop a representation of the time while interacting with the environment, allowing it to explicitly encode the time intervals.

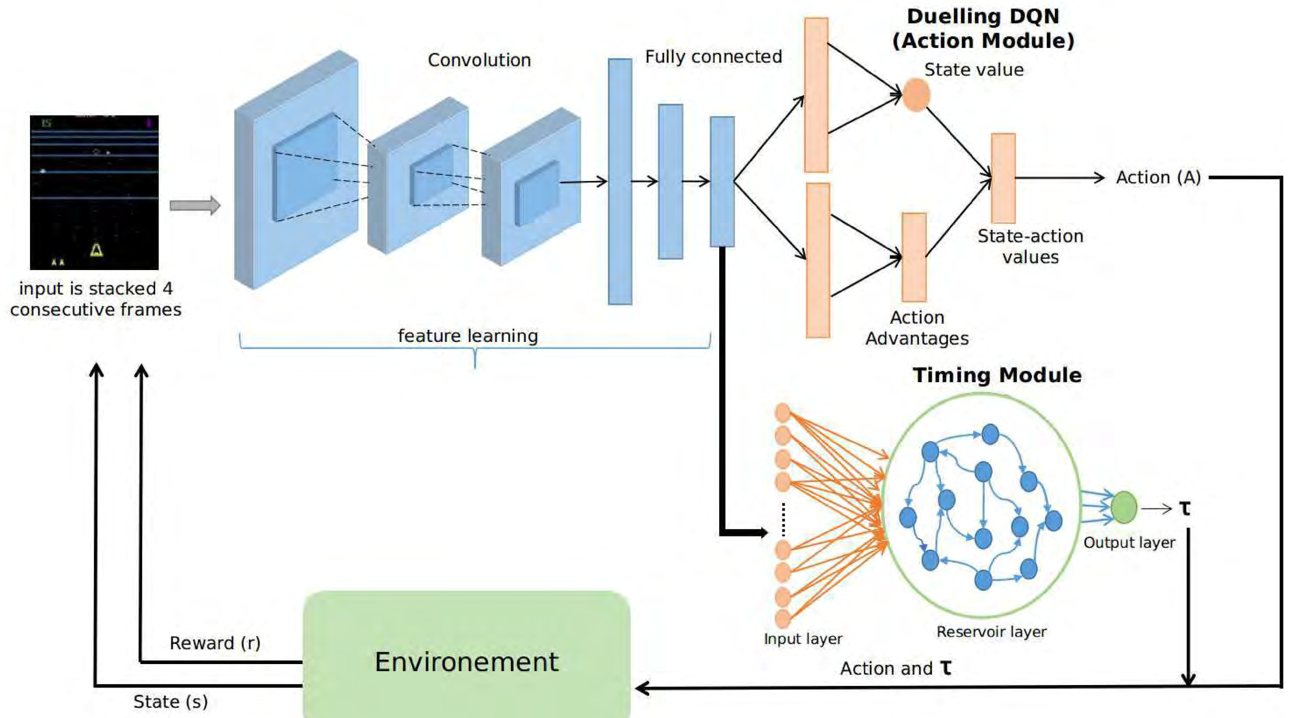
AIM

This research aims to investigate the learning of decision making and development of timing using a reinforcement learning framework. We achieve this by disentangling the process of learning time representation and optimal action into two modules. This research aims to address the following

1. Single agent learning to make an optimal decision at an optimal time interval.
2. Single agent learning to maintain an action for an optimal amount of time.
3. Multiple agent learning to maintain time synchrony in thier actions. (eg: Shape forming agents without centralized control)

METHODS

To disentangle the process of learning time representation and optimal action, we designed a Q-learning architecture that contains two modules 1. A timing module: reservoir computing [1] module, with recurrently connected non-linear firing rate neurons and one readout unit to learn the time representation 2. An action module: Dueling Deep Q Network [2]. At any time step *t* the agent receives a state s_t. State s_t is given as input to the left most layers (a feature learning module), and the output of feature learning module is a common input to both the timing module and action module.



RESULTS

We trained and evaluated our proposed method using Open AI gym environments on 6 Atari games.

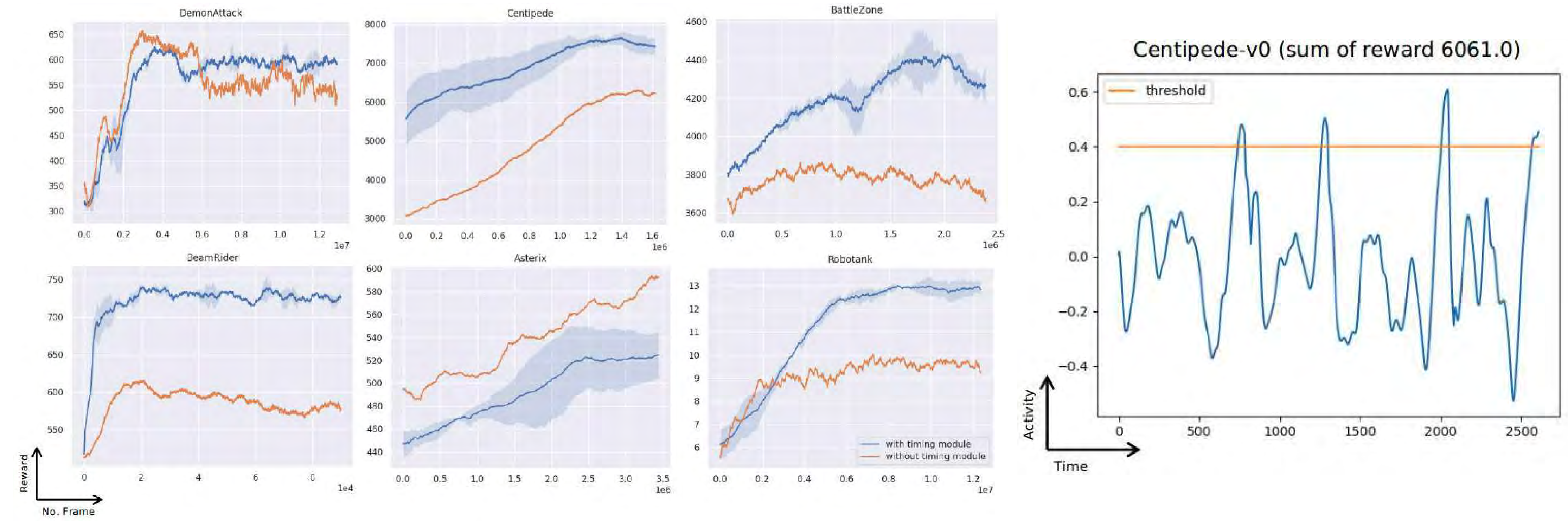


Figure: (Left) Comparison of Dueling DQN with timing module (blue) (average over 2 seeds) and without timing module (orange)on 6 Atari games. (Right) time representation for one episode of Centipede Game.

DISCUSSION

By disentangling the process of learning the temporal and spatial aspects of the action space into independent tasks. Through this strategy, the agent learned to an action and time representation simultaneously. We also observed several other interesting behaviours.Firstly, The agent learned to subdue its activity immediately after observing a new state. We interpreted this as the agent restarting its clock. A gradual increase in activity after receiving the state suggests that the learning might be comparable to a prior study where neuronal activity from the dorsolateral anterior striatum was recorded in rats, which showed a gradual increase in the firing rate up to the time of the expected reward.

CONCLUSIONS

Along with learning rules and objective functions, architectures of neural networks is also a crucial component in deep learning. This research, which used a reservoir of non-linear firing rate neurons gives crucial insights into the learning of temporal properties without discretizing time, thus opens the door to a new way of handling spatiotemporal computations in Reinforcement Learning. In the future, we should investigate the utility of such networks in the field of reinforcement learning for processing continuous-time inputs and actions in different environments and with different learning methods. Other temporal features that humans and non-human primates commonly exhibit, such as time-based prospective memory, and more generalised temporal scaling should also be further examined with reinforcement learning agents.

References

[1] Pathak, J., Hunt, B., Girvan, M., Lu, Z. and Ott, E., 2018. Model-free prediction of large spatiotemporally chaotic systems from data: A reservoir computing approach. Physical review letters, 120(2), p.024102.

[2] Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D. and Riedmiller, M., 2013. Playing atari with deep reinforcement learning. arXiv preprint arXiv:1312.5602.