Quentin Thomas

AIND

11 September 2017

# Research Analysis

## HRA for Reinforcement Learning

SUMMARY

Within reinforcement learning there exists a problem of getting the AI algorithms to generalize. In order to tackle this problem Microsoft teamed up with Mcgill University in order to develop a new approach known as Hybrid Reward Architecture for Reinforcement Learning.

This new algorithm was demonstrated on the arcade game MS. PACMAN. With this approach the Team in proposes to leverage an optimization function that replaces the normal reward function with an alternative reward that has a smoother optimal value. The architecture was based on the Horde Architecture which considers multiple reward functions of multiple agents which are trained in Parallel. The values for each agent is combined in an aggregator of the scores and the greedy optimal value is chosen. Utilizing this architecture the team was able to achieve way above human level performance on the game by gathering 23,377 points on a random start, while the best reported human could only get 15,375. The agent was able to generalize its task of fruit collection for higher point allocation during training in order to achieve higher scores than human players that previously played the game in the past.

# Works Cited

Hybrid Reward Architecture for Reinforcement Learning; June 13 2017 [Paper]

Horde: A Scalable Real-time Architecture for Learning Knowledge from Unsupervised Sensorimotor Interaction [Paper]