

# Homework 6 Write-up

## Conceptual Questions

**What are some of the differences between the REINFORCE algorithm (Monte-Carlo method) and the Advantage Actor Critic?**

The Advantage Actor Critic algorithm differs from the vanilla REINFORCE algorithm mainly in that it includes an extra network that learns the value function. It learns this network so that the loss can be a function with advantage instead of discounted future rewards. Since the advantage function is less erratic this reduces the models variance, which is favorable.

**In reinforcement learning, we often discuss the exploration-exploitation tradeoff. We must balance acquiring new knowledge about our environment (exploration) while simultaneously trying to maximize our return (utilizing all the knowledge we've collected thus far). How do policy gradient methods automatically perform this tradeoff?**

Policy gradient methods use random sampling to choose the action for a given state. This allows for both exploration and exploitation since sometimes the perceived ideal action will be taken, but sometimes non ideal actions can be made. Choosing the better choice proportionally more helps the model exploit the learned policy.

**What advantages do policy gradient methods have over deep Q learning methods? What advantages do deep Q learning methods have over policy gradient methods?**

In some scenarios the Q function can be too complex to learn. In these cases it is favorable to use policy gradient methods which train directly in the action space. But policy gradient methods can be wildly unstable (high variance) and often converge to local minima. When DQN methods can be applied to the task, they will often train with less noise, creating for a more stable training environment.

## **Ethical Implications**

**Do the potential economic benefits of autonomous vehicles outweigh the labor costs? Why or why not?**

The labor costs associated with employing autonomous vehicles are safety, efficiency, efficacy R&D and giving up the jobs that are currently filled by humans in driving, e.g. food delivery, transportation service, trucking. I believe that if there is a way to perform these tasks more efficiently we should because it's safer and more environmentally friendly.

**Do you agree or disagree with the People for the Ethical Treatment of Reinforcement Algorithms? Do we as humans currently have moral obligations to treat RL algorithms ethically? Why or why not? If not, what would change your mind?**

RL algorithms are becoming better at simulating reality. Assuming growth over time, eventually they will become real enough to take part in society in a meaningful way. In such a case that they would be able to impact society, they should be treated ethically and taught to be good because at that point they are affecting all of us.