A Monte Carlo AIXI Approximation

May 16, 2016

Abstract

We implemented the algorithm for learning and planning in partially observable Markov decision processes described in *A Monte Carlo AIXI Approximation*. Because this paper is highly focused on the theoretical aspect of the AIXI approximation, some details were omitted for ease of presentation. We used the following test domains from the paper to assess the performance of our replication,

- **1d-Maze**
- **Cheese Maze**
- **Biased Rock Paper Scissors**
- **Kuhn Poker**

We also implemented Active-LZ, a competing algorithm introduced by Farias et al (2010), which uses the Lempel-Ziv compression scheme for predicting future outcomes in unknown D-Markov environments. Our implementation attains near-optimal results for all four domains. This is encouraging as the domains differed in their complexity. 1d-maze is simpler than the other domains since it is deterministic, has a small state space, and the observations are uninformative. Cheese Maze is also deterministic, but it has a much larger state space. Rock paper scissors and Kuhn Poker are more complex as they are stochastic games, which requires the agent to spend more time gathering observations and rewards in order to learn the dynamics of the games.

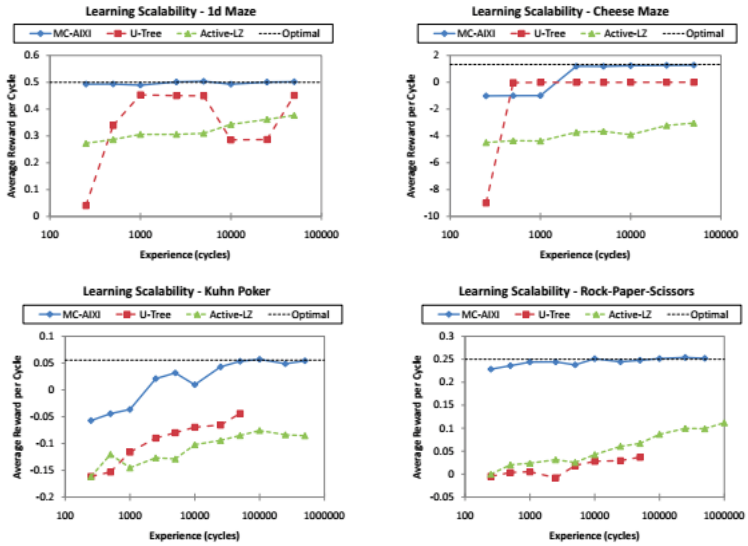

Figure 1: Graphs of the Results from *A Monte Carlo AIXI Approximation*

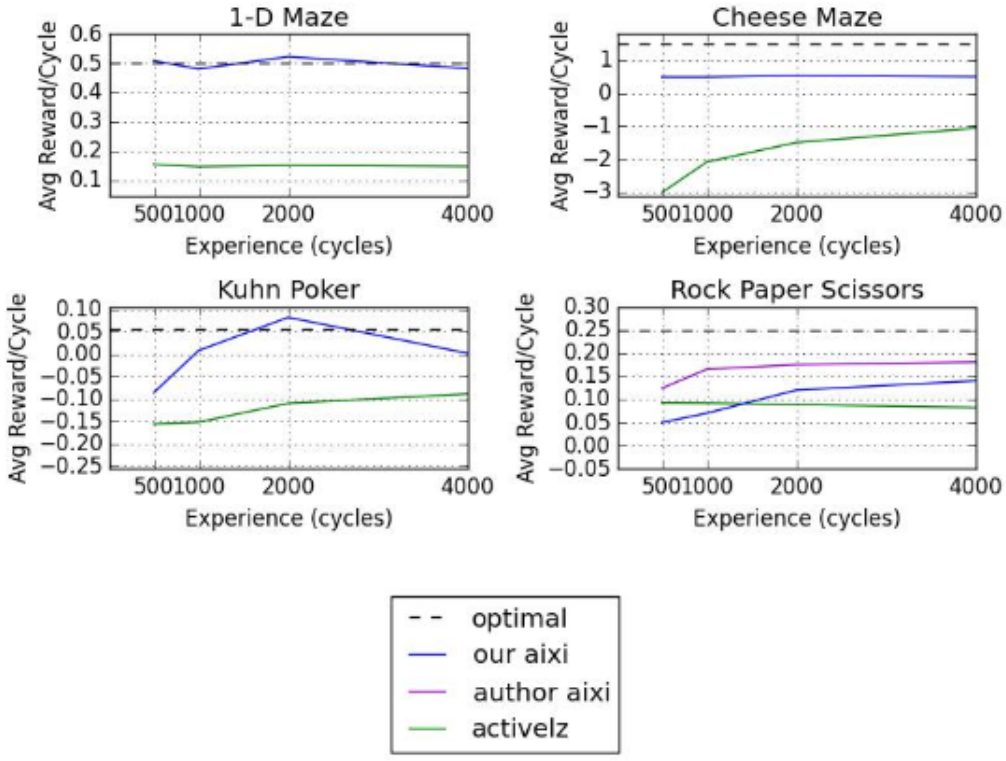

Figure 2: Our Results

Works Cited

Farias et al. 2010. Universal Reinforcement Learning. *IEEE Transactions on Information Theory*. 56(5).

Veness et al. 2011. A Monte Carlo AIXI Approximation. *Journal of Artificial Intelligence Research*. 40: 95-142.