

It has been theorized that giving reinforcement learning agents the ability to model each other would improve their performance in multi-agent tasks. We test this hypothesis by measuring the performance of two different types of agents in a series of competitive games. Specifically, we analyze the sample reward distribution from agents both with and without a “theory of mind” architecture that has been shown to be effective in modeling other agents’ behavior and beliefs. The sample reward distributions (i.e. the data to be analyzed) will be obtained by repeated simulation after the agents have been trained.