Hu, Junling; **Wellman, Michael P.**

Nash Q-learning for general-sum stochastic games. (English) Zbl 1094.68076

J. Mach. Learn. Res. 4, No. 6, 1039-1069 (2004).

Summary: We extend Q-learning to a noncooperative multiagent context, using the framework of general-sum stochastic games. A learning agent maintains Q-functions over joint actions, and performs updates based on assuming Nash equilibrium behavior over the current Q-values. This learning protocol provably converges given certain restrictions on the stage games (defined by Q-values) that arise during learning. Experiments with a pair of two-player grid games suggest that such restrictions on the game structure are not necessarily required. Stage games encountered during learning in both grid environments violate the conditions. However, learning consistently converges in the first grid game, which has a unique equilibrium Q-function, but sometimes fails to converge in the second, which has three different equilibrium Q-functions. In a comparison of offline learning performance in both games, we find agents are more likely to reach a joint optimal path with Nash Q-learning than with a single-agent Q-learning method. When at least one agent adopts Nash Q-learning, the performance of both agents is better than using single-agent Q-learning. We have also implemented an online version of Nash Q-learning that balances exploration with exploitation, yielding improved performance.

MSC:

68T05 Learning and adaptive systems in artificial intelligence
91A60 Probabilistic games; gambling

Keywords:

reinforcement learning; multiagent learning

Full Text: DOI