Abstract

Multi-agent reinforcement learning is an extension of reinforcement learning concept to multi-agent environments. Reinforcement learning allows to program agents by reward and punishment without specifying how to achieve the task. Formally agent-environment interaction in multi-agent reinforcement learning is presented as a discounted stochastic game. The task the agents are facing is formalized as the problem of finding Nash equilibria.

This thesis is devoted to development of multi-agent reinforcement learning algorithms. We propose an algorithm converging in self-play to Nash equilibria for high percentage of general-sum discounted stochastic games. The approach is based on generalization of replicator dynamics for discounted stochastic games. Before there was no algorithm that converged to Nash equilibria for general-sum discounted stochastic games (only for particular cases). The approach also outperforms the methods for solving general-sum discounted stochastic games: nonlinear optimization and stochastic tracing procedure. These algorithms function under the assumption that the games are known from the very beginning in contrast to reinforcement learning where the agent's task is to learn an optimal behavior in unknown environment. Another contribution is an algorithm that always converges to stationary policies and to best-response strategies against stationary opponents. Unlike the first algorithm it doesn't require that the opponents' rewards are observable. We give theoretical foundations for the convergence of the algorithms proposed in this thesis.

The possible application areas include traditional reinforcement learning tasks in multi-agent environments like robot soccer and development of trading agents along with numerous economic problems as a rule modeled as differential games in the field of capital accumulation, advertising, pricing, macroeconomics, warfare and resource economics. We propose to approximate the differential games with stochastic games and apply the developed solver.