Policy iteration for stochastic games

Marianne Akian
INRIA and CMAP, Ecole Polytechnique, CNRS
marianne.akian@inria.fr

joint work with *Stéphane Gaubert*

Abstract

Zero-sum stochastic games with finite state and action spaces, perfect information, and mean payoff criteria arise in particular from the monotone discretization of mean-payoff pursuit-evasion deterministic differential games. When the Markov chains associated to strategies are irreducible, the value of the game can be computed by using Hoffman and Karp policy iteration algorithm (1966), which is similar to the one introduced by Denardo (1967) for solving discounted games.

A feature of policy iteration is that the number of iterations is small in practice, whereas in general it can only be bounded by the number of strategies. Recently, Ye and Hansen, Miltersen and Zwick showed that policy iteration for one or two player zero-sum stochastic games, restricted to instances with a fixed discount rate, is strongly polynomial. We shall show that the Hoffman and Karp algorithm is also strongly polynomial for mean-payoff games with bounded first mean return time to a given state. The proof is based on methods of nonlinear Perron-Frobenius theory and on a reduction of the mean-payoff problem to a discounted problem with state dependent discount rate.

*

References

[1] M. Akian and S. Gaubert, *Policy iteration for perfect information stochastic mean payoff games with bounded first return times is strongly polynomial.* Preprint arXiv:1310.4953, 2013.