The Complexity of Ergodic Mean-payoff Games *

Krishnendu Chatterjee and Rasmus Ibsen-Jensen

IST Austria

Abstract. We study two-player (zero-sum) concurrent mean-payoff games played on a finite-state graph. We focus on the important sub-class of ergodic games where all states are visited infinitely often with probability 1. The algorithmic study of ergodic games was initiated in a seminal work of Hoffman and Karp in 1966, but all basic complexity questions have remained unresolved. Our main results for ergodic games are as follows: We establish (1) an optimal exponential bound on the patience of stationary strategies (where patience of a distribution is the inverse of the smallest positive probability and represents a complexity measure of a stationary strategy); (2) the approximation problem lies in \( \text{FNP} \); (3) the approximation problem is at least as hard as the decision problem for simple stochastic games (for which \( \text{NP} \cap \text{coNP} \) is the long-standing best known bound). We present a variant of the strategy-iteration algorithm by Hoffman and Karp; show that both our algorithm and the classical value-iteration algorithm can approximate the value in exponential time; and identify a subclass where the value-iteration algorithm is a \( \text{FPTAS} \). We also show that the exact value can be expressed in the existential theory of the reals, and establish square-root sum hardness for a related class of games.

1 Introduction

Concurrent Games. Concurrent games are played over finite-state graphs by two players (Player 1 and Player 2) for an infinite number of rounds. In every round, both players simultaneously choose moves (or actions), and the current state and the joint moves determine a probability distribution over the successor states. The outcome of the game (or a play) is an infinite sequence of states and action pairs. Concurrent games were introduced in a seminal work by Shapley [21], and they are the most well-studied game models in stochastic graph games, with many important special cases.

Mean-payoff (Limit-average) Objectives. The most fundamental objective for concurrent games is the limit-average (or mean-payoff) objective, where a reward is associated to every transition and the payoff of a play is the limit-inferior (or limit-superior) average of the rewards of the play. The original work of Shapley [21] considered discounted sum objectives (or games that stop with probability 1); and the class of concurrent games with limit-average objectives (or games that have zero stop probabilities) was introduced by Gillette in [14]. The Player-1 value \( \text{val}(s) \) of the game at a state \( s \) is the supremum value of the expectation that Player 1 can guarantee for the limit-average objective against all strategies of Player 2. The games are zero-sum, so the objective of

---

* The research was partly supported by FWF Grant No P 23499-N23, FWF NFN Grant No S11407-N23 (RiSE), ERC Start grant (279307: Graph Games), and Microsoft faculty fellows award.
† Full version available at [1].

J. Esparza et al. (Eds.): ICALP 2014, Part II, LNCS 8573, pp. 122–133, 2014. © Springer-Verlag Berlin Heidelberg 2014
Player 2 is the opposite. The study of concurrent mean-payoff games and its sub-classes have received huge attention over the last decades, both for mathematical results as well as algorithmic studies. Some key celebrated results are as follows: (1) the existence of values (or determinacy or equivalence of switching of strategy quantifiers for the players as in von-Neumann’s min-max theorem) for concurrent discounted games was established in [21]; (2) the existence of values for the celebrated game of Big-Match was established in [4]; and (3) developing on the results of [4] and on Puiseux series [3] the existence of values for concurrent mean-payoff games was established in [19].

**Sub-classes.** The general class of concurrent mean-payoff games is notoriously difficult for algorithmic analysis. The current best known solution for general concurrent mean-payoff games is achieved by a reduction to the theory of the reals over addition and multiplication with three quantifier alternations [7] (also see [16] for a better reduction for constant state spaces). The strategies that are required in general for concurrent mean-payoff games are infinite-memory strategies that depend in a complex way on the history of the game [19, 4], and analysis of such strategies make the algorithmic study complicated. Hence several sub-classes of concurrent mean-payoff games have been studied algorithmically both in terms of restrictions of the graph structure and restrictions of the objective. The three prominent restrictions in terms of the graph structure are as follows: (1) **Ergodic games (aka irreducible games)** where every state is visited infinitely often almost-surely. (2) **Turn-based stochastic games**, where in each state at most one player can choose between multiple moves. (3) **Deterministic games**, where the transition functions are deterministic. The most well-studied restriction in terms of objective is the **reachability objectives**. A reachability objective consists of a set $U$ of terminal states (absorbing or sink states that are states with only self-loops), such that the set $U$ is exactly the set of states where out-going transitions are assigned reward 1 and all other transitions are assigned reward 0. For all these sub-classes, except deterministic mean-payoff games (that is ergodic mean-payoff games, concurrent reachability games, and turn-based stochastic mean-payoff games) stationary strategies are sufficient, where a stationary strategy is independent of the past history of the game and depends only on the current state.

**An Example.** Consider the ergodic mean-payoff game shown in Figure 1. All transitions other than the dashed edges have probability 1, and each dashed edge has probability $1/2$. The transitions are annotated with the rewards. The stationary optimal strategy for both players is to play the first action ($a_1$ and $b_1$ for Player 1 and Player 2, respectively) with probability $4 - 2 \cdot \sqrt{3}$ in state $s$, and this ensures that the value is $\sqrt{3}$.

**Previous Results.** The decision problem of whether the value of the game at a state is at least a given threshold for turn-based stochastic reachability games (and also turn-based mean-payoff games with deterministic transition function) lie in $\text{NP} \cap \text{coNP}$ [8, 23]. They are among the rare and intriguing combinatorial problems that lie in $\text{NP} \cap \text{coNP}$, but not known to be in $\text{PTIME}$. The existence of polynomial-time algorithms for the above decision questions are long-standing open problems. The algorithmic solution for turn-based games that is most