Available online at www.sciencedirect.com
ScienceDirect
Information Processing Letters 106 (2008) 1–7
Information Processing Letters
www.elsevier.com/locate/ipl

Reduction of stochastic parity to stochastic mean-payoff games

Krishnendu Chatterjee [a,*], Thomas A. Henzinger [a,b]

[a] *EECS, UC Berkeley, USA*
[b] *EPFL, Switzerland*

Received 1 March 2007; received in revised form 20 August 2007; accepted 21 August 2007

Available online 18 October 2007

Communicated by K. Iwama

Abstract

A stochastic graph game is played by two players on a game graph with probabilistic transitions. We consider stochastic graph games with ω-regular winning conditions specified as parity objectives, and mean-payoff (or limit-average) objectives. These games lie in NP ∩ coNP. We present a polynomial-time Turing reduction of stochastic parity games to stochastic mean-payoff games.

Keywords: Formal verification and model checking; Stochastic games; Parity objectives; Mean-payoff objectives; Formal methods

1. Introduction

Graph games. A stochastic graph game [4] is played on a directed graph with three kinds of states: player-1, player-2, and probabilistic states. At player-1 states, player 1 chooses a successor state; at player-2 states, player 2 chooses a successor state; at probabilistic states, a successor state is chosen according to a given probability distribution. The outcome of playing the game forever is an infinite path through the graph. This path is called a *play*. If there are no probabilistic states, we refer to the game as a 2-*player graph game*; otherwise, as a $2\frac{1}{2}$-*player graph game*.

Parity objectives. The theory of graph games with ω-regular winning conditions is the foundation for modeling and synthesizing reactive processes with fairness constraints. In the case of $2\frac{1}{2}$-player graph games, the two players represent a reactive system and its environment, and the probabilistic states represent uncertainty. The *parity* objectives provide an adequate model, as the fairness constraints of reactive processes are ω-regular, and every ω-regular winning condition can be specified as a parity objective [10]. The solution problem for a $2\frac{1}{2}$-player game with parity objective Φ asks for each state s, for the maximal probability with which player 1 can ensure the satisfaction of Φ if the game is started from s. This probability is called the *value* of the game at s, and we refer to the problem of computing values at all states as the *value computation* problem. An *optimal strategy* for player 1 is a strategy that enables player 1 to win with that maximal probability. The existence of *pure memoryless* optimal strategies for $2\frac{1}{2}$-player games with parity objectives was established in [3]: a pure memoryless strategy chooses for each player-1 state a unique successor state, and the state chosen is indepen-

* Corresponding author.
E-mail addresses: c_krish@eecs.berkeley.edu (K. Chatterjee), tah@eecs.berkeley.edu (T.A. Henzinger).

0020-0190/$ – see front matter
doi:10.1016/j.ipl.2007.08.035

dent of the history of the play. The existence of pure memoryless optimal strategies implies that the solution problem for $2\frac{1}{2}$-player games with parity objectives lies in NP ∩ coNP.

Mean-payoff objectives. An important class of quantitative objectives is the class of mean-payoff (or limit-average) objectives. In case of mean-payoff objectives, there is a real-valued reward at each state, and the mean-payoff of player 1 for a play is the limit of the average of the rewards appearing in the play. The objective of player 1 is to maximize the mean-payoff, and value at a state s is the maximal expectation of the mean-payoff that player 1 can ensure if the game is started from s. In $2\frac{1}{2}$-player games with mean-payoff objectives, pure memoryless optimal strategies exist [7]. Again, the existence of pure memoryless optimal strategies implies that the solution problem for $2\frac{1}{2}$-player games with mean-payoff objectives lies in NP ∩ coNP.

Relationship. No polynomial-time algorithm is known for parity objectives and mean-payoff objectives, even in the case of 2-player games. A polynomial-time reduction of 2-player parity games to 2-player mean-payoff games was presented in [6]. A polynomial-time reduction of 2-player mean-payoff games to $2\frac{1}{2}$-player games with reachability objectives was presented in [12], and from the above reduction it is easy to obtain a polynomial-time reduction of 2-player mean-payoff games to $2\frac{1}{2}$-player games with parity objectives with only two parities.

Our result. We present a polynomial-time Turing reduction of $2\frac{1}{2}$-player parity games to $2\frac{1}{2}$-player mean-payoff games for value computation. The reduction generalizes the result of [6] from 2-player games to $2\frac{1}{2}$-player games. The proof of [6] depends on the analysis of graphs and cycles in a graph, whereas our proof depends on the analysis of Markov decision processes and closed connected recurrent set of states in Markov chains. Our proof proceeds in two steps: we first use the results of [2,6] to obtain a polynomial-time reduction of the problem of computing the set of states with value 1 in $2\frac{1}{2}$-player games with parity objectives to the problem of computing values in 2-player mean-payoff games; and then give a polynomial-time reduction of $2\frac{1}{2}$-player parity games to $2\frac{1}{2}$-player mean-payoff games for value computation. As a consequence of our reduction all algorithms for $2\frac{1}{2}$-player mean-payoff games can now be used for $2\frac{1}{2}$-player parity games (see [5] and the chapter by Raghavan in [9] for algorithms for $2\frac{1}{2}$-player mean-payoff games).

2. Definitions

We consider the class of turn-based probabilistic games and some of its subclasses.

Game graphs. A *turn-based probabilistic game graph* (*$2\frac{1}{2}$-player game graph*) $G = ((S, E), (S_1, S_2, S_P), \delta)$ consists of a directed graph (S, E), a partition (S_1, S_2, S_P) of the finite set S of states, and a probabilistic transition function $\delta : S_P \to \mathcal{D}(S)$, where $\mathcal{D}(S)$ denotes the set of probability distributions over the state space S. The states in S_1 are the *player*-1 states, where player 1 decides the successor state; the states in S_2 are the *player*-2 states, where player 2 decides the successor state; and the states in S_P are the *probabilistic* states, where the successor state is chosen according to the probabilistic transition function δ. We assume that for $s \in S_P$ and $t \in S$, we have $(s, t) \in E$ iff $\delta(s)(t) > 0$, and we often write $\delta(s, t)$ for $\delta(s)(t)$. For technical convenience we assume that every state in the graph (S, E) has at least one outgoing edge. For a state $s \in S$, we write $E(s)$ to denote the set $\{t \in S \mid (s, t) \in E\}$ of possible successors. The *turn-based deterministic game graphs (2-player game graphs)* are the special case of the $2\frac{1}{2}$-player game graphs with $S_P = \emptyset$. The *Markov decision processes* (*$1\frac{1}{2}$-player game graphs*) are the special case of the $2\frac{1}{2}$-player game graphs with $S_1 = \emptyset$ or $S_2 = \emptyset$. We refer to the MDPs with $S_2 = \emptyset$ as *player*-1 MDPs, and to the MDPs with $S_1 = \emptyset$ as *player*-2 MDPs.

Plays and strategies. An infinite path, or a *play*, of the game graph G is an infinite sequence $w = \langle s_0, s_1, s_2, \ldots\rangle$ of states such that $(s_k, s_{k+1}) \in E$ for all $k \in \mathbb{N}$. We write Ω for the set of all plays, and for a state $s \in S$, we write $\Omega_s \subseteq \Omega$ for the set of plays that start from the state s. A *strategy* for player 1 is a function $\sigma : S^* \cdot S_1 \to \mathcal{D}(S)$ that assigns a probability distribution to all finite sequences $\boldsymbol{w} \in S^* \cdot S_1$ of states ending in a player-1 state (the sequence represents a prefix of a play). Player 1 follows the strategy σ if in each player-1 move, given that the current history of the game is $\boldsymbol{w} \in S^* \cdot S_1$, she chooses the next state according to the probability distribution $\sigma(\boldsymbol{w})$. A strategy must prescribe only available moves, i.e., for all $\boldsymbol{w} \in S^*$, $s \in S_1$, and $t \in S$, if $\sigma(\boldsymbol{w} \cdot s)(t) > 0$, then $(s, t) \in E$. The strategies for player 2 are defined analogously. We denote by Σ and Π the set of all strategies for player 1 and player 2, respectively.

Once a starting state $s \in S$ and strategies $\sigma \in \Sigma$ and $\pi \in \Pi$ for the two players are fixed, the outcome of the game is a random walk $\omega_s^{\sigma,\pi}$ for which the probabilities of events are uniquely defined, where an *event* $\mathcal{A} \subseteq \Omega$ is a measurable set of plays. For a state $s \in S$ and an event $\mathcal{A} \subseteq \Omega$, we write $\Pr_s^{\sigma,\pi}(\mathcal{A})$ for the probability that a play belongs to $\mathcal{A}$ if the game starts from the state s and the players follow the strategies σ and π, respectively. For a measurable function $f : \Omega \to \mathbb{R}$ we denote by $\mathbb{E}_S^{\sigma,\pi}[f]$ the *expectation* of the function f under the probability measure $\Pr_s^{\sigma,\pi}(\cdot)$.

Strategies that do not use randomization are called pure. A player-1 strategy σ is *pure* if for all $\boldsymbol{w} \in S^*$ and $s \in S_1$, there is a state $t \in S$ such that $\sigma(\boldsymbol{w} \cdot s)(t) = 1$. A *memoryless* player-1 strategy does not depend on the history of the play but only on the current state; i.e., for all $\boldsymbol{w}, \boldsymbol{w}' \in S^*$ and for all $s \in S_1$ we have $\sigma(\boldsymbol{w} \cdot s) = \sigma(\boldsymbol{w}' \cdot s)$. A memoryless strategy can be represented as a function $\sigma : S_1 \to \mathcal{D}(S)$. A *pure memoryless strategy* is a strategy that is both pure and memoryless. A pure memoryless strategy for player 1 can be represented as a function $\sigma : S_1 \to S$. We denote by Σ^{PM} the set of pure memoryless strategies for player 1. The pure memoryless player-2 strategies Π^{PM} are defined analogously.

Given a pure memoryless strategy $\sigma \in \Sigma^{\mathrm{PM}}$, let G_σ be the game graph obtained from G under the constraint that player 1 follows the strategy σ. The corresponding definition G_π for a player-2 strategy $\pi \in \Pi^{\mathrm{PM}}$ is analogous, and we write $G_{\sigma,\pi}$ for the game graph obtained from G if both players follow the pure memoryless strategies σ and π, respectively. Observe that given a $2\frac{1}{2}$-player game graph G and a pure memoryless player-1 strategy σ, the result G_σ is a player-2 MDP. Similarly, for a player-1 MDP G and a pure memoryless player-1 strategy σ, the result G_σ is a Markov chain. Hence, if G is a $2\frac{1}{2}$-player game graph and the two players follow pure memoryless strategies σ and π, the result $G_{\sigma,\pi}$ is a Markov chain.

Qualitative objectives. We specify *qualitative* objectives for the players by providing a set of *winning* plays $\Phi \subseteq \Omega$ for each player. We say that a play ω *satisfies* the objective Φ if $\omega \in \Phi$. We study only zero-sum games, where the objectives of the two players are complementary; i.e., if player 1 has the objective Φ, then player 2 has the objective $\Omega \setminus \Phi$. We consider *ω-regular objectives* [10], specified as parity conditions. We also define reachability objectives, which is a special class of ω-regular objectives.

- *Reachability objectives.* Given a set $T \subseteq S$ of "target" states, the reachability objective requires that some state of T be visited. The set of winning plays is $\mathrm{Reach}(T) = \{\langle s_0, s_1, s_2, \ldots\rangle \in \Omega \mid s_k \in T$ for some $k \geqslant 0\}$.
- *Parity objectives.* For $c, d \in \mathbb{N}$, we write $[c..d] = \{c, c+1, \ldots, d\}$. Let $p : S \to [O..d]$ be a function that assigns a *priority* $p(s)$ to every state $s \in S$, where $d \in \mathbb{N}$. For a play $\omega = \langle s_0, s_1, \ldots\rangle \in \Omega$, we define $\mathrm{Inf}(\omega) = \{s \in S \mid s_k = s$ for infinitely many $k\}$ to be the set of states that occur infinitely often in ω. The *even-parity objective* is defined as $\mathrm{Parity}(p) = \{\omega \in \Omega \mid \max(p(\mathrm{Inf}(\omega)))$ is even$\}$, and the *odd-parity objective* as $\mathrm{coParity}(p) = \{\omega \in \Omega \mid \max(p(\mathrm{Inf}(\omega)))$ is odd$\}$. In other words, the even-parity objective requires that the maximum priority visited infinitely often is even, and the odd-parity objective is the dual. In sequel we will use Φ to denote parity objectives.

Quantitative objectives. A *quantitative* objective is specified as a measurable function $f : \Omega \to \mathbb{R}$. In zero-sum games the objectives of the players are functions f and $-f$, respectively. We consider a special class of quantitative objectives, namely, mean-payoff objectives. The definition of mean-payoff objectives is as follows.

- *Mean-payoff objectives.* Let $r : S \to \mathbb{R}$ be a real-valued reward function that assigns to every state s the reward $r(s)$. The *mean-payoff* objective *MeanPay* assigns to every play the "long-run" average of the rewards appearing in the play. Formally, for a play $\omega = \langle s_1, s_2, s_3, \ldots\rangle$ we have
$$\mathrm{MeanPay}(r)(\omega) = \liminf_{n\to\infty} \frac{1}{n} \cdot \sum_{i=1}^{n} r(s_i).$$
Note that the complementary objective $-\mathrm{MeanPay}$ is as follows
$$-\mathrm{MeanPay}(r)(\omega) = \limsup_{n\to\infty} \frac{1}{n} \cdot \sum_{i=1}^{n} -\bigl(r(s_i)\bigr).$$

Values and optimal strategies. Given objectives $\Phi \subseteq \Omega$ for player 1 and $\Omega \setminus \Phi$ for player 2, and measurable functions f and $-f$ for player 1 and player 2, respectively, we define the *value* functions $\langle\langle 1\rangle\rangle_{\mathrm{val}}$ and $\langle\langle 2\rangle\rangle_{\mathrm{val}}$ for the players 1 and 2, respectively, as the following functions from the state space S to the set $\mathbb{R}$ of reals: for all states $s \in S$, let

$$\langle\langle 1\rangle\rangle_{\mathrm{val}}(\Phi)(s) = \sup_{\sigma\in\Sigma} \inf_{\pi\in\Pi} \Pr_s^{\sigma,\pi}(\Phi);$$
$$\langle\langle 1\rangle\rangle_{\mathrm{val}}(f)(s) = \sup_{\sigma\in\Sigma} \inf_{\pi\in\Pi} \mathbb{E}_s^{\sigma,\pi}[f];$$
$$\langle\langle 2\rangle\rangle_{\mathrm{val}}(\Omega\setminus\Phi)(s) = \sup_{\pi\in\Pi} \inf_{\sigma\in\Sigma} \Pr_s^{\sigma,\pi}(\Omega\setminus\Phi);$$
$$\langle\langle 2\rangle\rangle_{\mathrm{val}}(-f)(s) = \sup_{\pi\in\Pi} \inf_{\sigma\in\Sigma} \mathbb{E}_s^{\sigma,\pi}[-f].$$

In other words, the values $\langle\langle 1\rangle\rangle_{\mathrm{val}}(\Phi)(s)$ and $\langle\langle 1\rangle\rangle_{\mathrm{val}}(f)(s)$ give the maximal probability and expectation with which player 1 can achieve her objectives Φ and f from state s, and analogously for player 2. The strategies that achieve the values are called optimal: a strategy σ for player 1 is *optimal* from the state s for the objective Φ if $\langle\langle 1\rangle\rangle_{\mathrm{val}}(\Phi)(s) = \inf_{\pi\in\Pi} \mathrm{Pr}_s^{\sigma,\pi}(\Phi)$; and σ is *optimal* from the state s for f if $\langle\langle 1\rangle\rangle_{\mathrm{val}}(f)(s) = \inf_{\pi\in\Pi} \mathbb{E}_s^{\sigma,\pi}[f]$. The optimal strategies for player 2 are defined analogously. We now state the classical determinacy results for $2\frac{1}{2}$-player parity and mean-payoff games.

Theorem 1 *(Quantitative determinacy). For all $2\frac{1}{2}$-player game graphs $G = ((S, E), (S_1, S_2, S_P), \delta)$, the following assertions hold*:

1. [7] *For all reward functions $r : S \to \mathbb{R}$ and all states s, we have $\langle\langle 1\rangle\rangle_{\mathrm{val}}(\mathrm{MeanPay}(r))(s) + \langle\langle 2\rangle\rangle_{\mathrm{val}} \times (-\mathrm{MeanPay}(r))(s) = 0$. Pure memoryless optimal strategies exist for both players from all states.*
2. [3,8,11] *For all parity objectives Φ and all states s, we have $\langle\langle 1\rangle\rangle_{\mathrm{val}}(\Phi)(s) + \langle\langle 2\rangle\rangle_{\mathrm{val}}(\Omega\setminus\Phi)(s) = 1$. Pure memoryless optimal strategies exist for both players from all states.*

Since in $2\frac{1}{2}$-player games with parity and mean-payoff objectives pure memoryless strategies suffice for optimality, in the sequel we consider only pure memoryless strategies.

Relationship between parity and mean-payoff objectives. The parity objectives lie in the intersection of the third level of the Borel hierarchy (i.e., in $\Sigma_3 \cap \Pi_3$) [10]. The mean-payoff objectives are complete for the third level of the Borel hierarchy: mean-payoff objectives are Π_3-complete (see [1] for a proof of Π_3-hardness, and they can also be shown to lie in Π_3). A polynomial-time reduction of 2-player parity games to 2-player mean-payoff games was presented in [6]. A polynomial-time reduction of 2-player mean-payoff games to $2\frac{1}{2}$-player games with reachability objectives was presented in [12], i.e., there is a polynomial-time reduction of 2-player mean-payoff games to $2\frac{1}{2}$-player games with parity objectives with only two priorities.

3. Reduction of $2\frac{1}{2}$-player parity to mean-payoff games

In this section we present a polynomial-time Turing reduction of $2\frac{1}{2}$-player parity games to $2\frac{1}{2}$-player mean-payoff games. The reduction will be obtained in two stages. The first stage consists of computation of set of states with value 1 for a parity objective. These states are called *almost-sure* winning states.

Almost-sure winning states. Given a $2\frac{1}{2}$-player game graph G with a parity objective Φ for player 1, we denote by

$$W_1^G(\Phi) = \{s \in S \mid \langle\langle 1\rangle\rangle_{\mathrm{val}}(\Phi)(s) = 1\};$$
$$W_2^G(\Omega\setminus\Phi) = \{s \in S \mid \langle\langle 2\rangle\rangle_{\mathrm{val}}(\Omega\setminus\Phi)(s) = 1\},$$

the sets of states such that the values for player 1 and player 2 are 1, respectively. These sets of states are also referred as the *almost-sure* winning states for the two players.

Reduction for almost-sure winning states. The results of [2] showed that the computation of almost-sure winning states in a $2\frac{1}{2}$-player game graph G with n states and a parity objective with d priorities, can be achieved by a reduction to a 2-player game graph with $n \cdot d$ states, and a parity objective with $d + 1$ priorities. The result of [6] established a polynomial-time reduction of 2-player games with parity objectives to 2-player games with mean-payoff objectives. The above two reductions ensure that the computation of almost-sure winning states in $2\frac{1}{2}$-player games with parity objectives can be reduced to the computation of values in 2-player games with mean-payoff objectives.

Theorem 2. *(See [2,6].) There is a polynomial-time algorithm that takes as input a $2\frac{1}{2}$-player game graph G and a parity objective Φ, and outputs a 2-player game graph G' with a reward function r' such that given the value function $\langle\langle 1\rangle\rangle_{\mathrm{val}}(\mathrm{MeanPay}(r'))$ in G', the almost-sure winning set $W_1^G(\Phi)$ can be computed in polynomial time.*

Reduction for value computation. We now present a reduction of $2\frac{1}{2}$-player parity games to $2\frac{1}{2}$-player mean-payoff games for value computation. Note that the computation of almost-sure winning states can be achieved by solving 2-player (and hence $2\frac{1}{2}$-player) mean-payoff games. Theorem 3 presents the reduction for value computation. We first present a lemma that will be used in the proof of Theorem 3. In sequel we will use the following terminology: for a Markov chain G, a set C is a *closed connected recurrent* set if C is a bottom strongly connected component in the graph of G.

Lemma 1. *Let C be a closed connected recurrent set of states in a Markov chain, and let $\delta_{\min} = \min\{\delta(s)(t) \mid s, t \in C, \delta(s)(t) > 0\}$. For two states $s, t \in C$, let*

$$\mathrm{freq}(s,t) = \liminf_{n\to\infty} \frac{1}{n} \cdot \sum_{j=0}^{n-1} \Pr_s(X_j = t),$$

where X_j is a random variable denoting the jth state of a play, i.e., $\mathrm{freq}(s,t)$ *denotes the "long-run" frequency of state t with starting state s. Then for all states $s, t \in C$ we have*

$$\mathrm{freq}(s,t) \geqslant \frac{1}{n} \cdot (\delta_{\min})^n,$$

where $n = |C|$.

Proof. For a state $t \in C$, let $\mathrm{In}(t) = \{s \in C \mid \delta(s)(t) > 0\}$ be the set of states with incoming edges to t. We start with two simple facts.

Fact 1. *For a state $t \in C$, for all $s \in C$, we have*

$$\mathrm{freq}(s,t) \geqslant \mathrm{freq}(s,t') \cdot \delta(t')(t) \geqslant \mathrm{freq}(s,t') \cdot \delta_{\min};$$

for $t' \in \mathrm{In}(t)$.

Fact 2. *For all states $s \in C$, we have* $\sum_{t\in C} \mathrm{freq}(s,t) = 1$.

The first fact relates the "long-run" frequency of a state to the "long-run" frequency of the predecessors, and since C is a closed connected recurrent set of states, the sum of the "long-run" frequencies of states in C is 1. Assume towards contradiction that there exist $s, t \in C$ such that $\mathrm{freq}(s,t) < (1/2) \cdot (\delta_{\min})^n$. It follows from Fact 1, that for all states $t' \in \mathrm{In}(t)$ we have $\mathrm{freq}(s,t') < (1/n) \cdot (\delta_{\min})^{n-1}$. Again for a state $t' \in \mathrm{In}(t)$, for all $t'' \in \mathrm{In}(t')$ we have $\mathrm{freq}(s,t'') < (1/n) \cdot (\delta_{\min})^{n-2}$, and so on. Since $|C| = n$, it follows that for all states $s' \in C$ we have $\mathrm{freq}(s,s') < \frac{1}{n}$. Again as $|C| = n$, this contradicts Fact 2 that $\sum_{s'\in C} \mathrm{freq}(s,s') = 1$. Hence the desired result follows. □

Theorem 3. *Let $G = ((S,E),(S_1,S_2,S_P),\delta)$ be a $2\frac{1}{2}$-player game graph. Let $p : S \to [0..d]$ be a priority function, and let $W_1 = W_1^G(\mathrm{Parity}(p))$ and $W_2 = W_2^G(\mathrm{coParity}(p))$ be the sets of almost-sure winning states for the two players. Let*

$$\delta_{\min} = \min\{\delta(s)(t) \mid s \in S_P, t \in S, \delta(s)(t) > 0\}.$$

Consider the reward function $r : S \to \mathbb{R}$ defined as follows:

$$r(s) = \begin{cases} 1 & \text{if } s \in W_1; \\ -1 & \text{if } s \in W_2; \\ (-1)^k \cdot (2 \cdot n)^k \cdot (\frac{1}{\delta_{\min}})^{n\cdot k}; & \\ \quad \text{if } p(s) = k \text{ and } s \in S\backslash(W_1 \cup W_2), \end{cases}$$

where $n = |S|$. Then for all states $s \in S\backslash(W_1 \cup W_2)$, we have

$$\langle\langle 1\rangle\rangle_{\mathrm{val}}\big(\mathrm{Parity}(p)\big)(s) = \frac{1}{2} \cdot \big(\langle\langle 1\rangle\rangle_{\mathrm{val}}\big(\mathrm{MeanPay}(r)\big)(s) + 1\big).$$

Proof. We prove the following two inequalities.

1. We first prove that for all $s \in S\backslash(W_1 \cup W_2)$ we have

$$\langle\langle 1\rangle\rangle_{\mathrm{val}}\big(\mathrm{Parity}(p)\big)(s) \leqslant \frac{1}{2} \cdot \big(\langle\langle 1\rangle\rangle_{\mathrm{val}}\big(\mathrm{MeanPay}(r)\big)(s) + 1\big).$$

Consider a pure memoryless optimal strategy σ for player 1 for the parity objective $\mathrm{Parity}(p)$. Fix the strategy in the mean-payoff game, and consider a pure memoryless counter-optimal strategy π for player 2 in the MDP G_σ (i.e., the strategy π is optimal in G_σ for the objective $-\mathrm{MeanPay}(r)$). We first show that

$$\Pr_s^{\sigma,\pi}\big(\mathrm{Reach}(W_2)\big) \leqslant \langle\langle 2\rangle\rangle_{\mathrm{val}}\big(\mathrm{coParity}(p)\big)(s) = 1 - \langle\langle 1\rangle\rangle_{\mathrm{val}}\big(\mathrm{Parity}(p)\big)(s).$$

Otherwise, if

$$\Pr_s^{\sigma,\pi}\big(\mathrm{Reach}(W_2)\big) > \langle\langle 2\rangle\rangle_{\mathrm{val}}\big(\mathrm{coParity}(p)\big)(s),$$

then player 2 plays π to reach W_2 and an almost-sure winning strategy for $\mathrm{coParity}(p)$ from W_2 to ensure that the probability to satisfy $\mathrm{coParity}(p)$ given σ is greater than $\langle\langle 2\rangle\rangle_{\mathrm{val}}(\mathrm{coParity}(p))(s)$; this contradicts that σ is optimal. Now consider the Markov chain $G_{\sigma,\pi}$. Let C be a closed connected recurrent set of states in $G_{\sigma,\pi}$. If $C \cap (S\backslash(W_1 \cup W_2)) \neq \emptyset$, then there is a state $s' \in C \cap (S\backslash(W_1 \cup W_2))$ with $\langle\langle 1\rangle\rangle_{\mathrm{val}}(\mathrm{Parity}(p))(s') > 0$. Since σ is optimal for player 1 for $\mathrm{Parity}(p)$ and in $G_{\sigma,\pi}$ from s' the set C is visited infinitely often with probability 1, it follows that $\max(p(C))$ is even. Let $z \in C$ be a state with $p(z) = \max(p(C))$. Then since the minimum transition probability is $\delta_{\min}$ and $|C| \leqslant |S|$, it follows from Lemma 1 that the long-run frequency for state z is at least $\frac{1}{n} \cdot (\delta_{\min})^n$. The reward assignment ensures that the long-run average for the closed connected recurrent set C is at least 1. This is obtained as follows. If $p(z) = 0$, then for all states $s \in C$ we must have $p(s) = p(z) = 0$, and then long-run average for C is $(2 \cdot n)^0 \cdot (1/\delta_{\min})^{n\cdot 0} = 1$. We consider the case with $p(z) \geqslant 2$ and then long-run average contribution by z is at least

$$\frac{1}{n}\cdot(\delta_{\min})^n\cdot(2\cdot n)^{p(z)}\cdot\left(\frac{1}{\delta_{\min}}\right)^{n\cdot p(z)} = 2\cdot\left((2\cdot n)^{p(z)-1}\cdot\left(\frac{1}{\delta_{\min}}\right)^{n\cdot(p(z)-1)}\right);$$

(this obtained by multiplying the long-run frequency of z along with its reward). Since $p(z)$ is the greatest priority appearing in C, the long-run average contribution of all the other states in C is at least

$$-\left((2\cdot n)^{p(z)-1}\cdot\left(\frac{1}{\delta_{\min}}\right)^{n\cdot(p(z)-1)}\right),$$

(in the worst case all other states have priority $p(z)-1$). Hence the long-run average in C is at least

$$\left((2\cdot n)^{p(z)-1}\cdot\left(\frac{1}{\delta_{\min}}\right)^{n\cdot(p(z)-1)}\right);$$

the claim follows. A lower bound on the long-run average payoff for player 1 is obtained as follows: we consider the maximum probability of reaching W_2 and consider the closed connected recurrent states C that intersect with W_2 is contained in W_2 (and the long-run average is -1 in this case) and with the rest of the probability the long-run average is at least 1. Hence we have

$$\begin{aligned}&\langle\langle 1\rangle\rangle_{\mathrm{val}}\big(\mathrm{MeanPay}(r)\big)(s)\\ &\quad\geqslant(-1)\cdot\big(1-\langle\langle 1\rangle\rangle_{\mathrm{val}}\big(\mathrm{Parity}(p)\big)(s)\big)\\ &\qquad+1\cdot\langle\langle 1\rangle\rangle_{\mathrm{val}}\big(\mathrm{Parity}(p)\big)(s)\\ &\quad=2\cdot\langle\langle 1\rangle\rangle_{\mathrm{val}}\big(\mathrm{Parity}(p)\big)(s)-1.\end{aligned}$$

2. We now prove that for all $s\in S\backslash(W_1\cup W_2)$ we have

$$\langle\langle 1\rangle\rangle_{\mathrm{val}}\big(\mathrm{Parity}(p)\big)(s) \geqslant\frac{1}{2}\cdot\big(\langle\langle 1\rangle\rangle_{\mathrm{val}}\big(\mathrm{MeanPay}(r)\big)(s)+1\big).$$

Consider a pure memoryless optimal strategy π for player 2 for the objective $\mathrm{coParity}(p)$. Fix the strategy in the mean-payoff game, and consider a pure memoryless counter-optimal strategy σ for player 1 in the MDP G_π (i.e., the strategy σ is optimal in G_σ for the objective $\mathrm{MeanPay}(r)$). We first show that

$$\mathrm{Pr}_s^{\sigma,\pi}\big(\mathrm{Reach}(W_1)\big)\leqslant\langle\langle 1\rangle\rangle_{\mathrm{val}}\big(\mathrm{Parity}(p)\big)(s).$$

Otherwise, if $\mathrm{Pr}_s^{\sigma,\pi}(\mathrm{Reach}(W_1))>\langle\langle 1\rangle\rangle_{\mathrm{val}}(\mathrm{Parity}(p))(s)$, then player 1 plays σ to reach W_1 and an almost-sure winning strategy for $\mathrm{Parity}(p)$ from W_1 to ensure that the probability to satisfy $\mathrm{Parity}(p)$ given π is greater than $\langle\langle 1\rangle\rangle_{\mathrm{val}}(\mathrm{Parity}(p))(s)$; this contradicts that π is optimal. Now consider the Markov chain $G_{\sigma,\pi}$. Let C be a closed connected recurrent set of states in $G_{\sigma,\pi}$. If $C\cap(S\backslash(W_1\cup W_2))=\emptyset$, then there is a state $s'\in C\cap(S\backslash(W_1\cup W_2))$ with $\langle\langle 2\rangle\rangle_{\mathrm{val}}(\mathrm{coParity}(p))(s')>0$. Since π is optimal for player 2 for $\mathrm{coParity}(p)$ and in $G_{\sigma,\pi}$ from s' the set C is visited infinitely often with probability 1, it follows that $\max(p(C))$ is odd. Let $z\in C$ be a state with $p(z)=\max(p(C))$. Then since the minimum transition probability is $\delta_{\min}$ and $|C|\leqslant|S|$, it follows from Lemma 1 that the long-run frequency for state z is at least $\frac{1}{n}\cdot(\delta_{\min})^n$. The reward assignment ensures that the long-run average for the closed connected recurrent set C is at most -1. This is obtained as follows: the long-run average contribution by z is at most

$$\frac{1}{n}\cdot(\delta_{\min})^n\cdot(-1)\cdot(2\cdot n)^{p(z)}\cdot\left(\frac{1}{\delta_{\min}}\right)^{n\cdot p(z)} = (-2)\cdot\left((2\cdot n)^{p(z)-1}\cdot\left(\frac{1}{\delta_{\min}}\right)^{n\cdot(p(z)-1)}\right);$$

(this obtained by multiplying the long-run frequency of z along with its reward). Since $p(z)$ is the greatest priority appearing in C, the long-run average contribution of all the other states in C is at most

$$\left((2\cdot n)^{p(z)-1}\cdot\left(\frac{1}{\delta_{\min}}\right)^{n\cdot(p(z)-1)}\right),$$

(in the worst case all other states have priority $p(z)-1$). Hence the long-run average in C is at most

$$-\left((2\cdot n)^{p(z)-1}\cdot\left(\frac{1}{\delta_{\min}}\right)^{n\cdot(p(z)-1)}\right);$$

the claim follows. An upper bound on the long-run average payoff for player 1 is obtained as follows: we consider the maximum probability of reaching W_1 and consider the closed connected recurrent states C that intersect with W_1 is contained in W_1 (and the long-run average is 1 in this case) and with the rest of the probability the long-run average is at most -1. Hence we have

$$\begin{aligned}&\langle\langle 1\rangle\rangle_{\mathrm{val}}\big(\mathrm{MeanPay}(r)\big)(s)\\ &\quad\leqslant 1\cdot\langle\langle 1\rangle\rangle_{\mathrm{val}}\big(\mathrm{Parity}(p)\big)(s)\\ &\qquad+(-1)\cdot\big(1-\langle\langle 1\rangle\rangle_{\mathrm{val}}\big(\mathrm{Parity}(p)\big)(s)\big)\\ &\quad=2\cdot\langle\langle 1\rangle\rangle\big(\mathrm{Parity}(p)\big)(s)-1.\end{aligned}$$

The desired result follows. □

Remark. In the proof of Theorem 3 we used the existence of pure memoryless optimal strategies in $2\frac{1}{2}$-player game graphs with parity objectives, and the existence of pure memoryless optimal strategies in MDPs with mean-payoff objectives. The proof does not rely on the existence of pure memoryless optimal strategies in $2\frac{1}{2}$-player game graphs with mean-payoff objectives.

Polynomial-time complexity of the reduction. The reduction of $2\frac{1}{2}$-player games with parity objectives to $2\frac{1}{2}$-player games with mean-payoff objectives is achieved by Theorems 2 and 3. We argue that the reduction is polynomial. The size of a game graph $G = ((S, E), (S_1, S_2, S_P), \delta)$ is

$$|G| = |S| + |E| + \sum_{t \in S} \sum_{s \in S_P} \left|\delta(s)(t)\right|,$$

where $|\delta(s)(t)|$ denotes the space to represent the transition probability $\delta(s)(t)$ in binary. The reduction of Theorem 3 is polynomial, since the reward at every state can be expressed in $n \cdot d \cdot |G| \cdot \log(n)$ bits, and $d \leqslant n$. Hence from Theorems 2 and 3 we obtain a polynomial-time Turing reduction of $2\frac{1}{2}$-player parity games to $2\frac{1}{2}$-player mean-payoff games.

Acknowledgement

This research was supported in part by the AFOSR MURI grant F49620-00-1-0327 and the NSF ITR grant CCR-0225610.

References

[1] K. Chatterjee, Concurrent games with tail objectives, Technical Report UC Berkeley, EECS-05-1385, 2005. Shorter version appeared in: CSL'06, in: LNCS, vol. 4207, Springer, 2006, pp. 256–270.
[2] K. Chatterjee, M. Jurdziński, T.A. Henzinger, Simple stochastic parity games, in: CSL'03, in: LNCS, vol. 2803, Springer, 2003, pp. 100–113.
[3] K. Chatterjee, M. Jurdziński, T.A. Henzinger, Quantitative stochastic parity games, in: SODA'04, SIAM, 2004, pp. 121–130.
[4] A. Condon, The complexity of stochastic games, Information and Computation 96 (1992) 203–224.
[5] J. Filar, K. Vrieze, Competitive Markov Decision Processes, Springer, 1997.
[6] M. Jurdzinski, Deciding the winner in parity games is in UP $\cap$ co-UP, Information Processing Letters 68 (1998) 119–124.
[7] T.A. Liggett, S.A. Lippman, Stochastic games with perfect information and time average payoff, SIAM Review 11 (1969) 604–607.
[8] A. McIver, C. Morgan, Games, probability, and the quantitative μ-calculus $q\mu$, in: LPAR'02, in: LNCS, vol. 2514, Springer, 2002, pp. 292–310.
[9] A. Neyman, S. Sorin, Stochastic Games and Applications, Kluwer Academic Publishers, 2003.
[10] W. Thomas, Languages, automata, and logic, in: G. Rozenberg, A. Salomaa (Eds.), Beyond Words, in: Handbook of Formal Languages, vol. 3, Springer, 1997, pp. 389–455 (Chapter 7).
[11] W. Zielonka, Perfect-information stochastic parity games, in: FoSSaCS'04, in: LNCS, vol. 2987, Springer, 2004, pp. 499–513.
[12] U. Zwick, M. Paterson, The complexity of mean-payoff games on graphs, Theoretical Computer Science 158 (1996) 343–359.