Games and Markov Decision Processes with Mean-payoff Parity and Energy Parity Objectives ⋆

Krishnendu Chatterjee[1] and Laurent Doyen[2]

[1] IST Austria (Institute of Science and Technology Austria)
[2] LSV, ENS Cachan & CNRS, France

Abstract. In this paper we survey results of two-player games on graphs and Markov decision processes with parity, mean-payoff and energy objectives, and the combination of mean-payoff and energy objectives with parity objectives. These problems have applications in verification and synthesis of reactive systems in resource-constrained environments.

1 Introduction

Two-player games on graphs and Markov decision processes (MDPs) are standard models for reactive systems. In the analysis of reactive systems, the problem may specify a goal as a set of desired traces (such as ω-regular specifications), or as a quantitative optimization objective for a payoff function on the traces. One classical example of quantitative objective is the mean-payoff objective [15, 19]. Recently, the energy objectives (corresponding to total-payoff functions) have also been considered in the design of resource-constrained embedded systems such as power-limited systems [2, 5], as well as in queueing processes, and gambling models (see also [3] and references therein). The energy objective requires that the sum of the rewards be always nonnegative along a trace. Energy objectives can be expressed in the setting of boundaryless one-counter systems [3]. In recent times, games and MDPs equipped with the combination of a parity objective (which is a canonical way to express the ω-regular conditions [21]), and a quantitative objective specified as either mean-payoff or energy condition have been considered [9, 5, 6]. Such combination of quantitative and qualitative objectives is crucial in the design of reactive systems with both resource constraints and functional requirements [4, 9, 2, 1]. Also the energy parity objective can be viewed as a natural extension of boundaryless one-counter systems with fairness conditions.

In this paper we summarize the main results about MDPs and games with parity, mean-payoff, energy, mean-payoff parity, and energy parity objectives. We also present an improved algorithm to solve MDPs with mean-payoff parity objectives.

2 Definitions

Probability distributions. A *probability distribution* over a finite set A is a function $\kappa : A \to [0, 1]$ such that $\sum_{a \in A} \kappa(a) = 1$. The *support* of κ is the set $\mathsf{Supp}(\kappa) = \{a \in A \mid \kappa(a) > 0\}$. We denote by $\mathcal{D}(A)$ the set of probability distributions on A.

⋆ This work was partially supported by FWF NFN Grant S11407-N23 (RiSE) and a Microsoft faculty fellowship.

Markov Decision Processes. A *Markov Decision Process* (MDP) $M = (Q, E, \delta)$ consists of a finite set Q of states partitioned into *player*-1 *states* Q_1 and *probabilistic states* Q_P (i.e., $Q = Q_1 \cup Q_P$ and $Q_1 \cap Q_P = \emptyset$), a set $E \subseteq Q \times Q$ of edges such that for all $q \in Q$, there exists (at least one) $q' \in Q$ such that $(q, q') \in E$, and a probabilistic transition function $\delta : Q_P \to \mathcal{D}(Q)$ such that for all $q \in Q_P$ and $q' \in Q$, we have $(q, q') \in E$ iff $\delta(q)(q') > 0$. We often write $\delta(q, q')$ for $\delta(q)(q')$. For a state $q \in Q$, we denote by $E(q) = \{q' \in Q \mid (q, q') \in E\}$ the set of possible successors of q. A *Markov chain* is a special case of MDP where $Q_1 = \emptyset$.

Plays. An MDP can be viewed as the arena of a game played for infinitely many rounds from a state $q_0 \in Q$ as follows. If the game is in a player-1 state q, then player 1 chooses a successor state from the set $E(q)$; otherwise the game is in a probabilistic state q, and the successor is chosen according to the probability distribution $\delta(q)$. This game results in a *play* from q_0, i.e., an infinite path $\rho = q_0 q_1 \dots$ such that $(q_i, q_{i+1}) \in E$ for all $i \geq 0$. The prefix of length n of ρ is denoted by $\rho(n) = q_0 \dots q_n$, the last state of $\rho(n)$ is $\mathsf{Last}(\rho(n)) = q_n$. We write Ω for the set of all plays.

Strategies. A *strategy* (for player 1) is a function $\sigma : Q^* Q_1 \to Q$ such that for all $\rho \in Q^*$, $q \in Q_1$, and $q' \in Q_P$, if $\sigma(\rho \cdot q) = q'$, then $(q, q') \in E$. We denote by Σ the set of all strategies. An *outcome* of σ from q_0 is a play $q_0 q_1 \dots$ where $q_{i+1} = \sigma(q_0 \dots q_i)$ for all $i \geq 0$ such that $q_i \in Q_1$.

Outcomes and measures. Once a starting state $q \in Q$ and a strategy $\sigma \in \Sigma$ are fixed, the outcome of the game is a random walk ω_q^σ for which the probabilities of every *event* $\mathcal{A} \subseteq \Omega$, which is a measurable set of plays, are uniquely defined [22]. For a state $q \in Q$ and an event $\mathcal{A} \subseteq \Omega$, we denote by $\mathbb{P}_q^\sigma(\mathcal{A})$ the probability that a play belongs to $\mathcal{A}$ if the game starts from the state q and player 1 follows the strategy σ. For a measurable function $f : \Omega \to \mathbb{R}$ we denote by $\mathbb{E}_q^\sigma[f]$ the *expectation* of the function f under the probability measure $\mathbb{P}_q^\sigma(\cdot)$.

Finite-memory strategies. A strategy uses *finite-memory* if it can be encoded by a deterministic transducer $\langle \mathsf{Mem}, m_0, \alpha_u, \alpha_n \rangle$ where Mem is a finite set (the memory of the strategy), $m_0 \in \mathsf{Mem}$ is the initial memory value, $\alpha_u : \mathsf{Mem} \times Q \to \mathsf{Mem}$ is an update function, and $\alpha_n : \mathsf{Mem} \times Q_1 \to Q$ is a next-move function. The *size* of the strategy is the number $|\mathsf{Mem}|$ of memory values. If the current state is a player-1 state q, and m is the current memory value, then the strategy chooses the next state q' according to $\alpha_n(m, q)$. If the current state is q, then the memory is updated to $\alpha_u(m, q)$. Formally, $\langle \mathsf{Mem}, m_0, \alpha_u, \alpha_n \rangle$ defines the strategy σ such that $\sigma(\rho \cdot q) = \alpha_n(\hat{\alpha}_u(m_0, \rho), q)$ for all $\rho \in Q^*$ and $q \in Q_1$, where $\hat{\alpha}_u$ extends α_u to sequences of states as expected. A strategy is *memoryless* if $|\mathsf{Mem}| = 1$, and a memoryless strategy $\sigma : Q_1 \to Q$ chooses one edge for every player 1 state. For a finite-memory strategy σ, M_σ denote the Markov chain obtained as the product of M with the transducer defining σ.

Two-player games. A *two-player game* is a graph $G = (Q, E)$ with the same assumptions as for MDP, except that the partition of Q is denoted (Q_1, Q_2) where Q_2 is the set of *player*-2 *states*. Player 2 plays the role of an adversary to player 1. The notions of play, strategies (in particular strategies for player 2), and outcome are analogous to the case of MDP (see [21] for standard definition of games).

Objectives. An *objective* for an MDP M (or game G) is a set $\phi \subseteq \Omega$ of infinite paths. Let $p : Q \to \mathbb{N}$ be a *priority function* and $w : E \to \mathbb{Z}$ be a *weight function*[3] where positive numbers represent rewards. We denote by W the largest weight (in absolute value) according to w. The *energy level* of a prefix $\gamma = q_0 q_1 \dots q_n$ of a play is $\mathsf{EL}(w, \gamma) = \sum_{i=0}^{n-1} w(q_i, q_{i+1})$, and the *mean-payoff value*[4] of a play $\rho = q_0 q_1 \dots$ is $\mathsf{MP}(w, \rho) = \liminf_{n\to\infty} \frac{1}{n} \cdot \mathsf{EL}(w, \rho(n))$. In the sequel, when the weight function w is clear from the context we omit it and simply write $\mathsf{EL}(\gamma)$ and $\mathsf{MP}(\rho)$. We denote by $\mathsf{Inf}(\rho)$ the set of states that occur infinitely often in ρ, and we consider the following objectives:

- *Parity objectives.* The *parity* objective $\mathsf{Parity}(p) = \{\rho \in \Omega \mid \min\{p(q) \mid q \in \mathsf{Inf}(\rho)\}$ is even $\}$ requires that the minimum priority visited infinitely often be even. The special cases of *Büchi* and *coBüchi* objectives correspond to the case with two priorities, $p : Q \to \{0, 1\}$ and $p : Q \to \{1, 2\}$ respectively.
- *Energy objectives.* Given an initial credit $c_0 \in \mathbb{N}$, the *energy* objective $\mathsf{PosEnergy}(c_0) = \{\rho \in \Omega \mid \forall n \geq 0 : c_0 + \mathsf{EL}(\rho(n)) \geq 0\}$ requires that the energy level be always positive.
- *Mean-payoff objectives.* Given a threshold $\nu \in \mathbb{Q}$, the *mean-payoff* objective $\mathsf{MeanPayoff}^{\geq\nu} = \{\rho \in \Omega \mid \mathsf{MP}(\rho) \geq \nu\}$ (resp. $\mathsf{MeanPayoff}^{>\nu} = \{\rho \in \Omega \mid \mathsf{MP}(\rho) > \nu\}$) requires that the mean-payoff value be at least ν (resp. strictly greater than ν).
- *Combined objectives.* The *energy parity* objective $\mathsf{Parity}(p) \cap \mathsf{PosEnergy}(c_0)$ and the *mean-payoff parity* objective $\mathsf{Parity}(p) \cap \mathsf{MeanPayoff}^{\sim\nu}$ (for $\sim \in \{\geq, >\}$) combine the requirements of parity and energy (resp., mean-payoff) objectives.

Almost-sure winning strategies. For MDPs, we say that a player-1 strategy σ is *almost-sure winning* in a state q for an objective ϕ if $\mathbb{P}_q^\sigma(\phi) = 1$. For two-player games, we say that a player-1 strategy σ is *winning* in a state q for an objective ϕ if all outcomes of σ starting in q belong to ϕ. For energy objectives with unspecified initial credit, we also say that a strategy is (almost-sure) winning if it is (almost-sure) winning for *some* finite initial credit.

Decision problems. We are interested in the following problems. Given an MDP M with weight function w and priority function p, and a state q_0:

- The *parity problem* asks whether there exists an almost-sure winning strategy for the parity objective from q_0.
- The *mean-payoff problem (resp. mean-payoff parity problem)* asks whether there exists an almost-sure winning strategy for the mean-payoff objective (resp. mean-payoff parity objective) with threshold 0 from q_0. Note that it is not restrictive to consider mean-payoff objectives with threshold 0 because for $\sim \in \{\geq, >\}$, we have $\mathsf{MP}(w, \rho) \sim \nu$ iff $\mathsf{MP}(w - \nu, \rho) \sim 0$, where $w - \nu$ is the weight function that assigns $w(e) - \nu$ to each edge $e \in E$.

[3] Sometimes we take the freedom to use rational weights (i.e., $w : E \to \mathbb{Q}$), while we always assume that weights are integers encoded in binary for complexity results.

[4] The results of this paper hold for the definition of mean-payoff value using $\limsup$ instead of $\liminf$.

– The *energy problem (resp. energy parity problem)* asks whether there exists a finite initial credit $c_0 \in \mathbb{N}$ and an almost-sure winning strategy for the energy objective (resp. energy parity objective) from q_0 with initial credit c_0.

The two-player game versions of these problems are defined analogously, and in two player games we are interested in the existence of winning strategies as compared to almost-sure winning strategies.

3 Games

In this section we summarize the results of games with all the objectives defined in the previous section. We first present the results for parity, mean-payoff and energy objectives, and then present the results for combined objectives.

3.1 Games with parity, mean-payoff, and energy objectives

In games with parity, mean-payoff, and energy objectives, memoryless winning strategies exist for both players. The result for parity objectives was established in [14] (also see [23]); for mean-payoff objectives it was shown in [18] (also see [13]); and for energy objectives it follows from [4]. The existence of memoryless optimal strategies, and the fact that the one-player version of the problems can be solved in polynomial time shows that the problems can be decided in NP $\cap$ coNP. For polynomial time algorithms for the one-player games, see [17] for parity objectives, [16] for mean-payoff objectives, and [4] for energy objectives. A major open question in this area of research is whether games with parity, mean-payoff, and energy objective can be solved in polynomial time. There is a close connection between mean-payoff and energy games, and the equivalence of the decision problems has been established in [2]. The main argument is as follows: by existence of memoryless strategies it follows that if the answer to the mean-payoff objectives with threshold $\nu = 0$ is true, then player 1 can fix a memoryless strategies such that in all cycles the sum of the rewards is non-negative, and this exactly coincides with the finite initial credit problem (where after a prefix, the sum of the rewards in cycles is non-negative). The results are summarized in the following theorem.

Theorem 1. *Memoryless winning strategies exist for both players in two-player games with parity, mean-payoff, and energy objectives, and the decision problems lie in NP $\cap$ coNP.*

3.2 Games with mean-payoff parity and energy parity objectives

Mean-payoff parity games were first studied in [9] and it was shown that winning strategies for player 1 require infinite memory in general. It follows from the results of [9] that memoryless winning strategies exist for player 2. In [5] games with energy parity objectives were studied and the following results were established: (1) winning strategies for player 1 with energy parity objectives require memory at least $2 \cdot (|Q| - 1) \cdot W + 1$ and memory of size $4 \cdot |Q| \cdot d \cdot W$ is sufficient; (2) memoryless winning strategies exist for player 2; and (3) the decision problem lie in NP $\cap$ coNP. It was also shown in [5] that the decision problem for mean-payoff parity objectives can be reduced to the decision problem for energy parity objectives, however, as winning strategies require infinite memory for mean-payoff parity objectives, the reduction is more involved than the reduction from mean-payoff to energy objectives. As a consequence it also follows that the decision problem for mean-payoff parity objectives lie in NP $\cap$ coNP.

Theorem 2. *In games with mean-payoff parity objectives, winning strategies for player 1 (with mean-payoff parity objective) require infinite memory in general, and memoryless winning strategies exist for player 2. In games with energy parity objectives, winning strategies for player 1 (with energy parity objective) require* $2 \cdot (|Q| - 1) \cdot W + 1$ *memory in general and memory of size* $4 \cdot |Q| \cdot d \cdot |W|$ *is sufficient, and memoryless winning strategies exist for player 2. The problems of deciding the winner in games with mean-payoff parity and games with energy parity objectives lie in NP* $\cap$ *coNP.*

4 Markov decision process

In this section we summarize the results for Markov decision processes (MDPs).

4.1 MDPs with parity, mean-payoff and energy objectives

As in the case of games, in MDPs memoryless almost-sure winning strategies exist for parity, mean-payoff and energy objectives. The result of existence of memoryless almost-sure winning strategies for MDPs with parity objectives follows from [12] (also see [10] for explicit proofs). The result of existence of memoryless almost-sure winning strategies for mean-payoff objectives follows from the result of [18] (also see [15, 19] for details of MDPs with mean-payoff objectives). Almost-sure winning in MDPs with energy objective is equivalent to a two-player energy game (where the probabilistic states are controlled by player 2). Indeed (1) a winning strategy in the game is trivially almost-sure winning in the MDP, and (2) if an almost-sure winning strategy σ in the MDP was not winning in the game, then for all initial credit c_0 there would exist an outcome ρ of σ such that $c_0 + \mathsf{EL}(\rho(i)) < 0$ for some position $i \geq 0$. The prefix $\rho(i)$ has a positive probability in the MDP, in contradiction with the fact that σ is almost-sure winning. As a consequence, it follows that memoryless almost-sure winning strategies exist for MDPs with energy objectives, and it also follows that solving MDPs with energy objectives is as hard as games with energy as well as games with mean-payoff objective. The results of [10] give a polynomial-time algorithm to solve almost-sure winning for MDPs with parity objective. For MDPs with mean-payoff objective, the almost-sure winning problem can be solved through linear programming (see [15, 19] for a linear program formulation to solve MDPs with mean-payoff objective). The problem of MDPs with energy objective lies in NP $\cap$ coNP due to the equivalence with games with energy objective.

Theorem 3. *Memoryless almost-sure winning strategies exist in MDPs with parity, mean-payoff and energy objectives. The decision problems of existence of almost-sure* *winning strategy can be solved in polynomial time for mean-payoff and parity objectives, and for energy objectives it belongs to NP $\cap$ coNP.*

4.2 MDPs with mean-payoff parity and energy parity objectives

MDPs with mean-payoff parity and energy parity objectives were considered in [6], and the following results were established. For mean-payoff parity objectives, almost-sure winning strategies require infinite memory in general, and the set of almost-sure winning states can be computed in polynomial time. For energy parity objectives, a polynomial time reduction was presented to games with energy Büchi objectives (parity objectives with two priorities), and it follows that the decision problem belongs to NP $\cap$ coNP.

Theorem 4. *In MDPs with mean-payoff parity objectives, almost-sure winning strategies require infinite memory in general. In MDPs with energy parity objectives, almost-sure winning strategies require $2 \cdot (|Q| - 1) \cdot W + 1$ memory in general and memory of size $2 \cdot |Q| \cdot |W|$ is sufficient. The set of almost-sure winning states can be computed in polynomial time for mean-payoff parity objectives. The decision problem of whether a state is almost-sure winning lies in NP $\cap$ coNP for energy parity objectives.*

4.3 New result: Improved algorithm for mean-payoff parity objectives

In this section we present an improved polynomial-time algorithm for the computation of the set of almost-sure winning states for mean-payoff parity objectives, using the results of [6] and hierarchical graph decomposition technique. We first define the basic notion of end-component for MDPs.

End-components. A set $U \subseteq Q$ is *δ-closed* if for all $q \in U \cap Q_P$ we have $\mathsf{Supp}(\delta(q)) \subseteq U$. The sub-MDP induced by a δ-closed set U is $M \upharpoonright U = (U, E \cap (U \times U), \delta)$. Note that $M \upharpoonright U$ is an MDP if for all $q \in U$ there exists $q' \in U$ such that $(q, q') \in E$. A *closed recurrent set* for a Markov chain is a δ-closed set $U \subseteq Q$ which is strongly connected. End-components in MDPs play a role equivalent to closed recurrent sets in Markov chains. Given an MDP $M = (Q, E, \delta)$ with partition (Q_1, Q_P), a set $U \subseteq Q$ of states is an *end-component* if U is δ-closed and the sub-MDP $M \upharpoonright U$ is strongly connected [11, 12]. We denote by $\mathcal{E}(M)$ the set of end-components of an MDP M.

End-component lemma. We now present an important lemma about end-components from [11, 12] that we use in the proofs of our result. It states that for arbitrary strategies (memoryless or not), with probability 1 the set of states visited infinitely often along a play is an end-component. This lemma allows us to derive conclusions on the (infinite) set of plays in an MDP by analyzing the (finite) set of end-components in the MDP.

Lemma 1 ([11, 12]). *Given an MDP M, for all states $q \in Q$ and all strategies $\sigma \in \Sigma$, we have $\mathbb{P}_q^\sigma(\{\omega \mid \mathsf{Inf}(\omega) \in \mathcal{E}(M)\}) = 1$.*

We now present the key lemma from [6] where it was shown that for an MDP that is an end-component such that the minimum priority is even, the mean-payoff parity objective $\mathsf{Parity}(p) \cap \mathsf{MeanPayoff}^{\geq \nu}$ is satisfied with probability 1 if the expected mean-payoff value is at least ν at some state (the result also holds for strict inequality). In other words, from the expected mean-payoff value of at least ν we ensure that both the mean-payoff and parity objective is satisfied with probability 1 from all states. For a state q, let $\mathsf{ValMP}(w)(q) = \sup_{\sigma \in \Sigma} \mathbb{E}_q^\sigma[\mathsf{MP}(w)]$ denote the expected mean-payoff value. The following lemma was established in [6].

Lemma 2 ([6]). *Consider an MDP M with state space Q, a priority function p, and weight function w such that (a) M is an end-component (i.e., Q is an end-component) and (b) the smallest priority in Q is even. If there is a state $q \in Q$ such that $\mathsf{ValMP}(w)(q) \geq \nu$ (resp. $\mathsf{ValMP}(w)(q) > \nu$), then there exists a strategy σ^* such that for all states $q \in Q$ we have $\mathbb{P}_q^{\sigma^*}(\mathsf{Parity}(p) \cap \mathsf{MeanPayoff}^{\geq \nu}) = 1$ (resp. $\mathbb{P}_q^{\sigma^*}(\mathsf{Parity}(p) \cap \mathsf{MeanPayoff}^{> \nu}) = 1$).*

Winning end-component. Given an MDP M with a parity objective $\mathsf{Parity}(p)$ and a mean-payoff objective $\mathsf{MeanPayoff}^{\geq \nu}$ for a weight function w, we call an end-component U *winning* if (a) $\min(p(U))$ is even; and (b) there exists a state with expected mean-payoff value at least ν in the sub-MDP induced by U, i.e., $\max_{q \in U} \mathsf{ValMP}(w)(q) \geq \nu$ in the sub-MDP induced by U. We denote by $\mathcal{W}$ the set of winning end-components, and let $\mathsf{Win} = \bigcup_{U \in \mathcal{W}} U$ be the union of the winning end-components.

Reduction to reachability of winning end-component. By Lemma 2 it follows that in every winning end-component the mean-payoff parity objective is satisfied with probability 1. Conversely, consider an end-component U that is not winning, then either the smallest priority is odd, or the maximal expected mean-payoff value that can be ensured for any state in U by staying in U is less than ν. Hence if only states in U are visited infinitely often, then with probability 1 (i) either the parity objective is not satisfied, or (ii) the mean-payoff objective is not satisfied. In other words, if an end-component that is not winning is visited infinitely often, then the mean-payoff parity objective is satisfied with probability 0. It follows that the almost-sure winning states can be computed by computing the set of almost-sure winning states for reachability objectives with the winning end-components as the target set (i.e., computing almost-sure reachability to the set Win). Since almost-sure winning states for MDPs with reachability objectives can be computed in polynomial time [10], it suffices to present a polynomial-time algorithm to compute Win in order to obtain a polynomial-time algorithm for MDPs with mean-payoff parity objectives.

Computing winning end-components. The computation of the winning end-components is done iteratively by computing winning end-components with smallest priority 0, then winning end-components with smallest priority 2, and so on. The computation of Win is as follows:

- For $i \geq 0$, let $\mathcal{W}_{2i}$ be the set of maximal end-components U with states with priority at least $2i$ and that contain at least one state with priority $2i$, i.e., U contains only states with priority at least $2i$, and contains at least one state with priority $2i$. Let $\mathcal{W}'_{2i} \subseteq \mathcal{W}_{2i}$ be the set of maximal end-components $U \in \mathcal{W}_{2i}$ such that there is a state $q \in U$ such that the expected mean-payoff value in the sub-MDP restricted to U is at least ν. Let $\mathsf{Win}_{2i} = \bigcup_{U \in \mathcal{W}'_{2i}} U$.

The set $\mathsf{Win} = \bigcup_{i=0}^{\lfloor d/2 \rfloor} \mathsf{Win}_{2i}$ is the union of the states of the winning end-components (formal pseudo-code in [7]).

Complexity of computing winning end-components. The winning end-component algorithm runs for $O(d)$ iterations and in each iteration requires to compute a maximal end-component decomposition (mec) and compute mean-payoff values of at most n end-components, where n is the number of states of the MDP. We now improve the $O(d)$ iterations to $O(\log d)$ using the hierarchical clustering technique of Tarjan [20].

Given a priority function $p : Q \to \{0, 1, \ldots, 2d\}$, for $0 \leq m \leq d$, let $Q_{\leq m} = \{q \in Q \mid p(q) \leq m\}$ denote the set of states with priority at most m. Given an MDP M, let M_i denote the MDP obtained by removing $Attr_R(Q_{\leq 2i-1})$ the set of states with priority less than $2i$ and its random attractor (random attractor denotes the probabilistic alternating reachability, for formal definition of random attractor see [8]). A mec C is a *winning mec* in M_i if there exists $u \in C$ such that $p(u) = 2i$ and there is some state $v \in C$ such that the expected mean-payoff value of v is greater than the given threshold ν. Let WE_i be the union of the vertices of winning mec in M_i, and let $\mathsf{WE} = \cup_{0 \leq i \leq d} \mathsf{WE}_i$.

Informal description of the new algorithm. If two states u, v belong to the same mec in M_i, they also belong to the same mec in M_{i-1}. Thus the mec's of M_i refine the ones of M_{i-1}, which can be exploited using the hierarchical clustering technique. Formally, we will compute WE by the recursive procedure $\textsc{WinMaxEC}(M, p, i, j)$. The procedure takes an MDP, and two indices i and j, and outputs $\bigcup_{i \leq 2k \leq j} \mathsf{WE}_{2k}$. To obtain WE we invoke $\textsc{WinMaxEC}(M, p, 0, 2d)$. Given the MDP M, and indices i, j, the procedure first computes the mec's of M_m, where $m = \lceil \frac{i+j}{2} \rceil$. If m is even, then the set WE_m of M_m is computed. Then we recursively call the procedures $\textsc{WinMaxEC}(M_u, p, m+1, j)$ and $\textsc{WinMaxEC}(M_\ell, p, i, m-1)$, where M_u is a sub-MDP containing only the edges *inside* the mec's of M_m and the MDP M_ℓ is obtained by collapsing each mec in M_m to a single vertex, thus containing only edges *outside* the mec's of M_m. The formal description of the algorithm is identical to the algorithm of Section 5 of [8] for almost-sure winning of MDPs with parity objectives, and the only change is while determining whether an end-component is winning along with the priority being even, we also check whether there is some state where the expected mean-payoff value is at least the given threshold. The correctness argument is essentially similar to the correctness of [8], and shows that $\mathsf{Win} = \mathsf{WE}$.

Running time analysis. Given a MDP M with n states, m edges and a parity objective with d priorities, let us denote by $T(m, n, d)$ the running time of $\textsc{WinMaxEC}$ on M. We observe that in E_u consists of edges with in mec's, and such edges are not present in E_ℓ. Thus we obtain the following recurrence relation for the running time $T(m, n, d)$ of $\textsc{WinMaxEC}$:

$$T(m, n, d) = T_M(m, n) + T(m_u, n, \lfloor \frac{d-1}{2} \rfloor) + T(m_\ell, n, \lceil \frac{d-1}{2} \rceil),$$

with $m_\ell + m_u \leq m$, and $T_M(m, n)$ denotes the time complexity of mec decomposition and MDPs with mean-payoff solving with m edges and n states. It is straightforward to show that $T(m, n, d) = O((\textsc{MaxEC}(n, m) + \textsc{MeanPayoffSolve}(n, m)) \cdot \log(d))$, where MAXEC and MEANPAYOFFSOLVE denote algorithms for computing the maximal end-component decomposition of an MDP and solving MDPs with mean-payoff objectives, respectively. Thus we obtain an improved algorithm to solve MDPs with mean-payoff parity objectives.

5 Conclusion

In this paper we summarized the main results of games and MDPs with parity, mean-payoff, energy, mean-payoff parity, and energy parity objectives. The major open questions are whether games with parity, mean-payoff, and energy objective can be solved in polynomial time.

References

1. R. Bloem, K. Chatterjee, T. A. Henzinger, and B. Jobstmann. Better quality in synthesis through quantitative objectives. In *CAV*, LNCS 5643, pages 140–156. Springer, 2009.
2. P. Bouyer, U. Fahrenberg, K. G. Larsen, N. Markey, and J. Srba. Infinite runs in weighted timed automata with energy constraints. In *Proc. of FORMATS*, LNCS 5215, pages 33–47. Springer, 2008.
3. Tomás Brázdil, Václav Brozek, Kousha Etessami, Antonín Kucera, and Dominik Wojtczak. One-counter Markov decision processes. In *Proc. of SODA*, pages 863–874. SIAM, 2010.
4. A. Chakrabarti, L. de Alfaro, T. A. Henzinger, and M. Stoelinga. Resource interfaces. In *Proc. of EMSOFT*, LNCS 2855, pages 117–133. Springer, 2003.
5. K. Chatterjee and L. Doyen. Energy parity games. In *Proc. of ICALP: Automata, Languages and Programming (Part II)*, LNCS 6199, pages 599–610. Springer, 2010.
6. K. Chatterjee and L. Doyen. Energy and mean-payoff parity Markov decision processes. In *MFCS*, pages 206–218, 2011.
7. K. Chatterjee and L. Doyen. Energy and mean-payoff parity Markov decision processes. Technical report, IST Austria, Feb, 2011. `http://pub.ist.ac.at/Pubs/TechRpts/2011/IST-2011-0001.pdf`.
8. K. Chatterjee and M. Henzinger. Faster and dynamic algorithms for maximal end-component decomposition and related graph problems in probabilistic verification. In *Proc. of SODA*. ACM SIAM, 2011.
9. K. Chatterjee, T. A. Henzinger, and M. Jurdziński. Mean-payoff parity games. In *Proc. of LICS*, pages 178–187. IEEE Computer Society, 2005.
10. K. Chatterjee, M. Jurdziński, and T. A. Henzinger. Quantitative stochastic parity games. In *Proc. of SODA: Symposium on Discrete Algorithms*, pages 114–123, 2004. Technical Report: UCB/CSD-3-1280 (October 2003).
11. C. Courcoubetis and M. Yannakakis. The complexity of probabilistic verification. *J. ACM*, 42(4):857–907, 1995.
12. L. de Alfaro. *Formal Verification of Probabilistic Systems*. PhD thesis, Stanford University, 1997.
13. A. Ehrenfeucht and J. Mycielski. Positional strategies for mean payoff games. *Int. Journal of Game Theory*, 8(2):109–113, 1979.
14. E.A. Emerson and C. Jutla. The complexity of tree automata and logics of programs. In *FOCS*, pages 328–337. IEEE Computer Society Press, 1988.

15. J. Filar and K. Vrieze. *Competitive Markov Decision Processes*. Springer, 1997.
16. R. M. Karp. A characterization of the minimum cycle mean in a digraph. *Discrete Mathematics*, 23(3):309–311, 1978.
17. V. King, O. Kupferman, and M. Y. Vardi. On the complexity of parity word automata. In *FoSSaCS*, pages 276–286. Springer, 2001.
18. T. A. Liggett and S. A. Lippman. Stochastic games with perfect information and time average payoff. *Siam Review*, 11:604–607, 1969.
19. M.L. Puterman. *Markov Decision Processes*. John Wiley and Sons, 1994.
20. R. E. Tarjan. A hierarchical clustering algorithm using strong components. *Inf. Process. Lett.*, 14(1):26–29, 1982.
21. W. Thomas. Languages, automata, and logic. In *Handbook of Formal Languages*, volume 3, Beyond Words, chapter 7, pages 389–455. Springer, 1997.
22. M.Y. Vardi. Automatic verification of probabilistic concurrent finite-state systems. In *FOCS'85*. IEEE Computer Society Press, 1985.
23. W. Zielonka. Infinite games on finitely coloured graphs with applications to automata on infinite trees. *Theor. Comput. Sci.*, 200:135–183, 1998.