Mean-Payoff Games with Partial-Observation*
(Extended Abstract)

Paul Hunter, Guillermo A. Pérez, and Jean-François Raskin

Département d'Informatique, Université Libre de Bruxelles (U.L.B.), Belgium
{paul.hunter,guillermo.perez,jraskin}@ulb.ac.be

Abstract. Mean-payoff games are important quantitative models for open reactive systems. They have been widely studied as games of perfect information. In this paper we investigate the algorithmic properties of several subclasses of mean-payoff games where the players have asymmetric information about the state of the game. These games are in general undecidable and not determined according to the classical definition. We show that such games are determined under a more general notion of winning strategy. We also consider mean-payoff games where the winner can be determined by the winner of a finite cycle-forming game. This yields several decidable classes of mean-payoff games of asymmetric information that require only finite-memory strategies, including a generalization of perfect information games where positional strategies are sufficient. We give an exponential time algorithm for determining the winner of the latter.

1 Introduction

Mean-payoff games (MPGs) are two-player, infinite duration, turn-based games played on finite edge-weighted graphs. The two players alternately move a token around the graph; and one of the players (Eve) tries to maximize the (limit) average weight of the edges traversed, whilst the other player (Adam) attempts to minimize the average weight. Such games are particularly useful in the field of verification of models of reactive systems, and the perfect information versions of these games have been extensively studied [4, 7, 8, 10]. One of the major open questions in the field of verification is whether the following decision problem, known to be in the intersection of the classes NP and coNP [10][1], can be solved in polynomial time: Given a threshold ν, does Eve have a strategy to ensure a mean-payoff value of at least ν?

In game theory the concepts of imperfect, partial and limited information indicate situations where players have asymmetric knowledge about the state of the game. In the context of verification games this partial knowledge is reflected in one or both players being unable to determine the precise location of the token amongst several equivalent vertices, and such games have also been extensively

* This work was supported by the ERC inVEST (279499) project.
[1] From results in [17] and [12] it follows that the problem is also in $\mathsf{UP} \cap \mathsf{coUP}$.

J. Ouaknine, I. Potapov, and J. Worrell (Eds.): RP 2014, LNCS 8762, pp. 163–175, 2014.

studied [2, 3, 9, 13, 16]. Adding partial-observation to verification games results in an enormous increase in complexity, both algorithmically and in terms of strategy synthesis. For example, it was shown in [9] that for MPGs with partial-observation, when the mean payoff value is defined using lim sup, the analogue of the above decision problem is undecidable; and whilst memoryless strategies suffice for MPGs with perfect information, infinite memory may be required. The first result of this paper is to show that this is also the case when the mean payoff value is defined using the stronger lim inf operator, closing two open questions posed in [9]. As a consequence, we generalize a result from [6] which uses the undecidability result from [9] to show several classical problems for mean-payoff automata are also undecidable.

To simplify our definitions and algorithmic results we initially consider a restriction on the set of observations which we term *limited-observation*. In games of limited-observation the current observation contains only those vertices consistent with the observable history, that is the observations are the *belief set of Eve* (see, e.g. [5]). This is not too restrictive as any MPG with partial-observation can be realized as a game of limited-observation via a subset construction. In Section 9 we consider the extension of our definitions to MPGs with partial-observation via this construction.

Our focus for the paper will be on games at the observation level, in particular we are interested in *observation-based strategies* for both players. Whilst observation-based strategies for Eve are usual in the literature, observation-based strategies for Adam have not, to the best of our knowledge, been considered. Such strategies are more advantageous for Adam as they encompass several simultaneous concrete strategies. Further, in games of limited-observation there is guaranteed to be at least one concrete strategy consistent with an observation-based strategy. Our second result is to show that although MPGs with partial-observation are not determined under the usual definition of strategy, they are determined when Adam can use an observation-based strategy.

In games of perfect information one aspect of MPGs that leads to simple (but not quite efficient) decision procedures is their equivalence to finite cycle-forming games. Such games are played as their infinite counterparts, however when the token revisits a vertex the game is stopped. The winner is determined by a finite analogue of the mean-payoff condition on the cycle now formed. Ehrenfeucht and Mycielski [10] and Björklund et al. [4][2] used this equivalence to show that positional strategies are sufficient to win MPGs with perfect information. Critically, a winning strategy in the finite game translates directly to a winning strategy in the MPG, so such games are especially useful for strategy synthesis.

We extend this idea to games of partial-observation by introducing a finite, perfect information, cycle-forming game played at the observation level. That is, the game finishes when an observation is revisited (though not necessarily the first time). In this reachability game winning strategies can be translated to finite-memory winning strategies in the MPG. This leads to a large, natural subclass of MPGs with partial-observation, *forcibly terminating* games, where

[2] A recent result of Aminof and Rubin [1] corrects some errors in [4].

Table 1. Summary of results for the classes of games studied.

	Forcibly Terminating	Forcibly FAC		FAC	
		limited-obs.	partial-obs.	limited-obs.	partial-obs.
Memory	Finite	Exponential	2-Exponential	Positional	Exponential
Class membership	Undecidable	PSPACE-complete	NEXP-hard, in EXPSPACE	coNP-complete	coNEXP-complete
Winner determination	R-complete	PSPACE-complete	EXP-complete	NP ∩ coNP	EXP-complete

determining the winner is decidable and finite memory observation-based strategies suffice.

Unfortunately, recognizing if an MPG is a member of this class is undecidable, and although determining the winner is decidable, we show that this problem is complete (under polynomial-time reductions) for the class of all decidable problems. Motivated by these negative algorithmic results, we investigate two natural refinements of this class for which winner determination and class membership are decidable. The first, *forcibly first abstract cycle* games (forcibly FAC games, for short), is the natural class of games obtained when our cycle-forming game is restricted to simple cycles. Unlike the perfect information case, we show that winning strategies in this finite simple cycle-forming game may still require memory, though this memory is at most exponential in the size of the game. The second refinement, *first abstract cycle* (FAC) games, is a further structural refinement that guarantees a winner in the simple cycle-forming game. We show that in this class of games positional observation-based strategies suffice.

Table 1 summarizes the results of this paper. For space reasons the full details of all proofs can be found in the technical report [11].

2 Preliminaries

Mean-payoff games. A *mean-payoff game (MPG) with partial-observation* is a tuple $G = \langle Q, q_I, \Sigma, \Delta, w, Obs \rangle$, where Q is a finite set of states, $q_I \in Q$ is the initial state, Σ is a finite set of actions, $\Delta \subseteq Q \times \Sigma \times Q$ is the transition relation, $w : \Delta \to \mathbb{Z}$ is the weight function, and $Obs \in \texttt{Partition}(Q)$ is a set of observations. We assume Δ is total. We say that G is a *mean-payoff game with limited-observation* if additionally, (1) $\{q_I\} \in Obs$, and (2) for each $(o, \sigma) \in Obs \times \Sigma$ the set $\{q' \mid \exists q \in o \text{ and } (q, \sigma, q') \in \Delta\}$ is a union of elements of Obs. If every element of Obs is a singleton, then we say G is a *mean-payoff game with perfect information.* For simplicity, we denote by $\texttt{post}_\sigma(s) = \{q' \in Q \mid \exists q \in s : (q, \sigma, q') \in \Delta\}$ the set of σ-successors of a set of states $s \subseteq Q$.

In this work, unless explicitly stated otherwise, we depict states from an MPG with partial-observation as circles and transitions as arrows labelled by an action-weight pair: σ, w. Observations are represented by dashed boxes.

Abstract and concrete paths. A *concrete path* in an MPG with partial-observation is a sequence $q_0\sigma_0q_1\sigma_1\ldots$ where for all $i \geq 0$ we have $q_i \in Q$, $\sigma_i \in \Sigma$ and $(q_i, \sigma_i, q_{i+1}) \in \Delta$. An *abstract path* is a sequence $o_0\sigma_0o_1\sigma_1\ldots$ where $o_i \in Obs$, $\sigma_i \in \Sigma$ and for all i there exists $q_i \in o_i$ and $q_{i+1} \in o_{i+1}$ with $(q_i, \sigma_i, q_{i+1}) \in \Delta$. Given an abstract path ψ, let $\gamma(\psi)$ be the (possibly empty) set of concrete paths that agree with the observation and action sequence. Note that in games of limited-observation this set is never empty. Also, given abstract (respectively concrete) path ρ, let $\rho[..n]$ represent the prefix of ρ up to the $(n+1)$-th observation (state), which we express as $\rho[n]$; similarly, we denote by $\rho[l..]$ the suffix of ρ starting from the l-th observation (state) and by $\rho[l..n]$ the finite subsequence starting and ending in the aforementioned locations. An *abstract (respectively concrete) cycle* is an abstract (concrete) path $\chi = o_0\sigma_0\ldots o_n$ where $o_0 = o_n$. We say χ is *simple* if $o_j \neq o_i$ for $0 \leq i < j < n$. Given $k \in \mathbb{N}$ define χ^k to be the abstract (concrete) cycle obtained by traversing χ k times. A *cyclic permutation* of χ is an abstract (concrete) cycle $o'_0\sigma'_0\ldots o'_n$ such that $o'_j = o_{j+k \pmod n}$ and $\sigma'_j = \sigma_{j+k \pmod n}$ for some k. If $\chi' = o'_0\sigma'_0\ldots o'_m$ is a cycle with $o'_0 = o_i$ for some i, the *interleaving* of χ and χ' is the cycle $o_0\sigma_0\ldots o_i\sigma'_0\ldots o'_m\sigma_i\ldots o_n$.

Given a concrete path $\pi = q_0\sigma_0q_1\sigma_1\ldots$, the *payoff* up to the $(n+1)$-th element is given by

$$w(\pi[..n]) = \sum_{i=0}^{n-1} w(q_i, \sigma_i, q_{i+1}).$$

If π is infinite, we define two *mean-payoff* values $\underline{MP}$ and $\overline{MP}$ as:

$$\underline{MP}(\pi) = \liminf_{n\to\infty} \frac{1}{n} w(\pi[..n]) \qquad \overline{MP}(\pi) = \limsup_{n\to\infty} \frac{1}{n} w(\pi[..n])$$

Plays and strategies. A play in an MPG with partial-observation G is an infinite abstract path starting at $o_I \in Obs$ where $q_I \in o_I$. Denote by $\mathtt{Plays}(G)$ the set of all plays and by $\mathtt{Prefs}(G)$ the set of all finite prefixes of such plays ending in an observation. Let $\gamma(\mathtt{Plays}(G))$ be the set of concrete paths of all plays in the game, and $\gamma(\mathtt{Prefs}(G))$ be the set of all finite prefixes of all concrete paths.

An *observation-based strategy for Eve* is a function from finite prefixes of plays to actions, i.e. $\lambda_\exists : \mathtt{Prefs}(G) \to \Sigma$. A play $\psi = o_0\sigma_0o_1\sigma_1\ldots$ is *consistent* with $\lambda_\exists$ if $\sigma_i = \lambda_\exists(\psi[..i])$ for all i. An *observation-based strategy for Adam* is a function $\lambda_\forall : \mathtt{Prefs}(G) \times \Sigma \to Obs$ such that for any prefix $\pi = o_0\sigma_0\ldots o_n \in \mathtt{Prefs}(G)$ and action σ, $\lambda_\forall(\pi, \sigma) \cap \mathtt{post}_\sigma(\pi[n]) \neq \emptyset$. A play $\psi = o_0\sigma_0o_1\sigma_1\ldots$ is consistent with $\lambda_\forall$ if $\lambda_\forall(\psi[..i], \sigma_i) = o_{i+1}$ for all i. A *concrete strategy for Adam* is a function $\mu_\forall : \gamma(\mathtt{Prefs}(G)) \times \Sigma \to Q$ such that for any concrete prefix $\pi = q_0\sigma_0\ldots q_n \in \gamma(\mathtt{Prefs}(G))$ and action σ, $\mu_\forall(\pi, \sigma) \in \mathtt{post}_\sigma(\{\pi[n]\})$. A play $\psi = o_0\sigma_0o_1\sigma_1\ldots$ is consistent with $\mu_\forall$ if there exists a concrete path $\pi \in \gamma(\psi)$ such that $\mu_\forall(\pi[..i], \sigma_i) = \pi[i+1]$ for all i.

We say an observation-based strategy for Eve $\lambda_\exists$ has *memory m* if there is a set M with $|M| = m$, an element $m_0 \in M$, and functions $\alpha_u : M \times Obs \to M$ and $\alpha_o : M \times Obs \to \Sigma$ such that for any play prefix $\rho = o_0\sigma_0\ldots o_n$ we have $\lambda_\exists(\rho) = \alpha_o(m_n, o_n)$, where m_n is defined inductively by $m_{i+1} = \alpha_u(m_i, o_i)$ for

$i \geq 0$. An observation-based strategy for Adam $\lambda_\forall$ has *memory* m if there is a set M with $|M| = m$, an element $m_0 \in M$, and functions $\alpha_u : M \times Obs \times \Sigma \to M$ and $\alpha_o : M \times Obs \times \Sigma \to Obs$ such that for any play prefix ending in an action $\rho = o_0\sigma_0 \ldots o_n\sigma_n$, we have $\lambda_\forall(\rho) = \alpha_o(m_n, o_n, \sigma_n)$, where m_n is defined inductively by $m_{i+1} = \alpha_u(m_i, o_i, \sigma_i)$. An observation-based strategy (for either player) with memory 1 is *positional*.

Note that for any concrete strategy μ of Adam there is a unique observation-based strategy λ_μ such that all plays consistent with μ are consistent with λ_μ. Conversely there may be several, but possibly no, concrete strategies that correspond to a single observation-based strategy. In games of limited-observation there is guaranteed to be at least one concrete strategy for every observation-based strategy.

Given a threshold $\nu \in \mathbb{R}$, we say a play ψ is *winning for Eve* if $\underline{MP}(\pi) \geq \nu$ for all concrete paths $\pi \in \gamma(\psi)$, otherwise it is *winning for Adam*. Given ν, one can construct an equivalent game in which Eve wins if and only if $\underline{MP}(\pi) \geq 0$ if and only if she wins the original game, so without loss of generality we will assume $\nu = 0$. A strategy λ is *winning* for a player if all plays consistent with λ are winning for that player. We say that a player *wins* G if (s)he has a winning strategy.

It was shown in [9] that in MPGs with partial-observation where finite memory strategies suffice Eve wins the $\underline{MP}$ version of the game if and only if she wins the $\overline{MP}$ version. As the majority of games considered in this paper only require finite memory, we can take either definition. For simplicity and consistency with Section 3 we will use $\underline{MP}$.

Reachability games. A *reachability game* $G = \langle Q, q_I, \Sigma, \Delta, \mathcal{T}_\exists, \mathcal{T}_\forall \rangle$ is a tuple where Q is a (not necessarily finite) set of states; Σ is a finite set of actions; $\Delta \subseteq Q \times \Sigma \times Q$ is a finitary transition function; $q_I \in Q$ is the initial state; and $\mathcal{T}_\exists, \mathcal{T}_\forall \subseteq Q$ are the terminating states. Notions of plays and strategies in the reachability game follow from the definitions for MPGs, however we extend plays to include finite paths that end in $\mathcal{T}_\exists$ or in $\mathcal{T}_\forall$. In the first case we declare Eve as the winner whereas the latter corresponds to Adam winning the game. In general, the game might not terminate. In this case we say neither player wins.

3 Undecidability of Liminf Games

Mean-payoff games with partial-observation were extensively studied in [9]. In that paper the authors show that, with the mean payoff condition defined using $\underline{MP}$ and $>$, determining whether Eve has a winning strategy is undecidable and when defined using $\overline{MP}$ and $\geq$, strategies with infinite memory may be necessary. The analogous, and more general, questions using $\underline{MP}$ and $\geq$ were left open. In this section we answer these questions, showing that both results still hold.

Proposition 1. *There exist MPGs with partial-observation for which Eve requires infinite memory observation-based strategies to ensure* $\underline{MP} \geq 0$.

Consider the game in Figure 2 and consider the strategy of Eve that plays (regardless of location) $aba^2ba^3ba^4b\ldots$ As b is played infinitely often in this strategy, the only concrete paths consistent with this strategy are $\pi = q_0q_1^\omega$ and $\pi = q_0q_1^kq_2^lq_3^\omega$ for non-negative integers k, l. In both cases $\underline{MP} \geq 0$, so the strategy is winning.

Against a finite memory strategy of Eve, Adam plays to ensure the game remains in $\{q_1, q_2\}$. As Eve's strategy has finite memory, her choice of actions must be ultimately periodic. Now there are two cases, if she plays a finite number of b's then Adam has a concrete winning strategy which consists in guessing when she will play the last b and moving to q_2. If, on the other hand, she plays b infinitely often then Adam can choose to stay in q_1 and again win the game.

Theorem 1. *Let G be an MPG with partial-observation. Determining whether Eve has an observation-based strategy to ensure $\underline{MP} \geq 0$ is undecidable.*

In [6], the authors present a reduction from blind MPGs to mean-payoff automata. This reduction, together with the undecidability result from [9], imply several classical automata-theoretical problems for mean-payoff automata are also undecidable. In [6], the authors study the non-strict $\geq$ relation between quantitative languages. It follows from the undecidability result presented above, that even when one considers the strict order, $>$, these problems remain undecidable.

Corollary 1. *The strict quantitative universality, and strict quantitative language inclusion problems are undecidable for non-deterministic and alternating mean-payoff automata.*

4 Observable Determinacy

One of the key features of MPGs with perfect information is that they are determined, that is, it is always the case that one player has a winning strategy. This is not true in games of partial or limited-observation as can be seen in Figure 1. Any concrete strategy of Adam reveals to Eve the successor of q_0 and

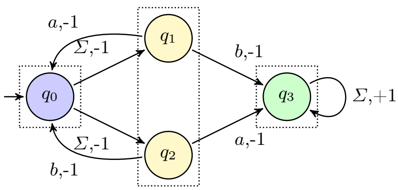

Fig. 1. A non-determined MPG with limited-observation ($\Sigma = \{a, b\}$).

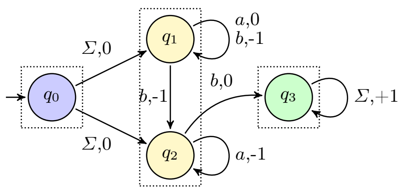

Fig. 2. An MPG with limited-observation which Eve requires infinite memory to win.

she can use this information to play to q_3. Conversely Adam can defeat any strategy of Eve by playing to whichever of q_1 or q_2 means the play returns to q_0 on Eve's next choice (recall Eve cannot distinguish q_1 and q_2 and must therefore choose an action to apply to the observation $\{q_1, q_2\}$). This strategy of Adam can be encoded as an observation-based strategy: "from $\{q_1, q_2\}$ with action a or b, play to $\{q_0\}$". It transpires that, under an assumption about large cardinals[3], any such counter-play by Adam is always encodable as an observable strategy.

Theorem 2 (Observable determinacy). *Assuming the existence of a measurable cardinal, one player always has a winning observation-based strategy in an MPG with limited-observation.*

The existence of a measurable cardinal implies Σ^1_1-Determinacy [14] – a weak form of the "Axiom of Determinacy". The observable determinacy of MPGs with limited-observation then follows from the following result:

Lemma 1. *The set of plays that are winning for Eve in an MPG with limited-observation is co-Suslin.*

5 Strategy Transfer

In this section we will construct a reachability game from an MPG with limited-observation in which winning strategies for either player are sufficient (but not necessary) for winning strategies in the original MPG.

Let us fix a mean-payoff game with limited-observation $G = \langle Q, q_I, \Sigma, \Delta, Obs, w\rangle$. We will define a reachability game on the weighted unfolding of G.

Let $\mathcal{F}$ be the set of functions $f : Q \to \mathbb{Z} \cup \{+\infty, \bot\}$. Our intention is to use functions in $\mathcal{F}$ to keep track of the minimum weight of all concrete paths ending in the given vertex. A function value of $\bot$ indicates that the given vertex is not in the current observation, and intuitively a function value of $+\infty$ is used to indicate to Eve that the token is not located at such a vertex. The added complication permits our winning condition to include games where Adam wins by ignoring paths going through the given vertex. The *support* of f is $\texttt{supp}(f) = \{q \in Q \mid f(q) \neq \bot\}$. We say that $f' \in \mathcal{F}$ is a σ-successor of $f \in \mathcal{F}$ if:

- $\texttt{supp}(f') \in Obs \wedge \texttt{supp}(f') \subseteq \texttt{post}_\sigma(\texttt{supp}(f))$; and
- for all $q \in \texttt{supp}(f')$, $f'(q)$ is either $\min\{f(q') + w(q', \sigma, q) \mid q' \in \texttt{supp}(f) \wedge (q', \sigma, q) \in \Delta\}$ or $+\infty$.

We define a family of partial orders, $\preceq_k$ ($k \in \mathbb{N}$), on $\mathcal{F}$ by setting $f \preceq_k f'$ if $\texttt{supp}(f) = \texttt{supp}(f')$ and $f(q) + k \leq f'(q)$ for all $q \in \texttt{supp}(f)$ (where $+\infty + k = +\infty$).

Denote by $\mathfrak{F}_G$ the set of all sequences $f_0\sigma_0 f_1 \ldots \sigma_{n-1} f_n \in (\mathcal{F} \cdot \Sigma)^* \mathcal{F}$ such that for all $0 \leq i < n$, f_{i+1} is a σ_i-successor of f_i. Observe that for each function-action sequence $\rho = f_0\sigma_0 \ldots f_n \in \mathfrak{F}_G$ there is a unique abstract path $\texttt{supp}(\rho) =$

[3] This assumption is independent of the theory of ZFC.

$o_0\sigma_0 \dots o_n$ such that $o_i = \mathtt{supp}(f_i)$ for all i. Conversely for each abstract path $\psi = o_0\sigma_0 \dots o_n$ there may be many corresponding function-action sequences in $\mathtt{supp}^{-1}(\psi)$.

The *reachability game associated with* G, i.e. $\Gamma_G = \langle \Pi_G, \Sigma, f_I, \delta, \mathcal{T}_\exists, \mathcal{T}_\forall \rangle$, is formally defined as follows: $f_I \in \mathcal{F}$ is the function for which $f(q) \mapsto 0$ if $q = q_I$ and $f(q) \mapsto \bot$ otherwise. Π_G is the subset of $\mathfrak{F}_G$ where for all $f_0\sigma_0 f_1 \dots \sigma_{n-1} f_n \in \Pi_G$ we have $f_0 = f_I$ and for all $0 \leq i < j < n$ we have $f_i \not\preceq_0 f_j$ and $f_j \not\preceq_1 f_i$; δ is the natural transition function, that is, if x and $x \cdot \sigma \cdot f$ are elements of Π_G then $(x, \sigma, x \cdot \sigma \cdot f) \in \delta$; $\mathcal{T}_\exists$ is the set of all $f_0\sigma_0 f_1 \dots \sigma_{n-1} f_n \in \Pi_G$ such that for some $0 \leq i < n$ we have $f_i \preceq_0 f_n$; and $\mathcal{T}_\forall$ is the set of all $f_0\sigma_0 f_1 \dots \sigma_{n-1} f_n \in \Pi_G$ such that for some $0 \leq i < n$ we have $f_n \preceq_1 f_i$ and $f_i(q) \neq +\infty$ for some $q \in \mathtt{supp}(f_i)$.

Note that the directed graph defined by Π_G and δ is a tree, but not necessarily finite. To gain an intuition about Γ_G, let us say an abstract cycle ρ is *good* if there exists $f_0\sigma_0 \dots f_n \in \mathtt{supp}^{-1}(\rho)$ such that $f_i(q) \neq +\infty$ for all q and all i and $f_0 \preceq_0 f_n$. Let us say ρ is *bad* if there exists $f_0\sigma_0 \dots f_n \in \mathtt{supp}^{-1}(\rho)$ such that $f_0(q) \neq +\infty$ for some $q \in \mathtt{supp}(f_0)$ and $f_n \preceq_1 f_0$. Then it is not difficult to see that Γ_G is essentially an abstract cycle-forming game played on G which is winning for Eve if a good abstract cycle is formed and winning for Adam if a bad abstract cycle is formed.

Theorem 3. *Let G be an MPG with limited-observation and let Γ_G be the associated reachability game. If Adam (Eve) has a winning strategy in Γ_G then (s)he has a finite-memory observation-based winning strategy in G.*

The idea behind the strategy for the mean-payoff game is straightforward. If Eve wins the reachability game then she can transform her strategy into one that plays indefinitely by returning, whenever the play reaches $\mathcal{T}_\exists$, to the natural previous position – namely the position that witnesses the membership of $\mathcal{T}_\exists$. By continually playing her winning strategy in this way Eve perpetually completes good abstract cycles and this ensures that all concrete paths consistent with the play have non-negative mean-payoff value. Likewise if Adam has a winning strategy in the reachability game, he can continually play his strategy by returning to the natural position whenever the play reaches $\mathcal{T}_\forall$. By doing this he perpetually completes bad abstract cycles and this ensures that there is a concrete path consistent with the play that has strictly negative mean-payoff value. The finiteness of the size of the memory required for this strategy follows from the following result.

Lemma 2. *If λ is a winning strategy for Adam or Eve in Γ_G, then there exists $N \in \mathbb{N}$ such that for all plays π consistent with λ, $|\pi| \leq N$.*

Although the following results are not used until Section 7, they give an intuition toward the correctness of the strategies described above.

Lemma 3. *Let ρ be an abstract cycle.*

(i) If ρ is good (bad) then an interleaving of ρ with another good (bad) cycle is also good (bad).

(ii) If ρ is good then for all k and all concrete cycles $\pi \in \gamma(\rho^k)$, $w(\pi) \geq 0$.
(iii) If ρ is bad then $\exists k \geq 0, \pi \in \gamma(\rho^k)$ such that $w(\pi) < 0$.

Corollary 2. *No cyclic permutation of a good abstract cycle is bad.*

6 Forcibly Terminating Games

The reachability game defined in the previous section gives a sufficient condition for determining the winner in an MPG with limited-observation. However, as there may be plays where no player wins, such games are not necessarily determined. The first subclass of MPGs with limited-observation we investigate is the class of games where the associated reachability game is determined.

Definition 1. *An MPG with limited-observation is* forcibly terminating *if in the corresponding reachability game Γ_G either Adam has a winning strategy to reach locations in $\mathcal{T}_\forall$ or Eve has a winning strategy to reach locations in $\mathcal{T}_\exists$.*

It follows immediately from Theorem 3 that finite memory strategies suffice for both players in forcibly terminating games. Note that an upper bound on the memory required is the number of vertices in the reachability game restricted to a winning strategy, and this is exponential in N, the bound obtained in Lemma 2.

Theorem 4 (Finite-memory determinacy). *One player always has a winning observation-based strategy with finite memory in a forcibly terminating MPG.*

We now consider the complexity of two natural decision problems associated with forcibly terminating games: the problem of recognizing if an MPG is forcibly terminating and the problem of determining the winner of a forcibly terminating game. Both results follow directly from the fact that we can accurately simulate a Turing Machine with an MPG with limited-observation.

Theorem 5. *Let M be a Deterministic Turing Machine. Then there exists an MPG with limited-observation G, constructible in polynomial time, such that Eve wins Γ_G if and only if M halts in the accept state and Adam wins Γ_G if and only if M halts in the reject state.*

Corollary 3 (Class membership). *Let G be an MPG with limited-observation. Determining if G is forcibly terminating is undecidable.*

Corollary 4 (Winner determination). *Let G be a forcibly terminating MPG. Determining if Eve wins G is* R*-complete.*

Proof. R-hardness follows from Theorem 5. For decidability, Lemma 2 implies that an alternating Turing Machine simulating a play on Γ_G will terminate.

7 Forcibly First Abstract Cycle Games

In this section and the next we consider restrictions of forcibly terminating games in order to find subclasses with more efficient algorithmic bounds. The negative algorithmic results from the previous section largely arise from the fact that the abstract cycles required to determine the winner are not necessarily simple cycles. Our first restriction of forcibly terminating games is the restriction of the abstract cycle-forming game to simple cycles.

More precisely, let G be an MPG with limited-observation and Γ_G be the associated reachability game. Define $\Pi'_G \subseteq \Pi_G$ as the set of all sequences $x = f_0\sigma_0 f_1\sigma_1 \ldots f_n \in \Pi_G$ such that $\mathtt{supp}(f_i) \neq \mathtt{supp}(f_j)$ for all $0 \leq i < j < n$ and denote by Γ'_G the reachability game $\langle \Pi'_G, \Sigma, f_I, \delta', \mathcal{T}'_\exists, \mathcal{T}'_\forall \rangle$ where δ' is δ restricted to Π'_G, $\mathcal{T}'_\exists = \mathcal{T}_\exists \cap \Pi'_G$ and $\mathcal{T}'_\forall = \mathcal{T}_\forall \cap \Pi'_G$.

Definition 2. *An MPG with limited-observation is* forcibly first abstract cycle *(or forcibly FAC) if in the associated reachability game Γ'_G either Adam has a winning strategy to reach locations in $\mathcal{T}'_\forall$ or Eve has a winning strategy to reach locations in $\mathcal{T}'_\exists$.*

One immediate consequence of the restriction to simple abstract cycles is that the bound in Lemma 2 is at most $|Obs|$. In particular an alternating Turing Machine can, in linear time, simulate a play of the reachability game and decide which player, if any, has a winning strategy. Hence the problems of deciding if a given MPG with partial-observation is forcibly FAC and deciding the winner of a forcibly FAC game are both solvable in PSPACE. The next results show that there is a matching lower bound for both these problems.

Theorem 6 (Class membership). *Let G be an MPG with limited-observation. Determining if G is forcibly FAC is* PSPACE*-complete.*

PSPACE-hardness follows from a reduction from the satisfiability of quantified boolean formulas. The construction is similar to the construction used to prove PSPACE-hardness for Generalized Geography in [15]. That is, the game proceeds through diamond gadgets – the choice of each player on which side to go through corresponds to the selection of the value for the quantified variable. The (abstract) play then passes through a gadget for the formula in the obvious way (Adam choosing for $\wedge$ and Eve choosing for $\vee$), returning to a diamond gadget when a variable is reached. If the variable has been seen before the cycle is closed and the game ends, otherwise the play proceeds to the bottom of the diamond gadget which has been seen before, thus ending the game one step later. We set up the concrete paths within the observations in such a way that if the cycle closes at the variable then it is good (and thus Eve wins) and if it closes at the bottom of the gadget then it is not good. Corollary 2 implies that the cycle closed is never bad, so either Eve wins and the game is forcibly FAC, or neither player wins and it is not forcibly FAC.

We can slightly modify the above construction in such a way that if the game does not finish when the play returns to a variable then Adam can close a bad cycle. This results in a forcibly FAC game that Eve wins if and only if the formula is satisfied. Hence,

Theorem 7 (Winner determination). *Let G be a forcibly FAC MPG. Determining if Eve wins G is* PSPACE*-complete.*

It also follows from the $|Obs|$ upper bound on plays in Γ'_G that there is an exponential upper bound on the memory required for a winning strategy for either player. Furthermore, we can show this bound is tight – the games constructed in the proof of Theorem 7 can be used to show that there are forcibly FAC games that require exponential memory for winning strategies.

Theorem 8 (Exponential memory determinacy). *One player always has a winning observation-based strategy with exponential memory in a forcibly FAC MPG. Further, for any $n \in \mathbb{N}$ there exists a forcibly FAC MPG, of size polynomial in n, such that any winning strategy has memory at least 2^n.*

8 First Abstract Cycle Games

We now consider a structural restriction that guarantees Γ'_G is determined.

Definition 3. *An MPG with limited-observation is a* first abstract cycle game *(FAC) if in the associated reachability game Γ'_G all leaves are in $\mathcal{T}'_\forall \cup \mathcal{T}'_\exists$.*

Intuitively, in an FAC game all simple abstract cycles (that can be formed) are either good or bad. It follows then from Corollary 2 that any cyclic permutation of a good cycle is also good and any cyclic permutation of a bad cycle is also bad. Together with Lemma 3, this implies the abstract cycle-forming games associated with FAC games can be seen to satisfy the following three assumptions: (1) A play stops as soon as an abstract cycle is formed, (2) The winning condition and its complement are preserved under cyclic permutations, and (3) The winning condition and its complement are preserved under interleavings. These assumptions correspond to the assumptions required in [1] for positional strategies to be sufficient for both players[4]. That is,

Theorem 9 (Positional determinacy). *One player always has a positional winning observation-based strategy in an FAC MPG.*

As we can check in polynomial time if a positional strategy is winning in an FAC MPG, we immediately have:

Corollary 5 (Winner determination). *Let G be an FAC MPG. Determining if Eve wins G is in* NP $\cap$ coNP.

A path in Γ'_G to a leaf not in $\mathcal{T}'_\forall \cup \mathcal{T}'_\exists$ provides a short certificate to show that an MPG with limited-observation is not FAC. Thus deciding if an MPG is FAC is in coNP. A matching lower bound can be obtained using a reduction from the complement of the Hamiltonian Cycle problem.

[4] These conditions supercede those of [4] which were shown in [1] to be insufficient for positional strategies.

Theorem 10 (Class membership). *Let G be an MPG with limited-observation. Determining if G is FAC is* coNP*-complete.*

9 MPGs with Partial-Observation

The translation from partial-observation to limited-observation games allows us to extend the notions of FAC and forcibly FAC games to the larger class of MPGs with partial-observation. In this section we will investigate the resulting algorithmic effect of this translation on the decision problems we have been considering.

We say an MPG with partial-observation is *(forcibly) first belief cycle*, or FBC, if the corresponding MPG with limited-observation is (forcibly) FAC.

9.1 FBC and Forcibly FBC MPGs

Our first observation is that FBC MPGs generalize the class of *visible weight games* studied in [9]. An MPG with partial-observation is considered a visible weights game if its weight function satisfies the condition that all σ-transitions between any pair of observations have the same weight. We base some of our results for FBC and forcibly FBC games on lower bounds established for problems on visible weights games.

Lemma 4. *Let G be a visible weights MPG with partial-observation. Then G is FBC.*

We now turn to the decision problems we have been investigating throughout the paper. Given the exponential blow-up in the construction of the game of limited-observation, it is not surprising that there is a corresponding exponential increase in the complexity of the class membership problem.

Theorem 11 (Class membership). *Let G be an MPG with partial-observation. Determining if G is FBC is* coNEXP*-complete and determining if G is forcibly FBC is in* EXPSPACE *and* NEXP*-hard.*

Somewhat surprisingly, for the winner determination problem we have an EXP-time algorithm to match the EXP-hardness lower bound from visible weights games. This is in contrast to the class membership problem in which an exponential increase in complexity occurs when moving from limited to partial-observation.

Theorem 12 (Winner determination). *Let G be a forcibly FBC MPG. Determining if Eve wins G is* EXP*-complete.*

Corollary 6. *Let G be an FBC MPG. Determining if Eve wins G is* EXP*-complete.*

References

1. Aminof, B., Rubin, S.: First cycle games. In: Mogavero, F., Murano, A., Vardi, M.Y. (eds.) SR. EPTCS, vol. 146, pp. 91–96 (2014)
2. Berwanger, D., Chatterjee, K., Doyen, L., Henzinger, T.A., Raje, S.: Strategy construction for parity games with imperfect information. In: van Breugel, F., Chechik, M. (eds.) CONCUR 2008. LNCS, vol. 5201, pp. 325–339. Springer, Heidelberg (2008)
3. Berwanger, D., Doyen, L.: On the power of imperfect information. In: FSTTCS, pp. 73–82 (2008)
4. Björklund, H., Sandberg, S., Vorobyov, S.: Memoryless determinacy of parity and mean payoff games: a simple proof. TCS 310(1), 365–378 (2004)
5. Chatterjee, K., Doyen, L.: Partial-observation stochastic games: How to win when belief fails. In: LICS, pp. 175–184. IEEE (2012)
6. Chatterjee, K., Doyen, L., Edelsbrunner, H., Henzinger, T.A., Rannou, P.: Mean-payoff automaton expressions. In: Gastin, P., Laroussinie, F. (eds.) CONCUR 2010. LNCS, vol. 6269, pp. 269–283. Springer, Heidelberg (2010)
7. Chatterjee, K., Doyen, L., Henzinger, T.A.: Quantitative languages. In: Kaminski, M., Martini, S. (eds.) CSL 2008. LNCS, vol. 5213, pp. 385–400. Springer, Heidelberg (2008)
8. Chatterjee, K., Doyen, L., Henzinger, T.A., Raskin, J.-F.: Generalized mean-payoff and energy games. In: FSTTCS, pp. 505–516 (2010)
9. Degorre, A., Doyen, L., Gentilini, R., Raskin, J.-F., Toruńczyk, S.: Energy and mean-payoff games with imperfect information. In: Dawar, A., Veith, H. (eds.) CSL 2010. LNCS, vol. 6247, pp. 260–274. Springer, Heidelberg (2010)
10. Ehrenfeucht, A., Mycielski, J.: Positional strategies for mean payoff games. International Journal of Game Theory 8, 109–113 (1979)
11. Hunter, P., Pérez, G.A., Raskin, J.-F.: Mean-payoff games with partial-observation (extended abstract). CoRR (2014)
12. Jurdziński, M.: Deciding the winner in parity games is in $\mathsf{UP} \cap \mathsf{coUP}$. IPL 68(3), 119–124 (1998)
13. Kupferman, O., Vardi, M.Y.: Synthesis with incomplete informatio. Advances in Temporal Logic 16, 109–127 (2000)
14. Martin, D.A., Steel, J.R.: Projective determinacy. Proceedings of the National Academy of Sciences of the United States of America 85(18), 6582 (1988)
15. Papadimitriou, C.H.: Computational complexity. John Wiley and Sons Ltd. (2003)
16. Reif, J.H.: The complexity of two-player games of incomplete information. Journal of Computer and System Sciences 29(2), 274–301 (1984)
17. Zwick, U., Paterson, M.: The complexity of mean payoff games on graphs. TCS 158(1), 343–359 (1996)