Energy and Mean-Payoff Games with Imperfect Information

Aldric Degorre[1], Laurent Doyen[2], Raffaella Gentilini[3], Jean-François Raskin[1], and Szymon Toruńczyk[2]

[1]Université Libre de Bruxelles (ULB), Belgium
[2]LSV, ENS Cachan & CNRS, France
[3]University of Perugia, Italy

Abstract. We consider two-player games with imperfect information and quantitative objective. The game is played on a weighted graph with a state space partitioned into classes of indistinguishable states, giving players partial knowledge of the state. In an energy game, the weights represent resource consumption and the objective of the game is to maintain the sum of weights always nonnegative. In a mean-payoff game, the objective is to optimize the limit-average usage of the resource. We show that the problem of determining if an energy game with imperfect information with *fixed* initial credit has a winning strategy is decidable, while the question of the existence of *some* initial credit such that the game has a winning strategy is undecidable. This undecidability result carries over to mean-payoff games with imperfect information. On the positive side, using a simple restriction on the game graph (namely, that the weights are visible), we show that these problems become EXPTIME-complete.

1 Introduction

Mean-payoff games (MPG) are two-player games of infinite duration played on (directed) weighted graphs. Two players move a pebble along the edges of the graph. The player owning the current state chooses an outgoing edge to a successor state, resulting in an infinite path in the graph called a play. The objective of Player 1 is to maximize the average weight of the edges traversed in the play, while Player 2 tries to minimize this value (hence, the game is zero-sum).

The algorithmics of MPG has been studied since the end of the seventies [12, 24] and still attracts a large interest for several reasons. First, the problem of deciding the winner in MPG has a remarkable status: it is known to be in the intersection of the classes NP and coNP [12], yet no polynomial algorithm is known for solving it. This problem has tight connection (through polynomial-time reductions) with important theoretical questions about logics and games, such as the μ-calculus model-checking, and solving parity games [13, 14, 17, 18]. Second, quantitative objectives in general are gaining interest in the specification and design of reactive systems [8, 5, 11], where the weights represent resource usage (e.g., energy consumption or network usage); the problem of controller synthesis with resource constraints requires the solution of quantitative games [20, 6, 1, 3]. Finally, mean-payoff games are log-space equivalent to energy games (EG) where the objective of Player 1 is to maintain the sum of the weight (called the energy level) positive, given a fixed initial credit of energy. This result leads to faster pseudo-polynomial time algorithms for solving mean-payoff games [4, 10].

The previous works on MPG and EG make the assumption that the game is of *perfect information*: the players have complete knowledge of the history of the play up to the current state. However, for applications in controller synthesis, it is often more appropriate that the players have partial knowledge (e.g., because the private variables of other players are not visible, or because the sensors have poor accuracy). In this paper, we consider mean-payoff and energy games with *imperfect information*. The game graph is given with a coloring of the state space that defines equivalence classes of indistinguishable states called observations [23, 7], and the strategies are observation-based (i.e., they rely on the past sequence of observations rather than states). Interestingly, we show that observation-based strategies may require infinite memory for mean-payoff games, while finite memory is sufficient for energy games, in contrast with games of perfect information where memoryless strategies suffice in both cases.

We consider the following decision problems: given a weighted colored graph, (*i*) for MPG, given a threshold ν, decide if Player 1 has an observation-based strategy to ensure a mean-payoff value at least ν; for EG, (*ii*) given a *fixed* initial credit, decide if Player 1 has an observation-based strategy to maintain the energy level always positive; and (*iii*) if the initial credit is *unknown*, decide if there exists an initial credit for which Player 1 wins. Our results can be summarized as follows.

First, we show that energy games with fixed initial credit are decidable. The argument is based on the existence of a well quasi order on the infinite set of play prefixes, which gives a bound on the depth of unravelling of the game graph needed to decide the problem. The size of this bound gives non-primitive recursive complexity to our procedure. Beside establishing the existence of finite-memory strategies for energy games, this result is also in contrast with the case of energy games with finite duration where the objective of Player 1 is to stop the game after finitely many steps with a positive energy level. There, even the fixed initial credit problem for blind games[1] is undecidable, using the note after Corollary 3.8 in [19].

Next, we show that for games with imperfect information (and even for blind games), both energy games with unknown initial credit, and mean-payoff games are undecidable, using reductions from the halting problem for two-counter machines. For energy games, since the problem with fixed initial credit is decidable, it shows that the problem with unknown initial credit is r.e. but not co-r.e. For mean-payoff games, we show that the problem is neither r.e. nor co-r.e. using a reduction from the complement of the halting problem for two-counter machines. This second reduction however requires at least two observations (or two colors) in the game. As a corollary of our results, we obtain the undecidability of univer-

[1] Blind games have a single observation (i.e., all states have the same color). They correspond to solving universality questions on nondeterministic automata.

sality and language inclusion for quantitative languages defined by mean-payoff conditions [6].

Finally, we identify a class of MPG and EG for which the decision problems are EXPTIME-complete. This class corresponds to the assumption that the weights are visible to the players, a reasonable restriction in the context of controller synthesis (if each process can observe its own energy level). The algorithmic solution of this class of games relies on a generalization of the classical subset construction that maintains the knowledge of Player 1 along the play [23, 7]. On the way, we also obtain that exponential memory is sufficient to win, and we recover the nice property that MPG and EG with imperfect information and visible weights are log-space equivalent.

2 Definitions

Games. A *weighted game with imperfect information* (or simply a *game*) is a tuple $G = \langle Q, q_0, \Sigma, \Delta, w, \mathsf{Obs}\rangle$, where Q is a finite set of states, $q_0 \in Q$ is the initial state, Σ is a finite alphabet, $\Delta \subseteq Q \times \Sigma \times Q$ is a labeled transition relation which is total, i.e. for all $q \in Q$ and $\sigma \in \Sigma$, there exists $q' \in Q$ such that $(q, \sigma, q') \in \Delta$; $w : \Delta \to \mathbb{Z}$ is a weight function, and $\mathsf{Obs} \subseteq 2^Q$ is a set of observations that partition the state space. For each state $q \in L$, we denote by $\mathsf{obs}(q)$ the unique observation $o \in \mathsf{Obs}$ such that $q \in o$; and for $s \subseteq L$ and $\sigma \in \Sigma$, we denote by $\mathsf{post}^G_\sigma(s) = \{q' \in Q \mid \exists q \in s : (q, \sigma, q') \in \Delta\}$ the set of σ-successors of s.

We consider the following special cases of interest. We say that the weights are *visible* if $w(q_1, \sigma, q_2) = w(q'_1, \sigma', q'_2)$ for all transitions $(q_1, \sigma, q_2), (q'_1, \sigma', q'_2) \in \Delta$ such that $\mathsf{obs}(q_1) = \mathsf{obs}(q'_1)$, $\mathsf{obs}(q_2) = \mathsf{obs}(q'_2)$, and $\sigma = \sigma'$. A game with *perfect information* is such that $\mathsf{Obs} = \{\{q\} \mid q \in Q\}$, i.e. each state is observable, and a *blind* game is such that $\mathsf{Obs} = \{Q\}$, i.e. all states are indistinguishable. We omit the set Obs in the definition of games of perfect information.

Games are played in rounds in which Player 1 chooses an action $\sigma \in \Sigma$, and Player 2 chooses a σ-successor of the current state.[2] The first round starts in the initial state q_0.

A *play* in G is an infinite sequence $\pi = q_0\sigma_0q_1\sigma_1\dots$ such that $(q_i, \sigma_i, q_{i+1}) \in \Delta$ for all $i \geq 0$. The prefix up to q_n of the play π is denoted by $\pi(n)$, and its last element is $\mathsf{Last}(\pi(n)) = q_n$. The set of plays in G is denoted $\mathsf{Plays}(G)$ and the set of corresponding prefixes is written as $\mathsf{Prefs}(G)$. The *observation sequence* of π is the sequence $\mathsf{obs}(\pi) = \mathsf{obs}(q_0)\sigma_0\mathsf{obs}(q_1)\sigma_1\dots$ and the (finite) observation sequence $\mathsf{obs}(\pi(n))$ of $\pi(n)$ is the prefix up to $\mathsf{obs}(q_n)$ of $\mathsf{obs}(\pi)$.

The *energy level* of a play prefix $\rho = q_0\sigma_0q_1\dots q_n$ is $\mathsf{EL}(\rho) = \sum_{i=0}^{n-1} w(q_i, \sigma_i, q_{i+1})$, and the *mean-payoff value* of a play π is either $\overline{\mathsf{MP}}(\pi) = \limsup_{n\to\infty} \frac{1}{n} \cdot \mathsf{EL}(\pi(n))$ or $\underline{\mathsf{MP}}(\pi) = \liminf_{n\to\infty} \frac{1}{n} \cdot \mathsf{EL}(\pi(n))$.

[2] For games of perfect information, this definition is equivalent to the classical setting where the state space is partitioned into Player 1 states and Player 2 states, and the player owning the current state chooses the successor.

Strategies. A *strategy* (for Player 1) in G is a function $\alpha : \mathsf{Prefs}(G) \to \Sigma$. A strategy α for Player 1 is called *observation-based* if for all prefixes $\rho, \rho' \in \mathsf{Prefs}(G)$, if $\mathsf{obs}(\rho) = \mathsf{obs}(\rho')$, then $\alpha(\rho) = \alpha(\rho')$, and it is called *memoryless* if $\alpha(\rho \cdot q) = \alpha(\rho' \cdot q)$ for all $\rho, \rho' \in (Q\Sigma)^*$ and $q \in Q$. Memoryless strategies are best viewed as functions $\alpha : Q \to \Sigma$. A prefix $\rho = q_0\sigma_0 q_1 \dots q_n \in \mathsf{Prefs}(G)$ is *consistent* with α if $\sigma_i = \alpha(\pi(i))$ for all $i \geq 0$. An *outcome* of α is a play whose all prefixes are consistent with α.

A strategy α has *finite-memory* if it can be encoded by a deterministic Moore machine $\langle M, m_0, \alpha_u, \alpha_n \rangle$ where M is a finite set of states (the memory of the strategy), $m_0 \in M$ is the initial state, $\alpha_u : M \times Q \to M$ is an update function, and $\alpha_n : M \to \Sigma$ is a next-action function. In state $m \in M$, the strategy plays the action $\alpha_n(m)$, and when Player 2 chooses the next state q of the game, the internal state is updated to $\alpha_u(m, q)$. Formally, the strategy α defined by $\langle M, m_0, \alpha_u, \alpha_n \rangle$ is such that $\alpha(\rho) = \alpha_n(\hat{\alpha}_u(m_0, q_0 \dots q_n))$ for all $\rho = q_0\sigma_0 q_1 \dots q_n \in \mathsf{Prefs}(G)$, where $\hat{\alpha}_u$ extends α_u to sequences of states as follows: $\hat{\alpha}_u(m, \epsilon) = m$ and $\hat{\alpha}_u(m, q_0 \dots q_n) = \hat{\alpha}_u(\alpha_u(m, q_0), q_1 \dots q_n)$ for all $m \in M$ and $q_1, \dots, q_n \in Q$.

Objectives. An *objective* for G is a set ϕ of infinite sequences of states and actions, that is, $\phi \subseteq (Q \times \Sigma)^\omega$. A strategy α of Player 1 is winning for ϕ if $\pi \in \phi$ for all outcomes π of α. We consider the following objectives.

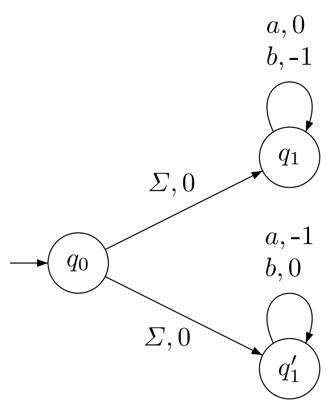

Fig. 1. A blind mean-payoff game.

- *Safety objectives.* Given a set $\mathcal{T} \subseteq Q$ of safe states, the safety objective $\mathsf{Safe}_G(\mathcal{T}) = \{q_0\sigma_0 q_1\sigma_1 \dots \in \mathsf{Plays}(G) \mid \forall n \geq 0 : q_n \in \mathcal{T}\}$ requires that only states in $\mathcal{T}$ be visited.
- *Energy objectives.* Given an initial credit $c_0 \in \mathbb{N}$, the energy objective $\mathsf{PosEnergy}_G(c_0) = \{\pi \in \mathsf{Plays}(G) \mid \forall n \geq 0 : c_0 + \mathsf{EL}(\pi(n)) \geq 0\}$ requires that the energy level be always positive.
- *Mean-payoff objectives.* Given a threshold $\nu \in \mathbb{Q}$, and $\sim \in \{>, \geq\}$, the mean-payoff objectives $\mathsf{MeanPayoffSup}_G^{\sim}(\nu) = \{\pi \in \mathsf{Plays}(G) \mid \overline{\mathsf{MP}}(\pi) \sim \nu\}$ and $\mathsf{MeanPayoffInf}_G^{\sim}(\nu) = \{\pi \in \mathsf{Plays}(G) \mid \underline{\mathsf{MP}}(\pi) \sim \nu\}$ require that the mean-payoff value be at least ν (resp., greater than ν).

When the game G is clear form the context, we omit the subscript in objective names. Mean-payoff objectives defined with $\sim \in \{<, \leq\}$ are obtained by duality since $\limsup_{i\to\infty} u_i = -\liminf_{i\to\infty} -u_i$.

We consider the following decision problems:

- The *fixed initial credit problem* asks, given an energy game G and an initial credit c_0, to decide whether there exists a winning observation-based strategy for the objective $\mathsf{PosEnergy}(c_0)$.
- The *unknown initial credit problem* asks, given an energy game G, to decide whether there exist an initial credit c_0 and a winning observation-based strategy for the objective $\mathsf{PosEnergy}(c_0)$.
- The *threshold problem* asks, given a mean-payoff game G, a threshold $\nu \in \mathbb{Q}$, and $\sim \in \{>, \geq\}$, to decide whether there exists a winning observation-based strategy for one of the mean-payoff objective $\mathsf{MeanPayoffSup}_G^{\sim}(\nu)$ or $\mathsf{MeanPayoffInf}_G^{\sim}(\nu)$.

Remark 1. The objectives $\mathsf{MeanPayoffSup}^{\sim}(\nu)$ and $\mathsf{MeanPayoffInf}^{\sim}(\nu)$ are equivalent for games with perfect information: Player 1 has a strategy to win according to $\mathsf{MeanPayoffSup}^{\sim}(\nu)$ if and only if he has a strategy to win according to $\mathsf{MeanPayoffInf}^{\sim}(\nu)$. However, the definitions are not equivalent for games with imperfect information.

As an example, consider the blind game of Fig. 1. All states are indistinguishable, and an initial nondeterministic choice determines the state q_1 or q_1' in which the game will loop forever.

We claim that Player 1 has an observation-based winning strategy for the objective $\mathsf{MeanPayoffSup}^{\geq}(0)$, but not for $\mathsf{MeanPayoffInf}^{\geq}(0)$. Note that in blind games, observation-based strategies can be viewed as infinite words. A winning strategy for the limsup version consists in playing sequences of a's and b's of increasing length in order to ensure a mean-payoff value $\overline{\mathsf{MP}}$ equal to 0 in both states q_1 and q_1'. For example, playing sequences of a's and b's such that the length of the i-th sequence is i times the length of the prefix played so far. This ensures that in q_1 and q_1', for all $i > 0$ there are infinitely many positions such that the average of the weights is greater than $-\frac{1}{i}$, showing that the limsup is 0 in all outcomes.

We show that for every word $w \in \{a, b\}^{\omega}$, the mean-payoff value according to $\underline{\mathsf{MP}}$ is at most $-\frac{1}{2}$. Let n_i and m_i be the numbers of a's and b's in the prefix of length i of w. Either $n_i \leq m_i$ for infinitely many i's, or $n_i \geq m_i$ for infinitely many i's. In the first case, the average of the weights (in state q_1) is infinitely often at most $-\frac{1}{2}$. The same holds in the second case using state q_1'. Therefore the $\liminf$ of the weight averages is at most $-\frac{1}{2}$, and Player 1 has no winning strategy for the mean-payoff objective defined using $\liminf$ and threshold 0.

Remark 2. Note that infinite memory is required to achieve mean-payoff value 0 (according to $\overline{\mathsf{MP}}$) in the game of Fig. 1. Indeed, for all finite-memory strategies (which can be viewed as ultimately periodic words), the mean-payoff value of an outcome is $\min\{-\frac{n}{n+m}, -\frac{m}{n+m}\} \leq -\frac{1}{2}$ where n and m are the numbers of a's and b's in the cycle of the strategy.

Note also that for finite-memory strategies, the mean-payoff objectives defined using $\overline{\mathsf{MP}}$ and $\underline{\mathsf{MP}}$ lead to equivalent games (see Theorem 6).

3 Energy Games with Imperfect Information

We present an algorithm for solving the fixed initial credit problem, and we show that the unknown initial credit problem is undecidable.

3.1 Fixed initial credit

Fix an energy game $G = \langle Q, q_0, \Sigma, \Delta, w, \mathsf{Obs}\rangle$ and an initial credit $c_0 \in \mathbb{N}$. To solve this energy game, we construct an equivalent safety game of perfect information using the following definitions.

Let $\mathcal{F}$ be the set of functions $f : Q \to \mathbb{Z} \cup \{\bot\}$. The *support* of f is $\mathsf{supp}(f) = \{q \in Q \mid f(q) \neq \bot\}$. A function $f \in \mathcal{F}$ stores the possible current states of the game G together with their energy level. We say that a function f is *nonnegative* if $f(q) \geq 0$ for all $q \in \mathsf{supp}(f)$. Initially, we set $f_{c_0}(q_0) = c_0$ and $f_{c_0}(q) = \bot$ for all $q \neq q_0$. The set $\mathcal{F}$ is ordered by the relation $\preceq$ such that $f_1 \preceq f_2$ if $\mathsf{supp}(f_1) = \mathsf{supp}(f_2)$ and $f_1(q) \leq f_2(q)$ for all $q \in \mathsf{supp}(f_1)$.

For $\sigma \in \Sigma$, we say that $f_2 \in \mathcal{F}$ is a σ-successor of $f_1 \in \mathcal{F}$ if there exists an observation $o \in \mathsf{Obs}$ such that $\mathsf{supp}(f_2) = \mathsf{post}^G_\sigma(\mathsf{supp}(f_1)) \cap o$ and $f_2(q) = \min\{f_1(q') + w(q', \sigma, q) \mid q' \in \mathsf{supp}(f_1) \wedge (q', \sigma, q) \in \Delta\}$ for all $q \in \mathsf{supp}(f_2)$. Given a sequence $x = f_0\sigma_0 f_1\sigma_1 \dots f_n$, let $f_x = f_n$ be the last function in x. Define the safety game $H = \langle Q^H, f_{c_0}, \Sigma, \Delta^H\rangle$ where Q^H is the smallest subset of $(\mathcal{F} \cdot \Sigma)^* \cdot \mathcal{F}$ such that

1. $f_{c_0} \in Q^H$, and
2. for each sequence $x \in Q^H$, if (i) f_x is nonnegative, and (ii) there is no strict prefix y of x such that $f_y \preceq f_x$, then $x \cdot \sigma \cdot f_2 \in Q^H$ for all σ-successors f_2 of f_x.

The transition relation Δ^H contains the corresponding triples $(x, \sigma, x \cdot \sigma \cdot f_2)$, and the game is made total by adding self-loops (x, σ, x) to sequences x without outgoing transitions. We call such sequences the *leaves* of H. Note that the game H is acyclic, except for the self-loops on the leaves.

By Dickson's lemma [9], the relation $\preceq$ on nonnegative functions is a *well quasi order*, i.e., for all infinite sequences $f_1 f_2 \dots$ of nonnegative functions, there exist two positions k, l such that $k < l$ and $f_k \preceq f_l$. Therefore, the state space Q^H is finite.

Lemma 1. *The game H has a finite state space.*

Proof. By contradiction, assume that Q^H is infinite. By König's lemma, since the transition relation of H is finitely branching, there exists an infinite sequence $f_0\sigma_0 f_1\sigma_1 \dots$ such that all its prefixes are in Q^H, and thus all f_i's are nonnegative and it is never the case that $f_k \preceq f_l$ for $k < l$. This is in contradiction with the fact that $\preceq$ is a well quasi order on nonnegative functions. □

Define the set of safe states in H as $\mathcal{T} = \{x \in Q^H \mid f_x \text{ is nonnegative}\}$. Intuitively, a winning strategy in the safety game H can be extended to an observation-based winning strategy in the energy game G because whenever a leaf of H is reached, there exists a $\preceq$-smaller ancestor that Player 1 can use to go on in G using the strategy played from the ancestor in H. The correctness argument is based on the fact that if Player 1 is winning from state f in H, then he is also winning from all $f' \succeq f$.

Lemma 2. *Let G be an energy game with imperfect information, and let $c_0 \in \mathbb{N}$ be an initial credit. There exists a winning observation-based strategy in G for the objective $\mathsf{PosEnergy}(c_0)$ if and only if there exists a winning strategy in H for the objective $\mathsf{Safe}(\mathcal{T})$.*

Proof. First, assume that α^o is a winning observation-based strategy in the game G for $\mathsf{PosEnergy}(c_0)$. We construct a function $\alpha : (\mathcal{F} \cdot \Sigma)^* \cdot \mathcal{F} \to \Sigma$ as follows. Given $\rho^H = f_0 \sigma_0 f_1 \dots f_n \in (\mathcal{F} \cdot \Sigma)^* \cdot \mathcal{F}$ such that $f_0 = f_{c_0}$ and f_{i+1} is a σ_i-successor of f_i for all $i \geq 0$, define the observation sequence $\mathsf{obs}(\rho^H) = o_0 \sigma_0 o_1 \dots o_n$ such that $\mathsf{supp}(f_i) \subseteq o_i$ for all $i \geq 0$. It is easy to see that $o_0 = \mathsf{obs}(q_0)$ and that this sequence exists and is unique.

We define $\alpha(\rho^H) = \alpha^o(\rho)$, where $\rho \in \mathsf{Prefs}(G)$ is such that $\mathsf{obs}(\rho) = \mathsf{obs}(\rho^H)$. The function α is well defined because such a prefix ρ always exists, and α^o is observation-based.

Consider an infinite sequence $\pi^H = f_0 \sigma_0 f_1 \dots$ consistent with α and such that $f_0 = f_{c_0}$ and f_{i+1} is a σ_i-successor of f_i for all $i \geq 0$. It is easy to show by induction on n that for all $q_n \in \mathsf{supp}(f_n)$, there exists a prefix $\rho = q_0 \sigma_0 q_1 \dots q_n \in \mathsf{Prefs}(G)$ consistent with α^o such that $\mathsf{obs}(\rho) = \mathsf{obs}(\pi^H(n))$, and $c_0 + \mathsf{EL}(\rho) = f_n(q_n)$ for all such prefixes ρ. Since $c_0 + \mathsf{EL}(\rho) \geq 0$ for all prefixes ρ of all outcomes of α^o in G, the previous properties imply that all functions f_i are nonnegative. since $\preceq$ is a well quasi order in nonnegative functions, there exist two positions k, l such that $k < l$ and $f_k \preceq f_l$.

Hence, if we define the memoryless strategy α^H in H such that $\alpha^H(\rho^H) = \alpha(\mathsf{Last}(\rho^H))$ for all prefixes $\rho^H \in \mathsf{Prefs}(H)$ that contain no leaf of H, then α^H is winning for the objective $\mathsf{Safe}(\mathcal{T})$ in H.

Second, assume that α^H is a winning strategy in the safety game H. We can assume that α^H is memoryless because memoryless strategies suffice for safety objectives. We construct a winning observation-based strategy α^o in G as follows. Given $\rho = q_0 \sigma_0 q_1 \dots q_n \in \mathsf{Prefs}(G)$, let $\rho^H = f_0 \sigma_0 f_1 \dots f_n$ such that $f_0 = f_{c_0}$ and f_{i+1} is a σ_i-successor of f_i with $q_{i+1} \in \mathsf{supp}(f_{i+1})$ for all $i \geq 0$. The *stack-prefix* $[\rho^H]$ of ρ^H is obtained as follows. We push the elements of ρ^H onto a stack, and whenever we push a function f_l such that there exists f_k in the stack with $k < l$ and $f_k \leq f_l$, we remove from the stack the sequence $\sigma_k f_{k+1} \dots f_l$, and we replace the suffix of ρ^H after f_l by the sequence $\sigma_l f'_{l+1} \sigma_{l+1} \dots f'_n$ such that $\mathsf{supp}(f'_{j+1}) = \mathsf{supp}(f_{j+1})$ and (assuming $f'_l = f_k$) f'_{j+1} is a σ_j-successor of f'_j for all $l \leq j < n$. Note that $f'_j \leq f_j$ for all $l \leq j < n$. The sequence on the stack at the end of this process is the stack-prefix $[\rho^H]$ of ρ^H.

It is easy to show by induction that if ρ is consistent with α^o, then $[\rho^H]$ is consistent with α^H, that $c_0 + \mathsf{EL}(\rho) \geq f_{[\rho^H]}(q_n)$, and that if $\mathsf{obs}(\rho_2) = \mathsf{obs}(\rho)$, then $[\rho_2^H] = [\rho^H]$. Therefore, the strategy α^o in G such that $\alpha^o(\rho) = \alpha^H([\rho^H])$ is observation-based.

Towards a contradiction, assume that α^o is not winning for $\mathsf{PosEnergy}(c_0)$ in G. Then, there exists a finite prefix ρ of an outcome of α^o in G such that $c_0 + \mathsf{EL}(\rho) < 0$ (take the shortest such $\rho = q_0\sigma_0 q_1 \dots q_n$). Let $[\rho^H]$ be the stack-prefix obtained by the above construction. It is easy to see that $[\rho^H] = f_0\sigma_0 f_1 \dots f_n$ occurs in some outcome of α^H, and therefore $f_n(q_n) \leq c_0 + \mathsf{EL}(\rho) < 0$, i.e., $[\rho^H] \notin \mathcal{T}$ is not a safe state, contradicting that α^H is a winning strategy in the safety game H. □

By Lemma 2, we get the decidability of the fixed initial credit problem, with non-primitive recursive upper bound since the size of the safety game H is bounded by the Ackermann function [22].

Theorem 1. *The fixed initial credit problem is decidable.*

Our algorithm for solving the initial credit problem has tight connections with the techniques used in the theory of well-structured transition systems [16]. In particular, the (finite) tree of reachable markings is used to decide the existence of an infinite execution in Petri nets. This problem can be reduced to the initial credit problem (with initial credit 0) as follows. Given a Petri net, construct a blind energy game in which Player 1 wins only by playing an infinite sequence of transitions of the Petri net. The game initially chooses a place of the Petri to monitor, and gives weight equal to the value of the initial marking in the monitored place. The effect of each transition on the monitored place is encoded by two edges in the game with weights corresponding to a decrement for the input tokens, followed by an increment for the output tokens in the place. Since the game is blind, Player 1 does not know which place is monitored, and therefore has to provide an infinite sequence of firable transitions (that never makes the energy level drop below zero in any place), ensuring that it corresponds to an infinite execution in the Petri net. This reduction gives EXPSPACE lower bound to the initial credit problem [15]. Unfortunately, it does not seem to exist an useful reduction of the initial credit problem to a decidable Petri net problem.

Note that from the proof of Lemma 2, it also follows that finite-memory strategies are sufficient to win energy games.

Corollary 1. *Finite-memory strategies are sufficient to win energy games with imperfect information.*

3.2 Unknown initial credit

We show that the unknown initial credit problem is undecidable using a reduction from the halting problem for deterministic 2-counter Minsky machines. A 2-*counter machine* M consists of a finite set of control states Q, an initial state $q_I \in Q$, a final state $q_F \in Q$, a set C of counters ($|C| = 2$) and a finite set δ_M of instructions manipulating two integer-valued counters. Instructions are of the form

$q: \ c := c + 1$ **goto** q'
$q:$ **if** $c = 0$ **then goto** q' **else** $c := c - 1$ **goto** q''.

Formally, instructions are tuples (q, α, c, q') where $q, q' \in Q$ are source and target states respectively, the action $\alpha \in \{inc, dec, 0?\}$ applies to the counter $c \in C$. We assume that M is deterministic: for every state $q \in Q$, either there is exactly one instruction $(q, \alpha, \cdot, \cdot) \in \delta_M$ and $\alpha = inc$, or there are two instructions $(q, dec, c, \cdot), (q, 0?, c, \cdot) \in \delta_M$.

A *configuration* of M is a pair (q, v) where $q \in Q$ and $v : C \to \mathbb{N}$ is a valuation of the counters. An *accepting run* of M is a finite sequence $\pi = (q_0, v_0)\delta_0(q_1, v_1)\delta_1 \dots \delta_{n-1}(q_n, v_n)$ where $\delta_i = (q_i, \alpha_i, c_i, q_{i+1}) \in \delta_M$ are instructions and (q_i, v_i) are configurations of M such that $q_0 = q_I$, $v_0(c) = 0$ for all $c \in C$, $q_n = q_F$, and for all $0 \le i < n$, we have $v_{i+1}(c) = v_i(c)$ for $c \ne c_i$, and (a) if $\alpha = inc$, then $v_{i+1}(c_i) = v_i(c_i) + 1$ (b) if $\alpha = dec$, then $v_i(c_i) \ne 0$ and $v_{i+1}(c_i) = v_i(c_i) - 1$, and (c) if $\alpha = 0?$, then $v_{i+1}(c_i) = v_i(c_i) = 0$. The corresponding *run trace* of π is the sequence of instructions $\bar{\pi} = \delta_0\delta_1 \dots \delta_{n-1}$.

The *halting problem* is to decide, given a 2-counter machine M, whether M has an accepting run. This problem is undecidable [21].

Theorem 2. *The unknown initial credit problem for energy games with imperfect information is undecidable, even for blind games.*

Proof. Given a (deterministic) 2-counter machine M, we construct a blind energy game G_M such that M has an accepting run if and only if there exists an initial credit $c_0 \in \mathbb{N}$ such that Player 1 has a winning strategy in G_M for $\mathsf{PosEnergy}(c_0)$. In particular, a strategy that plays a sequence $\#\bar{\pi}_0\#\bar{\pi}_1 \dots$ (where $\bar{\pi}_i$'s are run traces of M) is winning in G_M if and only if all but finitely many π_i's are accepting runs of M.

The alphabet of G_M is $\Sigma = \delta_M \cup \{\#\}$. The game G_M consists of an initial nondeterministic choice between several gadgets described below. Each gadget checks one property of the sequence of actions played in order to ensure that a trace of an accepting run in M is eventually played. Since the game is blind, it is not possible to see which gadget is executed, and therefore the strategy has to fulfill all properties simultaneously.

The gadget in Fig. 2 with $\sigma_1 = \#$ checks that the first symbol is a $\#$. If the first symbol is not $\#$, then the energy level drops below 0 no matter the initial credit. The gadget in Fig. 3 checks that a certain symbol σ_1 is always followed by a symbol σ_2, and it is used to ensure that $\#$ is followed by an instruction $(q_I, \cdot, \cdot, \cdot)$, and that every instruction $(q, \cdot, \cdot, q')$ is followed by an instruction $(q', \cdot, \cdot, q'')$, or by $\#$ if $q' = q_F$. The gadget in Fig. 4 ensures that $\#$ is played infinitely often. Otherwise, the gadget can guess the last $\#$ and jump to the middle state where no initial credit would allow to survive.

Finally, we use the gadget in Fig. 5 to check that the tests on counter c are correctly executed. It can accumulate in the energy level the increments and decrements of a counter c between the start of a run (i.e., when $\#$ occurs) and a zero test on c. A *positive cheat* occurs when $(\cdot, 0?, c, \cdot)$ is played while the counter c has positive value. Likewise, a *negative cheat* occurs when $(\cdot, dec, c, \cdot)$

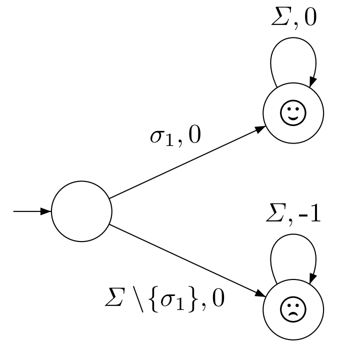

Fig. 2. Gadget to check that the first symbol is $\sigma_1 \in \Sigma$

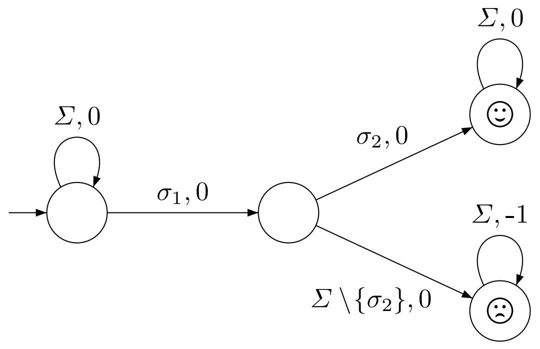

Fig. 3. Gadget to check that every symbol $\sigma_1 \in \Sigma$ is followed by $\sigma_2 \in \Sigma$.

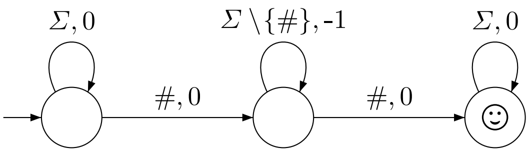

Fig. 4. Gadget to check that # is played infinitely often.

is played while the counter c has value 0. On reading the symbol #, the gadget can guess that there will be a positive or negative cheat by moving to the upper or lower state, respectively. In the upper state, the energy level simulates the operations on the counter c but with opposite effect, thus accumulating the opposite of its value changes. When a positive cheat occurs, the gadget returns to the initial state, thus decrementing the energy level. The lower state of the gadget is symmetric. A negative cheat costs one unit of energy. Note that the gadget has to go back to its initial state before the next #, as otherwise Player 1 wins.

The game G_M has such gadgets for each counter. Thus, a strategy in G_M which cheats infinitely often on a counter would not survive no matter the value of the initial credit.

The correctness of this construction is established as follows. First, assume that M has an accepting run π with trace $\bar{\pi}$. Then, the strategy playing $(\#\bar{\pi})^\omega$ is winning for $\mathsf{PosEnergy}(|\bar{\pi}|)$ because an initial credit $|\bar{\pi}|$ is sufficient to survive in the "∞-many #" gadget of Fig. 4, as well as in the zero-test gadget of Fig. 5 because all zero tests are correct in π and the counter values are bounded by $|\bar{\pi}|$. Second, if there exists a winning strategy in G_M with some finite initial credit, then the sequence played by this strategy can be decomposed into run traces separated by #, and since the strategy survived in the gadget of Fig. 5, there must be a point where all run traces played correspond to faithful simulation of M with respect to counter values, thus M has an accepting run. $\square$

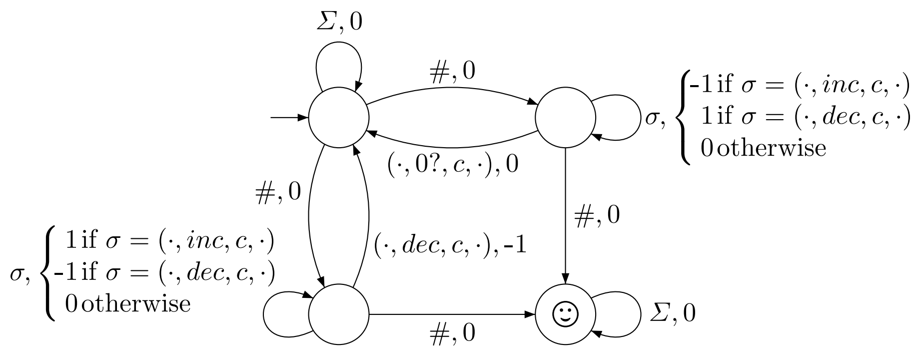

Fig. 5. Gadget to check the zero tests on counter c (assuming σ ranges over $\Sigma \setminus \{\#\}$).

4 Mean-payoff Games with Imperfect Information

In this section we consider games with imperfect information and mean-payoff objective. First, we recall that winning strategies may require infinite memory (see Remark 1 and Remark 2).

Theorem 3. *Finite-memory strategies are not sufficient in general to win mean-payoff games with imperfect information.*

We show that mean-payoff games with imperfect information are undecidable, which also shows that language inclusion and universality are undecidable for quantitative languages defined by mean-payoff functions [6].

Theorem 4. *The threshold problem for mean-payoff games with imperfect information and objective* $\mathsf{MeanPayoffSup}^{>}(0)$ *(or* $\mathsf{MeanPayoffInf}^{>}(0)$*) is undecidable (it is not co-r.e.), even for blind games.*

Proof. We give a reduction from the halting problem for two-counter machines which is similar to the proof of Theorem 2. We adapt the gadgets as follows. We increment by 1 the weights of the gadgets in Fig. 2, 3 and 4. This way, faithful simulations of the machine get mean-payoff value 1, while wrong simulations get value 0. In the gadget of Fig. 5, the symbol # gets always weight 1, thus rewarding finite run traces. The rest of the gadget is left unchanged.

We claim that Player 1 has a winning strategy if and only if the machine M halts. Indeed, if the machine halts and π is the accepting run, then playing $(\#\bar{\pi})^{\omega}$ is a winning strategy for Player 1: there is no cheat, so the zero-test gadget either stays in the initial state and performs an increment whenever it sees a # (i.e., every $|\pi|$ steps), or it tries to detect a cheat and would end up in the sink state. Therefore, this gadget gives a mean-payoff value at least $1/|\pi|$. The modified gadgets in Fig. 2, 3 and 4 also cannot punish Player 1 in any way and thus they give mean-payoff value 1.

To prove the other direction, assume that the machine M does not halt. The sequence played by Player 1 can be decomposed as a word $w = \#\bar{\pi}_1\#\bar{\pi}_2\#\bar{\pi}_3\dots$ If one of the words $\bar{\pi}_i$ is wrong with respect to the state transitions of the machine, or if there are finitely many #'s, then Player 2 uses one of the modified gadgets in Fig. 2, 3 or 4 to ensure mean-payoff value 0. Thus, we can assume that each word $\bar{\pi}_i$ is finite and respects the state transitions. Since M does not halt, in each of the sequences π_i there must be a cheat, so Player 2 can guess it and choose a run in the gadget of Fig. 5 with negative payoff that cancels out the positve payoff for #. Therefore, the outcome of the game cannot have positive mean-payoff value. □

The proof of Theorem 4 shows that the threshold problem for mean-payoff games is not co-r.e. We show that the problem is neither r.e. using a reduction from the complement of the halting problem for deterministic two-counter machines. However, the reduction requires at least two observations in the game, and it crucially relies on using non-strict inequality in the objective and on the mean-payoff value being defined using lim sup. The cases of lim inf and blind games remain open.

Theorem 5. *The threshold problem for mean-payoff games with imperfect information and objective* $\mathsf{MeanPayoffSup}^{\geq}(0)$ *is undecidable (it is not r.e.).*

In appendix B, we prove this theorem by using another reduction from two-counter machines. The main difference is that Player 1 wins in the constructed game if and only if the machine *does not* halt. Also, in contrary to the previous constructions, this one requires two different observations for Player 1.

We show that the result of Theorem 4 does not depend on the fact that winning observation-based strategies may require infinite memory: the question of the existence of a finite-memory winning strategy in MPG is also undecidable (but obviously r.e.).

Theorem 6. *Let G be a weighted game with imperfect information. A finite-memory strategy α is winning in G for* $\mathsf{MeanPayoffSup}^{\geq}(0)$ *(and also for* $\mathsf{MeanPayoffInf}^{\geq}(0)$*) if and only if there exists $c_0 \in \mathbb{N}$ such that α is winning in G for* $\mathsf{PosEnergy}(c_0)$.

Proof. ($\Leftarrow$) Let α be a finite-memory winning strategy in G defined by the Moore machine $\langle M, m_0, \alpha_u, \alpha_n\rangle$. Consider the graph G_α obtained as the product of G with M, where $(\langle m, q\rangle, \sigma, \langle m', q'\rangle)$ is a transition in G_α if $m' = \alpha_u(m, q)$, $\sigma = \alpha_n(m')$, and (q, σ, q') is a transition in G. All infinite paths in G_α from $\langle m_0, q_0\rangle$ are outcomes of α in G (if we project out the m-component) and therefore all cycles in G_α reachable from $\langle m_0, q_0\rangle$ have nonnegative sum of weights. Therefore, this ensures that an initial credit $c_0 = |M| \cdot |Q| \cdot W$ sufficient to win in G for $\mathsf{PosEnergy}(c_0)$, where W is the largest weight in G (in absolute value).
($\Rightarrow$) Let α be a finite-memory winning strategy in G for $\mathsf{PosEnergy}(c_0)$. Then $c_0 + \mathsf{EL}(\pi(n)) \geq 0$ for all outcomes π of α and all $n \geq 0$. This immediately entails that $\overline{\mathsf{MP}}(\pi) \geq 0$ (and $\underline{\mathsf{MP}}(\pi) \geq 0$). □

Corollary 2. *The problem of determining if there exists a finite-memory winning strategy in mean-payoff games with imperfect information is undecidable (it is not co-r.e.), even for blind games.*

5 Games with Visible Weights

In this section, we show that energy games with visible weights are decidable using a reduction to energy games of perfect information. The reduction is based on the subset construction used to solve imperfect information games with ω-regular objectives [7]. Intuitively, the subset construction monitors the knowledge of player 1 about the current state of the play. Because weights are visible, this knowledge is sufficient to define a weight function on the subset construction.

Given a game $G = \langle Q, q_0, \Sigma, \Delta, w, \mathsf{Obs} \rangle$ with visible weights, the *weighted subset construction* of G is the game structure of perfect information $G^{\mathsf{K}} = \langle Q^{\mathsf{K}}, q_0^{\mathsf{K}}, \Sigma, \Delta^{\mathsf{K}}, w^{\mathsf{K}} \rangle$ where:

- $Q^{\mathsf{K}} = \{s \subseteq Q \mid s \neq \emptyset \wedge \exists o \in \mathsf{Obs} : s \subseteq o\}$;
- $q_0^{\mathsf{K}} = \{q_0\}$;
- Δ^{K} contains the transitions (s_1, σ, s_2) such that $s_2 = \mathsf{post}_\sigma^G(s_1) \cap o \neq \emptyset$ for some observation $o \in \mathsf{Obs}$;
- w^{K} is such that $w^{\mathsf{K}}(s_1, \sigma, s_2) = w(q_1, \sigma, q_2)$ for all $(s_1, \sigma, s_2) \in \Delta^{\mathsf{K}}$ and $(q_1, \sigma, q_2) \in \Delta$ such that $q_1 \in s_1$ and $q_2 \in s_2$. This definition is meaningful because weights are visible.

The correctness of the weighted subset construction for solving energy games is justified by the following lemma.

Lemma 3. *Let G be an energy game with imperfect information and visible weights, and let $c_0 \in \mathbb{N}$ be an initial credit. There exists a winning observation-based strategy in G for the objective* $\mathsf{PosEnergy}(c_0)$ *if and only if there exists a winning strategy in G^{K} for the objective* $\mathsf{PosEnergy}(c_0)$.

It follows from Lemma 3 that energy games with visible weights can be solved in EXPTIME, using a pseudo-polynomial algorithm for energy games of perfect information [4, 10]. Such an algorithm is polynomial in the size of the game and in the value of largest weight. The EXPTIME upper bound follows from the fact that the size of G^{K} is exponential in the size of G, and the largest weight in G^{K} and G coincide. A matching lower bound can be obtained using a reduction from safety games. Deciding the winner of a safety game with imperfect information is EXPTIME-hard, even if the set of safe sates is a union of observations [2]. The reduction to energy games assigns weight 0 to all transitions, and adds self-loops on non-safe states labeled by every action of the alphabet with weight -1. It is easy to see that a strategy is winning in the safety game if and only if it is winning in the energy game with initial credit 0.

Theorem 7. *The fixed initial credit problem and the unknown initial credit problem for energy games with visible weights are EXPTIME-complete.*

The result of Lemma 3 can be generalized to any objective which depends only on the sequence of weights in the outcomes. In fact, the proof of Lemma 3 shows that there exist mappings from the set of observation-based strategies in G to the set of strategies in G^{K} (from α^o we construct α^{K}), and back (from α^{K} we construct α^o), such that the set of sequences of weights in the outcomes is the same for the two strategies. Therefore, the subset construction can also be used to solve mean-payoff games and we get the following result.

Theorem 8. *The threshold problem for mean-payoff games with visible weights is EXPTIME-complete.*

6 Conclusion

Energy and mean-payoff games are important classes of games that are already well-studied under the hypothesis of perfect information. This paper studies extensions of those games with imperfect information. We have shown that the picture is much more complex under that hypothesis. We have established that, although energy games with known initial credit and imperfect information remain decidable, but EXPSPACE-hard, energy games with unknown credit and mean-payoff games, even restricted to finite memory strategies, are undecidable. To regain decidability, we have shown that weights have to be visible. For this subclass the problems are EXPTIME-complete, furthermore energy games and mean-payoff games are log-space equivalent as in the simpler case of perfect information.

To complete the picture about different versions of mean-payoff games with imperfect information, an open question is the decidability of the thresholod problem for the objective $\mathsf{MeanPayoffInf}^{\geq}(0)$.

In games with imperfect information, randomized strategies are sometimes more powerful than pure strategies. For example, reachability objectives may be won with probability 1 (almost-sure winning) while no pure strategy is (surely) winning. Decision problems of interest include the threshold problem for mean-payoff games with imperfect information under almost-sure and positive winning conditions. Note that in the case of energy games, randomized strategies are not more powerful than pure strategies for almost-sure winning because the energy objectives define closed sets in the Cantor topology, and it is known that almost-sure winning coincides with sure winning for closed objectives (such as safety objectives).

References

1. R. Alur, A. Degorre, O. Maler, and G. Weiss. On omega-languages defined by mean-payoff conditions. In *Proc. of FOSSACS: Foundations of Software Science and Comp. Structures*, LNCS 5504, pages 333–347. Springer, 2009.
2. D. Berwanger and L. Doyen. On the power of imperfect information. In *Proc. of FSTTCS*, Dagstuhl Seminar Proceedings 08004. IBFI, 2008.

3. R. Bloem, K. Chatterjee, T. A. Henzinger, and B. Jobstmann. Better quality in synthesis through quantitative objectives. In *Proc. of CAV: Computer Aided Verification*, LNCS 5643, pages 140–156. Springer, 2009.
4. J. Chaloupka and L. Brim. Faster algorithm for mean-payoff games. In *Proc. of MEMICS: Math. and Engineering Methods in Comp. Science*, pages 45–53. NOVPRESS, 2009.
5. K. Chatterjee, L. de Alfaro, M. Faella, T. A. Henzinger, R. Majumdar, and M. Stoelinga. Compositional quantitative reasoning. In *Proc. of QEST: Quantitative Evaluaiton of Systems*, pages 179–188. IEEE Computer Society, 2006.
6. K. Chatterjee, L. Doyen, and T. A. Henzinger. Quantitative languages. In *Proceedings of CSL 2008: Computer Science Logic*, Lecture Notes in Computer Science 5213, pages 385–400. Springer-Verlag, 2008.
7. K. Chatterjee, L. Doyen, T. A. Henzinger, and J.-F. Raskin. Algorithms for omega-regular games of incomplete information. *Logical Methods in Computer Science*, 3(3:4), 2007.
8. L. de Alfaro. How to specify and verify the long-run average behavior of probabilistic systems. In *Proc. of LICS: Logic in Computer Science*, pages 454–465. IEEE Comp. Soc., 1998.
9. L. E. Dickson. Finiteness of the odd perfect and primitive abundant numbers with n distinct prime factors. *Am. J. of Mathematics*, 35(4):413–422, 1913.
10. L. Doyen, R. Gentilini, and J.-F. Raskin. Faster pseudopolynomial algorithms for meanpayoff games. Technical Report 2009.120, Université Libre de Bruxelles (ULB), Belgium, 2009.
11. M. Droste and P. Gastin. Weighted automata and weighted logics. *Th. C. Sci.*, 380(1-2):69–86, 2007.
12. A. Ehrenfeucht and J. Mycielski. International journal of game theory. *Positional Strategies for Mean-Payoff Games*, 8:109–113, 1979.
13. E. A. Emerson and C. Jutla. Tree automata, mu-calculus and determinacy. In *Proc. of FOCS*, pages 368–377. IEEE, 1991.
14. E. A. Emerson, C. S. Jutla, and A. P. Sistla. On model-checking for fragments of μ-calculus. In *Proc. of CAV: Computer Aided Verification*, LNCS 697, pages 385–396. Springer, 1993.
15. J. Esparza. Decidability and complexity of Petri net problems - An introduction. In *Lectures on Petri Nets I*, LNCS 1491, pages 374–428. Springer, 1998.
16. A. Finkel and P. Schnoebelen. Well-structured transition systems everywhere! *Th. Comp. Sc.*, 256(1-2):63–92, 2001.
17. E. Grädel, W. Thomas, and T. Wilke, editors. *Automata, Logics, and Infinite Games: A Guide to Current Research*, LNCS 2500. Springer, 2002.
18. Y. Gurevich and L. Harrington. Trees, automata, and games. In *Proc. of STOC*, pages 60–65. ACM Press, 1982.
19. D. Krob. The equality problem for rational series with multiplicities in the tropical semiring is undecidable. In *Proc. of ICALP: Automata, Languages and Programming*, LNCS 623, pages 101–112. Springer, 1992.
20. O. Kupferman and Y. Lustig. Lattice automata. In *Proc. 8th VMCAI: Verification, Model Checking, and Abstract Interpretation*, Lecture Notes in Computer Science 4349, pages 199–213. Springer-Verlag, 2007.
21. N. M. Minsky. *Finite and Infinite Machines*. Prentice-Hall, 1967.
22. C. Rackoff. The covering and boundedness problems for vector addition systems. *Th. Comp. Sc.*, 6:223–231, 1978.
23. J. H. Reif. The complexity of two-player games of incomplete information. *Journal of Computer and System Sciences*, 29(2):274–301, 1984.
24. U. Zwick and M. Paterson. The complexity of mean payoff games on graphs. *Th. Comp. Sc.*, 158:343–359, 1996.

A Proof of Lemma 3

Proof. ($\Leftarrow$) First, let α^o be a winning observation-based strategy in G for $\mathsf{PosEnergy}(c_0)$. We construct a winning strategy α^{K} in G^{K} as follows. Given $\rho^{\mathsf{K}} = s_0\sigma_0 s_1 \dots s_n \in \mathsf{Prefs}(G^{\mathsf{K}})$, the observation $\mathsf{obs}(\rho^{\mathsf{K}})$ of ρ^{K} is the sequence $o_0\sigma_0 o_1 \dots o_n$ such that $o_0 = \mathsf{obs}(q_0)$ and $s_i \subseteq o_i$ for all $i \geq 1$. It is easy to see that this sequence exists and is unique.

We define $\alpha^{\mathsf{K}}(\rho^{\mathsf{K}}) = \alpha^o(\rho)$, where $\rho \in \mathsf{Prefs}(G)$ is such that $\mathsf{obs}(\rho) = \mathsf{obs}(\rho^{\mathsf{K}})$. The strategy α^{K} is well defined because such a prefix ρ always exists, and α^o is observation-based. Towards a contradiction, assume that α^{K} is not winning for $\mathsf{PosEnergy}(c_0)$. Then, there exists a finite prefix ρ^{K} of an outcome of α^{K} in G^{K} such that $c + \mathsf{EL}(\rho^{\mathsf{K}}) < 0$. Let $\rho \in \mathsf{Prefs}(G)$ such that $\mathsf{obs}(\rho) = \mathsf{obs}(\rho^{\mathsf{K}})$. It is easy to see that ρ is a prefix of an outcome of α^o in G, and $\mathsf{EL}(\rho^{\mathsf{K}}) = \mathsf{EL}(\rho)$. Therefore, $c + \mathsf{EL}(\rho) < 0$, contradicting that α^o is winning in G for $\mathsf{PosEnergy}(c_0)$.

($\Rightarrow$) Second, let α^{K} be a winning strategy in G^{K} for $\mathsf{PosEnergy}(c_0)$. We construct an winning observation-based strategy α^o in G as follows. Given $\rho = q_0\sigma_0 q_1 \dots q_n \in \mathsf{Prefs}(G)$, let $\rho^{\mathsf{K}} = s_0\sigma_0 s_1 \dots s_n$ such that $s_0 = \{q_0\}$ and $s_{i+1} = \mathsf{post}^G_{\sigma_i}(s_i) \cap \mathsf{obs}(q_{i+1})$ for all $0 < i < n$.

It is easy to see that $\rho^{\mathsf{K}} \in \mathsf{Prefs}(G^{\mathsf{K}})$, that $\mathsf{EL}(\rho) = \mathsf{EL}(\rho^{\mathsf{K}})$, and that if $\mathsf{obs}(\rho_2) = \mathsf{obs}(\rho)$, then $\rho_2^{\mathsf{K}} = \rho^{\mathsf{K}}$. Therefore, the strategy α^o in G such that $\alpha^o(\rho) = \alpha^{\mathsf{K}}(\rho^{\mathsf{K}})$ is observation-based.

Towards a contradiction, assume that α^o is not winning for $\mathsf{PosEnergy}(c_0)$. Then, there exists a finite prefix ρ of an outcome of α^o in G such that $c + \mathsf{EL}(\rho) < 0$. Let ρ^{K} be the prefix obtained by the above construction. It is easy to show by induction that ρ^{K} is a prefix of an outcome of α^{K}. Therefore, $c + \mathsf{EL}(\rho^{\mathsf{K}}) < 0$, contradicting that α^{K} is winning in G^{K} for $\mathsf{PosEnergy}(c_0)$. $\square$

B Proof of Theorem 5

Proof. Given a (deterministic) 2-counter machine M, we construct a mean-payoff game G_M such that M does *not* halt if and only if Player 1 has a winning strategy in G_M for $\mathsf{MeanPayoffSup}^{\geq}(0)$. The game G_M consists of an initial nondeterministic choice in state q_0 between several gadgets, and some gadgets have the ability to "restart" the game by jumping back to the initial state q_0 (see illustration in Fig. 6). Player 1 needs to be informed when such a restart happens, and therefore the observation set in G_M is $\mathsf{Obs} = \{\{q_0\}, Q \setminus \{q_0\}\}$. If the machine M does not halt, then a winning strategy for Player 1 is to faithfully simulate M (and restart the simulation from the beginning whenever the game gets back in q_0).

The alphabet of G_M is $\Sigma = \delta_M \cup \{(last0, c) \mid c \in C\}$ where $(last0, c)$ is a special action used by Player 1 to announce that counter c will no longer be equal to 0 in the run of M. Details are given later. Below, we describe the gadgets, which are designed to make Player 1 loose if he visits the final state of M, or if he cheats in the simulation of M.

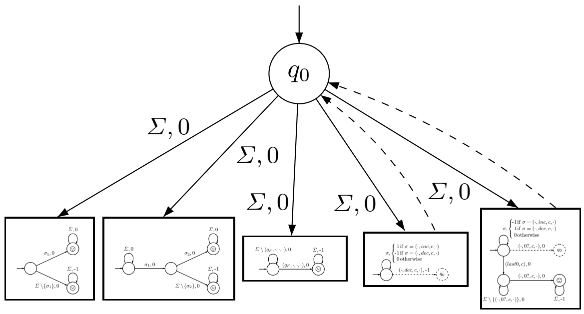

Fig. 6. The mean-payoff game G_M.

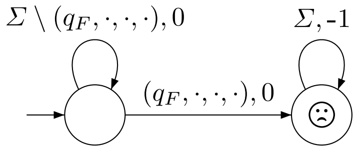

Fig. 7. Gadget to ensure that Player 1 looses if the final state is reached.

We instantiate the gadget of Fig. 2 with σ_1 of the form $(q_I, \cdot, \cdot, \cdot)$, and the gadget of Fig. 3 with σ_1 and σ_2 of the form $(q, \cdot, \cdot, q')$ and $(q', \cdot, \cdot, q'')$ respectively. These gadgets ensure that Player 1 provides a sequence of instructions corresponding to a path in the underlying graph of M. The gadget in Fig. 7 gives a negative mean-payoff value if player 1 simulates an accepting run.

The gadgets in Fig. 8 and Fig. 9 are used to check that the (zero and nonzero) tests on the counters are correctly executed. For nonzero tests, the gadget in Fig. 8 accumulates the increments and decrements of counter c, and if Player 1 cheats in decrementing a counter with value 0, then it can guess the cheat and jump back to the initial state q_0 of the game, giving a negative weight to this part of the play. Otherwise, if Player 1 is faithful, then the sum of weights would be nonnegative along a play in this gadget. For zero tests, the gadget in Fig. 9 accumulates the opposite of the value of counter c. Hence, the weight of the play drops by the value $v(c)$ of the counter c on a zero test, forcing Player 1 to ensure that $v(c) = 0$. Moreover, if Player 1 is faithful, we need to ensure that if the game remains in this gadget forever, then the average weight of the play is infinitely often 0 (so that the limsup of the average is at least 0). To do this, note that the counter c is either infinitely often equal to 0 in the run of M, or it eventually remains positive. In the first case, the mean-payoff value is nonnegative in the gadget, and in the second case, Player 1 uses the action $(last0, c)$ to announce that c will remain always positive. This implies that Player 1 no longer can use zero-tests from that point on, which is also checked by the gadget.

We can now prove that Player 1 has a winning strategy for $\mathsf{MeanPayoffSup}^{\geq}(0)$ in G_M if and only if M does not halt.

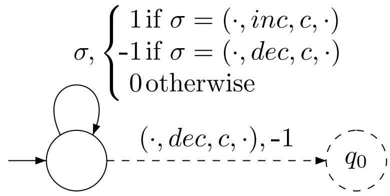

Fig. 8. Gadget to check nonzero tests on counter c.

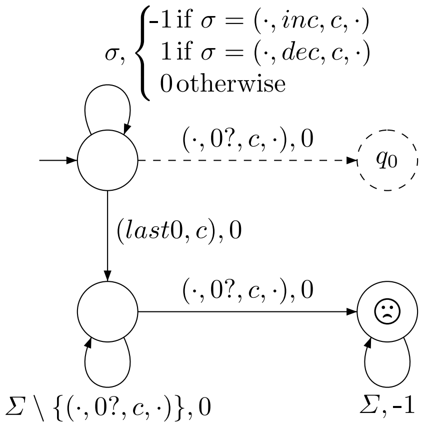

Fig. 9. Gadget to check zero tests on counter c.

($\Leftarrow$) Assume that M does not halt. We show that the strategy consisting in doing a faithful simulation of M whenever state q_0 is visited ensures a nonnegative mean-payoff value (assuming that $(last0, c)$ is played when a counter remain positive). Let π be an outcome of this strategy. We consider two cases.

1. If q_0 is visited finitely many times in π, then it is easy to see that the gadgets in Fig. 2, Fig. 3, and Fig. 7 give mean-payoff value 0. The same holds if after the last visit to q_0, the game is in the gadget of Fig. 8 because the sum of the weights in that gadget is equal to the value of the counter, thus nonnegative. Therefore, no matter the length of the finite prefix of π up to the last visit to q_0, the long-run average of the weight in π is nonnegative. Finally, if the game is in the gadget of Fig. 9 after the last visit to q_0, then either π can be decomposed into infinitely many cycles of weight 0 (between successive transitions labelled $(\cdot, 0?, c, \cdot)$), or $(last0, c)$ occurs and π has an (infinite) suffix with all weights 0. In the two cases[3], the mean-payoff value of π is nonnegative.
2. If q_0 is visited infinitely often in π, then π can be decomposed into cycles (between successive visits to q_0), and since Player 1 is always faithfully sim-

[3] Note that in the first case, we crucially need that $\overline{\mathsf{MP}}$ is defined using lim sup.

ulating M from q_0, the sum of the weights in every cycles is nonnegative no matter which gadget is executed (by similar arguments as above).

($\Rightarrow$) By contradiction, assume that M halts and let ℓ be the length of the (unique) accepting run of M. We show that all strategies α of Player 1 are loosing for $\mathsf{MeanPayoffSup}^{\geq}(0)$. The result is clear if α cheats in the sequence of control states of M, or in using the labels $(last0, c)$ because in those cases the mean-payoff value is -1 in one of the gadgets in Fig. 2, Fig. 3 or Fig. 9. Otherwise, we consider two cases.

1. If Player 1 eventually faithfully simulates[4] M during ℓ steps after a visit to q_0, then the final state q_F is reached and the strategy is losing in the gadget of Fig. 7.
2. If Player 1 always cheats on the counter tests within ℓ steps after a visit to q_0, then the sum of the weights in the gadget of Fig. 8 or Fig. 9 is at most -1 and the game gets back to q_0. Therefore, within every at most ℓ steps, the sum of the weights is -1, and the mean-payoff value of the outcome is at most $-\frac{1}{\ell}$.

□

[4] i.e., also according to the counter tests.