Rational Verification for Nash and Subgame-Perfect Equilibria in Graph Games

Léonard Brice ✉
Université Libre de Bruxelles, Belgium

Jean-François Raskin ✉
Université Libre de Bruxelles, Belgium

Marie van den Bogaard ✉
Univ Gustave Eiffel, CNRS, LIGM, F-77454 Marne-la-Vallée, France

Abstract

We study a natural problem about rational behaviors in multiplayer non-zero-sum sequential infinite duration games played on graphs: rational verification, that consists in deciding whether all the rational answers to a given strategy satisfy some specification. We give the complexities of that problem for two major concepts of rationality: Nash equilibria and subgame-perfect equilibria, and for three major classes of payoff functions: energy, discounted-sum, and mean-payoff.

2012 ACM Subject Classification Software and its engineering → Formal methods; Theory of computation → Logic and verification; Theory of computation → Solution concepts in game theory

Keywords and phrases Games on graphs, Nash equilibria, subgame-perfect equilibria

Digital Object Identifier 10.4230/LIPIcs.MFCS.2023.26

Related Version *Full Version*: https://arxiv.org/abs/2301.12913

1 Introduction

Formal methods are essential to guarantee the correctness of safety critical computer systems. Techniques like model-checking [13] or automated theorem proving [15] are now routinely used to develop systematically hardware pieces as well as embedded control systems. Nevertheless, there are contexts in which formal methods have not yet been applied successfully large-scale: that is the case of multi-agent systems, which still represent a challenge for formal verification techniques, because they are usually composed of *heterogeneous* components, ranging from traditional pieces of reactive code to wholly autonomous robots or human users. Producing operational model abstractions for this diversity of sub-systems is often challenging.

While it may be inconvenient, to say the least, to produce an operational model of the behavior of a human or a complex autonomous robot, identifying the high level objectives of those components may be easier. Taking into account those objectives is often key for reasoning about the correctness of a system that interacts with those components. Indeed, a system is usually not supposed to be correct in all circumstances, but only when agents in its environment behave in a way that concurs with their own objectives. In *rational verification* (RV), a system needs to enforce some property, not in all possible executions, but only in those in which the environment agents behave rationally with regards to their own objectives.

Rationality is the focus point of game theory, and can be formalized in several ways: for instance, with the notion of *Nash equilibrium* (NE) [24]. NEs have been used in a few promising contributions, like in verification of non-repudiation and fair exchange protocols [12, 20, 21], or planning of self-driving cars interacting with human drivers [26], etc. Nevertheless, those works do not propose a general framework for RV and their contributions are rather specific to their application domains. There is thus a need for more systematic study of formal frameworks for RV. Such a study has been started recently: for instance, the authors of [18]

48th International Symposium on Mathematical Foundations of Computer Science (MFCS 2023).
Editors: Jérôme Leroux, Sylvain Lombardy, and David Peleg; Article No. 26; pp. 26:1–26:15
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

study the automatic verification of an LTL specification in multi-agent systems that behave according to an NE, and in [11], the authors study a setting in which the environment has multiple objectives and only produces behaviors that are Pareto-optimal with regards to them. This work contributes to that line of research by considering a notion of rationality formalized by *subgame-perfect equilibria* (SPEs), a refinement of NEs that is better suited to formalize rationality in sequential games, in which NEs suffer from non-credible threats [25].

More precisely, we consider here the *rational verification problem*, which takes as inputs: (*i*) a multiplayer game graph with a designated player called *Leader*, (*ii*) a finite state description of a (potentially infinite) set of strategies for Leader, (*iii*) a description of the objective for Leader, and (*iv*) a description of the objectives of all the other players. It asks whether for all possible fixed strategies $\sigma_{\mathbb{L}}$ of Leader (defined by the finite state description), for all possible rational responses of the other agents, the generated outcome satisfies Leader's objective. That problem is well-suited to formalize the verification of correctness of a controller interacting with an environment composed of rational agents.

Contributions. To solve the RV problems, we provide a general construction, called the *product game* (Definition 5): we show that, given a game and a finite-state description of a set of Leader's strategies, one can incorporate the memory states of that finite-state description in the arena of the game in a way that Leader is implicitly forced to follow some strategy in the set. Thus, we show that the RV problem reduces in polynomial time to the *universal threshold problem*, a problem that is easier to study algorithmically: given a game, does every equilibrium satisfy a given specification? Also, some game classes we analyze have been addressed with slightly different definitions in previous literature. Interestingly, we provide a reduction in the opposite direction as well (Cor. 6).

We use that tool to prove the undecidability of RV in energy games (Th. 9 and 10); in the case of subgame-perfect RV, we show that undecidability holds even when Leader plays against only two players. We show that Nash RV is co-recursively enumerable in those games, and leave that question open for subgame-perfect RV – but contrary to the Nash setting, SPEs may require infinite memory to reach some payoffs (Prop. 11). In discounted-sum (DS) games, we show that the RV problems are at least as hard as the *target discounted-sum problem* (Th. 13), whose decidability is an open question. However, we prove that those problems are recursively enumerable (Th. 14). In the case of mean-payoff (MP) games, Cor. 8, combined with older results, entails that the RV problems are coNP-complete. But that case highlights a subtlety in the definition of RV: if one wants to check that a strategy is such that *every* rational response satisfies the specification, then when no such response exists, the strategy will be accepted. In the case of MP games, that leads to results that can be considered as counter-intuitive. We thus propose a stronger definition of the RV problem, called *achaotic RV*, to avoid that weakness: it consists in deciding whether a strategy satisfies the specification against every response that is *as rational as it can be*, using the notions of ε-NE and ε-SPE, that are quantitative relaxations of NE and SPE. We show that such a problem is P^{NP}-complete in MP games (Th. 19), and that in every other setting (Nash or subgame-perfect RV in the two other game classes), it coincides with RV, since rational responses always exist (Prop. 17). A synopsis of those results can be found in Table 1.

Related works. During the last decade, multiplayer games and their applications to reactive synthesis have raised a growing attention: the reader may refer to [3, 9, 10, 16, 22], and their references. The concept of *rational verification* appears in [19], where Gutierrez, Najib, Perelli, and Wooldridge give the complexity of several related problems. They use a definition

Table 1 Synopsis of our results.

	Nash RV	Ach. Nash RV	SP RV	Ach. SP RV
Energy	undecidable, co-RE		undecidable	
DS	TDS-hard, RE		TDS-hard, RE	
MP	coNP-complete		coNP-complete	P^{NP}-complete

that is slightly different from ours: their problem consists in deciding, given a game and a specification, whether all NEs (or one of them) in that game satisfy the specification, without any player representing the system (Leader in our setting). Still, as we show with Cor. 8, that problem is strongly related to ours. In [27], they also study if ω-regular properties are enforced by NEs induced by mean-payoff objectives. The objectives considered in those papers are only ω-regular objectives. Moreover, both in [19] and in [27] only NEs are considered, while our main contributions are about SPEs, that are arguably better suited for reasoning about sequential games [25], but also require substantially more complex techniques. In [14], Filiot, Gentilini, and Raskin study *Stackelberg values* of mean-payoff and discounted-sum two-player non-zero sum games, i.e. the payoff that Leader gets when the other player, *Follower*, plays the *best response* that is available with regards to his own objective. This is a synthesis problem while we consider a verification problem. They consider only one player in the environment while we consider the more general case of n players.

In [28], and later in [29], Ummels studies SPEs in parity games. He proves that they always exist, and that deciding whether there exists an SPE in a given parity game that generates a payoff vector between two given thresholds (the *constrained existence problem*, very close to the *universal threshold problem* studied in this paper) is EXPTIME-easy and NP-hard. In [8], Brihaye, Bruyère, Goeminne, Raskin, and van den Bogaard, study the same problem in quantitative reachability games, and prove that it is PSPACE-complete.

In [17], Flesch and Predtetchinski give a general procedure to characterize SPEs. In [4], Brice, Raskin, and van den Bogaard introduce the *negotiation function*, a tool that turns Flesch and Predtetchinski's procedure into effective algorithms for a large class of games. In [6], they use it to close the gap left by Ummels, proving that the constrained existence problem is NP-complete in parity games, with methods that they use later in [5] to prove that the same problem is also NP-complete in mean-payoff games. An alternative procedure to solve such SPE problems is proposed in [23], where Meunier constructs a two-player zero-sum game in which one player has a winning strategy if and only if there exists an SPE satisfying the desired constraint in the input game. That technique is nevertheless often costly, because the size of the constructed game is proportional to the number of possible payoff vectors; and for the same reason, it cannot be applied to games with infinite payoff spaces.

Energy objectives have also been widely studied, in connection with the study of vector additions systems with states and Petri nets, but almost always in a two-player zero-sum setting: see for instance [2, 22, 30]. As for discounted-sum objectives, they are defined for instance by Zwick and Paterson in [31], again in a two-player zero-sum setting. They are strongly related to the target discounted-sum problem, which is a long-standing open problem, as shown in [1] by Boker, Henzinger, and Otop. To the best of our knowledge, no algorithmic results are known for those classes of objectives in a multiplayer non-zero sum setting.

Structure of the paper. In Sec. 2, we introduce the necessary background. In Sec. 3, we present the product game. In Sec. 4, we exploit it to study energy games; in Sec. 5, DS games; and in Sec. 6, MP games. The complete proofs of our results, and additional results, are given in the complete version of this paper [7].

2 Background

Graphs, games and strategies. We call *graph* a finite directed graph, i.e. a pair (V, E) where V is a finite set of *vertices* and $E \subseteq V \times V$ is a set of *edges*. The edge (u, v), written uv, is an *outgoing edge* of u. A *path* in (V, E) is a finite or infinite sequence $\alpha = \alpha_0\alpha_1 \cdots \in V^* \cup V^\omega$ such that for every index k, we have $\alpha_k\alpha_{k+1} \in E$. We write $\mathsf{Occ}(\alpha)$ (resp. $\mathsf{Inf}(\alpha)$) for the set of vertices that occur (resp. that occur infinitely often) in α. For a given index k, we write $\alpha_{\leq k} = \alpha_{<k+1} = \alpha_0 \dots \alpha_k$, and $\alpha_{\geq k} = \alpha_{>k-1} = \alpha_k\alpha_{k+1} \dots$ A *cycle* is a finite path $c = c_0 \dots c_n$ with $c_n c_0 \in E$. A finite path α is *simple* if for every two indices $k \neq \ell$, we have $\alpha_k \neq \alpha_\ell$.

We call *non-initialized game* a tuple $\mathcal{G} = (\Pi, V, (V_i)_{i\in\Pi}, E, \mu)$, where:

- Π is a finite set of *players*;
- (V, E) is a graph, in which every vertex has at least one outgoing edge;
- $(V_i)_{i\in\Pi}$ is a partition of V, in which V_i is the set of vertices *controlled* by player i;
- a *play* (resp. *history*) in the game $\mathcal{G}$ is an infinite (resp. finite) path in the graph (V, E), and the set of plays (resp. histories) in $\mathcal{G}$ is denoted by $\mathsf{Plays}\mathcal{G}$ (resp. $\mathsf{Hist}\mathcal{G}$);
- the *payoff function* $\mu : \mathsf{Plays}\mathcal{G} \to \mathbb{R}^\Pi$ maps each play π to the tuple $\mu(\pi) = (\mu_i(\pi))_{i\in\Pi}$.

Given a set of players $P \subseteq \Pi$, we often write $V_P = \bigcup_{i\in P} V_i$. When i is a player and when the context is clear, we write $-i$ for the set $\Pi \setminus \{i\}$. We often assume that a special player, called *Leader* and denoted by the symbol $\mathbb{L}$, belongs to the set Π. An *initialized game* is a pair $(\mathcal{G}, v_0)$, often written $\mathcal{G}_{\upharpoonright v_0}$, where $\mathcal{G}$ is a non-initialized game and $v_0 \in V$ is a vertex called *initial vertex*. When the context is clear, we use the word *game* for both initialized and non-initialized games. A play (resp. history) in the initialized game $\mathcal{G}_{\upharpoonright v_0}$ is a play (resp. history) that has v_0 as first vertex. The set of plays (resp. histories) in $\mathcal{G}_{\upharpoonright v_0}$ is denoted by $\mathsf{Plays}\mathcal{G}_{\upharpoonright v_0}$ (resp. $\mathsf{Hist}\mathcal{G}_{\upharpoonright v_0}$). We also write $\mathsf{Hist}_i\mathcal{G}$ (resp. $\mathsf{Hist}_i\mathcal{G}_{\upharpoonright v_0}$) for the set of histories in $\mathcal{G}$ (resp. $\mathcal{G}_{\upharpoonright v_0}$) whose last vertex is controlled by player i.

A *strategy* for player i in the initialized game $\mathcal{G}_{\upharpoonright v_0}$ is a mapping $\sigma_i : \mathsf{Hist}_i\mathcal{G}_{\upharpoonright v_0} \to V$, such that $v\sigma_i(hv)$ is an edge of (V, E) for every hv. A history h is *compatible* with a strategy σ_i if and only if $h_{k+1} = \sigma_i(h_0 \dots h_k)$ for all k such that $h_k \in V_i$. This definition naturally extends to plays. A *strategy profile* for $P \subseteq \Pi$ is a tuple $\bar{\sigma}_P = (\sigma_i)_{i\in P}$, where each σ_i is a strategy for player i in $\mathcal{G}_{\upharpoonright v_0}$. A play, or a history, is *compatible* with $\bar{\sigma}_P$ if it is compatible with every σ_i for $i \in P$. Since the σ_i's domains are pairwise disjoint, we sometimes consider $\bar{\sigma}_P$ as one function: for $hv \in \mathsf{Hist}\mathcal{G}_{\upharpoonright v_0}$ such that $v \in \bigcup_{i\in P} V_i$, we liberally write $\bar{\sigma}_P(hv)$ for $\sigma_i(hv)$ with i such that $v \in V_i$. A *complete* strategy profile, usually written $\bar{\sigma}$, is a strategy profile for Π. Exactly one play is compatible with the strategy profile $\bar{\sigma}$: we call it its *outcome* and write $\langle\bar{\sigma}\rangle$ for it. When $\bar{\tau}_P$ and $\bar{\tau}'_Q$ are two strategy profiles with $P \cap Q = \emptyset$, we write $(\bar{\tau}_P, \bar{\tau}'_Q)$ for the strategy profile $\bar{\sigma}_{P\cup Q}$ such that $\sigma_i = \tau_i$ for $i \in P$, and $\sigma_i = \tau'_i$ for $i \in Q$.

Notable classes of games. Here, we will focus on three game classes. In those classes, each player i's payoff is based on a *reward mapping* $r_i : E \to \mathbb{Q}$. Intuitively, the reward mapping gives the (positive or negative) reward that player i gets for each action. The first class, energy games, is a class of *Boolean games*, i.e. games in which all payoffs are equal either to 0 or to 1. For such games, we say that player i *loses* the play π when $\mu_i(\pi) = 0$, and *wins* it when $\mu_i(\pi) = 1$. The other games are called *quantitative*. In energy games, the players seek to keep the aggregated sum of those rewards, their *energy level*, always nonnegative. That quantity symbolizes any resource that an agent could have to store: fuel, money, ...

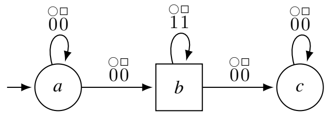


Figure 1 An example of mean-payoff game.

▶ **Definition 1** (Energy). *In a graph* (V, E)*, we associate to each reward mapping* r *the* energy level function $\mathsf{EL}_r : \mathsf{Hist}\mathcal{G} \to \mathbb{N} \cup \{\bot\}$ *defined by:*

- $\mathsf{EL}_r(h_0) = 0$*;*
- $\mathsf{EL}_r(h_{\leq n+1}) = \mathsf{EL}_r(h_{\leq n}) + r(h_n h_{n+1})$ *if* $\mathsf{EL}_r(h_{\leq n}) \neq \bot$*, and* $\mathsf{EL}_r(h_{\leq n}) + r(h_n h_{n+1}) \geq 0$*;*
- $\mathsf{EL}_r(h_{\leq n+1}) = \bot$ *otherwise.*

The game $\mathcal{G}$ *is an* energy game *if there exists a tuple* $(r_i)_{i \in \Pi}$ *of reward mappings such that for each* i *and every* π*, we have* $\mu_i(\pi) = 0$ *if* $\mathsf{EL}_{r_i}(\pi_{\leq n}) = \bot$ *for some* n*, and* $\mu_i(\pi) = 1$ *otherwise. When the context is clear, we write* EL_i *for* EL_{r_i}*.*

In discounted-sum games, each player's payoff is obtained by summing the rewards that the player obtains with some discount factor applied as the play goes along.

▶ **Definition 2** (Discounted-sum). *In a graph* (V, E)*, we define for each reward mapping* r *and each* discount factor $\lambda \in (0, 1)$ *the* discounted sum function $\mathsf{DS}_r^\lambda : h \mapsto \sum_k \lambda^k r(h_k h_{k+1})$*. Then, we write* $\mathsf{DS}_r^\lambda(\pi) = \lim_n \mathsf{DS}_r^\lambda(\pi_{\leq n})$*. The game* $\mathcal{G}$ *is a* discounted-sum game *(or DS game for short) if there exists a discount factor* $\lambda \in (0, 1) \cap \mathbb{Q}$ *and a tuple* $(r_i)_{i \in \Pi}$ *of reward mappings such that for each* i *and every* π*, we have* $\mu_i(\pi) = \mathsf{DS}_{r_i}^\lambda(\pi)$*. When the context is clear, we write* DS_i *for* $\mathsf{DS}_{r_i}^\lambda$*.*

In mean-payoff games, a players' payoff is equal to their asymptotic average reward.

▶ **Definition 3** (Mean-payoff). *In a graph* (V, E)*, we define for each reward mapping* r *the* mean-payoff function $\mathsf{MP}_r : h_0 \dots h_n \mapsto \frac{1}{n} \sum_k r(h_k h_{k+1})$*. Then, we write* $\underline{\mathsf{MP}}_r(\pi) = \liminf_n \mathsf{MP}_r(\pi_{\leq n})$*. The game* $\mathcal{G}$ *is a* mean-payoff game *(or MP game for short) if there exists a tuple* $(r_i)_{i \in \Pi}$ *of reward mappings, such that for each player* i*, we have* $\mu_i = \underline{\mathsf{MP}}_{r_i}$*. When the context is clear, we write* MP_i *for* MP_{r_i}*, and* $\underline{\mathsf{MP}}_i$ *for* $\underline{\mathsf{MP}}_{r_i}$*.*

Every game $\mathcal{G}$ from one of those three classes can be encoded with a finite number of bits. We write $\|\mathcal{G}\|$ for that number.

An example of MP game is given in Figure 1, with two players: player ○, who controls the vertices a and c, and player □, who controls the vertex b. The initial vertex is $v_0 = a$. We wrote above each edge the rewards that both players get when that edge is taken. Three types of plays are possible in that game: the one that loops on the vertex a gives both players the payoff 0; the ones that loop on the vertex b give both players the payoff 1; and the ones that loop on the vertex c give both players the payoff 0.

Equilibria and rational responses. In this paper, we study rational behaviors of players: we have, therefore, to define our rationality concepts. Let us start with the most classical one: *Nash equilibrium.* The strategy profile $\bar{\sigma}$ is a *Nash equilibrium* (resp. $\mathbb{L}$*-fixed Nash equilibrium*) – or ($\mathbb{L}$-fixed) *NE* for short – in $\mathcal{G}_{\upharpoonright v_0}$ if for each player i (resp. each player $i \neq \mathbb{L}$) and every strategy σ'_i, called *deviation of* σ_i, we have $\mu_i\left(\langle \sigma'_i, \bar{\sigma}_{-i} \rangle\right) \leq \mu_i\left(\langle \bar{\sigma} \rangle\right)$. When it is not the case, we call *profitable deviations* the deviations that do not satisfy that inequality.

As an example, in the game given in Figure 1, two types of NEs can be found: those that eventually loop on the vertex b, and give both players the payoff 1; and those that loop on a, but in which player ○ has no profitable deviation, because if she goes to the vertex b, player

□ threatens to go to the vertex c (and player □ has no profitable deviation, because he does never make any choice). However, player □'s threat is not *credible*, since going to the vertex c would give him the payoff 0, while he could stay on the vertex b and get the payoff 1. A stronger rationality concept, that avoids that phenomenon, is the one of *subgame-perfection*.

Let hv be a history in the game $\mathcal{G}$. The *subgame* of $\mathcal{G}$ after hv is the game $\mathcal{G}_{\upharpoonright hv} = \left(\Pi, V, (V_i)_i, E, \mu_{\upharpoonright hv}\right)_{\upharpoonright v}$, where $\mu_{\upharpoonright hv}$ maps each play π to its payoff in $\mathcal{G}$, assuming that the history hv has already been played, i.e. to the payoff $\mu_{\upharpoonright hv}(\pi) = \mu(h\pi)$. If σ_i is a strategy in $\mathcal{G}_{\upharpoonright v_0}$, its *substrategy* after hv is the strategy $\sigma_{i\upharpoonright hv} : h' \mapsto \sigma_i(hh')$ in the game $\mathcal{G}_{\upharpoonright hv}$.

The strategy profile $\bar{\sigma}$ is a *($\mathbb{L}$-fixed) subgame-perfect equilibrium* – or *($\mathbb{L}$-fixed) SPE* for short – in $\mathcal{G}_{\upharpoonright v_0}$ if and only if for every history h in $\mathcal{G}_{\upharpoonright v_0}$ (resp. every history h compatible with $\sigma_{\mathbb{L}}$), the strategy profile $\bar{\sigma}_{\upharpoonright h}$ is a ($\mathbb{L}$-fixed) Nash equilibrium in the subgame $\mathcal{G}_{\upharpoonright h}$.

NEs and SPEs entail two notions of rationality for the environment's responses to a strategy $\sigma_{\mathbb{L}}$ of Leader. A strategy profile $\bar{\sigma}_{-\mathbb{L}}$ is a *Nash response* to $\sigma_{\mathbb{L}}$ if the strategy profile $\bar{\sigma} = (\sigma_{\mathbb{L}}, \bar{\sigma}_{-\mathbb{L}})$ is an $\mathbb{L}$-fixed NE, and a *subgame-perfect response* if it is an $\mathbb{L}$-fixed SPE. The set of Nash (resp. subgame-perfect) responses to $\sigma_{\mathbb{L}}$ is written $\mathsf{NR}(\sigma_{\mathbb{L}})$ (resp. $\mathsf{SPR}(\sigma_{\mathbb{L}})$).

Finally, let $\rho \in \{\text{Nash}, \text{subgame-perfect}\}$. We call *$\rho$-equilibria* the NEs if $\rho = \text{Nash}$, and the SPEs if $\rho = \text{subgame-perfect}$. We will similarly talk about *$\mathbb{L}$-fixed ρ-equilibria*, and *ρ-responses*. We write $\rho\mathsf{R}(\sigma_{\mathbb{L}})$ for the set of ρ-responses to a strategy $\sigma_{\mathbb{L}}$.

Mealy machines. A *Mealy machine for player i* on a game $\mathcal{G}$ is a tuple $\mathcal{M} = (Q, q_0, \Delta)$, where Q is a finite set of *states*, where $q_0 \in Q$ is the *initial state*, and where $\Delta \subseteq (Q \times V_{-i} \times Q) \cup (Q \times V_i \times Q \times V)$ is a finite set of *transitions*, such that for every $(p, u, q, v) \in \Delta$, we have $uv \in E$, and such that for every $p \in Q$ and $u \in V$, there exists a transition (p, u, q) or $(p, u, q, v) \in \Delta$. Specialist readers will have noted that this definition is more general than the classical one, in which it is often assumed that for each p and u, there exists exactly one such transition: hereafter, such a machine will be called *deterministic*. Results about deterministic Mealy machines can be applied to *programs*, which are supposed to run deterministically; we chose to take a more general definition to capture also *protocols*, which may be given to an agent who would still have some room for manoeuvre in how they apply it.

A strategy σ_i in $\mathcal{G}_{\upharpoonright v_0}$ is *compatible* with $\mathcal{M}$ if there exists a mapping $h \mapsto q_h$ that maps every history h in $\mathcal{G}_{\upharpoonright v_0}$ to a state $q_h \in Q$, such that for every $hv \in \mathsf{Hist}_{-i}\mathcal{G}_{\upharpoonright v_0}$, we have $(q_h, v, q_{hv}) \in \Delta$, and for every $hv \in \mathsf{Hist}_i\mathcal{G}_{\upharpoonright v_0}$, we have $(q_h, v, q_{hv}, \sigma_i(hv)) \in \Delta$. The set of strategies in $\mathcal{G}_{\upharpoonright v_0}$ compatible with $\mathcal{M}$ is written $\mathsf{Comp}_{\upharpoonright v_0}(\mathcal{M})$. If $\mathcal{M}$ is deterministic, then there is exactly one strategy compatible with $\mathcal{M}$; we call it a *finite-memory* strategy.

Note that one can define analogously Mealy machines that capture a set of strategy profiles for several players, and even for the whole set Π. Note also that every Mealy machine $\mathcal{M}$ can be encoded with a finite number of bits: we write $\|\mathcal{M}\|$ for that number.

Figure 2 depicts a Mealy machine on the game of Figure 1. Each arrow from a state p to a state q labeled $u|v$ denotes the existence of a transition (p, u, q, v) (from the state p, the machine reads the vertex u, switches to the state q and outputs the vertex v). Each arrow from a state p to a state q labeled u denotes the existence of a transition (p, u, q) (from p, the machine reads u, switches to q and outputs nothing). It is a machine for player □, that is not deterministic: from the state q_0, reading the vertex b, the machine stays in q_0 but it can output either b or c. The strategies that are compatible with it can be described as follows: when player □ has to play, if the vertex a was seen an odd number of times, then he stays in b; in the opposite case, he can either stay in b or eventually go to c.

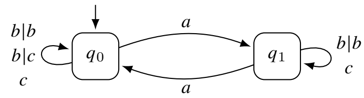


Figure 2 A non-deterministic one-player Mealy machine.

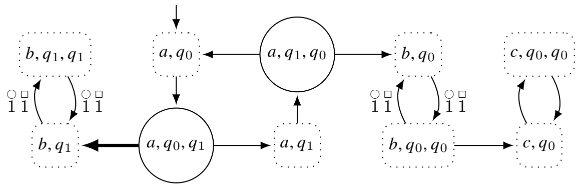


Figure 3 A product game.

Decision problem. Let us now define rational verification (RV). We define it for each game class $\mathcal{C}$, for each $\rho \in \{\text{Nash}, \text{subgame-perfect}\}$, and in both the deterministic and the non-deterministic setting.

▶ **Problem 4** ((Deterministic) ρ-rational verification problem in the class $\mathcal{C}$)**.** *Given a game $\mathcal{G}_{\restriction v_0} \in \mathcal{C}$, a threshold $t \in \mathbb{Q}$ and a (deterministic) Mealy machine $\mathcal{M}$ on $\mathcal{G}$, is every $\mathbb{L}$-fixed ρ-equilibrium $\bar{\sigma}$ with $\sigma_{\mathbb{L}} \in \mathsf{Comp}_{\restriction v_0}(\mathcal{M})$ such that $\mu_{\mathbb{L}}(\langle\bar{\sigma}\rangle) > t$?*

3 The product game

Although very intuitive, the RV problems are quite hard to study as they are. Indeed, their instances include two graph structures: a game and a Mealy machine. However, responding rationally to Leader's strategies that are compatible with $\mathcal{M}$ amounts to play rationally in a larger game, in which the machine $\mathcal{M}$ has been incorporated.

▶ **Definition 5** (Product game)**.** *Let $\mathcal{G}_{\restriction v_0}$ be a game, and let $\mathcal{M}$ be a Mealy machine for Leader in $\mathcal{G}$. Their* product game *is the game $\mathcal{G}_{\restriction v_0} \otimes \mathcal{M} = (\Pi \cup \{\mathbb{D}\}, V', (V'_i)_i, E', \mu')_{\restriction (v_0, q_0)}$ where the player $\mathbb{D}$, called* Demon*, chooses how the machine $\mathcal{M}$ will run. Formally:*

- *$V' = (V \times Q) \cup (V \times Q \times Q)$;*
- *$V'_{\mathbb{L}} = \emptyset$, $V'_i = V_i \times Q \times Q$ for every $i \in \Pi \setminus \{\mathbb{L}\}$, and $V'_{\mathbb{D}} = (V \times Q) \cup (V_{\mathbb{L}} \times Q \times Q)$;*
- *the set E' contains:*
  - *the edge $(u, p)(u, p, q)$ for each $(p, u, q) \in \Delta$ (if $u \notin V_{\mathbb{L}}$), or $(p, u, q, v) \in \Delta$ (if $u \in V_{\mathbb{L}}$);*
  - *the edge $(u, p, q)(v, q)$ for each $(p, u, q, v) \in \Delta$ (if $u \in V_{\mathbb{L}}$);*
  - *the edge $(u, p, q)(v, q)$ for each $(p, u, q) \in \Delta$, and each $uv \in E$ (if $u \notin V_{\mathbb{L}}$);*
- *each payoff function μ'_i maps every play $(\pi_0, q_0)(\pi_0, q_0, q_1)(\pi_1, q_1)\ldots$ to the payoff $\mu_i(\pi_0\pi_1\ldots)$ if $i \neq \mathbb{D}$, and to the payoff 0 if $i = \mathbb{D}$.*

Figure 3 depicts the game $\mathcal{G}_{\restriction v_0} \otimes \mathcal{M}$, when $\mathcal{G}_{\restriction v_0}$ is the game of Figure 1 and $\mathcal{M}$ the machine of Figure 2. Leader is then assimilated to player □, and Demon's vertices are represented by dotted boxes. The unreachable vertices have been omitted, and we have given only the non-zero rewards. Since, from the vertex (a, q_0, q_1), player ○ has always the possibility to go to the vertex (b, q_1) and to get the payoff 1, it can be shown that every NE and every SPE in that game gives player □ the payoff 1. As we will see now, that means that the strategies compatible with the machine $\mathcal{M}$ guarantee the payoff 1 to player □ against Nash-rational or subgame-perfect rational responses, i.e. that $\mathcal{G}_{\restriction v_0}, 1 - \varepsilon$, and $\mathcal{M}$, for every $\varepsilon > 0$, form a positive instance of the Nash and subgame-perfect RV problems.

▶ **Theorem 6.** *Let $\rho \in \{Nash, subgame\text{-}perfect\}$. Let $\mathcal{G}_{\upharpoonright v_0}$ be a game, let $\mathcal{M}$ be a Mealy machine for Leader in $\mathcal{G}$, and let $t \in \mathbb{Q}$. Then, every ρ-response $\bar{\sigma}_{-\mathbb{L}}$ to every strategy $\sigma_{\mathbb{L}} \in \mathsf{Comp}_{\upharpoonright v_0}(\mathcal{M})$ satisfies $\mu_{\mathbb{L}}(\langle\bar{\sigma}\rangle) > t$ if and only if every ρ-equilibrium $\bar{\tau}$ in the game $\mathcal{G}_{\upharpoonright v_0} \otimes \mathcal{M}$ satisfies $\mu'_{\mathbb{L}}(\langle\bar{\tau}\rangle) > t$.*

Thus, solving the ρ-RV problem in the game $\mathcal{G}_{\upharpoonright v_0}$ amounts to solve the *ρ-universal threshold problem* (ρ-UT problem) in $\mathcal{G}_{\upharpoonright v_0} \otimes \mathcal{M}$.

▶ **Problem 7** (ρ-universal threshold problem in the class C). *Given a game $\mathcal{G}_{\upharpoonright v_0} \in C$, a player $i \in \Pi$, and a threshold $t \in \mathbb{Q}$, is every ρ-equilibrium $\bar{\sigma}$ in $\mathcal{G}_{\upharpoonright v_0}$ such that $\mu_i(\langle\bar{\sigma}\rangle) > t$?*

Moreover, the size of the product game is bounded by a polynomial function of $\|\mathcal{G}\|$ and $\|\mathcal{M}\|$; and when the game $\mathcal{G}$ belongs to a class C among the three classes defined in Section 2, then all product games constructed from it also belong to C. Hence the following.

▶ **Corollary 8.** *Let C be a game class among energy games, DS games, and MP games. Then, in the class C, for a given $\rho \in \{Nash, subgame\text{-}perfect\}$, the ρ-UT problem, the ρ-RV problem, and the deterministic ρ-RV problem are reducible to each other in polynomial time.*

Proof.

- *The deterministic ρ-RV problem reduces to the ρ-RV problem,* because a non-deterministic Mealy machine is a Mealy machine.
- *The ρ-UT problem reduces to the deterministic ρ-RV problem.*
  Let $\mathcal{G}_{\upharpoonright v_0}$, i and t form an instance of the ρ-UT problem. We define the game $\mathcal{G}'_{\upharpoonright v_0}$ as equal to the game $\mathcal{G}_{\upharpoonright v_0}$, where Leader has been added to the player set, but controls no vertex. We define $\mu_{\mathbb{L}} = \mu_i$. If $\mathcal{G}$ belongs to the class C, so does $\mathcal{G}'$. Let $\mathcal{M}$ be the one-state deterministic Mealy machine on $\mathcal{G}'$ that never outputs anything. Then, a strategy profile $\bar{\sigma}$ in $\mathcal{G}'_{\upharpoonright v_0}$ is an $\mathbb{L}$-fixed ρ-equilibrium, if and only if it is an $\mathbb{L}$-fixed ρ-equilibrium with $\sigma_{\mathbb{L}} \in \mathsf{Comp}_{\upharpoonright v_0}(\mathcal{M})$, if and only if the strategy profile $\bar{\sigma}_{-\mathbb{L}}$ is a ρ-equilibrium in the game $\mathcal{G}_{\upharpoonright v_0}$. As a consequence $\mathcal{G}_{\upharpoonright v_0}$, i, and t form a positive instance of the ρ-UT problem, if and only if $\mathcal{G}'_{\upharpoonright v_0}$, $\mathcal{M}$, and t form a positive instance of the deterministic ρ-RV problem. Moreover, the latter can be constructed from the former in polynomial time.
- *The ρ-RV problem reduces to the ρ-UT problem,* by Th. 6, and since the product game $\mathcal{G}_{\upharpoonright v_0} \otimes \mathcal{M}$ can be constructed from $\mathcal{G}_{\upharpoonright v_0}$ and $\mathcal{M}$ in polynomial time. ◀

4 Energy games

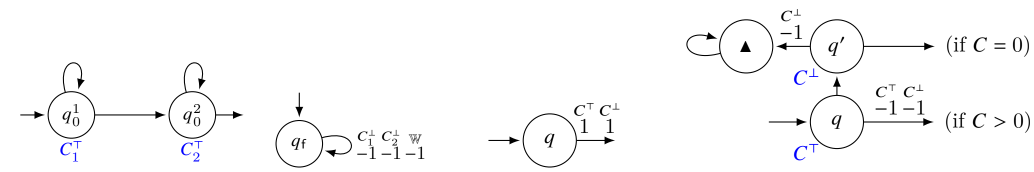
Let us now apply that result to our game classes: first, energy objectives.

Nash rational verification. RV problems are undecidable in this class, as we will show by reduction from the halting problem of two-counter machines (the reader who is not familiar with those machines may refer to [7]). However, Nash RV is co-recursively enumerable.

▶ **Theorem 9.** *In energy games, the Nash RV problem, deterministic or not, is undecidable and co-recursively enumerable.*

Proof sketch. We prove here that the Nash UT problem is undecidable and co-recursively enumerable. The theorem will follow by Cor. 8.

- *Undecidability.* We show undecidability by reduction from the halting problem of a two-counter machine. Let $\mathcal{K}$ be a two-counter machine. We define an energy game $\mathcal{G}_{\upharpoonright q_0^1}$ with five players – players $C_1^\top$, $C_1^\perp$, $C_2^\top$, $C_2^\perp$, and $\mathbb{W}$, called *Witness* – by assembling the



(a) Initial state. **(b)** Final state. **(c)** Incrementations. **(d)** Tests.

Figure 4 Gadgets.

gadgets presented in Figure 4 – the rewards that are not presented are equal to 0, and the players controlling relevant vertices are written in blue. Then, a play in $\mathcal{G}_{\upharpoonright v_0}$ that does not reach the vertex ▲ simulates a sequence of transitions of $\mathcal{K}$, that can be a valid run or not: at each step, the counter C_i is captured by the energy level of player $C_i^\top$, always equal to the energy level of player $C_i^\perp$. For each counter C_i, the player $C_i^\perp$ will have a profitable deviation if that play fakes a test to 0, by going to the vertex ▲; and the player $C_i^\top$ will lose, and therefore have a profitable deviation by staying in q_0^i if it fakes a positive test. Thus, as shown in the complete version of this proof, every NE outcome in the game $\mathcal{G}_{\upharpoonright q_0^1}$ is won by Witness if and only if the machine $\mathcal{K}$ does not terminate. As a consequence, the halting problem of two-counter machines reduces to the Nash UT problem in energy games, which is therefore undecidable.

- *Co-recursive enumerability.* As shown in the complete version of this proof, in an energy game $\mathcal{G}_{\upharpoonright v_0}$, if there exists an NE that makes some player i lose, then there exists a finite-memory one. Thus, a semi-algorithm that recognizes the negative instances of the UT problem consists in enumerating the finite-memory complete strategy profiles on $\mathcal{G}_{\upharpoonright v_0}$, and for each of them, to check (by diagonalization):
  - whether it is an NE: that is decidable (in polynomial time), by [7];
  - whether it makes player i lose: that is recursively enumerable, by constructing step by step its outcome and computing the energy levels on the fly.

  We have a negative instance of the UT problem if and only if at least one finite-memory strategy profile satisfies those two conditions. The Nash UT problem is therefore co-recursively enumerable. ◀

Subgame-perfect rational verification. In the subgame-perfect setting, the previous construction could also prove undecidability. But we choose to present a refinement of it, that proves a stronger result.

▶ **Theorem 10.** *In energy games, the subgame-perfect RV problem, deterministic or not, is undecidable, even when Leader plays against only two players.*

Again, the proof shows that, in particular, that problem is not recursively enumerable in energy games. It might still be the case that it is co-recursively enumerable. That would in particular be the case if finite memory was sufficient for an SPE to make any player i lose, when that is possible, as in the case of NEs. Unfortunately, one cannot follow this approach, because that statement is false: in order to be able to punish some player, without making another player lose, an SPE may have to memorize their energy levels, and therefore require infinite memory, as it will be the case in the example that follows. We leave therefore the question open.

▶ **Proposition 11.** *In the energy game presented in Figure 5, there exists an SPE that makes player □ lose, but no finite memory SPE can achieve that result.*

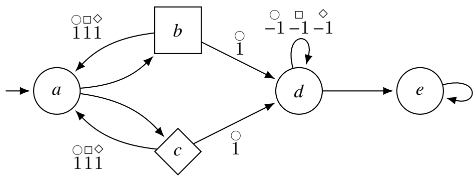


Figure 5 A game where infinite memory is necessary to make player □ lose.

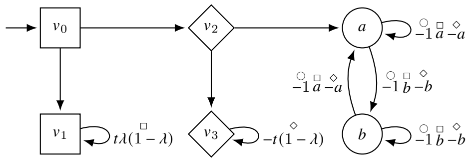


Figure 6 A game constructed from an instance of TDS.

5 Discounted-sum games

We will now move to DS objectives. First, let us define the following decision problem.

▶ **Problem 12** (Target discounted-sum problem). *Given four quantities $\lambda, a, b, t \in \mathbb{Q}$ with $0 < \lambda < 1$, is there a sequence $(u_n)_{n\in\mathbb{N}} \in \{a,b\}^\omega$ such that $\sum_{n\in\mathbb{N}} u_n\lambda^n = t$?*

Although it is a quite natural problem that appears in many different fields, the target discounted-sum (TDS) problem turns out to be surprisingly hard to solve, and its decidability status is still open. The interested reader may refer to [1] for more details. The following theorem shows that RV problems are at least as difficult.

▶ **Theorem 13.** *The TDS problem reduces to the complements of the (deterministic) Nash rational and subgame-perfect RV problems in discounted-sum games.*

Proof. We present here a reduction to the complements of the Nash universal and subgame-perfect UT problems; the result follows by Cor. 8. Let $a, b, t \in \mathbb{Q}$, let $\lambda \in \mathbb{Q} \cap (0,1)$, and let $\mathcal{G}_{\restriction v_0}$ be the DS game of Figure 6, with discount factor λ. In that game, there exists an NE $\bar{\sigma}$ with $\mu_\bigcirc(\langle\bar{\sigma}\rangle) < 0$, if and only if there exists an SPE $\bar{\sigma}$ with $\mu_\bigcirc(\langle\bar{\sigma}\rangle) < 0$, if and only if a, b, t, and λ form a positive instance of the TDS problem.

Indeed, if such an NE or SPE exists, it necessarily reaches the vertex a. But then, player □ must get at least the payoff $\mu_\Box(v_0v_1^\omega) = t\lambda^2$, and player ◇ the payoff $\mu_\Diamond(v_0v_2v_3^\omega) = -t\lambda^2$, otherwise they would have a profitable deviation. If such a play exists, then we have a positive instance of the TDS problem. Conversely, from such a positive instance, one can construct a play from v_0 in which player ○ gets the payoff $\frac{\lambda^2}{1-\lambda}$, player □ the payoff $t\lambda^2$, and player ◇ the payoff $-t\lambda^2$, and none of them has a profitable deviation in any subgame. ◀

The previous theorem suggests that finding algorithms solving those problems is a very ambitious objective. However, in the sequel, we will show that like the TDS problem, the RV problems are recursively enumerable. The key idea is the following: a property of DS objectives is that when a play gives to some player a payoff that is strictly smaller than some threshold, that can be seen on finite prefixes of those plays. Therefore, although strategy profiles are in general infinite objects that exist in uncountable number, profitable deviations can be found by analyzing their behaviors on a finite (but unbounded) number of histories.

▶ **Theorem 14.** *In DS games, the Nash rational and the subgame-perfect RV problems, deterministic or not, are recursively enumerable.*

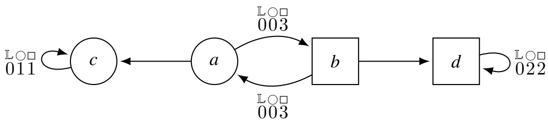


Figure 7 The temptation of chaos: an illustration.

6 Mean-payoff games

Classical rational verification. Let us now end with MP games. The reduction from RV problems to UT problems enables us to apply results and methods that already exist in the literature.

▶ **Theorem 15.** *In the class of MP games, the Nash rational and the subgame-perfect RV problems, deterministic or not, are* coNP*-complete.*

The temptation of chaos. It is now worth noting that the definition we gave of RV entails, in the case of MP games, results that may be considered as counter-intuitive. For instance, consider the game of Figure 7, where Leader owns no vertex, and consider the only (vacuous) strategy available for Leader. Does that strategy guarantee a payoff greater than 1? Intuitively, it does not, since Leader always receives the payoff 0. But still, that strategy, that game, and that threshold form a positive instance of subgame-perfect RV, because no $\mathbb{L}$-fixed SPE exists in that game (see [4]). More generally, the definition we give of RV considers that a *good* strategy for Leader is a strategy such that for every response of the environment that is rational, the generated outcome observes some specification. But a strategy is then good, in that sense, if *no* rational response of the environment exists: that is the phenomenon that we can call *temptation of chaos*. While that case does never occur in energy and DS games, where rational responses are always guaranteed to exist (as we will see below), it must be considered in MP games.

Achaotic rational verification. To avoid such phenomena, we introduce an alternative definition of RV, *achaotic RV*: a good strategy for Leader will be a strategy that guarantees the given threshold against every response that is *as rational as possible*. To define that problem, we need quantitative relaxations to the notions of NEs and SPEs. Let $\mathcal{G}_{\upharpoonright v_0}$ be a game. Let $\varepsilon \geq 0$. The strategy profile $\bar{\sigma}$ is an *ε-NE* (resp. *$\mathbb{L}$-fixed ε-NE*) in $\mathcal{G}_{\upharpoonright v_0}$ if and only if for each $i \in \Pi$ (resp. $\Pi \setminus \{\mathbb{L}\}$) and every deviation σ'_i of σ_i, the inequality $\mu_i\left(\langle \sigma'_i, \bar{\sigma}_{-i} \rangle\right) \leq \mu_i\left(\langle \bar{\sigma} \rangle\right) + \varepsilon$ holds: no deviation is profitable *by more than ε*. Note that 0-NEs coincide with NEs. We derive from that notion, as expected, the notions of ($\mathbb{L}$-fixed) ε-SPEs, ε-Nash and ε-subgame-perfect responses, and the notations $\varepsilon\mathsf{NR}(\sigma_{\mathbb{L}})$, $\varepsilon\mathsf{SPR}(\sigma_{\mathbb{L}})$, and $\varepsilon\rho\mathsf{R}(\sigma_{\mathbb{L}})$. We can now define our decision problem.

▶ **Problem 16** (Achaotic (deterministic) ρ-RV in the class $\mathcal{C}$)**.** *Given a game $\mathcal{G}_{\upharpoonright v_0} \in \mathcal{C}$, a threshold $t \in \mathbb{Q}$, and a Mealy machine (resp. a deterministic Mealy machine) $\mathcal{M}$ on $\mathcal{G}$, does there exist $\varepsilon \geq 0$ satisfying:*

- *$\varepsilon\rho\mathsf{R}(\sigma_{\mathbb{L}}) \neq \emptyset$ for some strategy $\sigma_{\mathbb{L}} \in \mathsf{Comp}_{\upharpoonright v_0}(\mathcal{M})$;*
- *and $\mu_{\mathbb{L}}(\langle \sigma_{\mathbb{L}}, \bar{\sigma}_{-\mathbb{L}} \rangle) > t$ for every $\sigma_{\mathbb{L}} \in \mathsf{Comp}_{\upharpoonright v_0}(\mathcal{M})$, and every $\bar{\sigma}_{-\mathbb{L}} \in \varepsilon\rho\mathsf{R}(\sigma_{\mathbb{L}})$?*

We will prove below that in mean-payoff games, there exists a least quantity $\varepsilon_{\min}$ such that $\varepsilon_{\min}\rho$-responses to a given strategy $\sigma_{\mathbb{L}}$ exist. For instance, in the example depicted by Figure 7, we have $\varepsilon_{\min} = 1$. Thus, we can rephrase the achaotic RV problems as follows: given a game $\mathcal{G}_{\upharpoonright v_0}$, a threshold $t \in \mathbb{Q}$ and a Mealy machine $\mathcal{M}$, do we have $\mu_{\mathbb{L}}(\langle \sigma_{\mathbb{L}}, \bar{\sigma}_{-\mathbb{L}} \rangle) > t$ for every $\sigma_{\mathbb{L}} \in \mathsf{Comp}_{\upharpoonright v_0}(\mathcal{M})$ and every $\bar{\sigma}_{-\mathbb{L}} \in \varepsilon_{\min}\rho\mathsf{R}(\sigma_{\mathbb{L}})$?

Among the problems we study here, this new definition is relevant in only one case: subgame-perfect RV in MP games. In all other cases, the RV problems are equivalent to their achaotic versions, because Nash and subgame-perfect responses are guaranteed to exist.

▶ **Proposition 17.** *Let $\mathcal{C}$ be a class of games, among the classes of energy games and DS games. Let $\rho \in \{Nash, subgame\text{-}perfect\}$. Then, the positive instances of the achaotic ρ-RV problem in $\mathcal{C}$ are exactly the positive instances of the ρ-RV problem. Similarly, the positive instances of the achaotic Nash-RV problem in MP games are exactly the positive instances of the ρ-RV problem.*

Now, an optimal algorithm for that problem in MP games requires the following lemma: in each game, there exists a least ε such that ε-SPEs exist, and it can be written with a polynomially bounded number of bits. To prove that, we need to use the notion of *negotiation function*, defined in [4]: a function from vertex labellings to vertex labellings whose least ε-fixed point (i.e., the least vertex labelling λ that is a fixed point of that function up to ε) characterizes ε-SPEs. Our result can be obtained by revisiting a proof of [5], that was designed to bound the number of bits required to write that least ε-fixed point, for a fixed ε. Hereafter, we write $\|\varepsilon\|$ for the number of bits required to write ε in a usual encoding.

▶ **Lemma 18.** *There exists a polynomial P_1 such that in every mean-payoff game $\mathcal{G}_{\upharpoonright v_0}$, there exists $\varepsilon_{\min}$ with $\|\varepsilon_{\min}\| \leq P_1(\|\mathcal{G}\|)$ such that $\varepsilon_{\min}$-SPEs exist in $\mathcal{G}_{\upharpoonright v_0}$, and ε-SPEs, for every $\varepsilon < \varepsilon_{\min}$, do not.*

We are now equipped to prove the following theorem.

▶ **Theorem 19.** *In the class of mean-payoff games, the achaotic subgame-perfect RV problem, deterministic or not, is P^{NP}-complete.*

Proof sketch. Using Lem. 18 and the same arguments as in the proof of Th. 6, those two problems are interreducible with the following one: given $\mathcal{G}_{\upharpoonright v_0}$ and a $t \in \mathbb{Q}$, does every $\varepsilon_{\min}$-SPE $\bar{\sigma}$ in $\mathcal{G}_{\upharpoonright v_0}$ satisfy $\mu_{\mathbb{L}}(\langle\bar{\sigma}\rangle) > t$? Let us prove P^{NP}-completeness for that problem.

- *Easiness.* By [5], there is an NP algorithm deciding, given ε and $\mathcal{G}_{\upharpoonright v_0}$, whether there is an ε-SPE in $\mathcal{G}_{\upharpoonright v_0}$, i.e. whether $\varepsilon \geq \varepsilon_{\min}$. Using Lem. 18 and the inequality $\varepsilon_{\min} \leq 2\max_{i,uv} |r_i(uv)|$, a dichotomous search can thus compute $\varepsilon_{\min}$ using polynomially many calls to that algorithm. Then, one last call can decide whether there exists an $\varepsilon_{\min}$-SPE $\bar{\sigma}$ such that $\mu_i(\langle\bar{\sigma}\rangle) \leq t$.
- *Hardness.* We proceed by reduction from the following P^{NP}-complete problem: given a Boolean formula φ in conjunctive normal form over the ordered variables $x_1, \dots, x_n$, is the lexicographically first valuation $\nu_{\min}$ satisfying φ such that $\nu_{\min}(x_n) = 1$? (and in particular, does such a valuation exist?) Let us write $\varphi = \bigwedge_{j=1}^{p} C_j$. We construct a game $\mathcal{G}_{\upharpoonright a}$, with a player called *Witness* and written $\mathbb{W}$, in which there exists an $\varepsilon_{\min}$-SPE $\bar{\sigma}$ such that $\mu_{\mathbb{W}}(\langle\bar{\sigma}\rangle) \leq 0$ if and only if φ is satisfiable and $\nu_{\min}(x_n) = 1$. That game, depicted in Figure 8 (unmentioned rewards are equal to 0, and we write $m = 2n + p$), has $2n + p + 4$ players: the literal players $x_1, \neg x_1, \dots, x_n, \neg x_n$; the clause players $C_1, \dots, C_p$; the player *Solver*, written $\mathbb{S}$; the player *Witness*, written $\mathbb{W}$; the player *Alice*, written $\mathbb{A}$; and the player *Bob*, written $\mathbb{B}$.
  This game is based on the classical example of MP game in which SPEs do not exist, already presented in Section 6. In the latter, from the vertex a, Alice can access a sink vertex, where Bob and her both get the payoff 1. Here, they access instead to a region where the choices of Solver define a valuation of $x_1, \dots, x_n$ – unless one of the literal players chooses to go to the sink vertex ▼, which will be a profitable deviation if Solver makes inconsistent choices (one literal and, later, its negation). That valuation ν

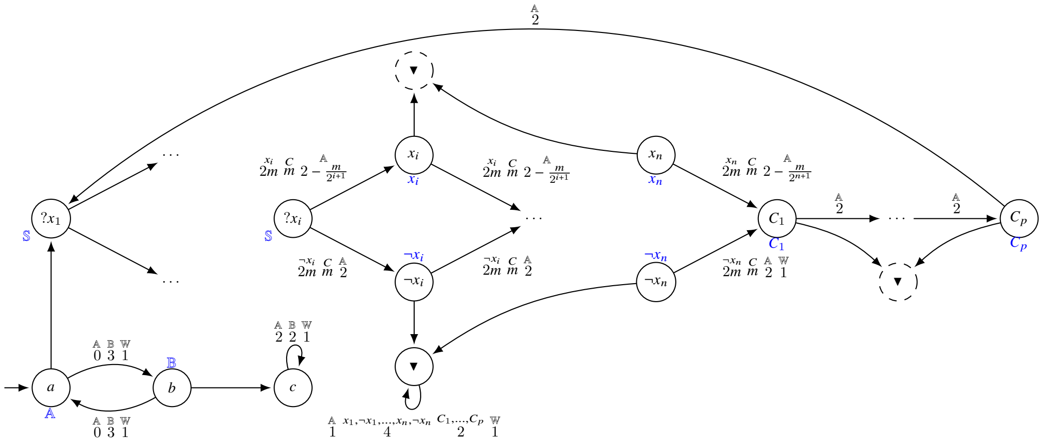


Figure 8 The game $\mathcal{G}_{\restriction a}$.

defines Alice's payoff $\mu_{\mathbb{A}}(\pi) = 2 - \sum_{i=1}^{n} \frac{\nu(x_i)}{2^i}$, and therefore defines how much deviating and reaching c is profitable for her. Consequently, as we show in the complete version of this proof, the valuation $\nu_{\min}$ is the binary encoding of the quantity $\varepsilon_{\min}$, and there is an $\varepsilon_{\min}$-SPE in which Witness gets the payoff 0 or less if and only if $\nu_{\min}(x_n) = 1$. ◀

References

1 Udi Boker, Thomas A. Henzinger, and Jan Otop. The target discounted-sum problem. In *30th Annual ACM/IEEE Symp. on Logic in Computer Science, LICS 2015, Kyoto, Japan, July 6-10, 2015*, pages 750–761. IEEE Computer Society, 2015. doi:10.1109/LICS.2015.74.

2 Patricia Bouyer, Ulrich Fahrenberg, Kim Guldstrand Larsen, Nicolas Markey, and Jirí Srba. Infinite runs in weighted timed automata with energy constraints. In Franck Cassez and Claude Jard, editors, *Formal Modeling and Analysis of Timed Systems, 6th Int. Conf., FORMATS 2008, Saint Malo, France, September 15-17, 2008. Proc.*, volume 5215 of *Lecture Notes in Computer Science*, pages 33–47. Springer, 2008. doi:10.1007/978-3-540-85778-5_4.

3 Romain Brenguier, Lorenzo Clemente, Paul Hunter, Guillermo A. Pérez, Mickael Randour, Jean-François Raskin, Ocan Sankur, and Mathieu Sassolas. Non-zero sum games for reactive synthesis. In *Language and Automata Theory and Applications – 10th Int. Conf., LATA 2016, Prague, Czech Republic, March 14-18, 2016, Proc.*, volume 9618 of *Lecture Notes in Computer Science*, pages 3–23. Springer, 2016.

4 Léonard Brice, Jean-François Raskin, and Marie van den Bogaard. Subgame-perfect equilibria in mean-payoff games. In Serge Haddad and Daniele Varacca, editors, *32nd Int. Conf. on Concurrency Theory, CONCUR 2021, August 24-27, 2021, Virtual Conference*, volume 203 of *LIPIcs*, pages 8:1–8:17. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021. doi:10.4230/LIPIcs.CONCUR.2021.8.

5 Léonard Brice, Jean-François Raskin, and Marie van den Bogaard. The complexity of SPEs in mean-payoff games. In Mikolaj Bojanczyk, Emanuela Merelli, and David P. Woodruff, editors, *49th Int. Colloquium on Automata, Languages, and Programming, ICALP 2022, July 4-8, 2022, Paris, France*, volume 229 of *LIPIcs*, pages 116:1–116:20. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022. doi:10.4230/LIPIcs.ICALP.2022.116.

6 Léonard Brice, Jean-François Raskin, and Marie van den Bogaard. On the complexity of SPEs in parity games. In Florin Manea and Alex Simpson, editors, *30th EACSL Annual Conf.on Computer Science Logic, CSL 2022, February 14-19, 2022, Göttingen, Germany (Virtual Conference)*, volume 216 of *LIPIcs*, pages 10:1–10:17. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022. doi:10.4230/LIPIcs.CSL.2022.10.

7 Léonard Brice, Jean-François Raskin, and Marie van den Bogaard. Rational verification and checking for nash and subgame-perfect equilibria in graph games. *CoRR*, abs/2301.12913, 2023. doi:10.48550/arXiv.2301.12913.

8 Thomas Brihaye, Véronique Bruyère, Aline Goeminne, Jean-François Raskin, and Marie van den Bogaard. The complexity of subgame perfect equilibria in quantitative reachability games. In Wan J. Fokkink and Rob van Glabbeek, editors, *30th Int. Conf. on Concurrency Theory, CONCUR 2019, August 27-30, 2019, Amsterdam, the Netherlands*, volume 140 of *LIPIcs*, pages 13:1–13:16. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2019. doi:10.4230/LIPIcs.CONCUR.2019.13.

9 Véronique Bruyère. Computer aided synthesis: A game-theoretic approach. In *Developments in Language Theory – 21st Int. Conf., DLT 2017, Liège, Belgium, August 7-11, 2017, Proc.*, volume 10396 of *Lecture Notes in Computer Science*, pages 3–35. Springer, 2017.

10 Véronique Bruyère. Synthesis of equilibria in infinite-duration games on graphs. *ACM SIGLOG News*, 8(2):4–29, 2021. doi:10.1145/3467001.3467003.

11 Véronique Bruyère, Jean-François Raskin, and Clément Tamines. Pareto-rational verification. In Bartek Klin, Slawomir Lasota, and Anca Muscholl, editors, *33rd Int. Conf. on Concurrency Theory, CONCUR 2022, September 12-16, 2022, Warsaw, Poland*, volume 243 of *LIPIcs*, pages 33:1–33:20. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022. doi:10.4230/LIPIcs.CONCUR.2022.33.

12 Krishnendu Chatterjee and Vishwanath Raman. Synthesizing protocols for digital contract signing. In Viktor Kuncak and Andrey Rybalchenko, editors, *Verification, Model Checking, and Abstract Interpretation – 13th Int. Conf., VMCAI 2012, Philadelphia, PA, USA, January 22-24, 2012. Proc.*, volume 7148 of *Lecture Notes in Computer Science*, pages 152–168. Springer, 2012. doi:10.1007/978-3-642-27940-9_11.

13 Edmund M. Clarke, Thomas A. Henzinger, Helmut Veith, and Roderick Bloem, editors. *Handbook of Model Checking.* Springer, 2018. doi:10.1007/978-3-319-10575-8.

14 Artur Czumaj, Anuj Dawar, and Emanuela Merelli, editors. *The Adversarial Stackelberg Value in Quantitative Games*, volume 168 of *Leibniz Int. Proc. in Informatics (LIPIcs)*, Dagstuhl, Germany, 2020. Schloss Dagstuhl–Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.ICALP.2020.127.

15 Gabriel Ebner, Sebastian Ullrich, Jared Roesch, Jeremy Avigad, and Leonardo de Moura. A metaprogramming framework for formal verification. *Proc. ACM Program. Lang.*, 1(ICFP):34:1–34:29, 2017. doi:10.1145/3110278.

16 Dana Fisman, Orna Kupferman, and Yoad Lustig. Rational synthesis. In Javier Esparza and Rupak Majumdar, editors, *Tools and Algorithms for the Construction and Analysis of Systems, 16th Int. Conf., TACAS 2010, Held as Part of the Joint European Conferences on Theory and Practice of Software, ETAPS 2010, Paphos, Cyprus, March 20-28, 2010. Proc.*, volume 6015 of *Lecture Notes in Computer Science*, pages 190–204. Springer, 2010. doi:10.1007/978-3-642-12002-2_16.

17 János Flesch and Arkadi Predtetchinski. A characterization of subgame-perfect equilibrium plays in borel games of perfect information. *Math. Oper. Res.*, 42(4):1162–1179, 2017. doi:10.1287/moor.2016.0843.

18 Julian Gutierrez, Muhammad Najib, Giuseppe Perelli, and Michael J. Wooldridge. Automated temporal equilibrium analysis: Verification and synthesis of multi-player games. *Artif. Intell.*, 287:103353, 2020. doi:10.1016/j.artint.2020.103353.

19 Julian Gutierrez, Muhammad Najib, Giuseppe Perelli, and Michael J. Wooldridge. On the complexity of rational verification. *CoRR*, abs/2207.02637, 2022. doi:10.48550/arXiv.2207.02637.

20 Steve Kremer and Jean-François Raskin. A game-based verification of non-repudiation and fair exchange protocols. In Kim Guldstrand Larsen and Mogens Nielsen, editors, *CONCUR 2001 – Concurrency Theory, 12th Int. Conf., Aalborg, Denmark, August 20-25, 2001, Proc.*, volume 2154 of *Lecture Notes in Computer Science*, pages 551–565. Springer, 2001. doi:10.1007/3-540-44685-0_37.

21 Steve Kremer and Jean-François Raskin. A game-based verification of non-repudiation and fair exchange protocols. *J. Comput. Secur.*, 11(3):399–430, 2003. URL: `http://content.iospress.com/articles/journal-of-computer-security/jcs185`, `doi:10.3233/jcs-2003-11307`.

22 Orna Kupferman, Giuseppe Perelli, and Moshe Y. Vardi. Synthesis with rational environments. *Ann. Math. Artif. Intell.*, 78(1):3–20, 2016. `doi:10.1007/s10472-016-9508-8`.

23 Noémie Meunier. *Multi-Player Quantitative Games: Equilibria and Algorithms*. PhD thesis, Université de Mons, 2016.

24 John F. Nash. Equilibrium points in n-person games. In *PNAS*, volume 36, pages 48–49. National Academy of Sciences, 1950.

25 Martin J. Osborne. *An introduction to game theory*. Oxford Univ. Press, 2004.

26 Dorsa Sadigh, Shankar Sastry, Sanjit A. Seshia, and Anca D. Dragan. Planning for autonomous cars that leverage effects on human actions. In David Hsu, Nancy M. Amato, Spring Berman, and Sam Ade Jacobs, editors, *Robotics: Science and Systems XII, University of Michigan, Ann Arbor, Michigan, USA, June 18 – June 22, 2016*, 2016. `doi:10.15607/RSS.2016.XII.029`.

27 Thomas Steeples, Julian Gutierrez, and Michael J. Wooldridge. Mean-payoff games with ω-regular specifications. In Frank Dignum, Alessio Lomuscio, Ulle Endriss, and Ann Nowé, editors, *AAMAS '21: 20th Int. Conf. on Autonomous Agents and Multiagent Systems, Virtual Event, United Kingdom, May 3-7, 2021*, pages 1272–1280. ACM, 2021. `doi:10.5555/3463952.3464099`.

28 Michael Ummels. Rational Behaviour and Strategy Construction in Infinite Multiplayer Games. Diploma thesis, RWTH Aachen, 2005. URL: `http://www.logic.rwth-aachen.de/~ummels/diplom.pdf`.

29 Michael Ummels. The complexity of nash equilibria in infinite multiplayer games. In Roberto M. Amadio, editor, *Foundations of Software Science and Computational Structures, 11th Int. Conf., FOSSACS 2008, Held as Part of the Joint European Conferences on Theory and Practice of Software, ETAPS 2008, Budapest, Hungary, March 29 – April 6, 2008. Proc.*, volume 4962 of *Lecture Notes in Computer Science*, pages 20–34. Springer, 2008. `doi:10.1007/978-3-540-78499-9_3`.

30 Yaron Velner, Krishnendu Chatterjee, Laurent Doyen, Thomas A. Henzinger, Alexander Moshe Rabinovich, and Jean-François Raskin. The complexity of multi-mean-payoff and multi-energy games. *Inf. Comput.*, 241:177–196, 2015.

31 Uri Zwick and Mike Paterson. *The complexity of mean payoff games*, volume 959, pages 1–10. Springer, April 2006. `doi:10.1007/BFb0030814`.