SUBGAME-PERFECT EQUILIBRIA IN MEAN-PAYOFF GAMES

LÉONARD BRICE ⓘ ^{a}, JEAN-FRANÇOIS RASKIN ⓘ ^{a}, AND MARIE VAN DEN BOGAARD ⓘ ^{b}

^{a} Université Libre de Bruxelles, Faculté des Sciences Campus de la Plaine - CP 212 Boulevard du Triomphe, ACC.2 1050 Bruxelles, Belgique
e-mail address: leonard.brice@ulb.be, jraskin@ulb.be

^{b} Université Gustave Eiffel, bâtiment Copernic, 5 boulevard Descartes 77420 Champs-sur-Marne, France
e-mail address: marie.van-den-bogaard@univ-eiffel.fr

ABSTRACT. In this paper, we provide an effective characterization of all the subgame-perfect equilibria in infinite duration games played on finite graphs with mean-payoff objectives. To this end, we introduce the notion of requirement, and the notion of negotiation function. We establish that the plays that are supported by SPEs are exactly those that are consistent with a fixed point of the negotiation function. Finally, we use that characterization to prove that the SPE threshold problem, whose status was left open in the literature, is decidable.

1. INTRODUCTION

The notion of Nash equilibrium (NE) is one of the most important and most studied solution concepts in game theory. A profile of strategies is an NE when no rational player has an incentive to change their strategy unilaterally, i.e. while the other players keep their strategies. Thus an NE models a stable situation. Unfortunately, it is well known that, in sequential games, NEs suffer from the problem of *non-credible threats*, see e.g. [Osb04]. In those games, some NEs only exist when some players do *not* play rationally in subgames and so use non-credible threats to force the NE. This is why, in sequential games, the stronger notion of *subgame-perfect equilibrium* is used instead: a profile of strategies is a subgame-perfect equilibrium (SPE) if it is an NE in all the subgames of the sequential game. Thus SPEs impose rationality even after a deviation has occured.

In this paper, we study sequential games that are infinite-duration games played on graphs with mean-payoff objectives, and focus on SPEs. While NEs are guaranteed to exist in infinite duration games played on graphs with mean-payoff objectives, it is known that it is not the case for SPEs, see e.g. [SV03, BBMR15]. We provide in this paper a constructive characterization of the entire set of SPEs, which allows us to decide, among others, the SPE

threshold problem. This problem was left open in previous contributions on the subject. More precisely, our contributions are described in the next paragraphs.

1.1. **Contributions.** First, we introduce two important new notions that allow us to capture NEs, and more importantly SPEs, in infinite duration games played on graphs with mean-payoff objectives[1]: the notion of *requirement* and the notion of *negotiation function*.

A requirement λ is a function that assigns to each vertex $v \in V$ of a game graph a value in $\mathbb{R} \cup \{-\infty, +\infty\}$. The value $\lambda(v)$ represents a requirement on any play $\rho = \rho_0 \rho_1 \ldots \rho_n \ldots$ that traverses this vertex: if one wants the player who controls the vertex v to follow ρ and to give up deviating from ρ, then the play must offer to that player a payoff that is at least $\lambda(v)$. An infinite play ρ is λ-*consistent* if, for each player i, the payoff of ρ for player i is larger than or equal to the largest value of λ on vertices occurring along ρ and controlled by player i.

We first use these notions to rephrase a classical result about NEs: if λ maps each vertex v to the largest value that the player who controls v can secure against a fully adversarial coalition of the other players, i.e. if $\lambda(v)$ is the zero-sum worst-case value, then the set of plays that are λ-consistent is exactly the set of plays that are supported by an NE (Theorem 3.8).

As SPEs are forcing players to play rationally in all subgames, we cannot rely on the zero-sum worst-case value to characterize them. Indeed, when considering the worst-case value, we allow adversaries to play fully adversarially after a deviation and so potentially in an irrational way w.r.t. their own objective. In fact, in an SPE, a player is refrained to deviate when opposed by a coalition of *rational adversaries*. To characterize this relaxation of the notion of worst-case value, we rely on our notion of *negotiation function*.

The negotiation function nego operates from the set of requirements into itself. To understand the purpose of the negotiation function, let us consider its application on the requirement λ that maps every vertex v to the worst-case value as above. Now, we can naturally formulate the following question: given v and λ, can the player who controls v improve the value that they can ensure against all the other players, if only plays that are consistent with λ are proposed by the other players? In other words, can this player enforce a better value when playing against the other players if those players are not willing to give away their own worst-case value? Clearly, securing this worst-case value can be seen as a minimal goal for any *rational* adversary. So $\text{nego}(\lambda)(v)$ returns this value; and this reasoning can be iterated. One of the contributions of this paper is to show that the least fixed point λ^* of the negotiation function is exactly characterizing the set of plays supported by SPEs (Theorem 4.4).

To turn this fixed point characterization of SPEs into algorithms, we additionally draw links between the negotiation function and two classes of zero-sum games, that are called *abstract* and *concrete* (see Theorem 6.3) negotiation games. The abstract negotiation game is conceptually simple but is played on an uncountably infinite graph, and therefore cannot be turned into an effective algorithm. However, it captures the intuition behind the concrete negotiation game, which is played on a finite graph. We show that in the concrete negotiation game, one of the players has a memoryless optimal strategy (Lemma 6.4), which can be used to solve it effectively. Thus, the negotiation function is computable. However, that is not

---

[1]A large part of our results apply to the larger class of games with prefix-independent objectives. For the sake of readability of this introduction, we focus here on mean-payoff games but the technical results in the paper usually cover broader classes of games.

sufficient to compute its least fixed point, because the sequence of Kleene-Tarski's iterations may require a transfinite number of steps to reach a fixed point (Theorem 5.3).

Nevertheless, we prove that the concrete negotiation game can be used to construct a geometrical representation of the fixed points of the negotiation function, from which one can extract its least fixed point (Theorem 7.3). Thus, the SPE threshold problem is decidable (Theorem 8.3).

All the previous results do also apply to ε-SPEs, a classical quantitative relaxation of SPEs — see for example [FP16]. In particular, Theorem 8.3 does also apply to the ε-SPE threshold problem.

1.2. **Related works.** Non-zero sum infinite duration games have attracted a large attention in recent years, with applications targeting reactive synthesis problems. We refer the interested reader to the following survey papers [BCH+16, Bru17] and their references for the relevant literature. We detail below contributions more closely related to the work presented here.

In [BDS13], Brihaye et al. offer a characterization of NEs in quantitative games for cost-prefix-linear reward functions based on the worst-case value. The mean-payoff is cost-prefix-linear. In their paper, the authors do not consider the stronger notion of SPE, which is the central solution concept studied in our paper. In [BMR14], Bruyère et al. study secure equilibria that are a refinement of NEs. Secure equilibria are not subgame-perfect and are, as classical NEs, subject to non-credible threats in sequential games.

In [Umm06], Ummels proves that there always exists an SPE in games with ω-regular objectives and defines algorithms based on tree automata to decide constrained SPE problems. Strategy logics, see e.g. [CHP10], can be used to encode the concept of SPE in the case of ω-regular objectives with application to the rational synthesis problem [KPV16] for instance. In [FKM+10], Flesch et al. show that the existence of ε-SPEs is guaranteed when the reward function is *lower-semicontinuous*. The mean-payoff reward function is neither ω-regular, nor lower-semicontinuous, and so the techniques defined in the papers cited above cannot be used in our setting. Furthermore, as already recalled above, see e.g. [VS03, BBMR15], contrary to the ω-regular case, SPEs in games with mean-payoff objectives may fail to exist.

In [BBMR15], Brihaye et al. introduce and study the notion of weak subgame-perfect equilibria, which is a weakening of the classical notion of SPE. This weakening is equivalent to the original SPE concept on reward functions that are *continuous*. This is the case for example for the quantitative reachability reward function, on which Brihaye et al. solve the SPE threshold problem in [BBG+19]. On the contrary, the mean-payoff cost function is not continuous and the techniques used in [BBMR15], and generalized in [BRPR17], cannot be used to characterize SPEs for the mean-payoff reward function.

In [Meu16], Meunier develops a method based on Prover-Challenger games to solve the problem of the existence of SPEs on games with a finite number of possible payoffs. This method is not applicable to the mean-payoff reward function, as the number of payoffs in this case is uncountably infinite.

In [FP17], Flesch and Predtetchinski present another characterization of SPEs on games with finitely many possible payoffs, based on a game structure that we will present here under the name of *abstract negotiation game*. Our contributions differ from this paper in two fundamental aspects. First, it lifts the restriction to finitely many possible payoffs. This is crucial as mean-payoff games violate this restriction. Instead, we identify a class of games, that we call *with steady negotiation*, that encompasses mean-payoff games and for which some

of the conceptual tools introduced in that paper can be generalized. Second, the procedure developed by Flesch and Predtetchinski is not an algorithm in computer science acceptation: it needs to solve infinitely many games that are not represented effectively, and furthermore it needs a transfinite number of iterations. On the contrary, our procedure is effective and leads to a complete algorithm in the classical sense: with guarantee of termination in finite time and applied on effective representations of games.

In [CDE+10], Chaterjee et al. study mean-payoff *automata*, and give a result that can be translated into an expression of all the possible payoff vectors in a mean-payoff game. In [BR15], Brenguier and Raskin give an algorithm to build the Pareto curve of a multi-dimensional two-player zero-sum mean-payoff game. Techniques defined in these papers are used in several technical steps of our algorithm.

1.3. **Structure of the paper.** In Section 2, we introduce the necessary background. Section 3 defines the notion of requirement and the negotiation function. Section 4 shows that the set of plays that are supported by an SPE are those that are λ-consistent, where λ is a fixed point of the negotiation function. Section 5 draws a link between the negotiation function and the abstract negotiation game. Section 6 shows that the abstract negotiation game can be transformed into a game on a finite graph, the concrete negotiation game, which can be solved to compute effectively the negotiation function. Section 7 uses the concrete negotiation game to prove that the negotiation function is a piecewise affine function, of which one can compute an effective representation. Finally, Section 8 applies these results to prove that the SPE threshold problem in mean-payoff games is decidable, 2**ExpTime**-easy and **NP**-hard.

2. Background

2.1. **Games, strategies, equilibria.** In all what follows, we will use the word *game* for the infinite duration turn-based quantitative games on finite graphs with complete information.

Definition 2.1 (Non-initialized game)**.** A *non-initialized game* — or *game* for short — is a tuple $G = (\Pi, V, (V_i)_{i \in \Pi}, E, \mu)$, where:

- Π is a finite set of *players*;
- (V, E) is a directed graph, called the *underlying graph* of G, whose vertices are sometimes called *states* and whose edges are sometimes called *transitions*, and in which every state has at least one outgoing transition. For the simplicity of writing, a transition $(v, w) \in E$ will often be written vw.
- $(V_i)_{i \in \Pi}$ is a partition of V, in which V_i is the set of states *controlled* by player i;
- $\mu : V^\omega \to \mathbb{R}^\Pi$ is an *payoff function*, that maps each infinite word ρ to the tuple $\mu(\rho) = (\mu_i(\rho))_{i \in \Pi}$ of the players' *payoffs*.

Definition 2.2 (Initialized game)**.** An *initialized game* — or game for short — is a tuple (G, v_0), often written $G_{\restriction v_0}$, where G is a non-initialized game and $v_0 \in V$ is a state called *initial state*.

When the context is clear, we often use the word *game* for both initialized and non-initialized games.

Definition 2.3 (Play, history). A *play* (resp. history) in the game G is an infinite (resp. finite) path in the graph (V, E). It is also a play (resp. history) in the initialized game $G_{\restriction v_0}$, where v_0 is its first vertex. The set of plays (resp. histories) in the game G (resp. the initialized game $G_{\restriction v_0}$) is denoted by $\mathsf{Plays}G$ (resp. $\mathsf{Plays}G_{\restriction v_0}, \mathsf{Hist}G, \mathsf{Hist}G_{\restriction v_0}$). We write $\mathsf{Hist}_i G$ (resp. $\mathsf{Hist}_i G_{\restriction v_0}$) for the set of histories in G (resp. $G_{\restriction v_0}$) of the form hv, where v is a vertex controlled by player i.

Given a play ρ (resp. a history h), we write $\mathsf{Occ}(\rho)$ (resp. $\mathsf{Occ}(h)$) the set of vertices that appear in ρ (resp. h), and $\mathsf{Inf}(\rho)$ the set of vertices that appear infinitely often in ρ. For a given index k, we write $\rho_{\leq k}$ (resp. $h_{\leq k}$), or $\rho_{<k+1}$ (resp. $h_{<k+1}$), the finite prefix $\rho_0 \ldots \rho_k$ (resp. $h_0 \ldots h_k$), and $\rho_{\geq k}$ (resp. $h_{\geq k}$), or $\rho_{>k-1}$ (resp. $h_{>k-1}$), the infinite (resp. finite) suffix $\rho_k \rho_{k+1} \ldots$ (resp. $h_k h_{k+1} \ldots h_{|h|-1}$). Finally, we write $\mathsf{first}(\rho)$ (resp. $\mathsf{first}(h)$) the first vertex of ρ, and $\mathsf{last}(h)$ the last vertex of h.

Definition 2.4 (Strategy, strategy profile). A *strategy* for player i in the initialized game $G_{\restriction v_0}$ is a function $\sigma_i : \mathsf{Hist}_i G_{\restriction v_0} \to V$, 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 $h_{k+1} = \sigma_i(h_0 \ldots h_k)$ for all k such that $h_k \in V_i$. A play ρ is compatible with σ_i if all its prefixes are.

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 $G_{\restriction v_0}$. A play or a history is *compatible* with $\bar{\sigma}_P$ if it is compatible with every σ_i for $i \in P$. 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 it $\langle \bar{\sigma} \rangle$.

When i is a player and when the context is clear, we will often write $-i$ for the set $\Pi \setminus \{i\}$. We will often refer to $\Pi \setminus \{i\}$ as the *environment* against player i. When $\bar{\tau}_P$ and $\bar{\tau}'_Q$ are two strategy profiles with $P \cap Q = \emptyset$, $(\bar{\tau}_P, \bar{\tau}'_Q)$ denotes the strategy profile $\bar{\sigma}_{P \cup Q}$ such that $\sigma_i = \tau_i$ for $i \in P$, and $\sigma_i = \tau'_i$ for $i \in Q$.

In a strategy profile $\bar{\sigma}_P$, the σ_i's domains are pairwise disjoint. Therefore, we can consider $\bar{\sigma}_P$ as one function: for $hv \in \mathsf{Hist}G_{\restriction v_0}$ 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 $v \in V_i$.

Before moving on to SPEs, let us recall the notion of Nash equilibrium.

Definition 2.5 (Nash equilibrium). Let $G_{\restriction v_0}$ be a game. A strategy profile $\bar{\sigma}$ is a *Nash equilibrium* — or *NE* for short — in $G_{\restriction v_0}$ if and only if for each player i and for every strategy σ'_i, called *deviation* of σ_i, we have the inequality $\mu_i(\langle \sigma'_i, \bar{\sigma}_{-i} \rangle) \leq \mu_i(\langle \bar{\sigma} \rangle)$.

To define SPEs, we need the notion of subgame.

Definition 2.6 (Subgame, substrategy). Let hv be a history in the game G. The *subgame* of G after hv is the game $(\Pi, V, (V_i)_i, E, \mu_{\restriction hv})_{\restriction v}$, where $\mu_{\restriction hv}$ maps each play to its payoff in G, assuming that the history hv has already been played: formally, for every $\rho \in \mathsf{Plays}G_{\restriction hv}$, we have $\mu_{\restriction hv}(\rho) = \mu(h\rho)$.

If σ_i is a strategy in $G_{\restriction v_0}$, its *substrategy* after hv is the strategy $\sigma_{i \restriction hv}$ in $G_{\restriction hv}$, defined by $\sigma_{i \restriction hv}(h') = \sigma_i(hh')$ for every $h' \in \mathsf{Hist}_i G_{\restriction hv}$.

Remark 2.7. The initialized game $G_{\restriction v_0}$ is also the subgame of G after the one-state history v_0.

Definition 2.8 (Subgame-perfect equilibrium). Let $G_{\restriction v_0}$ be a game. The strategy profile $\bar{\sigma}$ is a *subgame-perfect equilibrium* — or *SPE* for short — in $G_{\restriction v_0}$ if and only if for every history h in $G_{\restriction v_0}$, the strategy profile $\bar{\sigma}_{\restriction h}$ is a Nash equilibrium in the subgame $G_{\restriction h}$.

The notion of subgame-perfect equilibrium can be seen as a refinement of Nash equilibrium: it is a stronger equilibrium which excludes players resorting to non-credible threats.

Example 2.9. In the game represented in Figure 1a, where the square state is controlled by player $\square$ and the round states by player $\bigcirc$, if both players get the payoff 1 by reaching the state d and 0 in the other cases, there are actually two NEs: one, in blue, where $\square$ goes to the state b and then player $\bigcirc$ goes to d, and both win, and one, in red, where player $\square$ goes to c because player $\bigcirc$ was planning to go to e. However, only the blue one is an SPE, as moving from b to e is irrational for player $\bigcirc$ in the subgame $G_{\restriction ab}$.

An ε-SPE is a strategy profile which is *almost* an SPE: if a player deviates after some history, they will not be able to improve their payoff by more than a quantity $\varepsilon \geq 0$.

Definition 2.10 (ε-SPE). Let $G_{\restriction v_0}$ be a game, and $\varepsilon \geq 0$. A strategy profile $\bar{\sigma}$ from v_0 is an ε-SPE if and only if for every history hv, for every player i and every strategy σ_i', we have $\mu_i(\langle \bar{\sigma}_{-i\restriction hv}, \sigma_{i\restriction hv}' \rangle) \leq \mu_i(\langle \bar{\sigma}_{\restriction hv} \rangle) + \varepsilon$.

Remark 2.11. Note that a 0-SPE is an SPE, and conversely.

Hereafter, we focus on *prefix-independent* games, and in particular *mean-payoff* games.

2.2. Mean-payoff games.

Definition 2.12 (Mean-payoff, mean-payoff game). In a graph (V, E), we associate to each *reward function* $r : E \to \mathbb{Q}$ the *mean-payoff function*:

$$\mathsf{MP}_r : h_0 \ldots h_n \mapsto \frac{1}{n} \sum_{k=0}^{n-1} r_i(h_k h_{k+1}).$$

A game $G = (\Pi, V, (V_i)_i, E, \mu)$ is a *mean-payoff game* if its underlying graph is finite, and if there exists a tuple $(r_i)_{i \in \Pi}$ of reward functions, such that for each player i and every play ρ:

$$\mu_i(\rho) = \liminf_{n \to \infty} \mathsf{MP}_{r_i}(\rho_{\leq n}).$$

In a mean-payoff game, the quantities given by the function r_i represent the immediate reward that each action gives to player i. The final payoff of player i is their average payoff along the play, classically defined as the limit inferior over n (since the limit may not be defined) of the average payoff after n steps. When the context is clear, we liberally write $\mathsf{MP}_i(h)$ for $\mathsf{MP}_{r_i}(h)$, and $\mathsf{MP}(h)$ for $(\mathsf{MP}_i(h))_i$, as well as $r(uv)$ for $(r_i(uv))_i$.

Definition 2.13 (Prefix-independent game). A game G is *prefix-independent* if, for every history h and for every play ρ, we have $\mu(h\rho) = \mu(\rho)$. We also say, in that case, that the payoff function μ is prefix-independent.

Mean-payoff games are prefix-independent. A first important result that we need is the characterization of the set of possible payoffs in a mean-payoff game, which has been introduced in [CDE+10]. Given a graph (V, E), we write $\mathsf{SC}(V, E)$ the set of simple cycles it contains. Given a finite set D of dimensions and a set $X \subseteq \mathbb{R}^D$, we write $\mathsf{Conv}X$ the convex hull of X. We will often use the subscript notation $\mathsf{Conv}_{x \in X} f(x)$ for the set $\mathsf{Conv}f(X)$.

Definition 2.14 (Downward sealing). Given a set $Y \subseteq \mathbb{R}^D$, the *downward sealing* of Y is the set $\llcorner Y = \left\{ (\min_{\bar{z} \in Z} z_d)_{d \in D} \mid Z \text{ is a finite subset of } Y \right\}$.

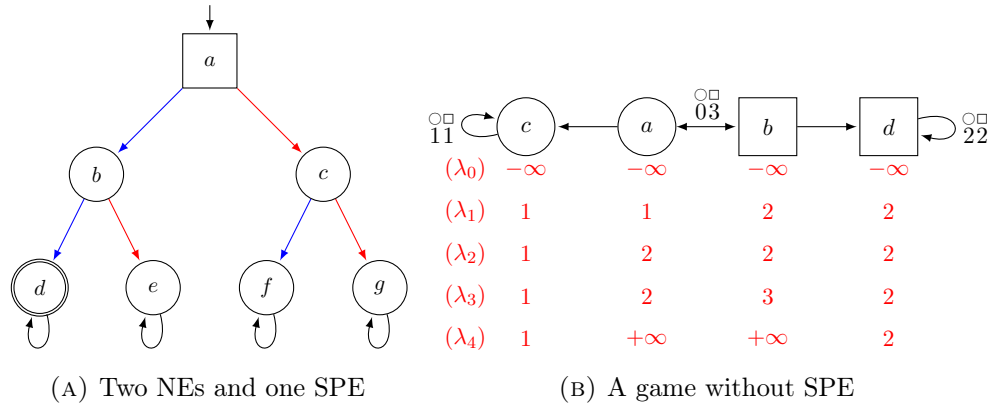(A) Two NEs and one SPE   (B) A game without SPE

FIGURE 1. Two examples of games

Example 2.15. In $\mathbb{R}^2$, if Y is the blue area in Figure 2b, then $^\llcorner Y$ is obtained by adding the gray area.

Lemma 2.16 [CDE$^+$10]**.** *Let G be a mean-payoff game, whose underlying graph is strongly connected. The set of the payoffs $\mu(\rho)$, where ρ is a play in G, is exactly the set:*

$$^\llcorner \left(\underset{c \in \mathsf{SC}(V,E)}{\mathsf{Conv}} \mathsf{MP}(c) \right).$$

2.3. **The ε-SPE threshold problem.** In the sequel, we prove the decidability of the ε-*SPE threshold problem*, which is a generalization of the *SPE threshold problem* (since SPEs are 0-SPEs and conversely, by Remark 2.11), defined as follows.

Definition 2.17 (ε-SPE threshold problem)**.** Given a rational number $\varepsilon \geq 0$, a mean-payoff game $G_{\restriction v_0}$ and two thresholds $\bar{x}, \bar{y} \in \mathbb{Q}^\Pi$, does there exist an ε-SPE $\bar{\sigma}$ in $G_{\restriction v_0}$ such that $\bar{x} \leq \mu(\langle \bar{\sigma} \rangle) \leq \bar{y}$?

That problem is illustrated by the two following examples.

Example 2.18 (A game without SPEs)**.** Let G be the mean-payoff game of Figure 1b, where each edge is labelled by the rewards $r_\bigcirc$ and $r_\square$. No reward is given for the edges ac and bd since they can be used only once, and therefore do not influence the final payoff. For now, the reader should not pay attention to the red labels below the states. As shown in [BRPR16], this game does not have any SPE, neither from the state a nor from the state b.

Indeed, the only NE outcomes from the state b are the plays where player $\square$ eventually leaves the cycle ab and goes to d: if he stays in the cycle ab, then player $\bigcirc$ would be better off leaving it, and if she does, player $\square$ would be better off leaving it before. From the state a, if player $\bigcirc$ knows that player $\square$ will leave, she has no incentive to do it before: there is no NE where $\bigcirc$ leaves the cycle and $\square$ plans to do it if ever she does not. Therefore, there is no SPE where $\bigcirc$ leaves the cycle. But then, after a history that terminates in b, player $\square$ has actually no incentive to leave if player $\bigcirc$ never plans to do it afterwards: contradiction.

Example 2.19 (A game where SPEs require infinite memory)**.** Let us now study the game of Figure 2a. By Lemma 2.16, the payoffs of possible plays in that game correspond

(A) The game G

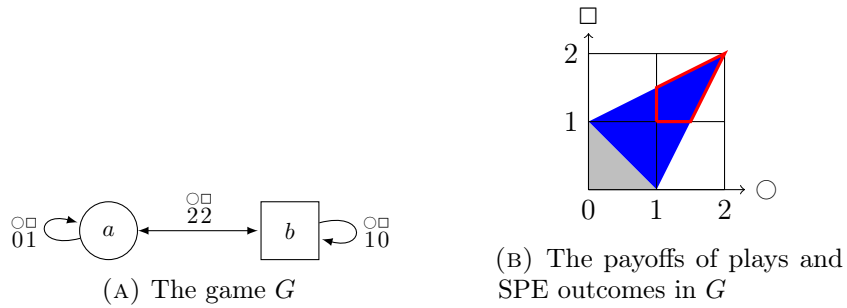(B) The payoffs of plays and SPE outcomes in G

Figure 2. A game with an infinity of SPEs

to the gray and blue areas in Figure 2b. Indeed, following exclusively one of the three simple cycles a, ab and b of the game graph during a play yields the payoffs $01, 10$ and 22, respectively. By combining those cycles with well chosen frequencies, one can obtain any payoff in the convex hull of those three points. Now, it is also possible to obtain the point 00 by using the properties of the limit inferior: it is for instance the payoff of the play $a^2 b^4 a^{16} b^{256} \ldots a^{2^{2^n}} b^{2^{2^{n+1}}} \ldots$. In fact, one can construct a play that yields any payoff in the convex hull of the four points $00, 10, 01$, and 22.

We claim that the payoffs of SPEs plays correspond to the red-circled area in Figure 2b: there exists an SPE $\bar{\sigma}$ in $G_{\restriction a}$ with $\langle \bar{\sigma} \rangle = \rho$ if and only if $\mu_\square(\rho), \mu_\bigcirc(\rho) \geq 1$. That statement will be a direct consequence of the results we show in the remaining sections, but let us give a first intuition: a play with such a payoff necessarily uses infinitely often both states. It is an NE outcome because none of the players can get a better payoff by looping forever on their state, and they can both force each other to follow that play, by threatening them to loop for ever on their state whenever they can. But such a strategy profile is clearly not an SPE.

It can be transformed into an SPE as follows: when a player deviates, say player $\square$, then player $\bigcirc$ can punish him by looping on a, not forever, but a large number of times, until player $\square$'s mean-payoff gets very close to 1. Afterwards, both players follow again the play that was initially planned. Since that threat is temporary, it does not affect player $\bigcirc$'s payoff on the long term, but it really punishes player $\square$ if that one tries to deviate infinitely often.

Not that such an SPE requires infinite memory.

2.4. **Two-player zero-sum games.** The concept of SPEs has been designed for non-zero-sum games with arbitrarily many players, but the methods we will present in the sequel will bring us back to the more classical framework of two-player zero-sum games, with more complex payoff functions. We will therefore need the following notions and results.

Definition 2.20 (Zero-sum game). A game G, with $\Pi = \{1, 2\}$, is *zero-sum* if $\mu_2 = -\mu_1$.

Definition 2.21 (Borel game). A game G is *Borel* if the function μ, from the set V^ω equipped with the product topology to the Euclidian space $\mathbb{R}^\Pi$, is Borel, i.e. if, for every Borel set $B \subseteq \mathbb{R}^\Pi$, the set $\mu^{-1}(B)$ is Borel.

Remark 2.22. Mean-payoff games are Borel.

Lemma 2.23 [Mar75]**.** *Let $G_{\restriction v_0}$ be a zero-sum Borel game, with $\Pi = \{1, 2\}$. Then, we have the following equality:*

$$\sup_{\sigma_1} \inf_{\sigma_2} \mu_1(\langle \bar{\sigma} \rangle) = \inf_{\sigma_2} \sup_{\sigma_1} \mu_1(\langle \bar{\sigma} \rangle).$$

That quantity is called *value* of $G_{\restriction v_0}$, denoted by $\mathsf{val}_1(G_{\restriction v_0})$. *Solving* a game $G_{\restriction v_0}$ means computing its value.

Definition 2.24 (Optimal strategy)**.** Let $G_{\restriction v_0}$ be a zero-sum Borel game, with $\Pi = \{1, 2\}$. The strategy σ_1 is *optimal* in $G_{\restriction v_0}$ if $\sup_{\sigma_2} \mu_1(\langle \bar{\sigma} \rangle) = \mathsf{val}_1(G_{\restriction v_0})$.

Now, let us define memoryless strategies, and state a condition under which they can be optimal.

Definition 2.25 (Memoryless strategy)**.** A strategy σ_i in a game $G_{\restriction v_0}$ is *memoryless* if for all vertices $v \in V_i$ and for all histories h and h', we have $\sigma_i(hv) = \sigma_i(h'v)$.

For every game $G_{\restriction v_0}$ and each player i, we write $\mathsf{ML}_i(G_{\restriction v_0})$, or $\mathsf{ML}(G_{\restriction v_0})$ when the context is clear, for the set of memoryless strategies for player i in $G_{\restriction v_0}$.

Definition 2.26 (Shuffling)**.** Let ρ, η and θ be three plays in a game G. The play θ is a *shuffling* of ρ and η if there exist two sequences of indices $k_0 < k_1 < \ldots$ and $\ell_0 < \ell_1 < \ldots$ such that $\eta_0 = \rho_{k_0} = \eta_{\ell_0} = \rho_{k_1} = \ldots$, and:

$$\theta = \rho_0 \ldots \rho_{k_0-1} \eta_0 \ldots \eta_{\ell_0-1} \rho_{k_0} \ldots \rho_{k_1-1} \eta_{\ell_0} \ldots \eta_{\ell_1-1} \ldots .$$

Definition 2.27 (Convexity, concavity)**.** A payoff function $\mu_i : \mathsf{Plays}\, G \to \mathbb{R}$ is *convex* if for every shuffling θ of two plays ρ and η, we have $\mu_i(\theta) \geq \min\{\mu_i(\rho), \mu_i(\eta)\}$. It is *concave* if $-\mu_i$ is convex.

Remark 2.28. Mean-payoff functions are convex.

Lemma 2.29. *In a two-player zero-sum game played on a finite graph, every player whose payoff function is concave has an optimal strategy that is memoryless.*

Proof. According to [Kop06], this result is true for qualitative objectives, i.e. when μ can only take the values 0 and 1. It follows that for every $\alpha \in \mathbb{R}$, if a player i, whose payoff function is concave, has a strategy that ensures $\mu_i(\rho) \geq \alpha$ (understood as a qualitative objective), then they have a memoryless one. Hence the equality:

$$\mathsf{val}_1(G_{\restriction v_0}) = \sup_{\sigma_1 \in \mathsf{ML}(G_{\restriction v_0})} \inf_{\sigma_2} \mu_1(\langle \bar{\sigma} \rangle).$$

Since the underlying graph (V, E) is assumed to be finite, there exists a finite number of memoryless strategies, hence the infimum above is realized by a memoryless strategy σ_1 that is, therefore, finite. $\qquad \square$

3. Requirements and negotiation

We will now see that SPEs are strategy profiles that respect some *requirements* about the payoffs, depending on the states they traverse. In this part, we develop the notions of *requirement* and *negotiation*.

3.1. **Requirement.** In the method we will develop further, we will need to analyze the players' behaviours when they have some *requirement* to satisfy. Intuitively, one can see requirements as *rationality constraints* for the players, that is, a threshold payoff value under which a player will not accept to follow a play. In all what follows, $\overline{\mathbb{R}}$ denotes the set $\mathbb{R} \cup \{\pm\infty\}$.

Definition 3.1 (Requirement). A *requirement* on the game G is a mapping $\lambda : V \to \overline{\mathbb{R}}$.

For a given state v, the quantity $\lambda(v)$ represents the minimal payoff that the player controlling v will require in a play beginning in v.

Definition 3.2 (λ-consistency). Let λ be a requirement on a game G. A play ρ in G is λ-*consistent* if and only if, for all $i \in \Pi$ and $n \in \mathbb{N}$ with $\rho_n \in V_i$, we have $\mu_i(\rho_{\geq n}) \geq \lambda(\rho_n)$. The set of the λ-consistent plays from a state v is denoted by $\lambda\mathsf{Cons}(v)$.

Definition 3.3 (λ-rationality). Let λ be a requirement on a mean-payoff game G. Let $i \in \Pi$. A strategy profile $\bar{\sigma}_{-i}$ is λ-*rational* if and only if there exists a strategy σ_i such that, for every history hv compatible with $\bar{\sigma}_{-i}$, the play $\langle \bar{\sigma}_{\restriction hv} \rangle$ is λ-consistent. We then say that the strategy profile $\bar{\sigma}_{-i}$ is λ-rational *assuming* σ_i. The set of λ-rational strategy profiles in $G_{\restriction v}$ is denoted by $\lambda\mathsf{Rat}(v)$.

Note that λ-rationality is a property of a strategy profile for all the players but one, player i. Intuitively, their rationality is justified by the fact that they collectively assume that player i will, eventually, play according to the strategy σ_i: if it is the case, then everyone gets their payoff satisfied.

Finally, let us define a particular requirement: the *vacuous requirement*, that requires nothing, and with which every play is consistent.

Definition 3.4 (Vacuous requirement). In any game, the *vacuous requirement*, denoted by λ_0, is the requirement constantly equal to $-\infty$.

3.2. **Negotiation.** We will show that SPEs in prefix-independent games are characterized by the fixed points of a function on requirements. That function captures a *negotiation* process: when a player has a requirement to satisfy, another player can hope a better payoff than what they can secure in general, and therefore update their own requirement. Note that we always use the convention $\inf \emptyset = +\infty$.

Definition 3.5 (Negotiation function, steady negotiation). Let G be a game. The *negotiation function* is the function that transforms any requirement λ on G into a requirement $\mathsf{nego}(\lambda)$ on G, such that for each $i \in \Pi$ and $v \in V_i$, we have:

$$\mathsf{nego}(\lambda)(v) = \inf_{\bar{\sigma}_{-i} \in \lambda\mathsf{Rat}(v)} \sup_{\sigma_i} \mu_i(\langle \bar{\sigma} \rangle).$$

If that infimum is realized for every λ, i and $v \in V_i$ such that $\lambda\mathsf{Rat}(v) \neq \emptyset$, then the game G is called a game *with steady negotiation*[2].

Remark 3.6. The negotiation function satisfies the following properties.

- It is monotone: if $\lambda \leq \lambda'$ (for the pointwise order, i.e. if for each v, $\lambda(v) \leq \lambda'(v)$), then $\mathsf{nego}(\lambda) \leq \mathsf{nego}(\lambda')$.

---

[2]The reader having read [BRvdB21] may note that the definition of steady negotiation has been relaxed here. The proof of Theorem 4.4 has been modified to fit with this new definition.

- It is also non-decreasing: for every λ, we have $\lambda \leq \mathsf{nego}(\lambda)$.
- There exists a λ-rational strategy profile from v against the player controlling v if and only if $\mathsf{nego}(\lambda)(v) \neq +\infty$.

In the general case, the quantity $\mathsf{nego}(\lambda)(v)$ represents the worst case value that the player controlling v can ensure, assuming that the other players play λ-rationally.

Example 3.7. Let us consider the game of Example 2.18: in Figure 1b, on the two first lines below the states, we present the requirements λ_0 and $\lambda_1 = \mathsf{nego}(\lambda_0)$, which is easy to compute since any strategy profile is λ_0-rational: for each v, $\lambda_1(v)$ is the classical *worst-case value* or *antagonistic value* of v, i.e. the best value the player controlling v can enforce against a fully hostile environment. Let us now compute the requirement $\lambda_2 = \mathsf{nego}(\lambda_1)$.

From c, there exists exactly one λ_1-rational strategy profile $\bar{\sigma}_{-\bigcirc} = \sigma_\square$, which is the empty strategy since player $\square$ has never to choose anything. Against that strategy, the best and the only payoff player $\bigcirc$ can get is 1, hence $\lambda_2(c) = 1$. For the same reasons, $\lambda_2(d) = 2$.

From b, player $\bigcirc$ can force $\square$ to get the payoff 2 or less, with the strategy profile $\sigma_\bigcirc : h \mapsto c$. Such a strategy is λ_1-rational, assuming the strategy $\sigma_\square : h \mapsto d$. Therefore, we have $\lambda_2(b) = 2$.

Finally, from a, player $\square$ can force $\bigcirc$ to get the payoff 2 or less, with the strategy profile $\sigma_\square : h \mapsto d$. Such a strategy is λ_1-rational, assuming the strategy $\sigma_\bigcirc : h \mapsto c$. But, he cannot force her to get less than the payoff 2, because she can force the access to the state b, and the only λ_1-consistent plays from b are the plays with the form $(ba)^k bd^\omega$. Therefore, $\lambda_2(a) = 2$.

It will be proved in Section 6 that mean-payoff games are with steady negotiation.

3.3. **Link with Nash equilibria.** Requirements and the negotiation function are able to capture Nash equilibria. Indeed, if λ_0 is the vacuous requirement, then $\mathsf{nego}(\lambda_0)$ characterizes the NE outcomes, in the following formal sense:

Theorem 3.8. *Let G be a game with steady negotiation. Then, a play ρ in G is an NE outcome if and only if ρ is $\mathsf{nego}(\lambda_0)$-consistent.*

Proof.
- Let $\bar{\sigma}$ be a Nash equilibrium in $G_{\restriction v_0}$, for some state v_0, and let $\rho = \langle \bar{\sigma} \rangle$: let us prove that the play ρ is $\mathsf{nego}(\lambda_0)$-consistent. Let $k \in \mathbb{N}$, let $i \in \Pi$ be such that $\rho_k \in V_i$, and let us prove that $\mu_i(\rho_{\geq k}) \geq \mathsf{nego}(\lambda_0)(\rho_k)$. For any deviation σ'_i of $\sigma_{i \restriction \rho_{\leq k}}$, by definition of NEs, we have $\mu_i\left(\langle \bar{\sigma}_{-i \restriction \rho_{\leq k}}, \sigma'_i \rangle\right) \leq \mu_i(\rho)$. Therefore, we have $\mu_i(\rho) \geq \sup_{\sigma'_i} \mu_i\left(\langle \bar{\sigma}_{-i \restriction \rho_{\leq k}}, \sigma'_i \rangle\right)$, hence $\mu_i(\rho) \geq \inf_{\bar{\tau}_{-i}} \sup_{\tau_i} \mu_i\left(\langle \bar{\tau}_{-i \restriction \rho_{\leq k}}, \tau_i \rangle\right)$, i.e. $\mu_i(\rho) \geq \mathsf{nego}(\lambda_0)(\rho_k)$.
- Conversely, let ρ be a $\mathsf{nego}(\lambda_0)$-consistent play from a state v_0. Let us define a strategy profile $\bar{\sigma}$ such that $\langle \bar{\sigma} \rangle = \rho$, by:
  - $\langle \bar{\sigma} \rangle = \rho$;
  - for each history of the form $\rho_0 \ldots \rho_k v$ with $v \neq \rho_{k+1}$, let i be the player controlling ρ_k. Since the game G is with steady negotiation, the infimum:

$$\inf_{\bar{\tau}_{-i} \in \lambda_0 \mathsf{Rat}(\rho_k)} \sup_{\tau_i} \mu_i(\langle \bar{\tau} \rangle) = \mathsf{nego}(\lambda_0)(v) \neq +\infty$$

  is a minimum. Let $\bar{\tau}^k_{-i}$ be λ_0-rational strategy profile from ρ_k realizing that minimum, and let τ^k_i be some strategy from ρ_k such that $\tau^k_i(\rho_k) = v$. Then, we define:

$$\langle \bar{\sigma}_{\restriction \rho_0 \ldots \rho_k v} \rangle = \langle \bar{\tau}^k_{\rho_k v} \rangle;$$

– for every other history h, the state $\bar{\sigma}(h)$ is defined arbitrarily.

   Let us prove that $\bar{\sigma}$ is an NE: let σ'_i be a deviation of σ_i, let $\rho' = \langle \bar{\sigma}_{-i}, \sigma'_i \rangle$ and let $\rho_0 \ldots \rho_k$ be the longest common prefix of ρ and ρ'. Let $v = \rho'_{k+1}$. Then, we have:

$$\mu_i(\rho') \leq \sup_{\tau^k_i} \mu_i\left(\langle \bar{\tau}^k \rangle\right) = \mathsf{nego}(\lambda_0)(\rho_k),$$

and since ρ is λ_0-consistent, we have $\mathsf{nego}(\lambda_0)(\rho_k) \leq \mu_i(\rho)$, hence $\mu_i(\rho') \leq \mu_i(\rho)$. $\qquad\square$

Example 3.9. Let us consider again the game of Example 2.18, with the requirement λ_1 given in Figure 1b. The only λ_1-consistent plays in this game, starting from the state a, are ac^ω, and $(ab)^k d^\omega$ with $k \geq 1$. One can check that those plays are exactly the NE outcomes in that game.

   In the following section, we will prove that as well as $\mathsf{nego}(\lambda_0)$ characterizes the NEs, the requirement that is the least fixed point of the negotiation function characterizes the SPEs.

4. Link between negotiation and SPEs

The notion of negotiation will enable us to find the SPEs, but also more generally the ε-SPEs, in a game. For that purpose, we need the notion of ε-fixed points of a function.

Definition 4.1 (ε-fixed point)**.** Let $\varepsilon \geq 0$, let D be a finite set and let $f : \overline{\mathbb{R}}^D \to \overline{\mathbb{R}}^D$ be a mapping. A tuple $\bar{x} \in \mathbb{R}^D$ is a ε-*fixed point* of f if for each $d \in D$, for $\bar{y} = f(\bar{x})$, we have $y_d \in [x_d - \varepsilon, x_d + \varepsilon]$.

Remark 4.2. A 0-fixed point is a fixed point, and conversely.

   By Tarski's fixed point theorem, the negotiation function, which is a monotone function from a complete lattice to itself, has a least fixed point. That result can be generalized to ε-fixed points.

Lemma 4.3. *Let $\varepsilon \geq 0$. On every game, the function nego has a least ε-fixed point.*

Proof. The following proof is a generalization of a classical proof of Tarski's fixed point theorem. Let Λ be the set of the ε-fixed points of the negotiation function. The set Λ is not empty, since it contains at least the requirement $v \mapsto +\infty$. Let λ^* be the requirement defined by:

$$\lambda^* : v \mapsto \inf_{\lambda \in \Lambda} \lambda(v).$$

   For every ε-fixed point λ of the negotiation function, we have then $\lambda^*(v) \leq \lambda(v)$ for each v, and then $\mathsf{nego}(\lambda^*)(v) \leq \mathsf{nego}(\lambda)(v)$ since nego is monotone; and therefore, we have $\mathsf{nego}(\lambda^*)(v) \leq \lambda(v) + \varepsilon$. As a consequence, we have:

$$\mathsf{nego}(\lambda^*)(v) \leq \inf_{\lambda \in \Lambda} \lambda(v) + \varepsilon = \lambda^*(v) + \varepsilon.$$

The requirement λ^* is an ε-fixed point of the negotiation function, and is therefore the least of them. $\qquad\square$

   In all what follows, for a given game G and a given $\varepsilon > 0$, we will write λ^* for the least ε-fixed point of the negotiation function. Intuitively, the requirement λ^* is such that, from every vertex v, the player i controlling v cannot enforce a payoff greater than $\lambda^*(v) + \varepsilon$ against a λ^*-rational behaviour. Therefore, the λ^*-consistent plays are such that if one player

tries to deviate, it is possible for the other players to prevent them improving their payoff by more than ε, while still playing rationally — which defines ε-SPE outcomes. Formally:

Theorem 4.4. *Let $G_{\restriction v_0}$ be a prefix-independent game played on a finite graph, and let $\varepsilon \geq 0$. Let θ be a play starting in v_0. If there exists an ε-SPE $\bar{\sigma}$ such that $\langle \bar{\sigma} \rangle = \theta$, then θ is λ^*-consistent. If G is also a game with steady negotiation, then conversely, if θ is λ^*-consistent, then it is an ε-SPE outcome.*

Proof.

- *If $\bar{\sigma}$ is an ε-SPE, then the play $\theta = \langle \bar{\sigma} \rangle$ is λ^*-consistent.*

  Let us define a requirement λ by, for each $i \in \Pi$ and $v \in V_i$:

$$\lambda(v) = \inf_{hv \in \mathsf{Hist}G_{\restriction v_0}} \mu_i(\langle \bar{\sigma}_{\restriction hv} \rangle).$$

  Then, for every history hv starting in v_0, the play $\langle \bar{\sigma}_{\restriction hv} \rangle$ is λ-consistent. In particular, the play θ is. Let us now prove that λ is an ε-fixed point of nego. We will then have $\lambda \geq \lambda^*$, which implies that the play θ is λ^*-consistent.

  Let $i \in \Pi$, let $v \in V_i$, and let us assume toward contradiction (since the negotiation function is non-decreasing) that $\mathsf{nego}(\lambda)(v) > \lambda(v) + \varepsilon$, that is to say:

$$\inf_{\bar{\tau}_{-i} \in \lambda\mathsf{Rat}(v)} \sup_{\tau_i} \mu_i(\langle \bar{\tau} \rangle) > \lambda(v) + \varepsilon = \inf_{hv \in \mathsf{Hist}G_{\restriction v_0}} \mu_i(\langle \bar{\sigma}_{\restriction hv} \rangle) + \varepsilon.$$

  Then, since all the plays generated by the strategy profile $\bar{\sigma}$ are λ-consistent, and therefore since any strategy profile of the form $\bar{\sigma}_{-i \restriction hv}$ is λ-rational, we have:

$$\inf_{hv} \sup_{\tau_i} \mu_i(\langle \bar{\sigma}_{-i \restriction hv}, \tau_i \rangle) > \inf_{hv} \mu_i(\langle \bar{\sigma}_{\restriction hv} \rangle) + \varepsilon.$$

  Therefore, there exists a history hv such that:

$$\sup_{\tau_i} \mu_i(\langle \bar{\sigma}_{-i \restriction hv}, \tau_i \rangle) > \mu_i(\langle \bar{\sigma}_{\restriction hv} \rangle) + \varepsilon,$$

  which is impossible if the strategy profile $\bar{\sigma}$ is an ε-SPE. Therefore, there is no such v, and the requirement λ is an ε-fixed point of the negotiation function.

- *If G is a game with steady negotiation and θ is λ^*-consistent, then θ is an ε-SPE outcome.*

  - *A particular case: if there exists v accessible from v_0 such that $\lambda^*(v) = +\infty$.*

    In that case, for each u such that $uv \in E$, if the player controlling u chooses to go to v, no λ^*-consistent play can be proposed to them from there, hence there is no λ^*-rational strategy profile against that player from u, and $\mathsf{nego}(\lambda^*)(u) = +\infty$. Since ε is finite and since λ^* is an ε-fixed point of the negotiation function, it follows that $\lambda^*(u) = +\infty$. Since v is accessible from v_0, we can repeat this argument and show that $\lambda^*(v_0) = +\infty$; in that case, there is no λ^*-consistent play θ from u, and then the proof is done.

    Therefore, for the rest of the proof, we assume that for all v, we have $\lambda^*(v) \neq +\infty$. As a consequence, since λ^* is an ε-fixed point of the function nego, for each v accessible from v_0, we have $\mathsf{nego}(\lambda^*)(v) \neq +\infty$; which implies that for each such v, there exists a λ^*-consistent strategy profile against the player controlling v, starting from v.

    The rest of the proof constructs the strategy profile $\bar{\sigma}$ and proves that it is an SPE. That construction is illustrated by Figure 3.

  - *Spare parts: the strategy profiles $\bar{\tau}^{v*}$.*

    Recall that since G is a game with steady negotiation, for every requirement λ^*, for every player i and for every state $v \in V_i$, since by the previous point we assume $\lambda^*\mathsf{Rat}(v) \neq \emptyset$,
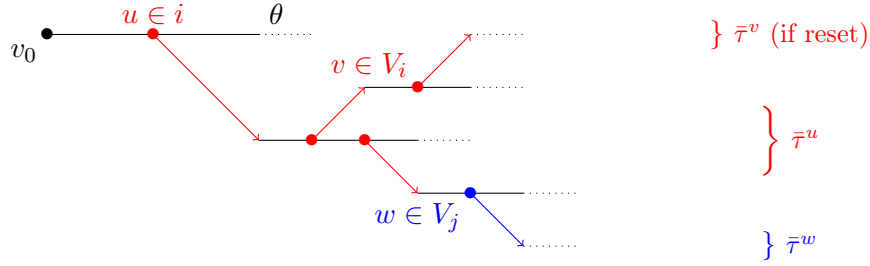
Figure 3. The construction of $\bar{\sigma}$

we know that there exists a strategy profile $\bar{\tau}^v_{-i}$ from v that is λ^*-rational assuming a strategy τ^v_i and that satisfies the inequality:

$$\sup_{\tau_i}\ \mu_i(\langle \bar{\tau}^v_{-i}, \tau_i\rangle) = \inf_{\bar{\tau}_{-i}\in\lambda\mathsf{Rat}(v)}\ \sup_{\tau_i}\ \mu_i(\langle\bar{\tau}\rangle) = \mathsf{nego}(\lambda^*)(v),$$

i.e. there exists a worst λ^*-rational strategy profile against player i from the state v, with regards to player i's payoff. Our goal in this part of the proof is to construct a strategy profile $\bar{\tau}^{v*}_{-i}$, that is λ^*-rational assuming a strategy τ^{v*}_i, and that will be used to punish player i when they deviate from $\bar{\sigma}$ until another player deviates.

The strategy profile $\bar{\tau}^v_{-i}$ and the strategy τ^v_i are not sufficient for that purpose, because if some history h compatible with $\bar{\tau}^v_{-i}$ is such that $\mu_i(\langle\bar{\tau}^v_{\restriction h}\rangle) < \mu_i(\langle\bar{\tau}^v\rangle)$, then in the corresponding subgame, it may be possible for player i to deviate and get a payoff that would be smaller than or equal to $\mu_i(\langle\bar{\tau}^v\rangle)$, but greater than $\mu_i(\langle\bar{\tau}^v_{\restriction h}\rangle)$. On the other hand, the construction of $\bar{\tau}^{v*}_{-i}$ will ensure that each time player i deviates, the other players punish them at least as harshly as they were planning to do before the deviation. Let us construct inductively the strategy profile $\bar{\tau}^{v*}$. We define it only on histories that are compatible with $\bar{\tau}^{v*}_{-i}$, since it can be defined arbitrarily on other histories. We proceed by assembling the strategy profiles of the form $\bar{\tau}^w$ for various $w \in V_i$, and the histories after which we follow a new $\bar{\tau}^w$ will be called the *resets* of $\bar{\tau}^{v*}$: they will be histories of the form hw', where h is empty or $\mathsf{last}(h) = w$.

First, we set $\langle\bar{\tau}^{v*}\rangle = \langle\bar{\tau}^v\rangle$: the one-state history v is then the first reset of $\bar{\tau}^{v*}_{-i}$.

Then, for every history hww' from v such that h is compatible with $\bar{\tau}^{v*}_{-i}$, that $w \in V_i$, and that $w' \neq \tau^{v*}_i(hw)$: let us decompose $hww' = h_1h_2$, so that the history $h_1\mathsf{first}(h_2)$ is the longest reset of $\bar{\tau}^{v*}_{-i}$ among the prefixes of hw. Or, in other words, so that the strategy profile $\bar{\tau}^{v*}_{\restriction h_1\mathsf{first}(h_2)}$ has been defined as equal to $\bar{\tau}^u$ over the prefixes of h_2 until w, where $u = v$ if h_1 is empty, or $u = \mathsf{last}(h_1)$ otherwise. By prefix-independence of G and by definition of $\bar{\tau}^u$ and $\bar{\tau}^w$, we have:

$$\inf_{\bar{\tau}_{-i}\in\lambda\mathsf{Rat}(w')}\ \sup_{\tau_i}\mu_i(\langle\bar{\tau}\rangle) \leq \sup_{\tau_i}\mu_i(\langle\bar{\tau}^w_{-i}, \tau_i\rangle) = \mathsf{nego}(\lambda^*)(w).$$

Let us now separate two cases.

∗ Suppose first that there is equality:

$$\inf_{\bar{\tau}_{-i}\in\lambda\mathsf{Rat}(w')}\ \sup_{\tau_i}\mu_i(\langle\bar{\tau}\rangle) = \mathsf{nego}(\lambda^*)(w).$$

Then, we choose $\langle\bar{\tau}^{v*}_{\restriction hww'}\rangle = \langle\bar{\tau}^u_{\restriction uh_2}\rangle$: the coalition of players against player i keeps following the same strategy profile.

* Suppose now that the inequality is strict:

$$\inf_{\bar{\tau}_{-i} \in \lambda\mathsf{Rat}(w')} \sup_{\tau_i} \mu_i(\langle \bar{\tau} \rangle) < \mathsf{nego}(\lambda^*)(w).$$

Then, we choose $\langle \bar{\tau}^{v*}_{\restriction hww'} \rangle = \langle \bar{\tau}^w_{\restriction ww'} \rangle$: player i has done something that lowers the payoff they can ensure, and therefore the other players have to update their strategy profile in order to punish them more. The history hw is a reset of $\bar{\tau}^{v*}_{-i}$.

Since there are finitely many histories of each length, this process completely defines $\bar{\tau}^{v*}$. Moreover, all the plays constructed are λ^*-consistent, hence the strategy profile $\bar{\tau}^{v*}_{-i}$ is λ^*-rational assuming τ^{v*}_i, as desired.

– *Construction of $\bar{\sigma}$.*
Let us now construct inductively the strategy profile $\bar{\sigma}$: we will prove in the next part of the proof that it is an ε-SPE. We proceed inductively, by defining all the plays $\langle \bar{\sigma}_{\restriction hv} \rangle$, for $hv \in \mathsf{Hist}(G_{v_0})$ with $v \neq \bar{\sigma}(h)$. We maintain the induction hypothesis that such a play is always λ^*-consistent.

* First, we choose $\langle \bar{\sigma} \rangle = \theta$, which satisfies the induction hypothesis.
* Let now huv be a history such that the strategy profile $\bar{\sigma}$ has been defined on all the prefixes of hu, which we now assume to be nonempty, but not on huv itself, and such that $v \neq \bar{\sigma}(hu)$. Let i be the player controlling the state u.
Then, we define $\langle \bar{\sigma}_{\restriction huv} \rangle = \langle \bar{\tau}^{u*}_{\restriction uv} \rangle$, and inductively, for every history $h'w$ starting from v and compatible with $\bar{\sigma}_{-i\restriction huv}$, we define $\langle \bar{\sigma}_{\restriction huh'w} \rangle = \langle \bar{\tau}^{u*}_{\restriction uh'w} \rangle$. The strategy profile $\bar{\sigma}_{\restriction huv}$ is then equal to $\bar{\tau}^{v*}_{\restriction uv}$ on any history compatible with $\bar{\tau}^{v*}_{-i}$.

Since there are finitely many histories of each length, this process completely defines $\bar{\sigma}$.

– *Such $\bar{\sigma}$ is an ε-SPE.*
Consider a history $h_0 w \in \mathsf{Hist}G_{\restriction v_0}$, a player $i \in \Pi$, and a deviation σ'_i of σ_i. Let $\rho = h_0 \langle \bar{\sigma}_{\restriction h_0 w} \rangle$, and let $\rho' = h_0 \langle \bar{\sigma}_{-i\restriction h_0 w}, \sigma'_{i\restriction h_0 w} \rangle$. We wish to prove that $\mu_i(\rho') \leq \mu_i(\rho) + \varepsilon$. First, if the play ρ' is compatible with σ_i, then $\rho' = \rho$ and the proof is immediate. Now, if it is not, we let $\rho'_{\leq n}$ denote the shortest prefix of ρ' such that $\rho'_{n-1} \in V_i$ and $\rho'_n \neq \sigma_i(\rho'_{<n})$, and such that $\rho'_{\geq n}$ is compatible with $\bar{\sigma}_{-i\restriction \rho'_{\leq n}}$. Thus, the transition $\rho'_{n-1}\rho'_n$ marks the time when player i begins to deviate unilaterally from σ_i. However, note that $\rho'_{\leq n}$ can be both longer or shorter than $h_0 w$: player i may have already deviated in $h_0 w$, or may wait afterwards to effectively deviate.

Be that as it may, the history $\rho'_{<n}$ is a common prefix of the plays ρ and ρ', and the substrategy profile $\bar{\sigma}_{\restriction \rho'_{\leq n}}$ has been defined during the construction of $\bar{\sigma}$ as equal to $\bar{\tau}^{v*}_{\restriction \rho'_{n-1}\rho'_n}$, where $v = \rho'_{n-1}$, on any history compatible with $\bar{\sigma}_{-i\restriction \rho'_{\leq n}}$.

By construction of $\bar{\tau}^{v*}$, the sequence $(\mathsf{nego}(\rho'_k))_{k \geq n-1, \rho'_k \in V_i}$ is non-increasing. It is therefore stationary (or finite), because it can take only a finite number of values. Consequently, there is a finite number of resets along the play $\rho'_{\geq n-1}$. Let $\rho'_{n-1} \dots \rho'_m$ be the last (longest) one. Afterwards, the play $\rho'_{\geq m}$ is compatible with the strategy profile $\bar{\tau}^{\rho'_{m-1}}_i$. By definition of that strategy profile, we have the inequality $\mu_i(\rho') \leq \mathsf{nego}(\lambda^*)(\rho'_{m-1})$. We need now to prove $\mathsf{nego}(\lambda^*)(\rho'_{m-1}) \leq \mu_i(\rho) + \varepsilon$.

Let $\rho_{\leq p} = \rho'_{\leq p}$ denote the longest common prefix of ρ and ρ' such that $\rho_p \in V_i$. Since player i does not control any vertex between ρ_p and ρ_{n-1}, and therefore cannot deviate, we have $\rho_{\geq p} = \langle \bar{\sigma}_{\restriction \rho_{\leq p}} \rangle$, which is λ^*-consistent. As a consequence, we have $\mu_i(\rho) \geq \lambda^*(\rho_p)$.

Finally, since the sequence of the quantities $\mathsf{nego}(\rho_k')$ with $\rho_k' \in V_i$ is non-increasing for $k \geq n - 1$, we also have $\mathsf{nego}(\lambda^*)(\rho_{m-1}') \leq \mathsf{nego}(\lambda^*)(\rho_p)$. Consequently, we have:

$$\mu_i(\rho') \leq \mathsf{nego}(\lambda^*)(\rho_{m-1}') \leq \mathsf{nego}(\lambda^*)(\rho_p) \leq \lambda^*(\rho_p) + \varepsilon \leq \mu_i(\rho) + \varepsilon.$$

The strategy profile $\bar{\sigma}$ is an ε-SPE. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

5. A first way to handle negotiation: the abstract negotiation game

5.1. **Informal definition.** We have now proved that SPEs are characterized by the requirements that are fixed points of the negotiation function; but we need to know how to compute, in practice, the quantity $\mathsf{nego}(\lambda)$ for a given requirement λ. We first define an *abstract negotiation game*, that is conceptually simple but not directly usable for an algorithmic purpose, because it is defined on an uncountably infinite state space.

A similar definition was given in [FP17], as a tool in a general method to compute SPE outcomes in games whose payoff functions have finite range, which is not the case of mean-payoff games. Here, linking that game with our concepts of requirements, negotiation function and steady negotiation enables us to present an effective algorithm in the case of mean-payoff games, by constructing a finite version of the abstract negotiation game, the *concrete negotiation game*.

The abstract negotiation game from a state $v_0 \in V_i$, with regards to a requirement λ, is denoted by $\mathsf{Abs}_{\lambda i}(G)_{\restriction v_0}$ and opposes two players, *Prover* and *Challenger*, with the following rules:

- first, Prover proposes a λ-consistent play ρ from v_0 (or loses, if she has no play to propose).
- Then, either Challenger accepts the play and the game terminates; or, he chooses an edge $\rho_k \rho_{k+1}$, with $\rho_k \in V_i$, from which he can make player i deviate, using another edge $\rho_k v$ with $v \neq \rho_{k+1}$: then, the game starts again from v instead of v_0.
- In the resulting play (either eventually accepted by Challenger, or constructed by an infinity of deviations), Prover wants player i's payoff to be low, and Challenger wants it to be high.

That game gives us the basis of a method to compute $\mathsf{nego}(\lambda)$ from λ: the maximal payoff that Challenger — or $\mathbb{C}$ for short — can ensure in $\mathsf{Abs}_{\lambda i}(G)_{\restriction [v_0]}$, with $v_0 \in V_i$, is also the maximal payoff that player i can ensure in $G_{\restriction v_0}$, against a λ-rational environment; hence the equality $\mathsf{val}_{\mathbb{C}}\left(\mathsf{Abs}_{\lambda i}(G)_{\restriction [v_0]}\right) = \mathsf{nego}(\lambda)(v_0)$. A proof of that statement, with a complete formalization of the abstract negotiation game, is presented in Appendix A.

Example 5.1. Let us consider again the game of Example 2.18: the requirement $\lambda_2 = \mathsf{nego}(\lambda_1)$, computed in Section 3.2, is also presented on the third line below the states in Figure 1b. Let us use the abstract negotiation game to compute the requirement $\lambda_3 = \mathsf{nego}(\lambda_2)$.

From a, Prover can propose the play abd^ω, and the only deviation Challenger can do is going to c; he has of course no incentive to do it. Therefore, $\lambda_3(a) = 2$. From b, whatever Prover proposes at first, Challenger can deviate and go to a. Then, from a, Prover cannot propose the play ac^ω, which is not λ_2-consistent: she has to propose a play beginning by ab, and to let Challenger deviate once more. He can then deviate infinitely often that way, and generate the play $(ba)^\omega$: therefore, $\lambda_3(b) = 3$. The other states keep the same values. Note
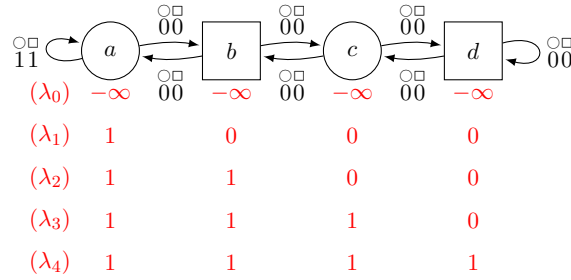
FIGURE 4. Iterations of the negotiation function

that there exists no λ_3-consistent play from a or b, hence $\mathsf{nego}(\lambda_3)(a) = \mathsf{nego}(\lambda_3)(b) = +\infty$. This proves that there is no SPE in that game.

5.2. **An imperfect method: the negotiation sequence.** A classical way to compute the least fixed point of a function is, as in the example above, to compute its iterations on the least element of the set we are considering until reaching a fixed point — which is, then, the least one. We call this sequence the *negotiation sequence*, and write it $(\lambda_n)_{n \in \mathbb{N}} = (\mathsf{nego}^n(\lambda_0))_n$. In many simple examples, in practice, computing the negotiation sequence, using the abstract negotiation game, is the way we will find the least fixed point of the negotiation function and solve SPE problems.

Example 5.2. Let G be the game of Figure 4, where each edge is labelled by the rewards $r_\bigcirc$ and $r_\square$. Below the states, we present the requirements $\lambda_0 : v \mapsto -\infty$, $\lambda_1 = \mathsf{nego}(\lambda_0)$, $\lambda_2 = \mathsf{nego}(\lambda_1)$, $\lambda_3 = \mathsf{nego}(\lambda_2)$, and $\lambda_4 = \mathsf{nego}(\lambda_3)$. Let us explicate those computations, using the abstract negotiation game. From λ_0 to λ_1: since every play is λ_0-consistent, Prover can always propose whatever she wants. From the state a, whatever she (trying to minimize player $\bigcirc$'s payoff) proposes, Challenger can always make player $\bigcirc$ deviate in order to loop on the state a. Then, in the game G, player $\bigcirc$ gets the payoff 1, hence $\lambda_1(a) = 1$. From the state b, Prover (trying to minimize player $\square$'s payoff) can propose the play $(bc)^\omega$. If Challenger makes player $\square$ deviate to go to the state a, then Prover can propose the play $a(bc)^\omega$. Even if Challenger makes player $\square$ deviate infinitely often, he cannot give him more than the payoff 0, hence $\lambda_1(b) = 0$. Similar situations happen from the states c and d, hence $\lambda_1(c) = \lambda_1(d) = 0$. From λ_1 to λ_2: now, from the state b, whatever Prover proposes at first, Challenger can make player $\square$ deviate and go to the state a. From there, since we have $\lambda_1(a) = 1$, Prover has to propose a play in which player $\bigcirc$ gets the payoff 1. The only such plays do also give the payoff 1 to player $\square$, hence $\lambda_2(b) = 1$. Similar situations explain $\lambda_3(c) = 1$, and $\lambda_4(c) = 1$. Finally, plays ending with the loop a^ω are all λ_4-consistent, hence Prover can always propose them, hence the requirement λ_4 is a fixed point of the negotiation function — and therefore the least.

The interested reader will find other such examples in Appendix B. However, this cannot be turned into an effective algorithm: the negotiation sequence is not always stationary.

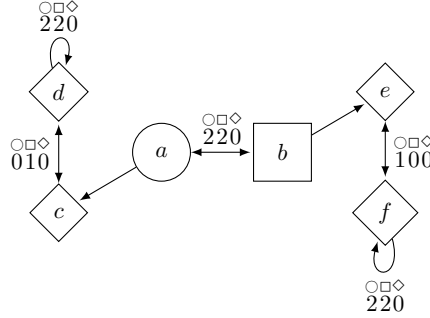**Theorem 5.3.** *There exists a mean-payoff game on which the negotiation sequence is not stationary.*

Figure 5. A game where the negotiation sequence is not stationary

Proof. Let G be the game of Figure 5. Since all the $\diamond$ rewards are equal to 0, for all $n > 0$, we have $\lambda_n(c) = \lambda_n(d) = \lambda_n(e) = \lambda_n(f) = 0$. Moreover, by symmetry of the game, we always have $\lambda_n(a) = \lambda_n(b)$. Therefore, to compute the negotiation sequence, it suffices to compute $\lambda_{n+1}(a)$ as a function of $\lambda_n(b)$, knowing that $\lambda_1(a) = \lambda_1(b) = 1$, and therefore that for all $n > 0$, $\lambda_n(a) = \lambda_n(b) \geq 1$.

From a, the worst play that player $\square$ could propose to player $\bigcirc$ would be a combination of the cycles cd and d giving her exactly 1. But then, player $\bigcirc$ will deviate to go to b, from which if player $\square$ proposes plays in the strongly connected component containing c and d, then player $\bigcirc$ will always deviate and generate the play $(ab)^\omega$, and then get the payoff 2.

Then, in order to give her a payoff lower than 2, player $\square$ has to go to the state e. Since player $\bigcirc$ does not control any state in that strongly connected component, the play he will propose will be accepted: he will, then, propose the worst possible combination of the cycles ef and f for player $\bigcirc$, such that he gets at least his requirement $\lambda_n(b)$. The payoff $\lambda_{n+1}(a)$ is then the minimal solution of the system:

$$\begin{cases} \lambda_{n+1}(a) = x + 2(1-x) \\ \quad 2(1-x) \geq \lambda_n(b) \\ \quad\quad 0 \leq x \leq 1 \end{cases}$$

that is to say $\lambda_{n+1}(a) = 1 + \frac{\lambda_n(b)}{2} = 1 + \frac{\lambda_n(a)}{2}$, and by induction, for all $n > 0$:

$$\lambda_n(a) = \lambda_n(b) = 2 - \frac{1}{2^{n-1}}$$

which converges to 2 but does never reach it. $\square$

6. A tool to compute negotiation: the concrete negotiation game

6.1. **Definition.** In the abstract negotiation game, Prover has to propose complete plays, on which we can make the hypothesis that they are λ-consistent. In practice, there will often be an infinity of such plays, and therefore it cannot be used directly for an algorithmic purpose. Instead, those plays can be given edge by edge, in a finite state game. Its definition is more technical, but it can be shown that it is equivalent to the abstract one.

Definition 6.1 (Concrete negotiation game). Let G be a prefix-independent game played on a finite graph, let $i \in \Pi$ and $v_0 \in V_i$, and let λ be a requirement on G. The *concrete negotiation game* of $G_{\restriction v_0}$ is the two-player zero-sum game $\mathsf{Conc}_{\lambda i}(G)_{\restriction s_0} = (\{\mathbb{P}, \mathbb{C}\}, S, (S_\mathbb{P}, S_\mathbb{C}), \Delta, \nu)_{\restriction s_0}$, defined as follows:

- player $\mathbb{P}$ is called *Prover*, and player $\mathbb{C}$ is called *Challenger*.
- The set of states controlled by Prover is $S_\mathbb{P} = V \times 2^V$, where the state $s = (v, M)$ contains the information of the current state v on which Prover has to define the strategy profile, and the *memory* M of the states that have been traversed so far since the last deviation, and that define the requirements Prover has to satisfy. The initial state is $s_0 = (v_0, \{v_0\})$.
- The set of states controlled by Challenger is $S_\mathbb{C} = E \times 2^V$, where in the state $s = (uv, M)$, the edge uv is the edge proposed by Prover.
- The set Δ contains three types of transitions: *proposals*, *acceptations* and *deviations*.
  - The proposals are transitions in which Prover proposes an edge of the game G:
  $$\mathsf{Prop} = \left\{ (v, M)(vw, M) \mid vw \in E, M \in 2^V \right\};$$
  - the acceptations are transitions in which Challenger accepts to follow the edge proposed by Prover (it is in particular his only possibility when that edge begins on a state that is not controlled by player i) — note that the memory is updated:
  $$\mathsf{Acc} = \left\{ (vw, M)(w, M \cup \{w\}) \mid j \in \Pi, w \in V_j \right\};$$
  - the deviations are transitions in which Challenger refuses to follow the edge proposed by Prover, as he can if that edge begins in a state controlled by player i — the memory is erased, and only the new state the deviating edge leads to is memorized:
  $$\mathsf{Dev} = \left\{ (uv, M)(w, \{w\}) \mid u \in V_i, w \neq v, uw \in E \right\}.$$
- Let $H = (h_0, M_0)(h_0 h_0', M_0) \dots (h_n h_n', M_n)$ be a history in $\mathsf{Conc}_{\lambda i}(G)$: the *projection* of the history H is the history $\dot{H} = h_0 \dots h_n$ in the game G. That definition is naturally extended to plays.
- The payoff function $\nu_\mathbb{C} = -\nu_\mathbb{P}$ measures player i's payoff, with a winning condition if the constructed strategy profile is not λ-rational, that is to say if after finitely many player i's deviations, it can generate a play which is not λ-consistent:
  - $\nu_\mathbb{C}(\pi) = +\infty$ if after some index $n \in \mathbb{N}$, the play $\pi_{\geq 2n}$ contains no deviation, and if the play $\dot{\pi}_{\geq n}$ is not λ-consistent[3];
  - $\nu_\mathbb{C}(\pi) = \mu_i(\dot{\pi})$ otherwise.

Like in the abstract negotiation game, the goal of Challenger is to find a λ-rational strategy profile that forces the worst possible payoff for player i, and the goal of Prover is to find a possibly deviating strategy for player i that gives them the highest possible payoff.

Remark 6.2. The concrete negotiation game has the following properties.
- If G is Borel, then $\mathsf{Conc}_{\lambda i}(G)$ is Borel.
- When $\pi_{\geq 2n}$ contains no deviation, the memory of its states is increasing, and therefore eventually equal to the memory $M = \mathsf{Occ}(\dot{\pi}_{\geq n})$. If it is the longest such suffix of π, it means that the play $\dot{\pi}_{\geq n}$ is λ-consistent if and only if for each player j and each vertex $v \in V_j$, we have $\mu_j(\dot{\pi}) \geq \lambda(v)$.

---

[3]When we combine the notations $\dot{\pi}$ and $\pi_{\geq n}$, the notation $\dot{\pi}$ is applied first; that is, the play $\dot{\pi}_{\geq n}$ is the projection of the play $\pi_{\geq 2n}$, not $\pi_{\geq n}$.

- When G is a mean-payoff game and when λ has finite values, the concrete negotiation game can be seen as a multidimensional two-player zero-sum mean-payoff game, with one dimension for each player, meant to control that each player gets the payoff they require, plus a special dimension $\star$, meant to measure player i's actual payoff. The rewards of the proposals are all equal to 0, and the rewards of acceptations and deviations are $\hat{r}_\star((uv, M)(v', N)) = 2r_i(uv')$, and $\hat{r}_j((uv, M)(v', N)) = 2r_j(uv') - 2\max\{\lambda(w) \mid w \in M \cap V_j\}$. The payoff $\nu_\mathbb{C}(\pi)$ equals then $+\infty$ for every π that contains finitely many deviations and such that for some $j \in \Pi$, the mean-payoff $\hat{\mu}_j(\pi)$ is negative, and $\nu_\mathbb{C}(\pi) = \hat{\mu}_\star(\pi)$ otherwise.

6.2. **Link with the negotiation function.** The concrete negotiation game is equivalent to the abstract one: the only differences are that the plays proposed by Prover are proposed edge by edge, and that their λ-consistency is not written in the rules of the game but in its payoff function.

Theorem 6.3. *Let G be a Borel prefix-independent game played on a finite graph. Let λ be a requirement, let i be a player and let $v_0 \in V_i$. Then, we have $\mathsf{val}_\mathbb{C}\left(\mathsf{Conc}_{\lambda i}(G)_{\restriction s_0}\right) = \mathsf{nego}(\lambda)(v_0)$. Moreover, if for each player i and every state $v_0 \in V_i$, Prover has an optimal strategy in $\mathsf{Conc}_{\lambda i}(G)_{\restriction(v_0, \{v_0\})}$, then G is a game with steady negotiation.*

Proof.
- *First direction:* $\mathsf{nego}(\lambda)(v_0) \leq \mathsf{val}_\mathbb{C}\left(\mathsf{Conc}_{\lambda i}(G)_{\restriction s_0}\right)$.

  Let $\tau_\mathbb{P}$ be a strategy such that $\sup_{\tau_\mathbb{C}} \nu_\mathbb{C}(\langle \bar{\tau} \rangle) \neq +\infty$, and let $\bar{\sigma}$ be the strategy profile defined by $\bar{\sigma}(\dot{H}) = w$ for every history H compatible with $\tau_\mathbb{P}$ (by induction, the projection is injective on the histories compatible with $\tau_\mathbb{P}$) with $\tau_\mathbb{P}(H) = (vw, \cdot)$, and arbitrarily defined on any other histories. We prove that the strategy profile $\bar{\sigma}_{-i}$ is λ-rational assuming the strategy σ_i, and that $\sup_{\sigma_i'} \mu_i(\langle \bar{\sigma}_{-i}, \sigma_i' \rangle) \leq \sup_{\tau_\mathbb{C}} \nu_\mathbb{C}(\langle \bar{\tau} \rangle)$.

  – The strategy profile $\bar{\sigma}_{-i}$ is λ-rational, assuming the strategy σ_i. Indeed, let us assume it is not. Then, there exists a history $h = h_0 \ldots h_n$ in $G_{\restriction v_0}$ compatible with $\bar{\sigma}_{-i}$ such that the play $\langle \bar{\sigma}_{\restriction h} \rangle$ is not λ-consistent. Then, let:
  $$Hs = (h_0, M_0)(h_0\bar{\sigma}(h_0), M_0) \ldots (h_n, M_n)$$
  be the only history in $\mathsf{Conc}_{\lambda i}(G)_{\restriction s_0}$ compatible with $\tau_\mathbb{P}$ such that $\dot{H} = h$. Let $\tau_\mathbb{C}$ be a strategy constructing the history h, defined by:
  $$\tau_\mathbb{C}(H_0 \ldots H_{2k-1}) = H_{2k}$$
  for every k, and:
  $$\tau_\mathbb{C}\left(H'(vw, M)\right) = (w, M \cup \{w\})$$
  for any other history $H'(vw, M)$. Then, the play $\pi = \langle \bar{\tau} \rangle$ contains finitely many deviations (Challenger stops the deviations after having constructed the history h), and the play $\dot{\pi}_{\geq n}$ is not λ-consistent. Therefore, we have $\nu_\mathbb{C}(\pi) = +\infty$, which is false by hypothesis.

  – Now, let us prove the inequality $\sup_{\sigma_i'} \mu_i(\langle \bar{\sigma}_{-i}, \sigma_i' \rangle) \leq \sup_{\tau_\mathbb{C}} \nu_\mathbb{C}(\langle \bar{\tau} \rangle)$. Let σ_i' be a strategy for player i, and let $\eta = \langle \bar{\sigma}_{-i}, \sigma_i' \rangle$. Let $\tau_\mathbb{C}$ be a strategy such that for every k:
  $$\tau_\mathbb{C}((\eta_0, \cdot)(\eta_0\cdot, \cdot) \ldots (\eta_k\cdot, \cdot)) = (\eta_{k+1}, \cdot),$$
  i.e. a strategy forcing η against $\tau_\mathbb{P}$. Then, since $\nu_\mathbb{C}(\langle \bar{\tau} \rangle) \neq +\infty$ by hypothesis on $\tau_\mathbb{P}$, we have $\mu_i(\eta) = \nu_\mathbb{C}(\langle \bar{\tau} \rangle)$, hence $\mu_i(\langle \bar{\sigma}_{-i}, \sigma_i' \rangle) \leq \sup_{\tau_\mathbb{C}} \nu_\mathbb{C}(\langle \bar{\tau} \rangle)$, hence the desired inequality.
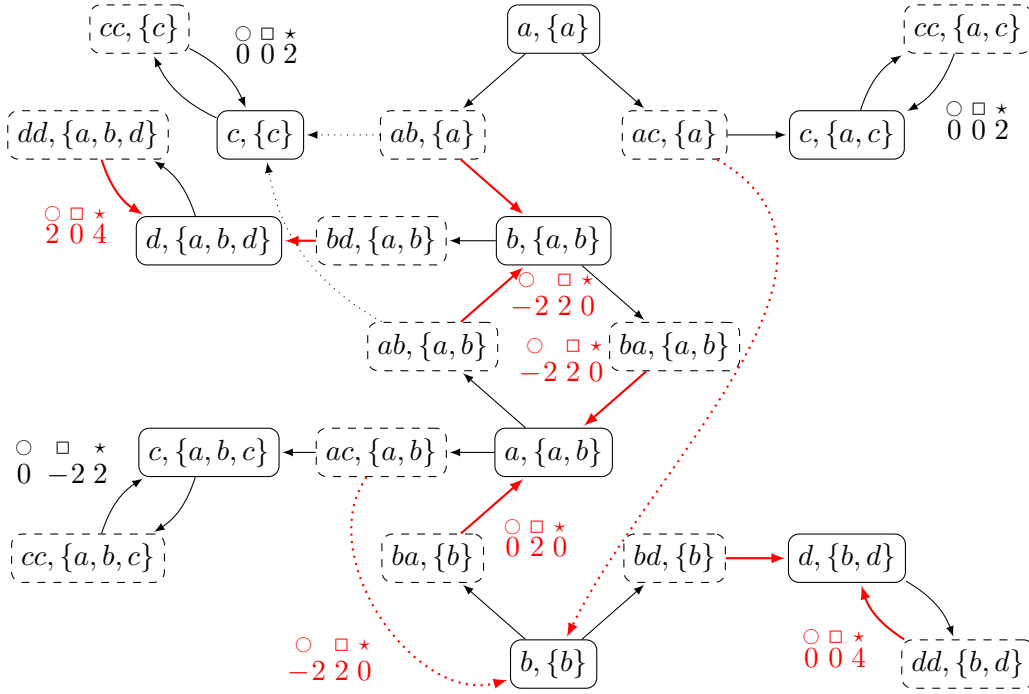
FIGURE 6. A concrete negotiation game

Moreover, if $\tau_{\mathbb{P}}$ is optimal, then the λ-rational strategy profile $\bar{\sigma}_{-i}$ realizes the infimum:

$$\inf_{\bar{\sigma}_{-i} \in \lambda \mathsf{Rat}(v_0)} \sup_{\sigma'_i} \mu_i(\langle \bar{\sigma}_{-i}, \sigma'_i \rangle),$$

hence if there exists such an optimal strategy for every vertex v_0, then the game G is with steady negotiation.

- *Second direction:* $\mathsf{val}_{\mathbb{C}} \left(\mathsf{Conc}_{\lambda i}(G)_{\restriction s_0} \right) \leq \mathsf{nego}(\lambda)(v_0)$.

  Let $\bar{\sigma}_{-i}$ be a λ-rational strategy profile from v_0, assuming the strategy σ_i; let us define a strategy $\tau_{\mathbb{P}}$, by $\tau_{\mathbb{P}}(H(v, \cdot)) = \left(v\bar{\sigma}(\dot{H}v), \cdot \right)$ for every history H and for every $v \in V$. Let us prove the inequality $\sup_{\tau_{\mathbb{C}}} \nu_{\mathbb{C}}(\langle \bar{\tau} \rangle) \leq \sup_{\sigma'_i} \mu_i(\langle \bar{\sigma}_{-i}, \sigma'_i \rangle)$.

  Let $\tau_{\mathbb{C}}$ be a strategy for Challenger, and let $\pi = \langle \bar{\tau} \rangle$. If $\nu_{\mathbb{C}}(\pi) = +\infty$, then there exists n such that the play $\pi_{\geq 2n}$ contains no deviation, i.e. $\dot{\pi}_{\geq n} = \langle \bar{\sigma}_{\restriction \dot{\pi}_{\leq n}} \rangle$, and that play is not λ-consistent, which is impossible. Therefore, we have $\nu_{\mathbb{C}}(\pi) \neq +\infty$, and as a consequence $\nu_{\mathbb{C}}(\pi) = \mu_i(\dot{\pi}) = \mu_i(\langle \bar{\sigma}_{-i}, \sigma'_i \rangle)$ for some strategy σ'_i, hence $\nu_{\mathbb{C}}(\pi) \leq \sup_{\sigma'_i} \mu_i(\langle \bar{\sigma}_{-i}, \sigma'_i \rangle)$, hence the desired inequality. □

6.3. **Example.** Let us consider again the game from Example 2.18. Figure 6 represents the game $\mathsf{Conc}_{\lambda_1 \circ}(G)$ (with $\lambda_1(a) = 1$ and $\lambda_1(b) = 2$), where the dashed states are controlled by Challenger, and the other ones by Prover.

The dotted arrows indicate the deviations, and the transitions have been labelled with the immediate rewards defined as in the remark above. The transitions that are not labelled are either zero for the three coordinates, or meaningless since they cannot be used more than

once. The red arrows indicate a (memoryless) optimal strategy for Challenger. Against that strategy, the lowest payoff Prover can ensure is 2. Therefore, we have $\mathsf{nego}(\lambda_1)(v_0) = 2$, in line with the abstract game in Example 5.1.

6.4. **Resolution.** We now know that $\mathsf{nego}(\lambda)(v)$, for a given requirement λ, a given player i and a given state $v \in V_i$, is the value of the concrete negotiation game $\mathsf{Conc}_{\lambda i}(G)_{\restriction(v,\{v\})}$. But we still do not know how to compute that value. We present here an important result for that purpose.

For any game $G_{\restriction v_0}$ and any memoryless strategy σ_i, we write $G_{\restriction v_0}[\sigma_i]$ the graph *induced* by σ_i, defined as the underlying graph of G where all the transitions that are not compatible with σ_i, and all the vertices that are then no longer accessible from v_0, have been omitted.

Lemma 6.4. *Let $G_{\restriction v_0}$ be a mean-payoff game, let i be a player, let λ be a requirement and let $\mathsf{Conc}_{\lambda i}(G)_{\restriction s_0}$ be the corresponding concrete negotiation game. There exists a memoryless strategy $\tau_{\mathbb{C}}$ that is optimal for Challenger.*

Proof. The structure of that proof is inspired from the proof of Lemma 14 in [VCD$^+$15].

Let $\nu'_{\mathbb{C}}$ be the payoff function defined by:
- $\nu'_{\mathbb{C}}(\pi) = +\infty$ if there exists n such that $\pi_{\geq 2n}$ contains no deviation, and such that the play $\dot{\pi}_{\geq n}$ is not λ-consistent.
- $\nu'_{\mathbb{C}}(\pi) = \limsup_n \mathsf{MP}_i(\dot{\pi}_{\leq n})$ otherwise.

The payoff function $\nu'_{\mathbb{C}}$ is then defined as $\nu_{\mathbb{C}}$, but with a limit superior instead of inferior. The payoff function $\nu'_{\mathbb{C}}$ is concave. Indeed, let π and χ be two plays in $\mathsf{Conc}_{\lambda i}(G)_{\restriction v_0}$, and let ξ be a shuffling of them. Let us check that $\nu'_{\mathbb{C}}(\xi) \leq \max\{\nu'_{\mathbb{C}}(\pi), \nu'_{\mathbb{C}}(\chi)\}$.

If either $\nu'_{\mathbb{C}}(\pi) = +\infty$ or $\nu'_{\mathbb{C}}(\chi) = +\infty$, it is immediate. Otherwise, we also have $\nu'_{\mathbb{C}}(\xi) \neq +\infty$: if either π or χ contains infinitely many deviations, then so does ξ. If both contain finitely many deviations, then so does ξ: the states of ξ have therefore eventually the same memory M, which is also the memory of, eventually, the states of both π and χ. Now, since $\nu'_{\mathbb{C}}(\pi), \nu'_{\mathbb{C}}(\chi) \neq +\infty$, we have $\mu_j(\dot{\pi}), \mu_j(\dot{\chi}) \geq \lambda(v)$ for each player j and every $v \in M \cap V_j$. Since mean-payoff functions are convex, it is also the case for the play $\dot{\xi}$, which is a shuffling of $\dot{\pi}$ and $\dot{\chi}$. Hence $\nu_{\mathbb{C}}(\xi) \neq +\infty$.

Therefore, we have $\nu_{\mathbb{C}}(\xi) = \limsup_n \mathsf{MP}_i(\dot{\xi}_{\leq n})$, as well as $\nu_{\mathbb{C}}(\pi) = \limsup_n \mathsf{MP}_i(\dot{\pi}_{\leq n})$ and $\nu_{\mathbb{C}}(\dot{\chi}) = \limsup_n \mathsf{MP}_i(\dot{\chi}_{\leq n})$. Since, as shown in [VCD$^+$15], mean-payoff functions defined with a limit superior are concave, it implies $\nu_{\mathbb{C}}(\xi) \leq \max\{\nu_{\mathbb{C}}(\pi), \nu_{\mathbb{C}}(\chi)\}$: the payoff function $\nu_{\mathbb{C}}$ is concave.

Therefore, by Lemma 2.29 Challenger has a memoryless strategy that is optimal with regards to the payoff function $\nu'_{\mathbb{C}}$: let us write it $\tau_{\mathbb{C}}$. Now, we want to prove that the memoryless strategy $\tau_{\mathbb{C}}$ is also optimal with regards to $\nu_{\mathbb{C}}$. Note that for every play π, we have $\nu_{\mathbb{C}}(\pi) \leq \nu'_{\mathbb{C}}(\pi)$, and therefore $\mathsf{val}_{\mathbb{C}}(\mathsf{Conc}_{\lambda i}(G)_{\restriction s_0}) \leq \alpha$, where α is the value of the game $\mathsf{Conc}_{\lambda i}(G)_{\restriction s_0}$ with the payoff function $\nu'_{\mathbb{C}}$ instead of $\nu_{\mathbb{C}}$. Therefore, we have proven that $\tau_{\mathbb{C}}$ is optimal with regards to $\nu_{\mathbb{C}}$ if we prove that $\inf_{\tau_{\mathbb{P}}} \nu_{\mathbb{C}}(\langle \bar{\tau} \rangle) \geq \alpha$.

Let π be a play compatible with $\tau_{\mathbb{C}}$, i.e. an infinite path from s_0 in the graph $\mathsf{Conc}_{\lambda i}(G)_{\restriction s_0}[\tau_{\mathbb{C}}]$. If $\nu_{\mathbb{C}}(\pi) = +\infty$, then clearly $\nu_{\mathbb{C}}(\pi) \geq \alpha$. Otherwise, we have $\nu_{\mathbb{C}}(\pi) = \mu_i(\dot{\pi})$, and by Lemma 2.16, we have:
$$\mu_i(\dot{\pi}) \geq \min_{c \in C} \mathsf{MP}_i(\dot{c}),$$
where C is the set of the simple cycles of the graph $\mathsf{Conc}_{\lambda i}(G)_{\restriction s_0}[\tau_{\mathbb{C}}]$. Now, for each such cycle, there exists a history H such that the play Hc^{ω} is compatible with the strategy $\tau_{\mathbb{C}}$,

and therefore satisfies $\nu'_{\mathbb{C}}(Hc^{\omega}) \geq \alpha$, and consequently $\mathsf{MP}_i(\dot{c}) \geq \alpha$. Therefore, we have $\mu_i(\dot{\pi}) \geq \alpha$, and the strategy $\tau_{\mathbb{C}}$ is optimal with regards to the payoff function $\nu_{\mathbb{C}}$. $\square$

Using this lemma, computing $\mathsf{nego}(\lambda)$ for any given λ amounts to looking for an optimal path for Prover in each graph $\mathsf{Conc}_{\lambda i}(G)_{\restriction s_0}[\tau_{\mathbb{C}}]$. When (V, E) is a graph, we write $\mathsf{SConn}(V, E)$ the set of its strongly connected components accessible from the vertex v.

Let K be a strongly connected subgraph of a concrete negotiation game. If K contains no deviation, then the states of K share all the same memory: let us us write it $\mathsf{Mem}(K)$. If K contains at least one deviation, we define $\mathsf{Mem}(K) = \emptyset$. Then, we write:

$$\mathsf{opt}(K) = \inf \left\{ x_i \;\middle|\; \begin{array}{c} \bar{x} \in {}^{\llcorner}\left(\underset{c \in \mathsf{SC}(K)}{\mathsf{Conv}} \mathsf{MP}_i(\dot{c}) \right), \\ \forall i \in \Pi, \forall v \in \mathsf{Mem}(K), x_i \geq \lambda(v) \end{array} \right\}.$$

The set in which the variable $\bar{x}$ evolves is the set of payoffs of plays that Prover can construct against Challenger, when she chooses to go in the strongly connected component K, and observes the requirements she has to observe — those stored in the common memory of the states of K if K contains no deviations, and none otherwise. Hence the following formal result.

Lemma 6.5. *Let G be a mean-payoff game, let λ be a requirement, let i be a player, and let $v \in V_i$. Then, we have:*

$$\mathsf{nego}(\lambda)(v) = \underset{\tau_{\mathbb{C}} \in \mathsf{ML}(\mathsf{Conc}_{\lambda i}(G))}{\sup} \quad \underset{K \in \mathsf{SConn}\left(\mathsf{Conc}_{\lambda i}(G)_{\restriction (v, \{v\})}[\tau_{\mathbb{C}}]\right)}{\inf} \mathsf{opt}(K).$$

Moreover, mean-payoff games are games with steady negotiation.

Proof. By Lemma 6.4, in the game $\mathsf{Conc}_{\lambda i}(G)_{\restriction s}$, where $s = (v, \{v\})$, there exists a memoryless strategy $\tau_{\mathbb{C}}$ which is optimal for Challenger. Therefore, the best payoff that Prover can ensure against every strategy of Challenger is the best payoff she can ensure against $\tau_{\mathbb{C}}$. It follows from Theorem 6.3 that the highest value player i can enforce against a hostile λ-rational environment is the minimal payoff of Challenger in a path in the graph $\mathsf{Conc}_{\lambda i}(G)_{\restriction s}[\tau_{\mathbb{C}}]$ starting from s. For any such path π, there exists a strongly connected component K of $\mathsf{Conc}_{\lambda i}(G)_{\restriction s}[\tau_{\mathbb{C}}]$ such that after a finite number of steps, the path π is a path in K. Let us now prove that the least payoff of Challenger in such a play is given by $\mathsf{opt}(K)$. Let us distinguish two cases.

- *If there is at least one deviation in K.*

  Then, for every play π in K, it is possible to transform π into a play π' with $\mu(\dot{\pi}') = \mu(\dot{\pi})$, which contains infinitely many deviations: it suffices to add round trips to a deviation, endlessly, but less and less often. Therefore, the outcomes $\nu_{\mathbb{C}}(\pi)$ of plays in K are exactly the mean-payoffs $\mu_i(\dot{\pi})$ of plays in K (plus possibly $+\infty$); and in particular, the lowest payoff Challenger can get in K is the quantity:

  $$\underset{c \in \mathsf{SC}(K)}{\min} \mathsf{MP}_i(\dot{c}),$$

  which is equal to $\mathsf{opt}(K)$ since $\mathsf{Mem}(K) = \emptyset$.

- *If there is no deviation in K.*

  By Lemma 2.16, the set of possible values of $\mu(\dot{\pi})$ for all plays π in K is exactly the set:

  $$X = {}^{\llcorner}\left(\underset{c \in \mathsf{SC}(K)}{\mathsf{Conv}} \hat{\mu}(c^{\omega}) \right).$$

Since all the plays in K contain finitely many deviations (actually none), for every path π in K, we have $\nu_{\mathbb{C}}(\pi) = +\infty$ if and only if there exists $j \in \Pi$ and $u \in V_j \cap \mathsf{Mem}(K)$ such that $\mu_j(\dot{\pi}) < \lambda(u)$. Then, the lowest outcome Prover can get in K is:

$$\inf\left\{x_i \mid \bar{x} \in X, \forall j \in \Pi, \forall u \in V_j \cap \mathsf{Mem}(K), x_j \geq \lambda(u)\right\},$$

that is to say $\mathsf{opt}(K)$.

Theorem 6.3 enables to conclude to the desired formula. Moreover, let us notice that in all cases, Prover can choose one optimal play against each memoryless strategy of Challenger. By determinacy of Borel games, it comes that Prover has an optimal strategy, hence by Theorem 6.3, mean-payoff games are games with steady negotiation. □

We are now able to compute $\mathsf{nego}(\lambda)$ for a given λ. However, because of Theorem 5.3, that is not sufficient to compute the least ε-fixed point λ^*, and then to decide the SPE threshold problem. Nevertheless, we will prove that λ^* can also be extracted from the concrete negotiation game.

7. Analysis of the negotiation function

Definition 7.1 (Piecewise affine function)**.** Let D be a finite set of dimensions. A function $f : \overline{\mathbb{R}}^D \to \overline{\mathbb{R}}^D$ is *piecewise affine* if for each $d \in D$, there exists a finite partition Φ_d of $\overline{\mathbb{R}}^D$, where every $P \in \Phi_d$ is a polyhedron, such that for each such P there exists $\bar{a}_P \in \mathbb{R}^D \setminus \{\bar{0}\}$ and $b_P \in \overline{\mathbb{R}}$ such that for every $\bar{x} \in P$, the vector $\bar{y} = f(\bar{x})$ satisfies:

$$y_d = \bar{a}_P \cdot \bar{x} + b_P,$$

where $\cdot$ is the canonical scalar product.

Remark 7.2. The function f is fully represented by the family $(\Phi_d, (\bar{a}_P, b_P)_{P \in \Phi_d})_{d \in D}$. That representation is finite if each polyhedron $P \in \Phi_d$, for each dimension d, is defined by rational equations, and if each $\bar{a}_P$ and each b_P has rational or infinite values.

Theorem 7.3. *Let us assimilate every requirement λ to the vector $\lambda = (\lambda(v))_{v \in V}$. Then, the negotiation function is piecewise affine, and a finite representation of it can be constructed in a time doubly exponential in the size of G.*

Proof. Let $i \in \Pi$, let $v \in V_i$, and let λ be a requirement. Let $\tau_{\mathbb{C}}$ be a memoryless strategy of Challenger in the game $\mathsf{Conc}_{\lambda i}(G)$, and let K be a strongly connected component of the graph $\mathsf{Conc}_{\lambda i}(G)_{\restriction s}$, where $s = (v, \{v\})$. The result will follow from the fact that the quantity $\mathsf{opt}(K)$ is, itself, a piecewise affine function of λ. Note that the underlying graph of the game $\mathsf{Conc}_{\lambda i}(G)$ does not depend on λ.

For a given λ, let us consider the polytope $Q = \mathsf{Conv}_{c \in \mathsf{SC}(K)} \mathsf{MP}(\dot{c}) \subseteq \mathbb{R}^\Pi$, and the polyhedron $R_\lambda = \{\bar{x} \in \mathbb{R}^\Pi \mid \forall i \in \Pi, \forall u \in \mathsf{Mem}(K), x_i \geq \lambda(u)\}$. Then, we have:

$$\mathsf{opt}(K) = \inf\{x_i \mid \bar{x} \in {}^\llcorner Q \cap R_\lambda\}$$

$$= \inf\left\{\min_{\bar{z} \in Z} z_i \;\middle|\; \begin{array}{l} Z \text{ is a finite subset of } Q, \\ \text{and } (\min_{\bar{z} \in Z} z_j)_j \in R_\lambda \end{array}\right\}$$

$$= \inf\left\{\min_{\bar{z} \in Z} z_i \;\middle|\; Z \text{ is a finite subset of } Q \cap R_\lambda\right\}$$

$$= \inf\left\{x_i \mid \bar{x} \in Q \cap R_\lambda\right\}.$$

From now, we can therefore drop the downward sealing. Moreover, let us note that the projection $\bar{x} \mapsto x_i$, as an affine mapping over the polytope $Q \cap R_\lambda$, finds its minimum on a vertex of that polytope. Each vertex $\{\bar{x}\}$ of $Q \cap R_\lambda$ is the intersection between a face F of the polytope Q, and a face F' of the polyhedron R_λ. Such a face F is of the form $\mathsf{Conv}_{c \in C}\mathsf{MP}(\dot{c})$, where C is a subset of $\mathsf{SC}(K)$; and such a face F' is of the form $\bigcap H_{\lambda w}$, where W is a subset of $\mathsf{Mem}(K)$, and where $H_{\lambda w}$ is the hyperplane $\{\bar{x} \mid x_j = \lambda(w)\}$ for each j and $w \in V_j$. Thus, if we define the set:

$$X = \left\{ \bar{x} \in R_\lambda \ \middle| \ \begin{array}{l} \exists C \subseteq \mathsf{SC}(K), \exists W \subseteq \mathsf{Mem}(K), \\ \mathsf{Conv}_{c \in C}\mathsf{MP}(\dot{c}) \cap \bigcap_{w \in W} H_{\lambda w} = \{\bar{x}\} \end{array} \right\},$$

it is included in the polytope $Q \cap R_\lambda$, and contains all its vertices; hence the equality $\mathsf{opt}(K) = \inf\{x_i \mid \bar{x} \in X\}$. We can therefore write:

$$\mathsf{opt}(K) = \inf_{C \subseteq \mathsf{SC}(K), W \subseteq \mathsf{Mem}(K)} f_{CW}(\lambda),$$

where f_{CW} is, for every C and W, the function defined by:

- $f_{CW}(\lambda) = x_i$ if the intersection $I = (\mathsf{Aff}_{c \in C}\mathsf{MP}(\dot{c})) \cap (\bigcap_{w \in W} H_{\lambda w})$, where $\mathsf{Aff}\,Z$ denotes the smallest affine space containing Z, is a singleton $\{\bar{x}\}$ with $\bar{x} \in \mathsf{Conv}_{c \in C}\mathsf{MP}(\dot{c}) \cap R_\lambda$,
- and $f_{CW}(\lambda) = +\infty$ otherwise.

Let us now study each of those functions f_{CW}. Given C, W and λ, we have $f_{CW}(\lambda) \neq +\infty$ if and only if the three following conditions are satisfied.

- First, the intersection I is a singleton. The elements of I are the points of the form $\bar{x} = \left(\sum_{c \in C} \alpha_c \mathsf{MP}_j(\dot{c})\right)_j$ with $\bar{\alpha} \in \mathbb{R}^C$, $\sum_c \alpha_c = 1$, and $x_j = \lambda(w)$ for each j and $w \in W \cap V_j$. The set I is therefore a singleton if and only if the matrix:

$$A = \left(\begin{cases} 1 & \text{if } w = \mathsf{sum} \\ \mathsf{MP}_j(\dot{c}) & \text{else, with } j \text{ s.t. } w \in V_j \end{cases} \right)_{w \in W \cup \{\mathsf{sum}\}, c \in C} \in \mathbb{R}^{(W \cup \{\mathsf{sum}\}) \times C}$$

is such that there exists exactly one vector $\bar{\alpha} \in \mathbb{R}^C$ satisfying:

$$A\bar{\alpha} = \left(\begin{cases} 1 & \text{if } j = \mathsf{sum} \\ \lambda(w) & \text{otherwise} \end{cases} \right)_{w \in W \cup \{\mathsf{sum}\}}.$$

That condition is satisfied if and only if A is invertible, which can be decided in a time polynomial in the size of A, and does actually not depend on λ: either it is not satisfied, and the function f_{CW} is constantly equal to $+\infty$, or it is, and only the following conditions must be considered.

- Second, the unique element of I belongs to $\mathsf{Conv}_{c \in C}\mathsf{MP}(\dot{c})$. That is the case if and only if the vector:

$$\bar{\alpha} = A^{-1} \left(\begin{cases} 1 & \text{if } j = \mathsf{sum} \\ \lambda(w) & \text{otherwise} \end{cases} \right)_{w \in W \cup \{\mathsf{sum}\}}$$

has only non-negative coordinates. Here we should decompose:

$$\left(\begin{cases} 1 & \text{if } j = \mathsf{sum} \\ \lambda(w) & \text{otherwise} \end{cases} \right)_{w \in W \cup \{\mathsf{sum}\}} = B\lambda + \bar{\beta}$$

where $B = \left(\begin{cases} 1 & \text{if } w = v \\ 0 & \text{otherwise} \end{cases} \right)_{w \in W \cup \{\mathsf{sum}\}, v \in V}$ and $\bar{\beta} = \left(\begin{cases} 1 & \text{if } w = \mathsf{sum} \\ 0 & \text{otherwise} \end{cases} \right)_{w \in W \cup \{\mathsf{sum}\}}.$

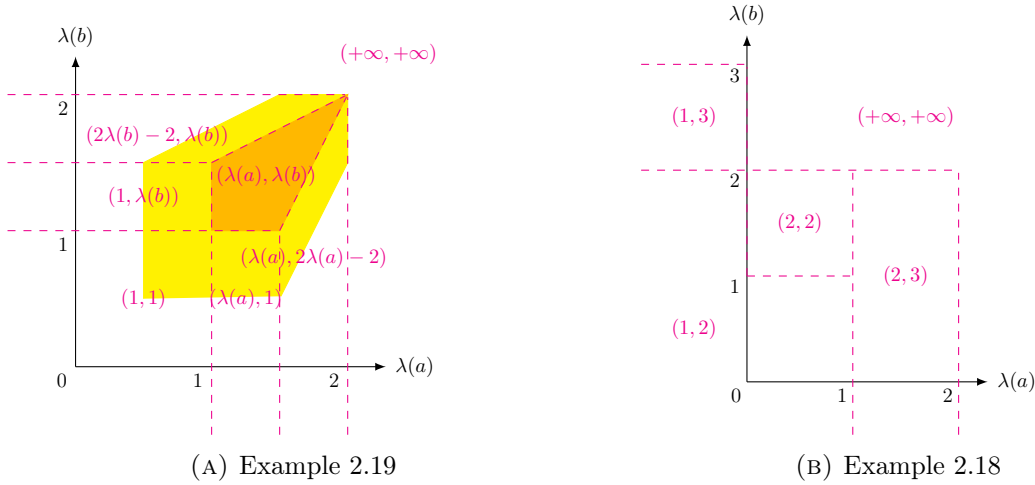(A) Example 2.19   (B) Example 2.18

Figure 7. The negotiation function on the games of Examples 2.19 and 2.18

Thus, the vector $\bar{\alpha} = A^{-1}(B\lambda + \bar{\beta})$ has non-negative coordinates if and only if λ belongs to the set:

$$P_0 = \left(\lambda' \mapsto A^{-1}(B\lambda' + \bar{\beta})\right)^{-1}\left([0, +\infty)^C\right)$$

which, as a pre-image of a polyhedron by an affine function, is itself a polyhedron, which can be constructed in a time polynomial in the size of A, B and β.

• Third, the vector:

$$\bar{x} = \left(\mathsf{MP}_j(\dot{c})\right)_{j\in\Pi, c\in C} \bar{\alpha}$$

is such that for each j and each $w \in \mathsf{Mem}(K)$ (not only in W), we have $x_j \geq \lambda(w)$. The set P_1 of requirements λ satisfying that condition can itself be written as the pre-image of a polyhedron by an affine function, and is therefore itself a polyhedron, which we can construct in a time polynomial in the size of A, B, β and $\left(\mathsf{MP}_j(\dot{c})\right)_{j\in\Pi, c\in C}$.

Therefore, the function f_{CW} is equal to $+\infty$ outside of the polyhedron $P_0 \cap P_1$, and satisfies:

$$f_{CW}(\lambda) = \left(\mathsf{MP}_i(\dot{c})\right)_{c\in C} \cdot A^{-1}(B\lambda + \bar{\beta})$$

inside it. It is therefore an affine function of which a representation can be constructed in a time polynomial in the size of K. Therefore, a representation of $\mathsf{opt}(K) = \inf_{C,W} f_{CW}(\lambda)$ as an affine function of λ can be constructed in a time exponential in the size of K, and the negotiation function, expressed by:

$$\mathsf{nego} : \lambda \mapsto \sup_{\tau_\mathbb{C}} \inf_K \inf_{C,W} f_{CW}(\lambda)$$

can be constructed in a time doubly exponential in the size of G.   □

Example 7.4. Let us consider the game of Example 2.19. If a requirement λ is represented by the tuple $(\lambda(a), \lambda(b))$, the function $\mathsf{nego} : \mathbb{R}^2 \to \mathbb{R}^2$ can be depicted by Figure 7a, where in any one of the regions delimited by the dashed lines, we wrote a formula for the couple $(\mathsf{nego}(\lambda)(a), \mathsf{nego}(\lambda)(b))$. The orange area indicates the fixed points of the function, and the yellow area the other $\frac{1}{2}$-fixed points.

Example 7.5. Now, let us consider the game of Example 2.18. Let us fix $\lambda(c) = 1$ and $\lambda(d) = 2$, and represent the requirements λ by the tuples $(\lambda(a), \lambda(b))$, as in the previous example. Then, the negotiation function is depicted by Figure 7b. One can check that there is no fixed point here, and even no $\frac{1}{2}$-fixed point — except $(+\infty, +\infty)$.

Consequently, the least ε-fixed point of the negotiation function can itself be computed using the classical linear algebra tool box.

8. Conclusion: algorithm and complexity

Thanks to all the previous results, we are now able to state the decidability of the ε-SPE threshold problem, and to bound its complexity. Let us start with a lower bound.

8.1. **Lower bound.**

Theorem 8.1. *The ε-SPE threshold problem is* **NP***-hard, even when ε is fixed equal to 0.*

Proof. We proceed by reduction from the **NP**-complete problem SAT. This proof is liberally inspired from the proof of the **NP**-hardness of the NE threshold problem in co-Büchi games by Michael Ummels, in [Umm08].

Let $\varphi = \bigwedge_{i=1}^{n} \bigvee_{j=1}^{m} L_{ij}$ be a formula from propositional logic, written in conjunctive normal form, over the finite variable set X. We construct a mean-payoff game $G^{\varphi}_{\restriction v_0}$ that admits an SPE where the player $\mathbb{S}$ gets the payoff 1, if and only if φ is satisfiable.

First, we define the set of players $\Pi = \{\mathbb{S}\} \cup X$: every variable of φ is a player and there is an additional special player $\mathbb{S}$, called *Solver*, who wants to prove that φ is satisfiable.

Then, let us define the state space: for each clause C_i, with $i \in \mathbb{Z}/n\mathbb{Z}$, of φ, we define a state C_i that is controlled by Solver, and for each litteral L_{ij} of C_i we define a state (C_i, L_{ij}), that is controlled by the player x such that $L_{ij} = x$ or $\neg x$. We add a transition from C_i to (C_i, L_{ij}), and another one from (C_i, L_{ij}) to C_{i+1}. Moreover, we add a sink state $\perp$, with a transition from it to itself, and transitions from all the states of the form $(C, \neg x)$ to it.

We define the reward function r on this game as follows:

- $r_{\mathbb{S}}(\perp\perp) = 0$, and $r_{\mathbb{S}}(uv) = 1$ for any other transition vw;
- for each player x, we have $r_x(uv) = 0$ for every transition leading to a state of the form $v = (C, x)$, and $r_x(uv) = 1$ for any other transition.

Note that Solver can only get the payoffs 0 (in a play that reaches $\perp$) or 1 (in any other play). Another player x gets the payoff 1 in a play that never visits (or finitely often, or infinitely often but with negligible frequence) a vertex of the form (C, x). Otherwise, he may get any payoff between 0.5 and 1, depending on the frequence with which such a state is visited. Finally, we initialize that game in $v_0 = C_1$.

Example 8.2. The game G^{φ}, when φ is the tautology $(x_1 \vee \neg x_1) \wedge \cdots \wedge (x_6 \vee \neg x_6)$, is represented by Figure 8. The rewards that are not written are equal to 1.

Now, let us prove that there is an SPE in $G^{\varphi}_{\restriction v_0}$ in which Solver gets the payoff 1, if and only if the formula φ is satisfiable, that is, if there exists a valuation $\nu : X \to \{0, 1\}$ that satisfies it.
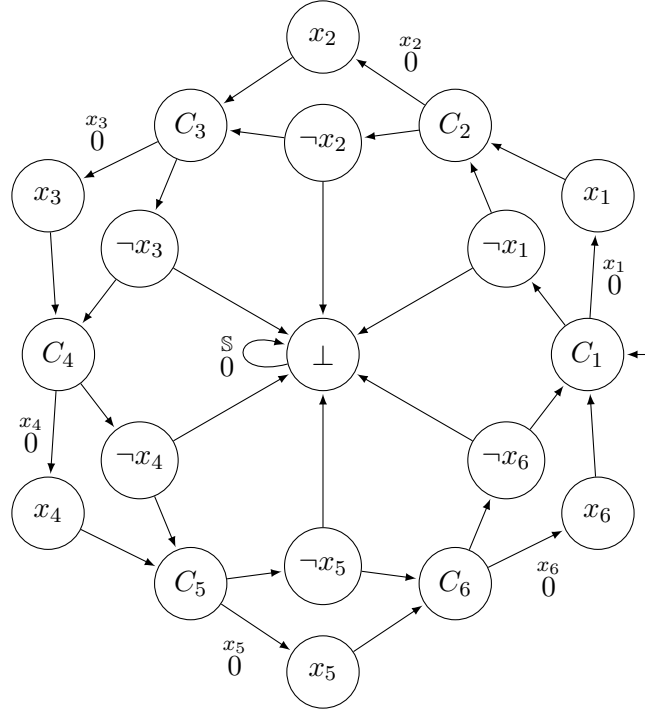
Figure 8. The game G^φ

- If such an SPE exists: let us write it $\bar{\sigma}$, and let $\rho = \langle \bar{\sigma} \rangle$. Since $\mu_\mathbb{S}(\rho) = 1$, the sink state $\bot$ is never visited. Let us define a valuation ν on X as follows: for each variable x, we have $\nu(x) = 1$ if and only if $\mu_x(\rho) < 1$.

  Now, let C be a clause of φ: since C, as a state, is necessarily visited infinitely often and with a fixed frequence in the play ρ (because no player ever go to the sink state $\bot$), one of its successors, say (C, L), is visited with a non-negligible frequence (more formally, the time between two occurrences of (C, L) is bounded). If L is a positive litteral, say x, then by definition of ν, we have $\nu(x) = 1$ and the clause C is satisfied.

  If L has the form $\neg x$, then each time the state $(C, \neg x)$ is traversed, player x has the possibility to deviate and to go to the sink state $\bot$, where he is sure to get the payoff 1. Since $\bar{\sigma}$ is an SPE, it means that he already gets the payoff 1 in the play ρ. By definition of ν, we then have $\nu(x) = 0$, hence the litteral $\neg x$ is satisfied, hence so is the clause C.

  The valuation ν satisfies all the clauses of φ, and therefore satisfies the formula φ itself.
- If φ is satisfied by some valuation ν: let us define a strategy profile $\bar{\sigma}$ by:
  - $\sigma_\mathbb{S}(hC) = (C, L)$ for each history hC where C is a clause of φ, where L is a litteral of C that is satisfied in the valuation ν;
  - and $\sigma_x(h(C, \neg x)) = \bot$ if and only if $\nu(x) = 1$ for each history $h(C, \neg x)$ where C is a clause of φ and x is a variable.

  Any other state has only one successor, hence we now have completely defined a strategy profile. Now, let us prove it is an SPE, in which Solver gets the payoff 1.

  Let hC be a history, where C is a clause of φ. We want to prove that $\bar{\sigma}_{\restriction hC}$ is a Nash equilibrium, in which Solver gets the payoff 1. Let $\rho = \langle \bar{\sigma}_{\restriction hC} \rangle$. If $\mu_\mathbb{S}(\rho) < 1$, i.e. if ρ is of the form $hD(D, \neg x)\bot^\omega$, then by definition of $\bar{\sigma}$ we have $\nu(x) = 0$. But then, we cannot

have $\sigma_{\mathbb{S}}(D) = (D, \neg x)$: contradiction. The play ρ never reaches the state $\bot$, and Solver gets the payoff 1, and as a consequence she does not have any profitable deviation.

Now, if another player x has a profitable deviation, it means that he does not get the payoff 1 in ρ, and therefore that some state of the form (D, x) is visited infinitely often. But then, if Solver choose to go to the state (D, x), it means that the litteral x is satisfied in ν, i.e. that $\nu(x) = 1$. In that case, if some clause D' contains the litteral $\neg x$, it is not a litteral satisfied by ν, and therefore the strategy $\sigma_{\mathbb{S}}$, as we defined it, never chooses the transition to the state $(D', \neg x)$, where player x could have the possibility to deviate from his strategy. Contradiction.

Finally, after a history of the form $h(C, L)$, either:
- $L = \neg x$ with $\nu(x) = 1$, and in that case, we have $\langle \bar{\sigma}_{\upharpoonright h(C,L)} \rangle = (C, L)\bot^{\omega}$, player x gets the payoff 1, and no player has a profitable deviation;
- or L is a positive litteral, and then there exists only one transition from the state (C, L) to another clause D, and we go back to the previous case;
- or $L = \neg x$ with $\nu(x) = 0$, and in that case, we have $\sigma_x(C, L) = D$ where D is the following clause, and by the first case the strategy profile $\bar{\sigma}_{\upharpoonright h(C,\neg x)D}$ is a Nash equilibrium. Moreover, since the litteral $\neg x$ is not satisfied in ν, the play $\langle \bar{\sigma}_{\upharpoonright h(C,\neg x)D} \rangle$ does never traverse again any state of the form $(D', \neg x)$, hence player x wins, and therefore has no profitable deviation: the strategy profile $\bar{\sigma}_{\upharpoonright h(C,\neg x)}$ is a Nash equilibrium.

The SPE threshold problem, and therefore the ε-SPE threshold problem, are **NP**-hard in mean-payoff games. $\qquad\square$

8.2. **Decidability and upper bound.** We can now present an algorithm that decides the ε-SPE threshold problem.

Theorem 8.3. *The ε-SPE threshold problem in mean-payoff games is decidable and* 2**ExpTime**-*easy.*

Proof. Given $G_{\upharpoonright v_0}, \bar{x}, \bar{y}$ and ε, by Theorem 7.3, an effective representation of the negotiation function can be computed in a time doubly exponential in the size of G; therefore, its least ε-fixed point λ^* can be computed in a time double exponential in the size of G and of ε, using the classical tools given by linear algebra.

Then, by Theorem 4.4 and since mean-payoff games are games with steady negotiation by Lemma 6.5, the tuple $(G_{\upharpoonright v_0}, \bar{x}, \bar{y}, \varepsilon)$ forms a positive instance of the ε-SPE threshold problem if and only if there exists a λ^*-consistent play ρ in $G_{\upharpoonright v_0}$ with $\bar{x} \le \mu(\rho) \le \bar{y}$.

The existence of such a play can be decided in doubly exponential time, as follows: for each subset $W \subseteq V$, construct the game G_W defined as the game G in which all the vertices that do not belong to W have been omitted. Then, for each connected component K of the underlying graph of G_W that is accessible from v_0, construct the polytope:

$$P = \;^{\llcorner}\left(\underset{c \in \mathsf{SC}(K)}{\mathsf{Conv}} \mathsf{MP}(c) \right) \cap \prod_i [x_i, y_i] \cap \prod_i \left[\max_{v \in W \cap V_i} \lambda^*(v), +\infty \right),$$

with the representations and algorithm given in [CDE$^+$10], which requires a time exponential in the size of K. If one of those exponentially many polytopes is nonempty, then using Lemma 2.16, there exists a play ρ in $G_{\upharpoonright v_0}$ with $\bar{x} \le \mu(\rho) \le \bar{y}$ and $\mathsf{Occ}(\rho) \subseteq W$ and $\mu_i(\rho) \ge \lambda^*(v)$ for each $i \in \Pi$ and $v \in V_i \cap W$, i.e. that is λ^*-consistent. Conversely, if such a play exists, then for $W = \mathsf{Occ}(\rho)$ and $K = \mathsf{Inf}(\rho)$, the polytope P will be nonempty. $\qquad\square$

References

[BBG⁺19]   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 *CONCUR*, volume 140 of *LIPIcs*, pages 13:1–13:16. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019.

[BBMR15]   Thomas Brihaye, Véronique Bruyère, Noémie Meunier, and Jean-François Raskin. Weak subgame perfect equilibria and their application to quantitative reachability. In *24th EACSL Annual Conference on Computer Science Logic, CSL 2015, September 7-10, 2015, Berlin, Germany*, volume 41 of *LIPIcs*, pages 504–518. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2015.

[BCH⁺16]   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 International Conference, LATA 2016, Prague, Czech Republic, March 14-18, 2016, Proceedings*, volume 9618 of *Lecture Notes in Computer Science*, pages 3–23. Springer, 2016.

[BDS13]   Thomas Brihaye, Julie De Pril, and Sven Schewe. Multiplayer cost games with simple Nash equilibria. In *Logical Foundations of Computer Science, International Symposium, LFCS 2013, San Diego, CA, USA, January 6-8, 2013. Proceedings*, volume 7734 of *Lecture Notes in Computer Science*, pages 59–73. Springer, 2013.

[BMR14]   Véronique Bruyère, Noémie Meunier, and Jean-François Raskin. Secure equilibria in weighted games. In *CSL-LICS*, pages 26:1–26:26. ACM, 2014.

[BR15]   Romain Brenguier and Jean-François Raskin. Pareto curves of multidimensional mean-payoff games. In Daniel Kroening and Corina S. Pasareanu, editors, *Computer Aided Verification - 27th International Conference, CAV 2015, San Francisco, CA, USA, July 18-24, 2015, Proceedings, Part II*, volume 9207 of *Lecture Notes in Computer Science*, pages 251–267. Springer, 2015. `doi:10.1007/978-3-319-21668-3\_15`.

[BRPR16]   Véronique Bruyère, Stéphane Le Roux, Arno Pauly, and Jean-François Raskin. On the existence of weak subgame perfect equilibria. *CoRR*, abs/1612.01402, 2016.

[BRPR17]   Véronique Bruyère, Stéphane Le Roux, Arno Pauly, and Jean-François Raskin. On the existence of weak subgame perfect equilibria. In *Foundations of Software Science and Computation Structures - 20th International Conference, FOSSACS 2017, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2017, Uppsala, Sweden, April 22-29, 2017, Proceedings*, volume 10203 of *Lecture Notes in Computer Science*, pages 145–161, 2017.

[Bru17]   Véronique Bruyère. Computer aided synthesis: A game-theoretic approach. In *Developments in Language Theory - 21st International Conference, DLT 2017, Liège, Belgium, August 7-11, 2017, Proceedings*, volume 10396 of *Lecture Notes in Computer Science*, pages 3–35. Springer, 2017.

[BRvdB21] Léonard Brice, Jean-François Raskin, and Marie van den Bogaard. Subgame-perfect equilibria in mean-payoff games. *CoRR*, 2021.

[CDE⁺10]   Krishnendu Chatterjee, Laurent Doyen, Herbert Edelsbrunner, Thomas A. Henzinger, and Philippe Rannou. Mean-payoff automaton expressions. In Paul Gastin and François Laroussinie, editors, *CONCUR 2010 - Concurrency Theory, 21th International Conference, CONCUR 2010, Paris, France, August 31-September 3, 2010. Proceedings*, volume 6269 of *Lecture Notes in Computer Science*, pages 269–283. Springer, 2010.

[CHP10]   Krishnendu Chatterjee, Thomas A. Henzinger, and Nir Piterman. Strategy logic. *Inf. Comput.*, 208(6):677–693, 2010. `doi:10.1016/j.ic.2009.07.004`.

[FKM⁺10]   János Flesch, Jeroen Kuipers, Ayala Mashiah-Yaakovi, Gijs Schoenmakers, Eilon Solan, and Koos Vrieze. Perfect-information games with lower-semicontinuous payoffs. *Math. Oper. Res.*, 35(4):742–755, 2010.

[FP16]   János Flesch and Arkadi Predtetchinski. On refinements of subgame perfect ϵ - equilibrium. *Int. J. Game Theory*, 45(3):523–542, 2016. `doi:10.1007/s00182-015-0468-8`.

[FP17]   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`.

[Kop06]   Eryk Kopczynski. Half-positional determinacy of infinite games. In Michele Bugliesi, Bart Preneel, Vladimiro Sassone, and Ingo Wegener, editors, *Automata, Languages and Programming, 33rd*

International Colloquium, ICALP 2006, Venice, Italy, July 10-14, 2006, Proceedings, Part II, volume 4052 of *Lecture Notes in Computer Science*, pages 336–347. Springer, 2006.

[KPV16]   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`.

[Mar75]   Donald A. Martin. Borel determinacy. *Annals of Mathematics*, pages 363–371, 1975.

[Meu16]   Noémie Meunier. *Multi-Player Quantitative Games: Equilibria and Algorithms*. PhD thesis, Université de Mons, 2016.

[Osb04]   Martin J. Osborne. *An introduction to game theory*. Oxford Univ. Press, 2004.

[SV03]   Eilon Solan and Nicolas Vieille. Deterministic multi-player Dynkin games. *Journal of Mathematical Economics*, 39(8):911–929, 2003.

[Umm06]   Michael Ummels. Rational behaviour and strategy construction in infinite multiplayer games. In *FSTTCS 2006: Foundations of Software Technology and Theoretical Computer Science, 26th International Conference, Kolkata, India, December 13-15, 2006, Proceedings*, volume 4337 of *Lecture Notes in Computer Science*, pages 212–223. Springer, 2006.

[Umm08]   Michael Ummels. The complexity of Nash equilibria in infinite multiplayer games. In Roberto M. Amadio, editor, *Foundations of Software Science and Computational Structures, 11th International Conference, 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. Proceedings*, volume 4962 of *Lecture Notes in Computer Science*, pages 20–34. Springer, 2008. `doi:10.1007/978-3-540-78499-9\_3`.

[VCD+15]   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.

[VS03]   Nicolas Vieille and Eilon Solan. Deterministic multi-player Dynkin games. *Journal of Mathematical Economics*, Vol.39,num. 8:pp.911–929, November 2003. `doi:10.1016/S0304-4068(03)00021-1`.

Appendix A. Abstract negotiation game

Definition A.1 (Abstract negotiation game). Let $G_{\restriction v_0}$ be a game, let $i \in \Pi$, and let λ be a requirement on G. The *abstract negotiation game* of $G_{\restriction v_0}$ for player i with requirement λ is the two-player zero-sum game:

$$\mathsf{Abs}_{\lambda i}(G)_{\restriction [v_0]} = (\{\mathbb{P}, \mathbb{C}\}, S, (S_{\mathbb{P}}, S_{\mathbb{C}}), \Delta, \nu)_{\restriction [v_0]},$$

where:

- $\mathbb{P}$ denotes the player *Prover* and $\mathbb{C}$ the player *Challenger*;
- the states of $S_{\mathbb{C}}$ are written $[\rho]$, where ρ is a λ-consistent play in G;
- the states of $S_{\mathbb{P}}$ are written $[hv]$, where hv is a history in G, with $h \in \mathsf{Hist}_i(G)$, or $[v]$ with $v \in V$, plus two additional states $\top$ and $\bot$;
- the set Δ contains the transitions of the form:
  - $[hv][v\rho]$, where $[hv] \in S_{\mathbb{P}}$ and $[v\rho] \in S_{\mathbb{C}}$ (Prover proposes a play);
  - $[\rho][\rho_0...\rho_n v]$, where $[\rho] \in S_{\mathbb{C}}, n \in \mathbb{N}, \rho_n \in V_i$, and $v \neq \rho_{n+1}$ (Challenger makes player i deviate);
  - $[\rho]\top$, where $[\rho] \in S_{\mathbb{C}}$ (Challenger accepts the proposed play);
  - $\top\top$ (the game is over);
  - $[hv]\bot$ (Prover has no more play to propose);
  - $\bot\bot$ (the game is over).
- ν is the payoff function defined by, for all $\rho^{(0)}, \rho^{(1)}, \ldots, h^{(1)}v_1, h^{(2)}v_2, \ldots, k, H$:

$$\nu_{\mathbb{C}} \left([v_0] \left[\rho^{(0)} \right] \left[h^{(1)} v_1 \right] \left[\rho^{(1)} \right] \ldots \left[h^{(k)} v_k \right] \left[\rho^{(k)} \right] \top^{\omega} \right)$$
$$= \qquad \mu_i \left(h^{(1)} \ldots h^{(k)} \rho^{(k)} \right),$$

$$\nu_{\mathbb{C}} \left([v_0] \left[\rho^{(0)} \right] \left[h^{(1)} v_1 \right] \left[\rho^{(1)} \right] \ldots \left[h^{(n)} v_n \right] \left[\rho^{(n)} \right] \ldots \right)$$
$$= \qquad \mu_i \left(h^{(1)} h^{(2)} \ldots \right),$$

$$\nu_{\mathbb{C}} \left(H \bot^{\omega} \right) = +\infty,$$

and by $\nu_{\mathbb{P}} = -\nu_{\mathbb{C}}$.

Remark A.2. If the game G is Borel, then so is the game $\mathsf{Abs}_{\lambda i}(G)$.

Proposition A.3. *Let $G_{\restriction v_0}$ be a Borel game, let λ be a requirement on G and let $i \in \Pi$ be such that $v_0 \in V_i$. Then, we have:*

$$\mathsf{val}_{\mathbb{C}} \left(\mathsf{Abs}_{\lambda i}(G)_{\restriction [v_0]} \right) = \mathsf{nego}(\lambda)(v_0).$$

Proof. Let $\alpha \in \mathbb{R}$, and let us prove that the following statements are equivalent:

(1) there exists a strategy $\tau_{\mathbb{P}}$ such that for every strategy $\tau_{\mathbb{C}}$, $\nu_{\mathbb{C}}(\langle \bar{\tau} \rangle) < \alpha$;
(2) there exists a λ-rational strategy profile $\bar{\sigma}_{-i}$ in the game $G_{\restriction v_0}$ such that for every strategy σ_i, we have $\mu_i \left(\langle \bar{\sigma} \rangle \right) < \alpha$.

- *(1) implies (2).*

  Let $\tau_{\mathbb{P}}$ be such that for every strategy $\tau_{\mathbb{C}}$, $\nu_{\mathbb{C}}(\langle \bar{\tau} \rangle) < \alpha$.

  In what follows, any history h compatible with an already defined strategy profile $\bar{\sigma}_{-i}$ in $G_{\restriction v_0}$ will be decomposed in:

$$h = v_0 h^{(0)} v_1 h^{(1)} \ldots h^{(n-1)} v_n h^{(n)},$$

so that there exist plays $\rho^{(0)}, \ldots, \rho^{(n-1)}, \eta$ and a history:

$$[v_0] \left[\rho^{(0)} \right] \left[v_1 h^{(1)} v_2 \right] \ldots \left[v_{n-1} h^{(n-1)} v_n \right] \left[v_n h^{(n)} \eta \right]$$

in the game $\mathsf{Abs}_{\lambda i}(G)$ compatible with $\tau_{\mathbb{P}}$: the existence and the unicity of that decomposition can be proved by induction. Intuitively, the history h is cut in histories which are prefixes of plays that can be proposed by Prover.

Then, let us define inductively the strategy profile $\bar{\sigma}_{-i}$ by, for every h such that $\bar{\sigma}_{-i}$ has been defined on the prefixes of h, and such that the last state of h is not controlled by player i, $\bar{\sigma}_{-i}(h) = \eta_0$ with η defined from h as higher. Let us prove that $\bar{\sigma}_{-i}$ is the desired strategy profile.

– *The strategy profile $\bar{\sigma}_{-i}$ is λ-rational.*

Let us define σ_i so that for every history hv compatible with $\bar{\sigma}_{-i}$, the play $\langle \bar{\sigma}_{\restriction hv} \rangle$ is λ-consistent.

For any history:

$$h = v_0 h^{(0)} v_1 h^{(1)} \ldots h^{(n-1)} v_n h^{(n)}$$

compatible with $\bar{\sigma}_{-i}$ and ending in V_i, let $\sigma_i(h) = \eta_0$ with η corresponding to the decomposition of h, so that by induction:

$$\langle \bar{\sigma}_{\restriction v_0 h^{(0)} v_1 h^{(1)} \ldots h^{(n-1)} v_n} \rangle = v_n h^{(n)} \eta.$$

Let now hv be a history in $G_{\restriction v_0}$, and let us show that the play $\langle \bar{\sigma}_{\restriction hv} \rangle$ is λ-consistent. If we decompose:

$$hv = v_0 h^{(0)} v_1 h^{(1)} \ldots h^{(n-1)} v_n h^{(n)}$$

with the same definition of η (note that the vertex v is now included in the decomposition), then $\langle \bar{\sigma}_{\restriction hv} \rangle = v\eta$, and by definition of the abstract negotiation game, $v_n h^{(n)} \eta$ is a λ-consistent play, and therefore so is $v\eta$.

– *The strategy profile $\bar{\sigma}_{-i}$ keeps player i's payoff under the value α.*

Let σ_i be a strategy for player i, and let $\rho = \langle \bar{\sigma} \rangle$. We want to prove that $\mu_i(\rho) < \alpha$.

Let us define two finite or infinite sequences $\left(\rho^{(k)} \right)_{k \in K}$ and $\left(h^{(k)} v_k \right)_{k \in K}$, where $K = \{1, \ldots, n\}$ or $K = \mathbb{N} \setminus \{0\}$, by for every $k \in K$:

$$\left[\rho^{(k)} \right] = \tau_{\mathbb{P}} \left([v_0] \left[\rho^{(0)} \right] \ldots \left[\rho^{(k-1)} \right] \left[h^{(k)} v_k \right] \right)$$

and so that for every k, the history $h^{(k)} v_k$ is the shortest prefix of ρ that is not a prefix of $h^{(1)} \ldots h^{(k-1)} \rho^{(k-1)}$ (or equivalently, the history $h^{(k)}$ is the longest common prefix of ρ and $h^{(1)} \ldots h^{(k-1)} \rho^{(k-1)}$).

Then, the length of the longest common prefix of $h^{(1)} \ldots h^{(k-1)} \rho^{(k)}$ and ρ increases with k, and the set K is finite if and only if there exists n such that $h^{(1)} \ldots h^{(n-1)} \rho^{(n)} = \rho$. In the infinite case, let:

$$\chi = [v_0] \left[\rho^{(0)} \right] \left[h^{(1)} v_1 \right] \ldots \left[\rho^{(k)} \right] \left[h^{(k)} v_k \right] \ldots.$$

The play χ is compatible with $\tau_{\mathbb{P}}$, hence $\nu_{\mathbb{C}}(\chi) < \alpha$, that is to say:

$$\mu_i \left(h^{(1)} h^{(2)} \ldots \right) < \alpha,$$

ie. $\mu_i(\rho) < \alpha$.

In the finite case, let:

$$\chi = [v_0]\left[\rho^{(0)}\right]\left[h^{(1)}v_1\right]\ldots\left[\rho^{(n)}\right]\top^\omega.$$

For the same reason, $\nu_\mathbb{C}(\chi) < \alpha$, that is to say $\mu_i\left(h^{(1)}\ldots h^{(n)}\rho^{(n)}\right) = \mu_i(\rho) < \alpha$.

- *(2) implies (1).*

  Let $\bar\sigma_{-i}$ be a strategy profile keeping player i's payoff below α, λ-rational assuming a strategy σ_i. Let us define a strategy $\tau_\mathbb{P}$ for Prover in the abstract negotiation game.

  Let $H = [v_0]\left[\rho^{(0)}\right]\left[h^{(1)}v_1\right]\left[\rho^{(1)}\right]\ldots\left[h^{(n)}v_n\right]$ be a history in the abstract game, ending in $S_\mathbb{P}$. Then, we define:

  $$\tau_\mathbb{P}(H) = \left[\langle\bar\sigma_{\restriction h^{(1)}\ldots h^{(n)}v_n}\rangle\right].$$

  If H is a history ending in $\top$, then $\tau_\mathbb{P}(H) = \top$, and in the same way if H ends in $\bot$, then $\tau_\mathbb{P}(H) = \bot$.

  Let us show that $\tau_\mathbb{P}$ is the strategy we were looking for. Let χ be a play compatible with $\tau_\mathbb{P}$, and let us note that the state $\bot$ does not appear in χ. Then, the play χ can only have two forms:

  − If $\chi = [v_0]\left[\rho^{(0)}\right]\left[h^{(1)}v_1\right]\ldots\left[\rho^{(n)}\right]\top^\omega$, then we have:

  $$\rho^{(n)} = \langle\bar\sigma_{\restriction h^{(1)}\ldots h^{(n)}v_n}\rangle,$$

  and the history $h^{(1)}\ldots h^{(n)}v_n$ in the game $G_{\restriction v_0}$ is compatible with $\bar\sigma_{-i}$. By hypothesis, we have:

  $$\mu_i\left(h^{(1)}\ldots h^{(n)}\rho^{(n)}\right) < \alpha,$$

  hence $\nu_\mathbb{C}(\chi) < \alpha$.

  − If $\chi = [v_0]\left[\rho^{(0)}\right]\ldots\left[h^{(n)}v_n\right]\left[\rho^{(n)}\right]\ldots$, then the play $\rho = h^{(1)}h^{(2)}\ldots$ is compatible with $\bar\sigma_{-i}$, and by hypothesis $\mu_i(\rho) < \alpha$, hence $\nu_\mathbb{C}(\chi) < \alpha$. □

Appendix B. Some examples of negotiation sequences

We gather in this section some examples that could be interesting for the reader who would want to get a full overall view on the behaviour of the negotiation function on the mean-payoff games.

Example B.1. Let us take again the game of Example 2.18: let us give (in red) the values of $\lambda_1 = \mathsf{nego}(\lambda_0)$, which correspond to the antagonistic values.
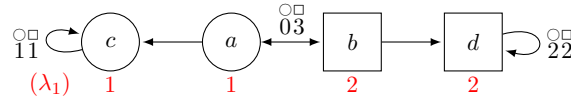

At the second step, let us execute the abstract game on the state a, with the requirement λ_1: whatever Prover proposes at first, Challenger has the possibility to deviate and to reach the state b. Then, Prover has to propose a λ_1-consistent play from the state b, i.e. a play in which player ○ gets at least the payoff 2: such a play necessarily ends in the state d, and gives player □ the payoff 2.

The other states keep the same values.

(λ_2)
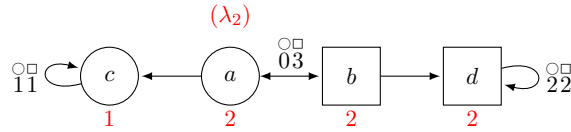
But then, at the third step, from the state b: whatever Prover proposes at first, Challenger can deviate to reach the state a. Then, Prover has to propose a λ_2-consistent play from a, i.e. a play in which player $\bigcirc$ gets at least the payoff 2: such a play necessarily end in the state d, i.e. after possibly some prefix, Prover proposes the play abd^ω. But then, Challenger can always deviate to go back to the state a; and the play which is thus created is $(ab)^\omega$ which gives player $\square$ the payoff 3.

(λ_3)
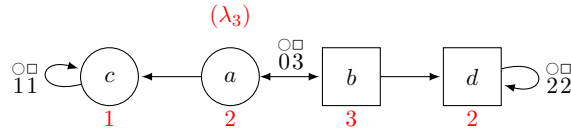
Finally, from the states a and b, there exists no λ_3-consistent play, and therefore no λ-rational strategy profile.

(λ_4)
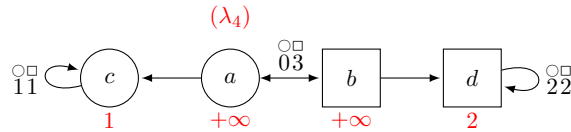
and for all $n \geq 4$, $\lambda_n = \lambda_4$.

Example B.2. In all the previous examples, all the games whose underlying graphs were strongly connected contained SPEs. Here is an example of game with a strongly connected underlying graph that does not contain SPEs. This game is similar to the game of Example 2.18, hence we do not give the details of the computation of the negotiation sequence.

(λ_1)
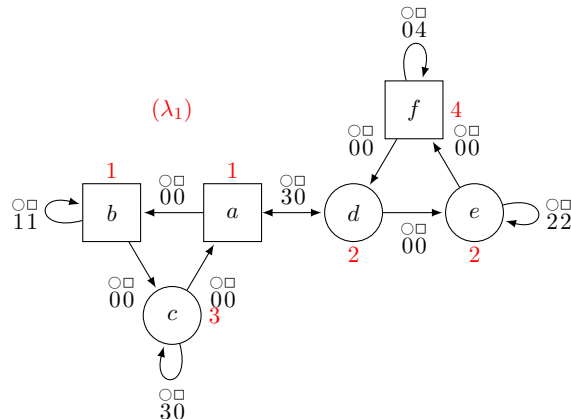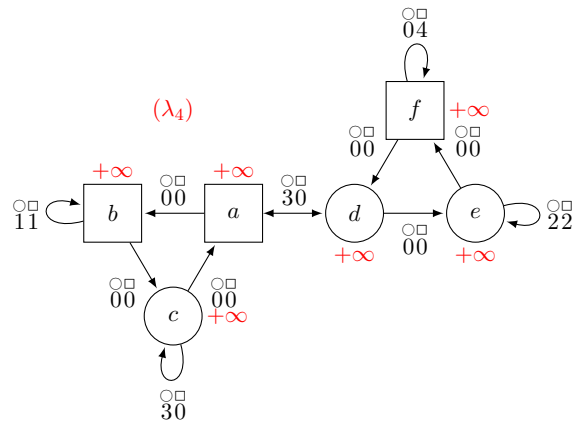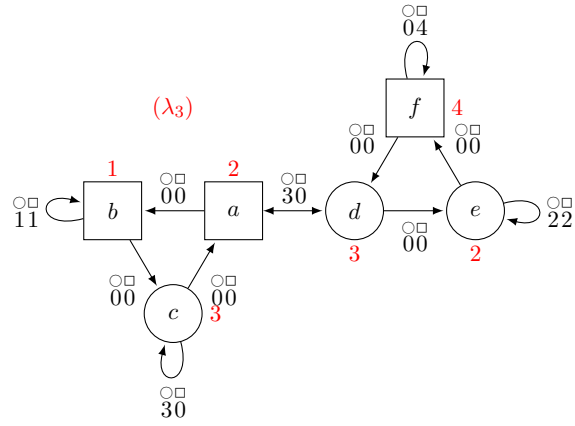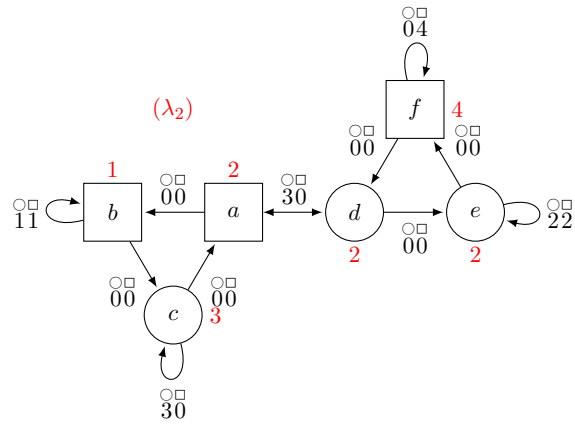
Example B.3. This example shows how a new requirement can emerge from the combination of several cycles.

Let G be the following game:
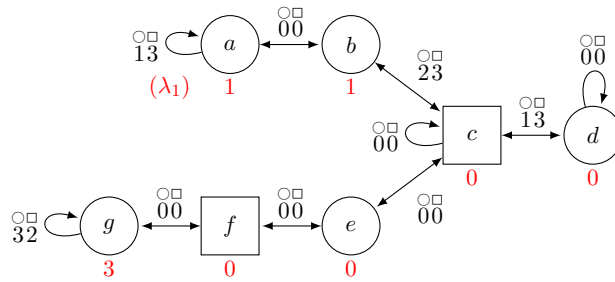
At the first step, the requirement λ_1 captures the antagonistic values.
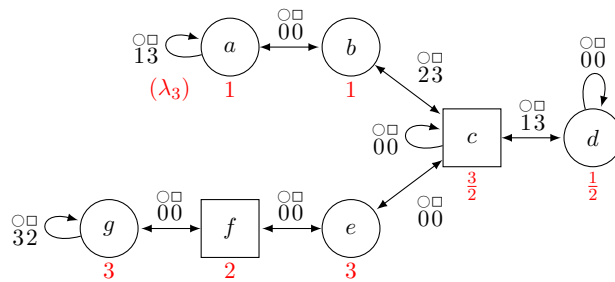
Then, from the state c, if player $\square$ forces the access to the state b, then player $\bigcirc$ must get at least 1: the worst play that can be proposed to player $\square$ is then $(babc)^\omega$, which gives player $\square$ the payoff $\frac{3}{2}$.

From the state f, if player $\square$ forces the access to the state g, then the worst play that can be proposed to them is g^ω.
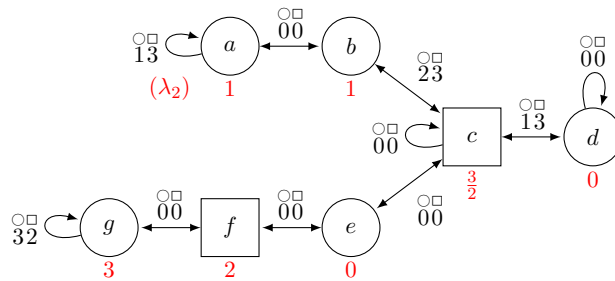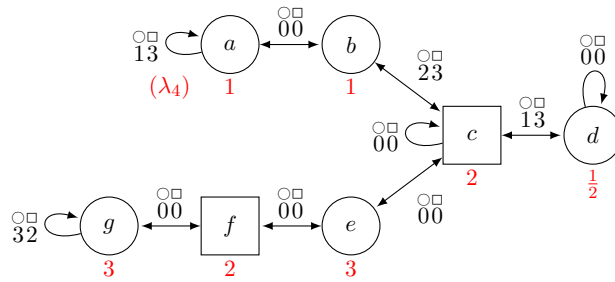

Then, from the state d, if player $\bigcirc$ forces the access to the state c, then player $\square$ must get at least $\frac{3}{2}$: the worst play that can be proposed to player $\bigcirc$ is then $(cccd)^\omega$, which gives player $\bigcirc$ the payoff $\frac{1}{2}$.
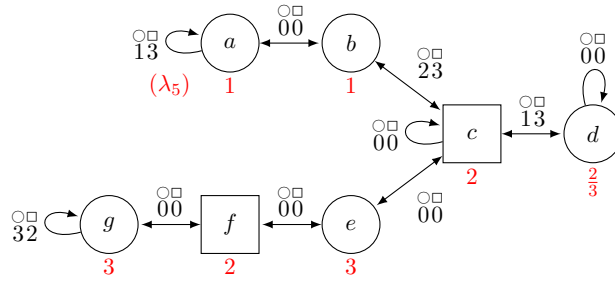
At the same time, from the state e, player $\bigcirc$ can now force the acces to the state f: then, the worst play that can be proposed to them is fg^ω.

But then, from the state c, player $\square$ can now force the access to the state e: then, the worst play that can be proposed to them is efg^ω.

And finally, from that point, if from the state d player $\bigcirc$ forces the access to the state c, then player $\square$ must have at least the payof 2; and therefore, the worst play that can be proposed to player $\bigcirc$ is now $(ccd)^\omega$, which gives her the payoff $\frac{2}{3}$.

The requirement λ_5 is a fixed point of the negotiation function.