

Equilibria in quantitative reachability games^{*}

Thomas Brihaye, Véronique Bruyère, and Julie De Pril^{**}

University of Mons – Belgium

Abstract. In this paper, we study turn-based quantitative multiplayer non zero-sum games played on finite graphs with reachability objectives. In this framework each player aims at reaching his own goal as soon as possible. We prove existence of finite-memory Nash (resp. secure) equilibria in multiplayer (resp. two-player) games.

1 Introduction

General framework. The construction of correct and efficient computer systems (hardware or software) is recognized as an extremely difficult task. To support the design and verification of such systems, mathematical logic, automata theory [9] and more recently model-checking [6] have been intensively studied. The model-checking approach, which is now an important part of the design cycle in industries, have proved its efficiency when applied to systems which can be accurately modeled as a finite-state automaton. In contrast, the application of these techniques to computer software, complex systems like embedded systems or distributed systems has been less successful. This could be partly explained by the following reasons: classical automata-based models do not faithfully capture the complex interactive behavior of modern computational systems that are usually composed of several interacting components, also interacting with an environment that is only partially under control. Recent research works show that it is suitable to generalize automata models used in the classical approach to verification, with the more flexible and mathematically deeper game-theoretic framework [12, 13].

Game theory meets automata theory. The basic framework that extends computational models with concepts from game theory is the so-called two-player zero-sum games played on graphs [7]. Many problems in verification and design of reactive systems can be modeled with this approach, like modeling controller-environment interactions. Given a model of a system interacting with an hostile environment, given a control objective (like preventing the system to reach some bad configurations), the controller synthesis problem asks to build a controller ensuring that the control objective is enforced whatever the environment will do. Two-players zero-sum games played on graphs are adequate models to solve this

^{*} This work has been partly supported by the ESF project GASICS and a grant from the National Bank of Belgium.

^{**} This author is supported by a grant from L’Oreal-UNESCO/F.R.S.-FNRS.

problem [14]. Moves of Player 1 model actions of the controller whereas moves of Player 2 model the uncontrollable actions of the environment, and a winning strategy for Player 1 is an abstract form of a control program that enforces the control objective.

The controller synthesis problem is suitable to model purely antagonist interactions between a controller and an hostile environment. However in order to study more complex systems with more than two components whose objectives are not necessarily antagonists, we need multiplayer and non zero-sum games to model them adequately. Moreover, we are not looking for winning strategies, but rather try to find relevant notions of equilibria, for instance the famous notion of Nash equilibria [12]. On the other hand, only qualitative objectives have been considered so far to specify, for example, that a player must be able to reach a target set of states in the underlying game graph. But, in line with the previous point, we also want to express and solve games for quantitative objectives such as forcing the game to reach a particular set of states within a given time bound, or within a given energy consumption limit. In summary, we need to study *equilibria* for *multiplayer-player non zero-sum* games played on graphs with *quantitative* objectives. This article provides some new results in this research direction.

Related work. Several recent papers have considered two-player zero-sum games played on finite graphs with regular objectives enriched by some *quantitative aspects*. Let us mention some of them: games with *finitary objectives* [5], games with *prioritized requirements* [1], *request-response games* where the waiting times between the requests and the responses are minimized [10, 15], and games whose winning conditions are expressed via *quantitative languages* [2].

Other works concern qualitative non zero-sum games. The notion of secure equilibrium, an interesting refinement of Nash equilibrium, has been introduced in [4]. It has been proved that a unique secure equilibrium always exists for two-player non zero-sum games with regular objectives. In [8], general criteria ensuring existence of Nash equilibria, subgame perfect equilibria (resp. secure equilibria) are provided for n -player (resp. 2-player) games, as well as complexity results.

Finally, we mention reference [3] that combines both the quantitative and the non zero-sum aspects. It is maybe the nearest related work compared to us, however the framework and the objectives are pretty different. In [3], the authors study games played on graphs with terminal vertices where quantitative payoffs are assigned to the players. These games may have cycles but all the infinite plays form a single outcome (like in chess where every infinite play is a draw). In that paper, criteria are given that ensure existence of Nash (resp. subgame perfect) equilibria in pure and memoryless strategies.

Our contribution. We here study turn-based quantitative multiplayer non zero-sum games played on finite graphs with reachability objectives. In this framework each player aims at reaching his own goal as soon as possible. We focus on

existence results for two solution concepts: Nash equilibrium and secure equilibrium. We prove existence of Nash (resp. secure) equilibria in n -player (resp. 2-player) games. Moreover, we show that these equilibria can be chosen with finite memory. Our results are not a direct consequence of the existing results in the qualitative framework, they require some new proof techniques. To the best of our knowledge, this is the first general result about existence of equilibria in quantitative multiplayer player games played on graphs.

Organization of the paper. Section 2 is dedicated to definitions. We present the games and the equilibria we study. In Section 3 we first prove an existence result for Nash equilibria and provide the finite-memory characterization. Existence of secure equilibria in two-player games is then established. Detailed proofs and examples can be found in the Appendix.

2 Preliminaries

2.1 Definitions

We consider here *quantitative* games played on a graph where all the players have *reachability objectives*. It means that, given a certain set of vertices Goal_i , each player i wants to reach one of these vertices as soon as possible.

This section is mainly inspired by reference [8].

Definition 1. An infinite turn-based quantitative multiplayer reachability game is a tuple $\mathcal{G} = (\Pi, V, (V_i)_{i \in \Pi}, v_0, E, (\text{Goal}_i)_{i \in \Pi})$ where

- Π is a finite set of players,
- $G = (V, (V_i)_{i \in \Pi}, v_0, E)$ is a finite directed graph where V is the set of vertices, $(V_i)_{i \in \Pi}$ is a partition of V into the state sets of each player, $v_0 \in V$ is the initial vertex, and $E \subseteq V \times V$ is the set of edges, and
- $\text{Goal}_i \subseteq V$ is the goal set of player i .

We assume that each vertex has at least one outgoing edge. A *play* $\rho \in V^\omega$ (respectively a *history* $h \in V^+$) of \mathcal{G} is an *infinite* (respectively a *finite*) path through the graph G starting from vertex v_0 . Remark that a history is always non empty because it starts with vertex v_0 . The set $H \subseteq V^+$ is made up of all the histories of \mathcal{G} . A *prefix* (respectively *proper prefix*) \mathbf{p} of a history $h = h_0 \dots h_k$ is a finite sequence $h_0 \dots h_l$, with $l \leq k$ (respectively $l < k$), denoted by $\mathbf{p} \leq h$ (respectively $\mathbf{p} < h$). We similarly consider a prefix \mathbf{p} of a play ρ , denoted by $\mathbf{p} < \rho$.

We say that a play $\rho = \rho_0 \rho_1 \dots$ *visits* a set $S \subseteq V$ (respectively a vertex $v \in V$) if there exists $l \in \mathbb{N}$ such that ρ_l is in S (respectively $\rho_l = v$). The same terminology also stands for a history h . Similarly, we say that ρ *visits* S *after* (respectively *in*) a prefix $\rho_0 \dots \rho_k$ if there exists $l > k$ (respectively $l \leq k$) such that ρ_l is in S . For any play ρ we denote by $\text{Visit}(\rho)$ the set of $i \in \Pi$ such that ρ visits Goal_i . The set $\text{Visit}(h)$ for a history h is defined similarly. The function Last

returns, given a history $h = h_0 \dots h_k$, the last vertex h_k of h , and the *length* $|h|$ of h is the number k of its *edges*¹.

For any play $\rho = \rho_0 \rho_1 \dots$ of \mathcal{G} , we note $\text{Payoff}_i(\rho)$ the *payoff* of player i , defined by:

$$\text{Payoff}_i(\rho) = \begin{cases} l & \text{if } l \text{ is the least index such that } \rho_l \in \text{Goal}_i, \\ +\infty & \text{otherwise.} \end{cases}$$

We note $\text{Payoff}(\rho) = (\text{Payoff}_i(\rho))_{i \in \Pi}$ the *payoff profile* for the play ρ . The aim of each player i is to *minimize* his payoff, i.e. reach his goal set Goal_i as soon as possible.

A *strategy* of player i in \mathcal{G} is a function $\sigma : V^* \times V_i \rightarrow V$ assigning to each history hv ending in a vertex v of player i a next vertex $\sigma(hv)$ such that $(v, \sigma(hv))$ belongs to E . We say that a play $\rho = \rho_0 \rho_1 \dots$ of \mathcal{G} is *consistent* with a strategy σ of player i if $\rho_{k+1} = \sigma(\rho_0 \dots \rho_k)$ for all $k \in \mathbb{N}$ such that $\rho_k \in V_i$. The same terminology is used for a history h of \mathcal{G} . A *strategy profile* of \mathcal{G} is a tuple $(\sigma_i)_{i \in \Pi}$ where σ_i is a strategy for player i . It determines a unique play of \mathcal{G} consistent with each strategy σ_i , called the *outcome* of $(\sigma_i)_{i \in \Pi}$ and denoted by $\langle (\sigma_i)_{i \in \Pi} \rangle$.

A strategy σ of player i is *memoryless* if σ depends only on the current vertex, i.e. $\sigma(hv) = \sigma(v)$ for all $h \in H$ and $v \in V_i$. More generally, σ is a *finite-memory strategy* if the equivalence relation \approx_σ on H defined by $h \approx_\sigma h'$ if $\sigma(h\delta) = \sigma(h'\delta)$ for all $\delta \in V^*V_i$ has finite index. In other words, a finite-memory strategy is a strategy that can be implemented by a finite automaton with output. A strategy profile $(\sigma_i)_{i \in \Pi}$ is called *memoryless* or *finite-memory* if each σ_i is a memoryless or a finite-memory strategy, respectively.

For a strategy profile $(\sigma_i)_{i \in \Pi}$ with outcome ρ and a strategy σ'_j of player j ($j \in \Pi$), we say that *player j deviates from ρ after a prefix h of ρ* if there exists a prefix h' of ρ such that $h \leq h'$, h' is consistent with σ'_j and $\sigma'_j(h') \neq \sigma_j(h')$. We also say *player j deviates from ρ just after a prefix h of ρ* if h is consistent with σ'_j and $\sigma'_j(h) \neq \sigma_j(h)$.

We now introduce the notion of *Nash equilibrium* and *secure equilibrium*.

Definition 2. A strategy profile $(\sigma_i)_{i \in \Pi}$ of a game \mathcal{G} is a *Nash equilibrium* if for all player $j \in \Pi$ and for all strategy σ'_j of player j , we have:

$$\text{Payoff}_j(\rho) \leq \text{Payoff}_j(\rho')$$

where $\rho = \langle (\sigma_i)_{i \in \Pi} \rangle$ and $\rho' = \langle \sigma'_j, (\sigma_i)_{i \in \Pi \setminus \{j\}} \rangle$.

This definition means that player j (for all $j \in \Pi$) has no incentive to deviate since he increases his payoff when using σ'_j instead of σ_j . A strategy σ'_j such that $\text{Payoff}_j(\rho) > \text{Payoff}_j(\rho')$ is called a *profitable deviation* for player j with respect to $(\sigma_i)_{i \in \Pi}$. In this case either player j gets an infinite payoff for ρ and a finite payoff for ρ' (ρ' visits Goal_j , but ρ does not), or player j gets a finite payoff for ρ and a strictly lower payoff for ρ' (ρ' visits Goal_j earlier than ρ does).

¹ Note that the length is not defined as the number of vertices.

In order to define the notion of secure equilibrium² we first need to associate an appropriate binary relation \prec_j on the payoff profiles with each player $j \in \Pi$. Given two payoff profiles $(x_i)_{i \in \Pi}$ and $(y_i)_{i \in \Pi}$:

$$(x_i)_{i \in \Pi} \prec_j (y_i)_{i \in \Pi} \quad \text{iff} \quad (x_j > y_j) \vee (x_j = y_j \wedge \forall k \ x_k \leq y_k \wedge \exists k \ x_k < y_k).$$

We then say that *player j prefers $(y_i)_{i \in \Pi}$ to $(x_i)_{i \in \Pi}$* . In other words, player j prefers a payoff profile to another either if he can decrease his own payoff, or if he can increase the payoff of all his opponents, among which one is strictly increased, while keeping his own payoff.

Definition 3. *A strategy profile $(\sigma_i)_{i \in \Pi}$ of a game \mathcal{G} is a secure equilibrium if for all players $j \in \Pi$, there does not exist a strategy σ'_j of player j such that:*

$$\text{Payoff}(\rho) \prec_j \text{Payoff}(\rho')$$

where $\rho = \langle (\sigma_i)_{i \in \Pi} \rangle$ and $\rho' = \langle \sigma'_j, (\sigma_i)_{i \in \Pi \setminus \{j\}} \rangle$.

In other words, player $j \in \Pi$ (for all $j \in \Pi$) has no incentive to deviate, with respect to the relation \prec_j . Note that any secure equilibrium is a Nash equilibrium. A strategy σ'_j such that $\text{Payoff}(\rho) \prec_j \text{Payoff}(\rho')$ is called a \prec_j -profitable deviation for player j with respect to $(\sigma_i)_{i \in \Pi}$.

Definition 4. *The type of a Nash or a secure equilibrium $(\sigma_i)_{i \in \Pi}$ in a reachability game \mathcal{G} is the set of players $j \in \Pi$ such that the outcome ρ of $(\sigma_i)_{i \in \Pi}$ visits Goal_j . It is denoted by $\text{Type}((\sigma_i)_{i \in \Pi})$.*

In other words, $\text{Type}((\sigma_i)_{i \in \Pi}) = \text{Visit}(\rho)$.

The previous definitions are illustrated on a simple two-player game in the Appendix (Section A, Example 14).

The questions studied in this article are the following ones:

Problem 1 *Given \mathcal{G} a quantitative multiplayer reachability game, does there exist a Nash equilibrium (respectively a secure equilibrium) in \mathcal{G} ?*

Problem 2 *Given a Nash equilibrium (respectively a secure equilibrium) in a quantitative multiplayer reachability game \mathcal{G} , does there exist a memoryless or a finite-memory Nash equilibrium (respectively secure equilibrium) with the same type?*

We provide partial positive answers in Section 3. These problems have been investigated in the qualitative framework (see [8]). We show in the Appendix that Problems 1 and 2 can not be reduced to problems on qualitative games (see Section B).

² Our definition naturally extends the notion of *secure equilibrium* proposed in [4] to the quantitative reachability framework. A longer discussion comparing the two notions can be found in Section B.

2.2 Unraveling

In the proofs of this article we need to unravel the graph $G = (V, (V_i)_{i \in \Pi}, v_0, E)$ from the initial vertex v_0 , which ends up in an *infinite tree*, denoted by T . This tree can be seen as a new graph where the set of vertices is the set H of histories of \mathcal{G} , the initial vertex is v_0 , and a pair $(hv, hvv') \in H \times H$ is an edge of T if $(v, v') \in E$. A history h is a vertex of player i in T if $\text{Last}(h) \in V_i$, and it belongs to the goal set of player i if $\text{Last}(h) \in \text{Goal}_i$.

We denote by \mathcal{T} the related game. This game \mathcal{T} played on the unraveling T of G is equivalent to the game \mathcal{G} played on G in the following sense. A play $(\rho_0)(\rho_0\rho_1)(\rho_0\rho_1\rho_2)\dots$ in \mathcal{T} induces a unique play $\rho = \rho_0\rho_1\rho_2\dots$ in \mathcal{G} , and conversely. Thus, we denote a play in \mathcal{T} by the respective play in \mathcal{G} . The bijection between plays of \mathcal{G} and plays of \mathcal{T} allows us to use the same payoff function Payoff , and to transform easily strategies in \mathcal{G} to strategies in \mathcal{T} (and conversely).

We also need to study the tree T limited to a certain depth $d \geq 0$: we note $\text{Trunc}_d(T)$ the *truncated tree of T of depth d* and $\text{Trunc}_d(\mathcal{T})$ the *finite game played on $\text{Trunc}_d(T)$* . More precisely, the set of vertices of $\text{Trunc}_d(T)$ is the set of histories $h \in H$ of length $\leq d$; the edges of $\text{Trunc}_d(T)$ are defined in the same way as for T except that for the histories h of length d , there exists no edge (h, hv) . A play ρ in $\text{Trunc}_d(\mathcal{T})$ corresponds to a history of \mathcal{G} of length *equal to d* . The notions of payoff and strategy are defined exactly like in the game \mathcal{T} , but limited to the depth d . For instance, a player gets an infinite payoff for a play ρ (of length d) if his goal set is not visited by ρ .

3 Nash equilibria and secure equilibria

From now on we will often use the term *game* to denote a quantitative multiplayer reachability game according to Definition 1.

3.1 Existence of a Nash equilibrium

In this section we positively solve Problem 1 for Nash equilibria.

Theorem 5. *In every quantitative multiplayer reachability game, there exists a finite-memory Nash equilibrium.*

The proof of this theorem is based on the following ideas. By Kuhn's theorem (Theorem 6), there exists a Nash equilibrium in the game $\text{Trunc}_d(\mathcal{T})$ played on the finite tree $\text{Trunc}_d(T)$, for any depth d . By choosing an adequate depth d , Proposition 8 will enable to extend this Nash equilibrium to a Nash equilibrium in the infinite tree T , and thus in \mathcal{G} . Let us detail these ideas.

We first recall Kuhn's theorem [11]. A *preference relation* is a total reflexive transitive binary relation.

Theorem 6 (Kuhn's theorem). *Let Γ be a finite tree and \mathcal{G} a game played on Γ . For each player $i \in \Pi$, let \succsim_i be a preference relation on payoff profiles. Then there exists a strategy profile $(\sigma_i)_{i \in \Pi}$ such that for every player $j \in \Pi$ and every strategy σ'_j of player j in \mathcal{G} we have*

$$\text{Payoff}(\rho') \succsim_j \text{Payoff}(\rho)$$

where $\rho = \langle (\sigma_i)_{i \in \Pi} \rangle$ and $\rho' = \langle \sigma'_j, (\sigma_i)_{i \in \Pi \setminus \{j\}} \rangle$.

Corollary 7. *Let \mathcal{G} be a game and T be the unraveling of G . Let $\text{Trunc}_d(\mathcal{T})$ be the game played on the truncated tree of T of depth d , with $d \geq 0$. Then there exists a Nash equilibrium in $\text{Trunc}_d(\mathcal{T})$.*

Proof. For each player $j \in \Pi$, we define the relation \succsim_j on payoff profiles in the following way: let $(x_i)_{i \in \Pi}$ and $(y_i)_{i \in \Pi}$ be two payoff profiles, we say that $(x_i)_{i \in \Pi} \succsim_j (y_i)_{i \in \Pi}$ iff $x_j \leq y_j$. It is clearly a preference relation which captures the Nash equilibrium. The strategy profile $(\sigma_i)_{i \in \Pi}$ of Kuhn's theorem is then a Nash equilibrium in $\text{Trunc}_d(\mathcal{T})$. \square

The next proposition states that it is possible to extend a Nash equilibrium in $\text{Trunc}_d(\mathcal{T})$ to a Nash equilibrium in the game \mathcal{T} , if the depth d is equal to $(|\Pi| + 1) \cdot 2 \cdot |V|$. We obtain Theorem 5 as a consequence of Corollary 7 and Proposition 8.

Proposition 8. *Let \mathcal{G} be a game and T be the unraveling of G . Let $\text{Trunc}_d(\mathcal{T})$ be the game played on the truncated tree of T of depth $d = (|\Pi| + 1) \cdot 2 \cdot |V|$. If there exists a Nash equilibrium in the game $\text{Trunc}_d(\mathcal{T})$, then there exists a finite-memory Nash equilibrium in the game \mathcal{T} .*

The proof of Proposition 8 roughly works as follows. Let $(\sigma_i)_{i \in \Pi}$ be a Nash equilibrium in $\text{Trunc}_d(\mathcal{T})$. A well-chosen prefix $\alpha\beta$, with β being a cycle, is first extracted from the outcome ρ of $(\sigma_i)_{i \in \Pi}$. The outcome of the required Nash equilibrium $(\tau_i)_{i \in \Pi}$ in \mathcal{T} will be equal to $\alpha\beta^\omega$. As soon as a player deviates from this play, all the other players form a coalition against him to punish him in a way that this deviation is not profitable for him. These ideas are detailed in the next two lemmas whose complete proofs can be found in the Appendix (Section C).

In Lemma 10 we need to consider the qualitative two-player zero-sum game \mathcal{G}_j played on the graph G , where player j plays in order to reach his goal set Goal_j , against the coalition of all other players that wants to prevent him from reaching his goal set. Player j plays on the vertices from V_j and the coalition on $V \setminus V_j$. We have the following proposition (see [7]).

Proposition 9. *Let $\mathcal{G}_j = (V, V_j, V \setminus V_j, E, \text{Goal}_j)$ be the qualitative two-player zero-sum reachability game associated to player j . Then player j has a memoryless strategy ν_j that enables him to reach Goal_j within $|V| - 1$ edges from each vertex v from which he wins the game \mathcal{G}_j . On the contrary, the coalition has a memoryless strategy ν_{-j} that forces the play to stay in $V \setminus \text{Goal}_j$ from each vertex v from which it wins the game \mathcal{G}_j .*

The play ρ of Lemma 10 is illustrated in Figure 1.

Lemma 10. *Let $d \geq 0$. Let $(\sigma_i)_{i \in \Pi}$ be a Nash equilibrium in $\text{Trunc}_d(\mathcal{T})$ and ρ the (finite) outcome of $(\sigma_i)_{i \in \Pi}$. Suppose that ρ has a prefix $\alpha\beta\gamma$, where β contains at least one vertex, such that*

$$\begin{aligned}\text{Visit}(\alpha) &= \text{Visit}(\alpha\beta\gamma) \\ \text{Last}(\alpha) &= \text{Last}(\alpha\beta) \\ |\alpha\beta| &\leq l \cdot |V| \\ |\alpha\beta\gamma| &= (l+1) \cdot |V|\end{aligned}$$

for some $l \geq 1$.

Let $j \in \Pi$ be such that α does not visit Goal_j . Consider the qualitative two-player zero-sum game $\mathcal{G}_j = (V, V_j, V \setminus V_j, E, \text{Goal}_j)$. Then for all histories hu of \mathcal{G} consistent with $(\sigma_i)_{i \in \Pi \setminus \{j\}}$ and such that $|hu| \leq |\alpha\beta|$, the coalition of the players $i \neq j$ wins the game \mathcal{G}_j from u .

Condition $\text{Visit}(\alpha) = \text{Visit}(\alpha\beta\gamma)$ means that if Goal_i is visited by $\alpha\beta\gamma$, it has already been visited by α . Condition $\text{Last}(\alpha) = \text{Last}(\alpha\beta)$ means that β is a cycle.

This lemma means in particular that the players $i \neq j$ can play together to prevent player j from reaching his goal set Goal_j , in case he deviates from the play $\alpha\beta$ (as $\alpha\beta$ is consistent with $(\sigma_i)_{i \in \Pi \setminus \{j\}}$). We denote by ν_{-j} the memoryless winning strategy of the coalition. For each player $i \neq j$, let $\nu_{i,j}$ be the memoryless strategy of player i in \mathcal{G} induced by ν_{-j} .

Lemma 11 states that one can define a Nash equilibrium $(\tau_i)_{i \in \Pi}$ in the game \mathcal{T} , based on the Nash equilibrium $(\sigma_i)_{i \in \Pi}$ in the game $\text{Trunc}_d(\mathcal{T})$.

Lemma 11. *Let $d \geq 0$. Let $(\sigma_i)_{i \in \Pi}$ be a Nash equilibrium in $\text{Trunc}_d(\mathcal{T})$ and $\alpha\beta\gamma$ be a prefix of $\rho = \langle (\sigma_i)_{i \in \Pi} \rangle$ as defined in Lemma 10. Then there exists a Nash equilibrium $(\tau_i)_{i \in \Pi}$ in the game \mathcal{T} . Moreover $(\tau_i)_{i \in \Pi}$ is finite-memory, and $\text{Type}((\tau_i)_{i \in \Pi}) = \text{Visit}(\alpha)$.*

Proof. Let $\Pi = \{1, \dots, n\}$. As α and β end in the same vertex, we can consider the infinite play $\alpha\beta^\omega$ in the game \mathcal{T} . Without loss of generality we can order the players so that

$$\begin{aligned}\forall i \leq k & \quad \text{Payoff}_i(\alpha\beta^\omega) < +\infty & (\alpha \text{ visits } \text{Goal}_i) \\ \forall i > k & \quad \text{Payoff}_i(\alpha\beta^\omega) = +\infty & (\alpha \text{ does not visit } \text{Goal}_i)\end{aligned}$$

where $1 \leq k \leq n$. In the second case, notice that ρ could visit Goal_i (but after the prefix $\alpha\beta\gamma$).

The Nash equilibrium $(\tau_i)_{i \in \Pi}$ required by Lemma 11 is intuitively defined as follows. First the outcome of $(\tau_i)_{i \in \Pi}$ is exactly $\alpha\beta^\omega$. Secondly the first player j who deviates from $\alpha\beta^\omega$ is punished by the coalition of the other players in the following way. If $j \leq k$ and the deviation occurs in the tree $\text{Trunc}_d(\mathcal{T})$, then the coalition plays according to $(\sigma_i)_{i \in \Pi \setminus \{j\}}$ in this tree. It prevents player j from reaching his goal set Goal_j faster than in $\alpha\beta^\omega$. And if $j > k$, the coalition plays

according to $(\nu_i)_{i \in \Pi \setminus \{j\}}$ (given by Lemma 10) so that player j does not reach his goal set at all.

We begin by defining a punishment function P on the vertex set H of T such that $P(h)$ indicates the first player j who has deviated from $\alpha\beta^\omega$, with respect to h . We write $P(h) = \perp$ if no deviation has occurred. For $h \in V^*$ and $v \in V_i$ we let:

$$P(hv) = \begin{cases} \perp & \text{if } P(h) = \perp \text{ and } hv < \alpha\beta^\omega, \\ i & \text{if } P(h) = \perp \text{ and } hv \not< \alpha\beta^\omega, \\ P(h) & \text{otherwise } (P(h) \neq \perp). \end{cases}$$

The Nash equilibrium $(\tau_i)_{i \in \Pi}$ is then defined as follows: let h be a history ending in a vertex of V_i ,

$$\tau_i(h) = \begin{cases} v & \text{if } P(h) = \perp \text{ (} h < \alpha\beta^\omega \text{); such that } hv < \alpha\beta^\omega, \\ \text{arbitrary} & \text{if } P(h) = i, \\ \nu_{i,P(h)}(h) & \text{if } P(h) \neq \perp, i \text{ and } P(h) > k, \\ \sigma_i(h) & \text{if } P(h) \neq \perp, i, P(h) \leq k \text{ and } |h| \leq d, \\ \text{arbitrary} & \text{otherwise } (P(h) \neq \perp, i, P(h) \leq k \text{ and } |h| > d) \end{cases} \quad (1)$$

where *arbitrary* means that the next vertex is chosen arbitrarily (in a memoryless way). Clearly the outcome of $(\tau_i)_{i \in \Pi}$ is the play $\alpha\beta^\omega$, and $\text{Type}((\tau_i)_{i \in \Pi})$ is equal to $\text{Visit}(\alpha)$ ($= \text{Visit}(\alpha\beta)$).

It remains to prove that $(\tau_i)_{i \in \Pi}$ is a finite-memory Nash equilibrium in the game \mathcal{T} . The end of this proof can be found in the Appendix (Section C). \square

We can now proceed to the proof of Proposition 8.

Proof (of Proposition 8). Let $\Pi = \{1, \dots, n\}$ and $d = (n+1) \cdot 2 \cdot |V|$. Let $(\sigma_i)_{i \in \Pi}$ be a Nash equilibrium in the game $\text{Trunc}_d(\mathcal{T})$ and ρ its outcome.

To be able to use Lemmas 10 and 11, we consider the prefix \mathbf{pq} of ρ of minimal length such that

$$\begin{aligned} \exists l \geq 1 \quad & |\mathbf{p}| = (l-1) \cdot |V| \\ & |\mathbf{pq}| = (l+1) \cdot |V| \\ & \text{Visit}(\mathbf{p}) = \text{Visit}(\mathbf{pq}). \end{aligned} \quad (2)$$

The following statements are true.

- (i) $l \leq 2 \cdot n + 1$.
- (ii) If $\text{Visit}(\mathbf{p}) \not\subseteq \text{Visit}(\rho)$, then $l < 2 \cdot n + 1$.

Indeed the first statement results from the fact that in the worst case, the play ρ visits the goal set of a new player in each prefix of length $i \cdot 2 \cdot |V|$, $1 \leq i \leq n$, i.e. $|\mathbf{p}| = n \cdot 2 \cdot |V|$. It follows that \mathbf{pq} exists as a prefix of ρ , because the length d of ρ is equal to $(n+1) \cdot 2 \cdot |V|$ by hypothesis. Thus $\text{Visit}(\mathbf{p}) \subseteq \text{Visit}(\rho)$. Suppose that there exists $i \in \text{Visit}(\rho) \setminus \text{Visit}(\mathbf{p})$, then ρ visit Goal_i after the prefix \mathbf{pq} by Equation (2). The second statement follows easily.

Given the length of \mathbf{q} , one vertex of V is visited at least twice by \mathbf{q} . More precisely, we can write

$$\begin{aligned} \mathbf{p}\mathbf{q} &= \alpha\beta\gamma \quad \text{with} \quad \text{Last}(\alpha) = \text{Last}(\alpha\beta) \\ |\alpha| &\geq (l-1) \cdot |V| \\ |\alpha\beta| &\leq l \cdot |V|. \end{aligned} \tag{3}$$

In particular, $|\mathbf{p}| \leq |\alpha|$. See Figure 1. We have $\text{Visit}(\alpha) = \text{Visit}(\alpha\beta\gamma)$, and $|\alpha\beta\gamma| = (l+1) \cdot |V|$.

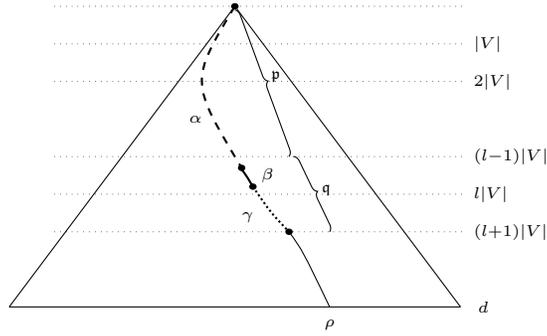


Fig. 1. Slicing of the play ρ in the tree $\text{Trunc}_d(T)$.

As the hypotheses of Lemmas 10 and 11 are verified, we can apply them in this context to get a Nash equilibrium $(\tau_i)_{i \in \Pi}$ in the game \mathcal{T} such that $\text{Type}((\tau_i)_{i \in \Pi}) = \text{Visit}(\alpha)$. \square

Proposition 8 asserts that given a game \mathcal{G} and the game $\text{Trunc}_d(\mathcal{T})$ played on the truncated tree of T of a well-chosen depth d , one can lift any Nash equilibrium $(\sigma_i)_{i \in \Pi}$ of $\text{Trunc}_d(\mathcal{T})$ to a Nash equilibrium $(\tau_i)_{i \in \Pi}$ of \mathcal{G} . The proof of Proposition 8 states that the type of $(\tau_i)_{i \in \Pi}$ is equal to $\text{Visit}(\alpha)$. We give in the Appendix an example (see Example 17) that shows that, with this approach, it is impossible to preserve the type of the lifted Nash equilibrium $(\sigma_i)_{i \in \Pi}$.

3.2 Nash equilibrium with finite memory

In this section we study the kind of strategies we can impose for a Nash equilibrium in a quantitative multiplayer reachability game. We show that given a Nash equilibrium, we can construct another Nash equilibrium with the same type such that all its strategies are finite-memory. We then answer to Problem 2 for Nash equilibria.

Theorem 12. *Let $(\sigma_i)_{i \in \Pi}$ be a Nash equilibrium in a quantitative multiplayer reachability game \mathcal{G} . Then there exists a finite-memory Nash equilibrium of the same type in \mathcal{G} .*

The proof is based on two steps. The first step constructs from $(\sigma_i)_{i \in \Pi}$ another Nash equilibrium $(\tau_i)_{i \in \Pi}$ with the same type such that the play $\langle (\tau_i)_{i \in \Pi} \rangle$ is of the form $\alpha\beta^\omega$ with $\text{Visit}(\alpha) = \text{Type}((\sigma_i)_{i \in \Pi})$. This is possible thanks to two lemmas (Lemmas 18 and 19 given in the Appendix), by first eliminating unnecessary cycles in the play $\langle (\sigma_i)_{i \in \Pi} \rangle$ and then locating a prefix $\alpha\beta$ such that β is a cycle that can be infinitely repeated.

The second step transforms the Nash equilibrium $(\tau_i)_{i \in \Pi}$ into a finite-memory one thanks to Lemmas 10 and 11 given in Section 3.1. For that, we consider the strategy profile $(\tau_i)_{i \in \Pi}$ limited to the tree T truncated at a well-chosen depth.

The detailed proof of Theorem 12 can be found in the Appendix (Section D).

3.3 Existence of a secure equilibrium

In this section we positively answer to Problem 1 for secure equilibria in *two-player* games.

Theorem 13. *In every quantitative two-player reachability game, there exists a finite-memory secure equilibrium.*

The proof of this theorem is based on the same ideas as for the proof of Theorem 5 (existence of a Nash equilibrium). By Kuhn’s theorem (Theorem 6), there exists a secure equilibrium in the game $\text{Trunc}_d(\mathcal{T})$ played on the finite tree $\text{Trunc}_d(T)$, for any depth d . By choosing an adequate depth d , Proposition 23 given in the Appendix enables to extend this secure equilibrium to a secure equilibrium in the infinite tree T , and thus in \mathcal{G} . The details of the proof are given in the Appendix (Section E).

4 Conclusion and perspectives

In this paper, we proved the existence of finite-memory Nash (resp. secure) equilibria for quantitative multiplayer (resp. two-player) reachability games played on finite graphs. We do believe that our results remain true when the model is enriched by allowing positive weights on edges (instead of weight 1 on each edge). Indeed the idea is to replace any edge with a weight $c \geq 1$ by a path of length c composed of c new edges, and use the results proved in this article.

There are several interesting directions for further research. First, we intend to investigate the existence of secure equilibria in the n -player framework. Secondly, we would like to check whether our results remain true when the model is enriched by allowing a n -tuple of non-negative weights on edges (one weight by player). Then, we will also investigate deeper the size of the memory needed in the equilibria. This could be a first step towards a study of the complexity of computing equilibria with certain requirements, in the spirit of [8]. We also intend to look for existence results for *subgame perfect equilibria*. Finally we would like to address these questions for other objectives such as Büchi or request-response.

Acknowledgments The authors are grateful to Jean-François Raskin and Hugo Gimbert for useful discussions.

References

1. R. Alur, A. Kanade, and G. Weiss. Ranking automata and games for prioritized requirements. In *Computer Aided Verification, 20th International Conference, CAV 2008*, volume 5123 of *LNCS*, pages 240–253. Springer, 2008.
2. R. Bloem, K. Chatterjee, T. Henzinger, and B. Jobstmann. Better quality in synthesis through quantitative objectives. In *CAV: Computer-Aided Verification*, volume 5643 of *LNCS*, pages 140–156. Springer, 2009.
3. E. Boros and V. Gurvich. Why chess and back gammon can be solved in pure positional uniformly optimal strategies. Rutcor Research Report 21-2009, Rutgers University, 2009.
4. K. Chatterjee, T. Henzinger, and M. Jurdziński. Games with secure equilibria. *Theoretical Computer Science*, 365(1-2):67–82, 2006.
5. K. Chatterjee and T. A. Henzinger. Finitary winning in omega-regular games. In *TACAS*, volume 3920 of *LNCS*, pages 257–271. Springer, 2006.
6. E. Clarke, O. Grumberg, and D. Peled. *Model Checking*. MIT Press, Cambridge, MA, 2000.
7. E. Grädel, W. Thomas, and T. Wilke. Automata, logics, and infinite games. volume 2500 of *LNCS*. Springer, 2002.
8. E. Grädel and M. Ummels. Solution concepts and algorithms for infinite multi-player games. In K. Apt and R. van Rooij, editors, *New Perspectives on Games and Interaction*, volume 4 of *Texts in Logic and Games*, pages 151–178. Amsterdam University Press, 2008.
9. J. E. Hopcroft and J. D. Ullman. *Introduction to automata theory, languages, and computation*. Addison-Wesley Publishing Co., Reading, Mass., 1979. Addison-Wesley Series in Computer Science.
10. F. Horn, W. Thomas, and N. Wallmeier. Optimal strategy synthesis in request-response games. In *Automated Technology for Verification and Analysis, 6th International Symposium, ATVA 2008*, volume 5311 of *LNCS*, pages 361–373. Springer, 2008.
11. H. Kuhn. Extensive games and the problem of information. *Classics in Game Theory*, pages 46–68, 1953.
12. J. Nash. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences of the United States of America*, 36(1):48–49, 1950.
13. M. Osborne and A. Rubinstein. *A course in game theory*. MIT Press, Cambridge, MA, 1994.
14. W. Thomas. On the synthesis of strategies in infinite games. In *STACS 95 (Munich, 1995)*, volume 900 of *LNCS*, pages 1–13. Springer, Berlin, 1995.
15. M. Zimmermann. Time-optimal winning strategies for poset games. In *CIAA*, volume 5642 of *LNCS*, pages 217–226. Springer, 2009.

Technical Appendix

A Examples of Nash and secure equilibria

Example 14. Let $\mathcal{G} = (V, V_1, V_2, v_0, E, \text{Goal}_1, \text{Goal}_2)$ be the two-player game depicted in Figure 2. The states of player 1 (resp. 2) are represented by circles (resp. squares)³. Thus, according to Figure 2, $V_1 = \{A, C, D\}$ and $V_2 = \{B\}$, the initial vertex v_0 is the vertex A , and we set $\text{Goal}_1 = \{C\}$ and $\text{Goal}_2 = \{D\}$.

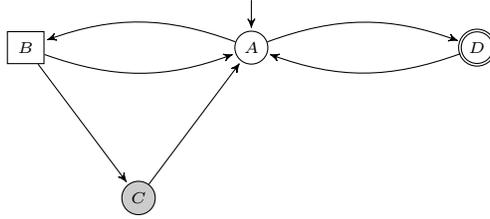


Fig. 2. A two-player game with $\text{Goal}_1 = \{C\}$ and $\text{Goal}_2 = \{D\}$.

An example of play in \mathcal{G} is given by $\rho = (AD)^\omega$, which visits Goal_2 but not Goal_1 , leading to the payoff profile $\text{Payoff}((AD)^\omega) = (+\infty, 1)$. The play ρ is, among others, the outcome of the strategy⁴ profile (σ_1, σ_2) where $\sigma_1(hA) = D$ and $\sigma_2(hB) = C$.

Let us show that the strategy profile (σ_1, σ_2) is not a Nash equilibrium, by proving that player 1 has a profitable deviation σ'_1 in which he manages to decrease his own payoff. With σ'_1 defined by $\sigma'_1(hA) = B$, we get the play $\langle (\sigma'_1, \sigma_2) \rangle = (ABC)^\omega$ such that $\text{Payoff}((ABC)^\omega) = (2, +\infty)$, and in particular $\text{Payoff}_1((ABC)^\omega) < \text{Payoff}_1(\rho)$.

On the opposite, one can show that (σ'_1, σ_2) is a Nash equilibrium. However (σ'_1, σ_2) is not a secure equilibrium. Indeed, player 2 has a \prec_2 -profitable deviation in which he can increase player 1's payoff without modifying his own payoff. With σ'_2 the strategy of player 2 defined by $\sigma'_2(hB) = A$, we get the play $\langle (\sigma'_1, \sigma'_2) \rangle = (AB)^\omega$ such that $\text{Payoff}((AB)^\omega) = (+\infty, +\infty)$, and $\text{Payoff}(\langle (\sigma'_1, \sigma_2) \rangle) \prec_2 \text{Payoff}(\langle (\sigma'_1, \sigma'_2) \rangle)$.

Notice that all strategies discussed so far are memoryless. In order to obtain a Nash equilibrium of type $\{1, 2\}$, finite-memory strategies are necessary. We define the following finite-memory strategy profile (τ_1, τ_2) :

$$\tau_1(hA) = \begin{cases} D & \text{if } h = \epsilon \\ B & \text{if } h \neq \epsilon \end{cases} ; \quad \tau_2(hB) = \begin{cases} C & \text{if } D \text{ has been visited by } h \\ A & \text{otherwise.} \end{cases}$$

³ We will keep this convention through the Appendix.

⁴ Note that player 1 has no choice in vertices C and D , that is $\sigma_1(hv)$ is necessarily equal to A for $v \in \{C, D\}$.

The outcome $\pi = \langle (\tau_1, \tau_2) \rangle$ is equal to $AD(ABC)^\omega$ and has payoff $(4, 1)$. In order to prove that (τ_1, τ_2) is a Nash equilibrium, we prove that no player has a profitable deviation. For player 2 it is clearly impossible to get a payoff less than 1. To try to get a payoff less than 4, player 1 must use a strategy τ'_1 such that $\tau'_1(A) = B$. But then player 2 chooses $\tau_2(AB) = A$. The prefix ABA of the outcome of (τ'_1, τ_2) shows that player 1 will increase his payoff of 4.

However (τ_1, τ_2) is not a secure equilibrium since player 2 has a \prec_2 -profitable deviation τ'_2 such that $\tau'_2(hB) = A$ for all histories h . One can show that, in this example, there is no secure equilibrium of type $\{1, 2\}$.

B Qualitative games vs quantitative games

Given a quantitative multiplayer reachability game \mathcal{G} , one can naturally define a *qualitative* version of \mathcal{G} , denoted by $\overline{\mathcal{G}}$, such that the payoffs are *qualitative*. Given a play ρ of \mathcal{G} , the qualitative payoff of player i is defined by:

$$\overline{\text{Payoff}}_i(\rho) = \begin{cases} \text{Win} & \text{if } \text{Payoff}_i(\rho) \text{ is finite} \\ \text{Lose} & \text{otherwise.} \end{cases}$$

We note $\overline{\text{Payoff}}(\rho) = (\overline{\text{Payoff}}_i(\rho))_{i \in \Pi}$ the *qualitative* payoff profile for the play ρ . In this framework, player i aims at reaching his own objective, i.e. at obtaining payoff Win. With this idea in mind, one can naturally adapt the notion of Nash (resp. secure) equilibrium to the qualitative framework.

The existence of Nash (resp. secure) equilibria in n -player (resp. 2-player) qualitative games $\overline{\mathcal{G}}$ has been proved in [8, Corollary 12] (resp. [4, Theorem 2]) for reachability objectives, and more generally for Borel objectives.

The next example illustrates that lifting Nash equilibria in $\overline{\mathcal{G}}$ to Nash equilibria in \mathcal{G} does not work. We developed new ideas in Sections 3.1 and 3.3 to solve Problem 1.

Example 15. Let us consider the two-player game \mathcal{G} depicted in Figure 3, such that $\text{Goal}_1 = \{B, E\}$ and $\text{Goal}_2 = \{C\}$. Notice that only player 1 effectively plays in this game. We are going to exhibit a secure (and thus Nash) equilibrium (σ_1, σ_2) in the qualitative game $\overline{\mathcal{G}}$ that can not be lifted neither to a secure nor to a Nash equilibrium in the quantitative game \mathcal{G} . The strategy profile (σ_1, σ_2) is defined such that $\langle (\sigma_1, \sigma_2) \rangle = ADE^\omega$. It is a secure equilibrium in $\overline{\mathcal{G}}$ with the qualitative payoff profile (Win, Lose). However (σ_1, σ_2) is not a Nash (and thus not a secure) equilibrium in \mathcal{G} . Indeed, the play ABC^ω provides a smaller payoff to player 1, i.e. $\text{Payoff}_1(ABC^\omega) < \text{Payoff}_1(ADE^\omega)$. Notice that for this example, there is no equilibrium in \mathcal{G} of type $\{1\}$.

The next proposition shows that on the opposite, any Nash equilibrium in a quantitative game \mathcal{G} can be lifted to a Nash equilibrium in the qualitative game $\overline{\mathcal{G}}$.

Proposition 16. *If $(\sigma_i)_{i \in \Pi}$ is a Nash equilibrium in a quantitative multiplayer reachability game \mathcal{G} , then $(\sigma_i)_{i \in \Pi}$ is also a Nash equilibrium in $\overline{\mathcal{G}}$.*

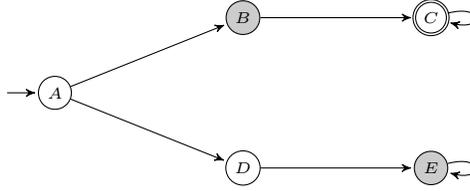


Fig. 3. A game \mathcal{G} with an equilibrium in $\overline{\mathcal{G}}$ that can not be lifted to \mathcal{G} .

Proof. For a contradiction, assume that in $\overline{\mathcal{G}}$, player j has a profitable deviation σ'_j w.r.t. $(\sigma_i)_{i \in \Pi}$. This is only possible if $\text{Payoff}_j(\langle (\sigma_i)_{i \in \Pi} \rangle) = \text{Lose}$ and $\text{Payoff}_j(\langle \sigma'_j, (\sigma_i)_{i \in \Pi \setminus \{j\}} \rangle) = \text{Win}$. Thus when playing σ'_j against $(\sigma_i)_{i \in \Pi \setminus \{j\}}$, player j manages to visit Goal_j . Clearly enough, σ'_j would also be a profitable deviation w.r.t. $(\sigma_i)_{i \in \Pi}$ in \mathcal{G} , contradicting the hypothesis. \square

Note that Proposition 16 is false for secure equilibria. To see that, let us come back to the game \mathcal{G} of Figure 3. The strategy profile (σ_1, σ_2) such that $\langle (\sigma_1, \sigma_2) \rangle = ABC^\omega$ is a secure equilibrium in the quantitative game \mathcal{G} but not in the qualitative game $\overline{\mathcal{G}}$.

C Existence of a Nash equilibrium

Proof (of Lemma 10). By contradiction suppose that player j wins the game \mathcal{G}_j from u . By Proposition 9 player j has a memoryless winning strategy ν_j which enables him to reach his goal set Goal_j within at most $|V| - 1$ edges from u . We show that ν_j leads to a profitable deviation for player j w.r.t. $(\sigma_i)_{i \in \Pi}$ in the game $\text{Trunc}_d(\mathcal{T})$, which is impossible by hypothesis.

Let ρ' be a play in $\text{Trunc}_d(\mathcal{T})$ such that hu is a prefix of ρ' , and from u , player j plays according to the strategy ν_j and the other players $i \neq j$ continue to play according to σ_i . As the play ρ' is consistent with the memoryless winning strategy ν_j from u , it visits Goal_j and we have

$$\begin{aligned}
 \text{Payoff}_j(\rho') &\leq |hu| + |V| && \text{(by Proposition 9)} \\
 &\leq (l+1) \cdot |V| && \text{(by hypothesis)} \\
 &\leq d && \text{(as } \alpha\beta\gamma \leq \rho \text{)}.
 \end{aligned}$$

We consider the following two cases. If $\text{Payoff}_j(\rho) = +\infty$ (i.e. ρ does not visit Goal_j), we have

$$\text{Payoff}_j(\rho') < \text{Payoff}_j(\rho) = +\infty.$$

On the contrary, if $\text{Payoff}_j(\rho) < +\infty$ (i.e. ρ visits Goal_j , but after the prefix $\alpha\beta\gamma$ by hypothesis), then we have

$$\text{Payoff}_j(\rho') < \text{Payoff}_j(\rho)$$

as $\text{Payoff}_j(\rho) > (l+1) \cdot |V|$.

Since ρ' is consistent with $(\sigma_i)_{i \in \Pi \setminus \{j\}}$, the strategy of player j induced by the play ρ' is a profitable deviation for player j w.r.t. $(\sigma_i)_{i \in \Pi}$, in both cases. \square

Proof (End of the proof of Lemma 11). We first show that the strategy profile $(\tau_i)_{i \in \Pi}$ defined in Equation (1) is a Nash equilibrium in the game \mathcal{T} .

Let τ'_j be a strategy of player j . We show that this is not a profitable deviation for player j w.r.t. $(\tau_i)_{i \in \Pi}$. We distinguish the following two cases:

(i) $j \leq k$ ($\text{Payoff}_j(\alpha\beta^\omega) < +\infty$, α visits Goal_j).

To improve his payoff, player j has no incentive to deviate after the prefix α . Thus we assume that the strategy τ'_j causes a deviation from a vertex visited by α . By Equation (1) the other players first play according to $(\sigma_i)_{i \in \Pi \setminus \{j\}}$ in $\text{Trunc}_d(\mathcal{T})$, and then in an arbitrary way.

Suppose that τ'_j is a profitable deviation for player j w.r.t. $(\tau_i)_{i \in \Pi}$ in the game \mathcal{T} . Let $\pi = \langle (\tau_i)_{i \in \Pi} \rangle$ and $\pi' = \langle \tau'_j, (\tau_i)_{i \in \Pi \setminus \{j\}} \rangle$. Then

$$\text{Payoff}_j(\pi') < \text{Payoff}_j(\pi).$$

In the other hand we know that

$$\text{Payoff}_j(\pi) = \text{Payoff}_j(\rho) \leq |\alpha|.$$

So if we limit the play π' in \mathcal{T} to its prefix of length d , we get a play ρ' in $\text{Trunc}_d(\mathcal{T})$ such that

$$\text{Payoff}_j(\rho') = \text{Payoff}_j(\pi') < \text{Payoff}_j(\rho).$$

As the play ρ' is consistent with the strategies $(\sigma_i)_{i \in \Pi \setminus \{j\}}$ by Equation (1), the strategy τ'_j restricted to the tree $\text{Trunc}_d(\mathcal{T})$ is a profitable deviation for player j w.r.t. $(\sigma_i)_{i \in \Pi}$ in the game $\text{Trunc}_d(\mathcal{T})$. This is impossible.

(ii) $j > k$ ($\text{Payoff}_j(\alpha\beta^\omega) = +\infty$, $\alpha\beta^\omega$ does not visit Goal_j).

If player j deviates from $\alpha\beta^\omega$ (with the strategy τ'_j), by Equation (1) the other players combine against him and play according to ν_{-j} . By Lemma 10 this coalition wins the game \mathcal{G}_j from any vertex visited by $\alpha\beta^\omega$. So the strategy ν_{-j} of the coalition keeps the play $\langle \tau'_j, (\tau_i)_{i \in \Pi \setminus \{j\}} \rangle$ away from the set Goal_j whatever player j does. Therefore τ'_j is not a profitable deviation for player j w.r.t. $(\tau_i)_{i \in \Pi}$ in the game \mathcal{T} .

We now prove that $(\tau_i)_{i \in \Pi}$ is a finite-memory strategy profile. According to the definition of finite-memory strategy (see Section 2) we have to prove that each relation \approx_{τ_i} on H has finite index (recall that $h \approx_{\tau_i} h'$ if $\tau_i(h\delta) = \tau_i(h'\delta)$ for all $\delta \in V^*V_i$). In this aim we define for each player i an equivalence relation \sim_{τ_i} with finite index such that

$$\forall h, h' \in H, \quad h \sim_{\tau_i} h' \Rightarrow h \approx_{\tau_i} h'.$$

We first define an equivalence relation \sim_P with finite index related to the punishment function P . For all prefixes h, h' of $\alpha\beta^\omega$, i.e. such that no player is punished, this relation does not distinguish two histories that are identical except for a certain number of cycles β . For the other histories it just has to remember the first player, say i , who has deviated. The definition of \sim_P is as follows:

$$\begin{aligned} h \sim_P h' & \quad \text{if } h = \alpha\beta^l\beta', h' = \alpha\beta^m\beta', \beta' < \beta, l, m \geq 0 \\ hv \sim_P h'v' & \quad \text{if } v, v' \in V_i, h, h' < \alpha\beta^\omega, \text{ but } hv, h'v' \not< \alpha\beta^\omega \\ hv \sim_P hv\delta & \quad \text{if } h < \alpha\beta^\omega, hv \not< \alpha\beta^\omega, \delta \in V^*. \end{aligned}$$

The relation \sim_P is an equivalence relation on H with finite index.

We now turn to the definition of \sim_{τ_i} . It is based on the definition of τ_i (given in (1)) and \sim_P . To get an equivalence with finite index we proceed as follows. Recall that each strategy $\nu_{i,P(h)}$ is memoryless and when a player plays arbitrarily, his strategy is also memoryless. Furthermore notice that, in the definition of τ_i , the strategy σ_i is only applied to histories h with length $|h| \leq d$. For histories h such that $\tau_i(h) = v$ with $hv < \alpha\beta^\omega$, it is enough to remember information with respect to $\alpha\beta$ as already done for \sim_P . Therefore for $h, h' \in H$ we define \sim_{τ_i} in the following way:

$$\begin{aligned} h \sim_{\tau_i} h' & \text{ if } h \sim_P h' \text{ and } (P(h) = \perp \\ & \text{or } P(h) = i \text{ and } \text{Last}(h) = \text{Last}(h') \\ & \text{or } P(h) \neq \perp, i, P(h) > k \text{ and } \text{Last}(h) = \text{Last}(h') \\ & \text{or } P(h) \neq \perp, i, P(h) \leq k, |h|, |h'| > d \text{ and} \\ & \text{Last}(h) = \text{Last}(h')). \end{aligned}$$

Notice that this relation satisfies

$$h \sim_{\tau_i} h' \Rightarrow \tau_i(h) = \tau_i(h') \text{ and } \text{Last}(h) = \text{Last}(h')$$

and has finite index. Therefore if $h \sim_{\tau_i} h'$, then $h \approx_{\tau_i} h'$ and the relation \approx_{τ_i} has finite index. □

Example 17. Let us consider the two-player game \mathcal{G} depicted in Figure 4 with $\text{Goal}_1 = \{C\}$, $\text{Goal}_2 = \{E\}$. One can show that \mathcal{G} admits only Nash equilibria of type $\{2\}$ or \emptyset . Indeed, on one hand, there is no play of \mathcal{G} where both goals are visited, and on the other hand given a strategy profile $(\sigma_i)_{i \in \Pi}$ such that $\langle (\sigma_i)_{i \in \Pi} \rangle$ visits Goal_1 , (i.e. $\langle (\sigma_i)_{i \in \Pi} \rangle$ is of the form A^+BC^ω), playing D instead of C is clearly a profitable deviation for player 2.

We will now see that for each $d \geq 2$ the game played on $\text{Trunc}_d(T)$ admits a Nash equilibrium of type $\{1\}$. From the above discussion, this equilibrium can not be lifted to an equilibrium of the same type in \mathcal{G} . A truncated tree $\text{Trunc}_d(T)$ is depicted in Figure 5. One can show that the strategy profile leading to the outcome $A^{d-1}BC$ (depicted in bold in the figure) is a Nash equilibrium in $\text{Trunc}_d(T)$

of type $\{1\}$. Following the lines of the proof of Proposition 8, we see that this Nash equilibrium is lifted to a Nash equilibrium of \mathcal{G} with outcome A^ω and type \emptyset .

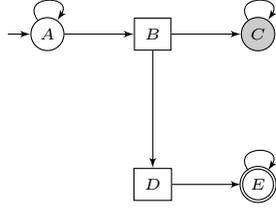


Fig. 4. A game \mathcal{G} .

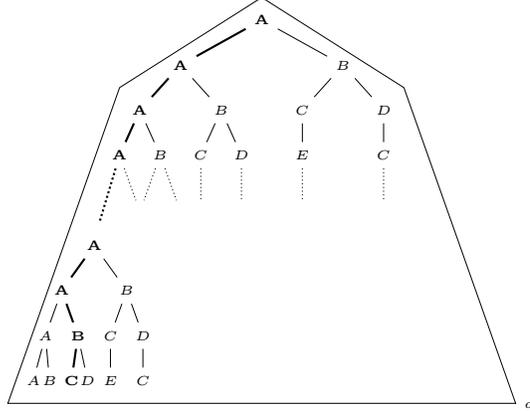


Fig. 5. The truncated tree $\text{Trunc}_d(T)$.

D Nash equilibria with finite memory

The next lemma indicates how to eliminate a cycle in the outcome of a strategy profile.

Lemma 18. *Let $(\sigma_i)_{i \in \Pi}$ be a strategy profile in a game \mathcal{G} and $\rho = \langle (\sigma_i)_{i \in \Pi} \rangle$ its outcome. Suppose that $\rho = \mathfrak{p}\mathfrak{q}\tilde{\rho}$, where \mathfrak{q} contains at least one vertex, such that*

$$\text{Visit}(\mathfrak{p}) = \text{Visit}(\mathfrak{p}\mathfrak{q})$$

$$\text{Last}(\mathfrak{p}) = \text{Last}(\mathfrak{p}\mathfrak{q}).$$

We define a strategy profile $(\tau_i)_{i \in \Pi}$ as follows:

$$\tau_i(h) = \begin{cases} \sigma_i(h) & \text{if } \mathfrak{p} \not\preceq h, \\ \sigma_i(\mathfrak{p}\mathfrak{q}\delta) & \text{if } h = \mathfrak{p}\delta \end{cases}$$

where h is a history of \mathcal{G} with $\text{Last}(h) \in V_i$. We get the outcome $\langle (\tau_i)_{i \in \Pi} \rangle = \mathfrak{p}\tilde{\rho}$. If a strategy τ'_j is a profitable deviation for player j w.r.t. $(\tau_i)_{i \in \Pi}$, then there exists a profitable deviation σ'_j for player j w.r.t. $(\sigma_i)_{i \in \Pi}$.

Proof. Let $\Pi = \{1, \dots, n\}$. We write

$$\rho = \langle (\sigma_i)_{i \in \Pi} \rangle \text{ of payoff profile } (x_1, \dots, x_n),$$

$$\pi = \langle (\tau_i)_{i \in \Pi} \rangle \text{ of payoff profile } (y_1, \dots, y_n).$$

We observe that as $\rho = \mathbf{p}q\tilde{\rho}$, we have $\pi = \mathbf{p}\tilde{\rho}$ (see Figures 6 and 7). It follows that

$$\forall i \in \Pi, \quad y_i \leq x_i. \quad (4)$$

More precisely,

$$\text{- if } x_i = +\infty, \text{ then } y_i = +\infty; \quad (5)$$

$$\text{- if } x_i < +\infty \text{ and } i \in \text{Visit}(\mathbf{p}), \text{ then } y_i = x_i;$$

$$\text{- if } x_i < +\infty \text{ and } i \notin \text{Visit}(\mathbf{p}), \text{ then } y_i = x_i - (|\mathbf{q}| + 1). \quad (6)$$

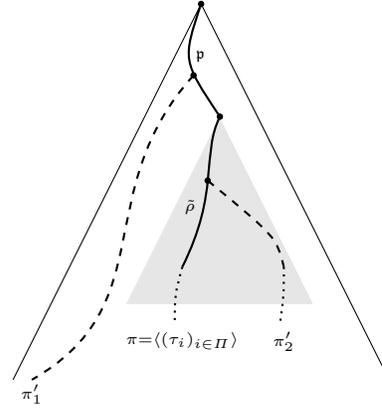
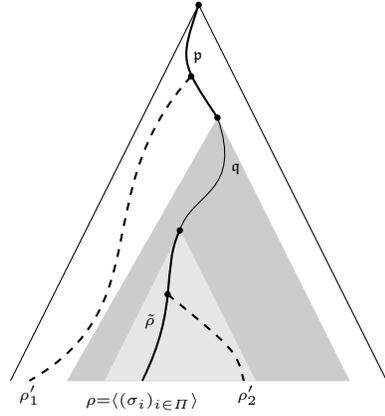


Fig. 6. Play ρ and possible deviations. **Fig. 7.** Play π and possible deviations.

Let τ'_j be a profitable deviation for player j w.r.t. $(\tau_i)_{i \in \Pi}$, and π' be the outcome of the strategy profile $(\tau'_j, (\tau_i)_{i \in \Pi \setminus \{j\}})$. Then

$$\text{Payoff}_j(\pi') < y_j.$$

We show how to construct a profitable deviation σ'_j for player j w.r.t. $(\sigma_i)_{i \in \Pi}$. Two cases occur:

- (i) player j deviates from π just after a proper prefix h of \mathbf{p} (like for the play π'_1 in Figure 7).

We define $\sigma'_j = \tau'_j$ and we denote by ρ' the outcome of $(\sigma'_j, (\sigma_i)_{i \in \Pi \setminus \{j\}})$. Given the definition of the strategy profile $(\tau_i)_{i \in \Pi}$, one can verify that $\rho' = \pi'$ (see the play ρ'_1 in Figure 6). Thus

$$\text{Payoff}_j(\rho') = \text{Payoff}_j(\pi') < y_j \leq x_j$$

by Equation (4), which implies that σ'_j is a profitable deviation of player j w.r.t. $(\sigma_i)_{i \in \Pi}$.

(ii) player j deviates from π after the prefix \mathbf{p} (π and π' coincide at least on \mathbf{p}).

This case is illustrated by the play π'_2 in Figure 7. We define for all histories h ending in a vertex of V_j :

$$\sigma'_j(h) = \begin{cases} \sigma_j(h) & \text{if } \mathbf{pq} \not\preceq h, \\ \tau'_j(\mathbf{p}\delta) & \text{if } h = \mathbf{pq}\delta. \end{cases}$$

Let $\rho' = \langle \sigma'_j, (\sigma_i)_{i \in \Pi \setminus \{j\}} \rangle$. As player j deviates after \mathbf{p} with the strategy τ'_j , one can prove that

$$\pi' = \mathbf{p}\tilde{\pi}' \quad \text{and} \quad \rho' = \mathbf{pq}\tilde{\pi}'$$

by definition of $(\tau_i)_{i \in \Pi}$ (see the play ρ'_2 in Figure 6). As $\text{Payoff}_j(\pi') < y_j$, it means that $j \notin \text{Visit}(\mathbf{p})$ (otherwise the deviation would not be profitable for player j). Since $\text{Visit}(\mathbf{p}) = \text{Visit}(\mathbf{pq})$, we also have

$$\text{Payoff}_j(\pi') + (|\mathbf{q}| + 1) = \text{Payoff}_j(\rho').$$

By Equations (5) and (6), we get

$$\begin{aligned} & \text{either } x_j = y_j = +\infty \text{ and } \text{Payoff}_j(\rho') < x_j, \\ & \text{or } x_j = y_j + (|\mathbf{q}| + 1) \text{ and } \text{Payoff}_j(\rho') < x_j, \end{aligned}$$

which proves that σ'_j is a profitable deviation for player j w.r.t. $(\sigma_i)_{i \in \Pi}$. \square

While Lemma 18 deals with elimination of unnecessary cycles, Lemma 19 deals with repetition of a useful cycle.

Lemma 19. *Let $(\sigma_i)_{i \in \Pi}$ be a strategy profile in a game \mathcal{G} and $\rho = \langle (\sigma_i)_{i \in \Pi} \rangle$ its outcome. We assume that $\rho = \mathbf{pq}\tilde{\rho}$, where \mathbf{q} contains at least one vertex, such that*

$$\begin{aligned} \text{Visit}(\mathbf{p}) &= \text{Visit}(\rho) \\ \text{Last}(\mathbf{p}) &= \text{Last}(\mathbf{pq}). \end{aligned}$$

We define a strategy profile $(\tau_i)_{i \in \Pi}$ as follows:

$$\tau_i(h) = \begin{cases} \sigma_i(h) & \text{if } \mathbf{p} \not\preceq h, \\ \sigma_i(\mathbf{p}\delta) & \text{if } h = \mathbf{pq}^k\delta, k \in \mathbb{N}, \text{ and } \mathbf{q} \not\preceq \delta \end{cases}$$

where h is a history of \mathcal{G} with $\text{Last}(h) \in V_i$. We get the outcome $\langle (\tau_i)_{i \in \Pi} \rangle = \mathbf{pq}^\omega$. If a strategy τ'_j is a profitable deviation for player j w.r.t. $(\tau_i)_{i \in \Pi}$, then there exists a profitable deviation σ'_j for player j w.r.t. $(\sigma_i)_{i \in \Pi}$.

Proof. We use the same notations as in the proof of Lemma 18. Here we have $x_i = y_i$ for all $i \in \Pi$ since $\text{Visit}(\mathbf{p}) = \text{Visit}(\rho)$. One can prove that $\pi = \mathbf{pq}^\omega$ (see Figures 8 and 9).

We show how to define a profitable deviation σ'_j from the deviation τ'_j . We distinguish the following two cases:

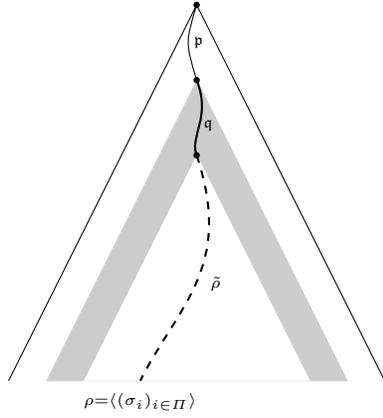


Fig. 8. Play ρ and its prefix $\mathfrak{p}\mathfrak{q}$.

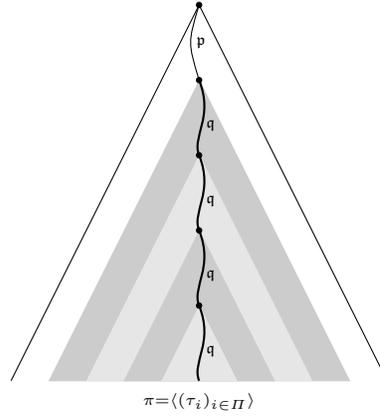


Fig. 9. Play $\pi = \mathfrak{p}\mathfrak{q}^\omega$.

- (i) player j deviates from π just after a proper prefix h of $\mathfrak{p}\mathfrak{q}$.

We define $\sigma'_j = \tau'_j$. As in the first case of the proof of Lemma 18, we have $\text{Payoff}_j(\rho') < x_j$, which implies that σ'_j is a profitable deviation of player j w.r.t. $(\sigma_i)_{i \in \Pi}$.

- (ii) player j deviates from π after the prefix $\mathfrak{p}\mathfrak{q}$, i.e. after a prefix $\mathfrak{p}\mathfrak{q}^k$ and strictly before the prefix $\mathfrak{p}\mathfrak{q}^{k+1}$ ($k \geq 1$).

We define for all histories h ending in a vertex of V_j :

$$\sigma'_j(h) = \begin{cases} \sigma_j(h) & \text{if } \mathfrak{p} \not\preceq h, \\ \tau'_j(\mathfrak{p}\mathfrak{q}^k\delta) & \text{if } h = \mathfrak{p}\delta. \end{cases}$$

One can prove that

$$\pi' = \mathfrak{p}\mathfrak{q}^k \tilde{\pi}' \quad \text{and} \quad \rho' = \mathfrak{p}\tilde{\pi}'.$$

And then, in the point of view of payoffs we have

$$\text{Payoff}_j(\rho') < \text{Payoff}_j(\pi') < y_j = x_j,$$

which proves that σ'_j is a profitable deviation for player j w.r.t. $(\sigma_i)_{i \in \Pi}$. \square

The next proposition achieves the first step of the proof of Theorem 12 as mentioned in Section 3.2. It shows that one can construct from a Nash equilibrium another Nash equilibrium with the same type and with an outcome of the form $\alpha\beta^\omega$. Its proof uses Lemmas 18 and 19.

Proposition 20. *Let $(\sigma_i)_{i \in \Pi}$ be a Nash equilibrium in a game \mathcal{G} . Then there exists a Nash equilibrium $(\tau_i)_{i \in \Pi}$ with the same type such that $\langle (\tau_i)_{i \in \Pi} \rangle = \alpha\beta^\omega$, where $\text{Visit}(\alpha) = \text{Type}((\sigma_i)_{i \in \Pi})$ and $|\alpha\beta| < (|\Pi| + 1) \cdot |V|$.*

Proof. Let $\Pi = \{1, \dots, n\}$ and $\rho = \langle (\sigma_i)_{i \in \Pi} \rangle$. Without loss of generality suppose that

$$\begin{aligned} \text{Payoff}(\rho) = (x_1, \dots, x_n) \quad & \text{such that } x_1 \leq \dots \leq x_k < +\infty \\ & \text{and } x_{k+1} = \dots = x_n = +\infty \end{aligned}$$

where $1 \leq k \leq n$. We consider two cases:

(i) $x_1 \geq |V|$.

Then, there exists a prefix \mathbf{pq} of ρ , with \mathbf{q} containing at least one vertex, such that

$$\begin{aligned} |\mathbf{pq}| &< x_1 \\ \text{Visit}(\mathbf{p}) &= \text{Visit}(\mathbf{pq}) = \emptyset \\ \text{Last}(\mathbf{p}) &= \text{Last}(\mathbf{pq}). \end{aligned}$$

We define the strategy profile $(\tau_i)_{i \in \Pi}$ as proposed in Lemma 18. By this lemma it is actually a Nash equilibrium in \mathcal{G} . With $\pi = \langle (\tau_i)_{i \in \Pi} \rangle$, we have

$$\rho = \mathbf{pq}\tilde{\rho} \quad \text{and} \quad \pi = \mathbf{p}\tilde{\rho}.$$

Thus if the payoff profile for the play π is (y_1, \dots, y_n) , we have

$$\begin{aligned} y_1 &< x_1, \dots, y_k < x_k \\ y_{k+1} &= x_{k+1} = +\infty, \dots, y_n = x_n = +\infty. \end{aligned}$$

(ii) $(x_{l+1} - x_l) \geq |V|$ for $1 \leq l \leq k - 1$.

Then, there exists a prefix \mathbf{pq} of ρ , with \mathbf{q} containing at least one vertex, such that

$$\begin{aligned} x_l &< |\mathbf{pq}| < x_{l+1} \\ \text{Visit}(\mathbf{p}) &= \text{Visit}(\mathbf{pq}) = \{1, \dots, l\} \\ \text{Last}(\mathbf{p}) &= \text{Last}(\mathbf{pq}). \end{aligned}$$

We define the strategy profile $(\tau_i)_{i \in \Pi}$ given in Lemma 18. It is then a Nash equilibrium in \mathcal{G} , and for $\pi = \langle (\tau_i)_{i \in \Pi} \rangle$, we have

$$\rho = \mathbf{pq}\tilde{\rho} \quad \text{and} \quad \pi = \mathbf{p}\tilde{\rho}.$$

Hence if the payoff profile for the play π is (y_1, \dots, y_n) , we have

$$\begin{aligned} y_1 &= x_1, \dots, y_l = x_l; \\ y_{l+1} &< x_{l+1}, \dots, y_k < x_k; \\ y_{k+1} &= x_{k+1} = +\infty, \dots, y_n = x_n = +\infty. \end{aligned}$$

From the two previous cases, we can assume without loss of generality that $(\sigma_i)_{i \in \Pi}$ is a Nash equilibrium with a payoff profile (x_1, \dots, x_n) such that

$$\begin{aligned} x_i &< i \cdot |V| && \text{for } i \leq k; \\ x_i &= +\infty && \text{for } i > k. \end{aligned}$$

Let us go further. We can write $\rho = \alpha\beta\tilde{\rho}$ such that

$$\begin{aligned} \text{Visit}(\alpha) &= \text{Visit}(\rho) \\ \text{Last}(\alpha) &= \text{Last}(\alpha\beta) \\ |\alpha\beta| &< (k+1) \cdot |V| \leq (n+1) \cdot |V|. \end{aligned}$$

Indeed, the prefix h of ρ of length $(n+1) \cdot |V|$ visits each goal set Goal_i , with $i \leq k$, and after the last visited Goal_k , there remains enough vertices to observe a cycle. Notice that $\text{Visit}(\alpha) = \text{Visit}(\alpha\beta) = \text{Visit}(\rho) (= \text{Type}((\sigma_i)_{i \in \Pi}))$.

If we define the strategy profile $(\tau_i)_{i \in \Pi}$ like in Lemma 19, we get a Nash equilibrium in \mathcal{G} with outcome $\alpha\beta^\omega$ and the same type as $(\sigma_i)_{i \in \Pi}$. \square

We are now ready to prove Theorem 12.

Proof (of Theorem 12). Let $\Pi = \{1, \dots, n\}$. Let $(\sigma_i)_{i \in \Pi}$ be a Nash equilibrium in the game \mathcal{G} . The first step consists in constructing a Nash equilibrium as in Proposition 20. Let us denote it again by $(\sigma_i)_{i \in \Pi}$. Let $\rho = \langle (\sigma_i)_{i \in \Pi} \rangle = \alpha\beta^\omega$ such that $\text{Visit}(\alpha) = \text{Type}((\sigma_i)_{i \in \Pi})$ and $|\alpha\beta| < (n+1) \cdot |V|$. The strategy profile $(\sigma_i)_{i \in \Pi}$ is also a Nash equilibrium in the game \mathcal{T} played on the unravelling T of G .

For the second step we consider $\text{Trunc}_d(T)$ the truncated tree of T of depth $d = (n+2) \cdot |V|$. It is clear that the strategy profile $(\sigma_i)_{i \in \Pi}$ limited to this tree is also a Nash equilibrium of $\text{Trunc}_d(\mathcal{T})$.

We know that $|\alpha\beta| < (n+1) \cdot |V|$ and we set γ such that $\alpha\beta\gamma$ is a prefix of ρ and $|\alpha\beta\gamma| = (n+2) \cdot |V|$. Furthermore we have $\text{Last}(\alpha) = \text{Last}(\alpha\beta)$ and $\text{Visit}(\alpha) = \text{Visit}(\alpha\beta\gamma)$ (since $\text{Visit}(\alpha) = \text{Type}(\rho)$). Then this prefix $\alpha\beta\gamma$ satisfies the properties described in Lemma 10 (by setting $l = (n+1) \cdot |V|$). By Lemma 11 we conclude that there exists a Nash equilibrium $(\tau_i)_{i \in \Pi}$ with finite memory such that $\text{Type}((\tau_i)_{i \in \Pi}) = \text{Visit}(\alpha)$, that is with the same type as the initial Nash equilibrium $(\sigma_i)_{i \in \Pi}$. \square

E Existence of a secure equilibrium

The notion of secure equilibrium is based on the binary relations \prec_j of Definition 3. One can easily see that \prec_j is not reflexive. Moreover, \prec_j is not total as soon as at least three players are involved (for instance $(1, 2, 3) \not\prec_1 (1, 3, 2)$ and $(1, 3, 2) \not\prec_1 (1, 2, 3)$). To be able to apply Kuhn's theorem, it is more convenient to define secure equilibria via a preference relation. In this aim, we first define an equivalence relation \sim_j on payoff profiles for each player. Given two payoff profiles $(x_i)_{i \in \Pi}$ and $(y_i)_{i \in \Pi}$:

$$(x_i)_{i \in \Pi} \sim_j (y_i)_{i \in \Pi} \quad \text{iff} \quad (\forall k \ x_k = y_k) \vee (x_j = y_j \wedge \exists k \ x_k < y_k \wedge \exists k \ x_k > y_k).$$

The idea behind the introduction of \sim_j is that player j is *indifferent* to any payoff profiles \sim_j -equivalent. We can now define a preference relation, denoted by \succsim_j , for each player j . Given two payoff profiles $(x_i)_{i \in \Pi}$ and $(y_i)_{i \in \Pi}$:

$$(x_i)_{i \in \Pi} \succsim_j (y_i)_{i \in \Pi} \quad \text{iff} \quad (x_i)_{i \in \Pi} \sim_j (y_i)_{i \in \Pi} \quad \vee \quad (x_i)_{i \in \Pi} \prec_j (y_i)_{i \in \Pi}.$$

One can verify that \succsim_j is a preference relation. We can now provide an equivalent definition of secure equilibrium.

Proposition 21. *A strategy profile $(\sigma_i)_{i \in \Pi}$ of a game \mathcal{G} is a secure equilibrium iff for all players $j \in \Pi$ and for all strategies σ'_j of player j in \mathcal{G} , we have:*

$$\text{Payoff}(\rho') \succsim_j \text{Payoff}(\rho)$$

where $\rho = \langle (\sigma_i)_{i \in \Pi} \rangle$ and $\rho' = \langle \sigma'_j, (\sigma_i)_{i \in \Pi \setminus \{j\}} \rangle$.

Since each \succsim_i is a preference relation, we get the next corollary by Kuhn's theorem (in a multiplayer framework).

Corollary 22. *Let \mathcal{G} be a quantitative multiplayer reachability game and T be the unraveling of \mathcal{G} . Let $\text{Trunc}_d(\mathcal{T})$ be the game played on the truncated tree of T of depth d , with $d \geq 0$. Then there exists a secure equilibrium in $\text{Trunc}_d(\mathcal{T})$.*

The next proposition states that it is possible to extend a secure equilibrium in $\text{Trunc}_d(\mathcal{T})$ to a secure equilibrium in the game \mathcal{T} , if the depth d is equal to $(|\Pi| + 1) \cdot 2 \cdot |V|$ and there are only two players.

Proposition 23. *Let \mathcal{G} be a quantitative two-player reachability game and T be the unraveling of \mathcal{G} . Let $\text{Trunc}_d(\mathcal{T})$ be the game played on the truncated tree of T of depth $d = (|\Pi| + 1) \cdot 2 \cdot |V|$.*

If there exists a secure equilibrium (σ_1, σ_2) in the game $\text{Trunc}_d(\mathcal{T})$, then there exists a finite-memory secure equilibrium in the game \mathcal{T} .

The proof of Proposition 23 works pretty much as the one of Proposition 8. A well-chosen prefix $\alpha\beta$, with β being a cycle, is first extracted from the outcome ρ of the secure equilibrium (σ_1, σ_2) of $\text{Trunc}_d(\mathcal{T})$. The outcome of the required secure equilibrium of \mathcal{T} will be equal to $\alpha\beta^\omega$. As soon as a player deviates from this play, the other player punishes him, but the way to define punishment is here more involved than in the proof of Proposition 8. Notice that Proposition 23 stands for two-player games because its proof uses Lemma 24 that has been proved only for two players.

We begin with Lemma 24 whose hypotheses are the same as in Lemma 10. Recall that Lemma 10 states that for all $j \in \Pi$ such that α does not visit his goal set Goal_j , the players $i \neq j$ can play together to prevent player j from reaching his goal set Goal_j from any history hu consistent with $(\sigma_i)_{i \in \Pi \setminus \{j\}}$ and such that $|hu| \leq |\alpha\beta|$. We denote by ν_{-j} the memoryless winning strategy of the coalition, and for each player $i \neq j$, $\nu_{i,j}$ the memoryless strategy of player i in \mathcal{G} induced by ν_{-j} .

Lemma 24 states that if α visits Goal_1 for example, then α visits Goal_2 or ρ does not visit Goal_2 . It is given for two-player games only.

Lemma 24. *Let $d \geq 0$. Let (σ_1, σ_2) be a secure equilibrium in $\text{Trunc}_d(\mathcal{T})$ and $\rho = \langle (\sigma_1, \sigma_2) \rangle$ its outcome. Suppose that ρ has a prefix $\alpha\beta\gamma$, where β contains at least one vertex, such that*

$$\begin{aligned}\text{Visit}(\alpha) &= \text{Visit}(\alpha\beta\gamma) \\ \text{Last}(\alpha) &= \text{Last}(\alpha\beta) \\ |\alpha\beta| &\leq l \cdot |V| \\ |\alpha\beta\gamma| &= (l+1) \cdot |V|\end{aligned}$$

for some $l \geq 1$. Then we have

$$(\text{Visit}(\alpha) \neq \emptyset \vee \text{Visit}(\rho) \neq \{1, 2\}) \Rightarrow \text{Visit}(\alpha) = \text{Visit}(\rho).$$

Proof. By contradiction, assume that $2 \in \text{Visit}(\rho) \setminus \text{Visit}(\alpha)$. The hypothesis implies that $1 \in \text{Visit}(\alpha)$ or $1 \notin \text{Visit}(\rho)$.

By Lemma 10, player 1 wins the game \mathcal{G}_2 from $\text{Last}(\alpha)$, that is has a memoryless winning strategy $\nu_{1,2}$ from this vertex. Then if player 1 plays according to σ_1 until depth $|\alpha|$, and then switches to $\nu_{1,2}$ from $\text{Last}(\alpha)$, this strategy is a \prec_1 -profitable deviation for player 1 w.r.t. (σ_1, σ_2) . Indeed, if $1 \in \text{Visit}(\alpha)$, player 1 manages to increase player 2's payoff while keeping his own payoff. On the other hand, if $1 \notin \text{Visit}(\rho)$, either player 1 succeeds in either reaching his goal set (i.e. strictly decreasing his payoff), or he does not reach it (then gets the same payoff as in ρ) but succeeds in increasing player 2's payoff. Thus we get a contradiction. \square

Lemma 25 says that one can define a secure equilibrium (τ_1, τ_2) in the game \mathcal{T} , from the secure equilibrium (σ_1, σ_2) in the game $\text{Trunc}_d(\mathcal{T})$. The definition of the strategy profile (τ_1, τ_2) is a little different from the one in the proof of Lemma 11 because here, if player 1 deviates (for example), then player 2 has to prevent him from reaching his goal set Goal_1 (faster), or having the same payoff but succeeding in increasing player 2's payoff.

Lemma 25. *Let $d \geq 0$. Let (σ_1, σ_2) be a secure equilibrium in $\text{Trunc}_d(\mathcal{T})$ and $\alpha\beta\gamma$ be a prefix of $\rho = \langle (\sigma_1, \sigma_2) \rangle$ as defined in Lemma 24. Then there exists a secure equilibrium (τ_1, τ_2) in the game \mathcal{T} . Moreover (τ_1, τ_2) is finite-memory and $\text{Type}((\tau_1, \tau_2)) = \text{Visit}(\alpha)$.*

Proof. Like in the proof of Lemma 11 we consider the infinite play $\alpha\beta^\omega$ in the game \mathcal{T} . The basic idea of the strategy profile (τ_1, τ_2) is the same as for the Nash equilibrium case: player 1 (resp. 2) plays according to $\alpha\beta^\omega$ and punishes player 2 (resp. 1) if he deviates from $\alpha\beta^\omega$, in the following way. Suppose that player 2 deviates (the case for player 1 is similar). Then player 1 plays according to σ_1 until depth $|\alpha|$, and after that, he plays arbitrarily if α visits Goal_2 , otherwise he plays according to $\nu_{1,2}$.

We define the same punction function P as in the proof of Lemma 11: $P(v_0) = \perp$, and for $h \in V^*$ and $v \in V_i$ ($i = 1, 2$),

$$P(hv) = \begin{cases} \perp & \text{if } P(h) = \perp \text{ and } hv < \alpha\beta^\omega, \\ i & \text{if } P(h) = \perp \text{ and } hv \not< \alpha\beta^\omega, \\ P(h) & \text{otherwise } (P(h) \neq \perp). \end{cases}$$

The definition of the secure equilibrium (τ_1, τ_2) is as follows:

$$\tau_i(h) = \begin{cases} v & \text{if } P(h) = \perp \text{ (} h < \alpha\beta^\omega \text{); such that } hv < \alpha\beta^\omega, \\ \text{arbitrary} & \text{if } P(h) = i, \\ \sigma_i(h) & \text{if } P(h) \neq \perp, i \text{ and } |h| \leq |\alpha|, \\ \nu_{i,P(h)}(h) & \text{if } P(h) \neq \perp, i, |h| > |\alpha| \text{ and } \alpha \text{ does not visit } \text{Goal}_{P(h)}, \\ \text{arbitrary} & \text{otherwise (} P(h) \neq \perp, i, |h| > |\alpha| \text{ and } \alpha \text{ visits } \text{Goal}_{P(h)}) \end{cases}$$

where $i = 1, 2$, and *arbitrary* means that the next vertex is chosen arbitrarily (in a memoryless way). Clearly the outcome of (τ_1, τ_2) is the play $\alpha\beta^\omega$, and $\text{Type}((\tau_1, \tau_2))$ is equal to $\text{Visit}(\alpha)$ ($= \text{Visit}(\alpha\beta)$). Moreover, as done in the proof of Lemma 11, (τ_1, τ_2) is a finite-memory strategy profile.

It remains to show that (τ_1, τ_2) is a secure equilibrium in the game \mathcal{T} . Assume that there exists a \prec_1 -profitable deviation τ'_1 for player 1 w.r.t. (τ_1, τ_2) . The case of a \prec_2 -profitable deviation τ'_2 for player 2 is similar. We construct a play ρ' in $\text{Trunc}_d(\mathcal{T})$ as follows: player 1 plays according to the strategy τ'_1 restricted to $\text{Trunc}_d(\mathcal{T})$ (denoted by σ'_1) and player 2 plays according to σ_2 . Thus the play ρ' coincide with the play $\pi' = \langle \tau'_1, \tau_2 \rangle$ at least until depth $|\alpha|$ (by definition of τ_2); it can differ afterwards. We have:

$$\begin{aligned} \rho &= \langle \sigma_1, \sigma_2 \rangle && \text{of payoff profile } (x_1, x_2) \\ \rho' &= \langle \sigma'_1, \sigma_2 \rangle && \text{of payoff profile } (x'_1, x'_2) \\ \pi &= \langle \tau_1, \tau_2 \rangle && \text{of payoff profile } (y_1, y_2) \\ \pi' &= \langle \tau'_1, \tau_2 \rangle && \text{of payoff profile } (y'_1, y'_2). \end{aligned}$$

The situation is depicted in Figure 10.

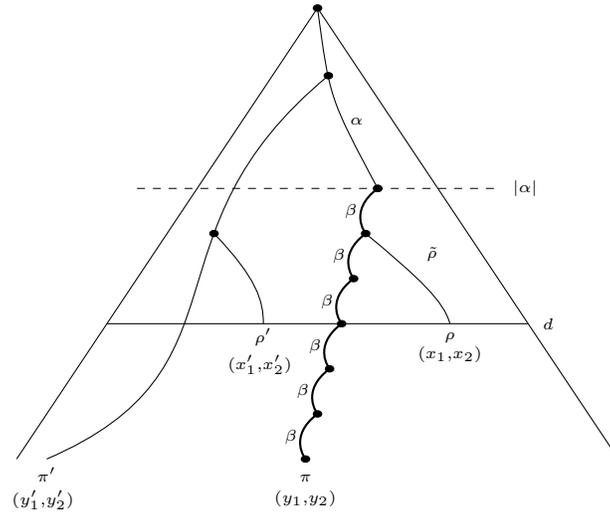


Fig. 10. Plays ρ and π , and their respective deviations ρ' and π' .

We are going to show that $(x_1, x_2) \prec_1 (x'_1, x'_2)$, meaning that σ'_1 is a \prec_1 -profitable deviation for player 1 w.r.t. (σ_1, σ_2) in $\text{Trunc}_d(\mathcal{T})$. This will lead to the contradiction. As τ'_1 is a \prec_1 -profitable deviation w.r.t. (τ_1, τ_2) , one of the following three cases stands.

(i) $y'_1 < y_1 < +\infty$.

As $\pi = \alpha\beta^\omega$, it means that α visits F_1 , and then:

$$y'_1 < y_1 = x_1 \leq |\alpha|.$$

As $y'_1 < |\alpha|$, we have $x'_1 = y'_1$ (as ρ' and π' coincide until depth $|\alpha|$). Therefore $x'_1 < x_1$.

(ii) $y'_1 < y_1 = +\infty$.

If $y'_1 \leq |\alpha|$, we have $x'_1 = y'_1$ (by the same argument as before). Furthermore $x_1 > |\alpha|$ since $y_1 = +\infty$ and α is a common prefix of ρ and π . So we have $x'_1 \leq |\alpha| < x_1$.

We show that the case $y'_1 > |\alpha|$ is impossible. By definition of τ_2 the play π' is consistent with σ_2 until depth $|\alpha|$, and then with $\nu_{2,1}$ (as $y_1 = +\infty$). By Lemma 10 the play π' can not visit Goal_1 after a depth $> |\alpha|$.

(iii) $y_1 = y'_1$ and $y_2 < y'_2$.

Note that this implies $y_2 < +\infty$ and $x_2 = y_2$ (as $\pi = \alpha\beta^\omega$). Since ρ' and π' coincide until depth $|\alpha|$, $y_2 < y'_2$ and $x_2 = y_2 \leq |\alpha|$, we have

$$x_2 = y_2 < x'_2$$

showing that the payoff of player 2 is increased. It remains to consider the case of player 1, that is, showing that either he keeps the same payoff, or he decreases his payoff.

If $y'_1 = y_1 < +\infty$, it follows as in the first case that:

$$y_1 = x_1 \leq |\alpha| \quad \text{and} \quad x'_1 = y'_1.$$

Therefore $x_1 = x'_1$, i.e. player 1 has the same payoff in ρ and ρ' .

On the contrary, if $y'_1 = y_1 = +\infty$, it follows that $1 \notin \text{Visit}(\alpha)$. As $y_2 < +\infty$, we have $2 \in \text{Visit}(\alpha)$. By Lemma 24, we know that $1 \notin \text{Visit}(\rho)$, i.e. $x_1 = +\infty$. If ρ' visits F_1 , player 1 gets a payoff $x'_1 < +\infty$ (then $x'_1 < x_1$). Otherwise, he has the same payoff $x'_1 = x_1 = +\infty$ as in ρ . □

We can now complete the proof of Proposition 23.

Proof (of Proposition 23). Let (σ_1, σ_2) be a secure equilibrium in $\text{Trunc}_d(\mathcal{T})$, and ρ its outcome. We define the prefixes \mathfrak{pq} and $\alpha\beta\gamma$ as in the proof of Proposition 8 (see Figure 1).

As the hypotheses of Lemmas 10, 24 and 25 are verified, we can apply them in this context to get a secure equilibrium (τ_1, τ_2) with finite memory and such that $\text{Type}((\tau_1, \tau_2)) = \text{Visit}(\alpha)$. □