Thèse de Doctorat de l'Université de Mons

Spécialité :
Sciences Mathématiques

Equilibria in Multiplayer Cost Games

présentée par

Julie De Pril

pour l'obtention du grade de Docteur en Sciences.

Thèse soutenue le 25 avril 2013 devant le jury composé de :

Pr. Thomas Brihaye, Université de Mons (Directeur)
Pr. Véronique Bruyère, Université de Mons (Co-directeur)
Pr. Krishnendu Chatterjee, IST Austria
Pr. Hadrien Mélot, Université de Mons
Pr. Jean-François Raskin, Université Libre de Bruxelles
Pr. Cédric Rivière, Université de Mons
Pr. Wolfgang Thomas, RWTH Aachen

Année académique 2012–2013

Acknowledgements

I would like to thank all the people who made this thesis possible, in one way or another.

I am especially grateful to my supervisors, Thomas Brihaye and Véronique Bruyère, for their assistance, advice and encouragement throughout the course of my thesis.

I would like to thank each member of my thesis committee for his reading of this document.

I am thankful to my co-authors Thomas Brihaye, Véronique Bruyère, Hugo Gimbert and Sven Schewe. I also thank János Flesch, Jeroen Kuipers, Gijs Schoenmakers and Koos Vrieze for interesting discussions.

A special thank you to Marc Labie and Anaïs Périlleux for letting me discover the field of microfinance with so much enthusiasm.

I also thank the F.R.S.-FNRS, as well as the L'Oréal-Unesco Foundation and the GASICS project, for financial support.

Finally, I say some words in French to thank my friends and family.

Merci à Aaron, Marc et Mickael pour les blagues et les rires en conférence et en séminaire.

Merci à Quentin, notamment pour les truffes, pralines et oeufs en chocolat qui ont rendu la rédaction de ma thèse plus douce.

Un merci très spécial à Marc pour l'agréable co-habitation dans notre bureau au cours de ces quelques années de thèse. Le moins que je puisse dire, c'est que ce fut extrêmement... Cocasse !

Merci à mes ami(e)s pour leur bienveillance. Je ne citerai pas leurs noms, certaine qu'ils se reconnaîtront (et, il faut l'avouer, un peu angoissée d'oublier malencontreusement un nom).

Je remercie tout particulièrement mes parents de m'avoir toujours soutenue et encouragée. Et d'être là, indéfectiblement. Les mots ne seront pas suffisants pour décrire à quel point je leur suis reconnaissante.

Un énorme merci à mon frère Simon, dont je suis très fière. Il voulait que je le remercie de m'avoir inspiré les résultats de cette thèse... Mais, ce n'est pas réellement le cas. Ceci dit, il m'a insufflé de bien plus importantes et plus belles idées concernant la Vie.

Je remercie également ma marraine, mon Pilo, Marine, Camille, Bernard, Annick, Céline et Luc pour tous les précieux moments passés en famille, ainsi que pour leur présence constante. Un merci plus particulier à Céline et Jean-Phi pour le merveilleux cadeau qu'ils me feront très bientôt.

Thomas, chaque instant partagé avec toi m'enrichit. Je te dis mille et un mercis de m'offrir ces perles de vie... Faisant sourire mon coeur de pur bonheur...

Contents

Chapter 1

Introduction

We here present the context of this thesis, which is on the edge between game theory and model-checking. We then give our main contributions and discuss related work.

Game theory. Game theory [OR94] is a branch of mathematics which has been successfully applied in various domains including economics, biology, and computer science. Game theory attempts to model strategic situations where several individuals are interacting and try to predict what will be the decisions taken by individuals in a given situation, assuming rationality of those individuals. In this framework, a strategic situation is called a *game* and each individual taking part to this game is called a *player*, their decisions are called *strategies*. One could say that the systematic study of game theory started with the 1944 book "*Theory of Games and Economic Behaviour*" by John von Neumann and Oskar Morgenstern [vNM44]. This book mainly focused on strictly competitive situations in which only two individuals interact, also known as *zero-sum games*. Another important step forward in the development of game theory, around 1950, has been the introduction of the concept of *Nash equilibrium* together with a proof of its existence [Nas50]. This important notion allows us to study *multiplayer non-zero-sum games*. Following the seminal work of Nash, several other important notions have been introduced including the *subgame perfect equilibrium* of Reinhard Selten [Sel65], the concepts of *incomplete information* and *Bayesian games* of John Harsanyi [Har68]. Let us mention that Nash, Selten and Harsanyi became Economics Nobel laureates in 1994 for their contributions to game theory and their applications to economics. Repeated games (see for example [A+81, BPS07]) and stochastic games (introduced by Lloyd Shapley [Sha53], Economics Nobel laureates in 2012, together with Alvin E. Roth) are also two examples of games intensively studied and largely applied in economics.

Algorithmic game theory. More recently a large current of work in game theory has arisen under the driving force of the computer-science community. This led to the new subfield of *algorithmic game theory* [NRTV07]. Nash theorem is a purely existential result, i.e., it ensures the existence of Nash equilibria but gives no clue on how we could compute them and how difficult this computation would be. The needs for computational and complexity information in this domain are well illustrated by a famous quotation from Kamal Jain (Microsoft Research), popularised by Christos Papadimitriou (Berkeley): "If your laptop can't find the equilibrium, then neither can the market.". In the general case of non-zero-sum games, results concerning the complexity of finding Nash equilibria have only been obtained recently by Papadimitriou et al [DGP09]. They have introduced a new complexity class, called PPAD, to characterise the hardness to compute such equilibria. Beside those important initial results, the computational aspects of game theoretical concepts are not yet well studied, but there is currently a large international research effort to better understand them.

Model-checking. *Computer aided verification*, and more specifically *model-checking*, is a branch of computer science which offers techniques to check automatically that a given computer system satisfies a given specification [CGP01]. The model-checking techniques apply on a model or an abstraction of the system (such as a finite automaton) together with a translation of the specification into a logical formula (of a temporal logic for instance). The *classical model-checking* approach intends to develop efficient algorithms to provide a boolean (yes or no) answer to the question: "Does the system satisfy the specification?". Computer aided verification is essential when considering computer systems responsible of critical tasks such as air traffic management, control of nuclear power plant,... Nowadays, this technology is an important part of the design cycle in companies such as Intel or IBM.

Unfortunately, classical model-checking techniques do not trivially extend to complex systems, such as embedded or distributed systems. A main reason for this is that such systems often consist of multiple independent components with individual objectives. These components can be viewed as selfish agents that may cooperate and compete at the same time. It is difficult to model the interplay between these components with traditional finite state machines, as they cannot reflect the intricate quantitative valuation of an agent on how well he has met his goal (in term of quality of service, time and/or energy consumption,...). In particular, it is not realistic to assume that these components are always cooperating to satisfy a common goal, as it is, e.g., assumed in works that distinguish between an environment and a system. We argue that it is more realistic to assume that all components act like selfish agents that try to achieve their own objectives and are either unconcerned about the effect this has on the other components or consider this effect to be secondary. It is indeed a recent trend to enhance the system models used in the classical approach of verification by quantitative cost and gain functions, and to exploit the well established game-theoretic framework [Nas50, OR94] for their formal analysis.

The first steps towards the extension of computational models with concepts from classical game theory were done by advancing from boolean to general infinite two-player zero-sum games played on graphs [GTW02]. Like their qualitative counterparts, those games are adequate to model interaction problems between a controller and an environment [Tho95, Tho08]. As usual in control theory, one can distinguish between moves of a control player, who plays actions to control a system to meet a control objective, and an antagonistic environment player. In the classical setting, the control player has a qualitative objective—he might, for example, try to enforce a temporal specification—whereas the environment tries to prevent this. In the extension to quantitative games, the controller instead tries to maximise its gain, while the environment tries to minimise it. This extension lifts the controller synthesis problem from a constructive extension of a decision problem to a classical optimisation problem.

However, the previous extension has not lifted the restriction to purely antagonist interactions between a controller and a hostile environment. In order to study more complex systems with more than two components, and with objectives that are not necessarily antagonist, new *multi-agent models*, inspired from game theory, have been considered: *multi-player non-zero-sum games*. In this context, *equilibria* take the place that winning and optimal strategies take in qualitative and quantitative two-player games zero-sum games, respectively. Depending on the kind of the agents' rational behaviour, several sorts of equilibria can be considered, including Nash equilibrium, subgame perfect equilibrium, or secure equilibrium (recent concept of equilibrium particularly suitable in the context of controller synthesis, introduced by K. Chatterjee, T. A. Henzinger and M. Jurdzinski [CHJ04]). Surprisingly, qualitative objectives have so far prevailed in the study of equilibria for distributed systems. However, we argue that quantitative objectives—such as reaching a set of target states quickly or with a minimal consumption of energy—are natural objectives that ought to be studied alongside (or instead of) traditional qualitative objectives.

In summary, in order to study complex interactive computer systems with *more than two components*, and with *quantitative objectives* that are not necessarily antagonist, we resort to *multiplayer non-zero-sum quantitative games played on graphs* (also called *multiplayer cost games* in this document). Our thesis follows this research direction.

Our contribution. Most of our results concern existence of several kinds of equilibria in (turn-based) multiplayer cost games. We study Nash equilibria and more refined notions of equilibria. As Nash equilibria do not exist in every multiplayer cost game, we focus on a particular subclass of cost games where each player has a quantitative reachability objective, and we also define large subclasses of cost games where simple [1] Nash equilibria exist.

We here give our main results, listed in chronological order of their discovery. Our contributions are detailed in Chapter 3.

In collaboration with Thomas Brihaye and Véronique Bruyère, we study *multiplayer quantitative reachability games*. In this framework, each player has a goal set of vertices of the graph, and aims at reaching his own goal set as soon as possible. We focus on existence results for two solution concepts: *Nash equilibrium* and *secure equilibrium*. We show the existence of Nash (resp. secure) equilibria in multiplayer (resp. two-player) games, and also show that these equilibria can be chosen with finite memory (see [BBD10, BBD12] and Chapters 4, 5). Moreover, we prove that given a Nash (resp. secure) equilibrium of a multiplayer (resp. two-player) game, we can build a finite-memory Nash (resp. secure) equilibrium of the same type, i.e. preserving the set of players achieving their reachability objectives (see [BBD12] and Chapters 4, 5).

Together with Thomas Brihaye, Véronique Bruyère and Hugo Gimbert, we consider alternative solution concepts to the classical notion of Nash (or secure) equilibria. In particular, in the present framework of games on graphs, it is very natural to consider the notion of *subgame perfect equilibrium*. Indeed if the initial state or the initial history of the system is not known, then a robust controller should be subgame perfect. We prove that there exists a subgame perfect equilibrium in every multiplayer quantitative reachability game (see [BBDG12, BBDG13] and Chapter 6).

We also introduce an even stronger solution concept with the notion of *subgame perfect secure equilibrium*, which gathers both the sequential nature of subgame perfect equilibria and the verification-oriented aspects of secure equilibria. We show that there exists a subgame perfect secure equilibrium in every two-player quantitative reachability game (see [BBDG12, BBDG13] and Chapter 7).

1. By *simple* Nash equilibrium, we mean that each of its strategies can be represented by a finite automaton of reasonable size.

Moreover, we provide an algorithm that decides in ExpSpace whether there exists a secure equilibrium (with players' costs below some thresholds) in a multiplayer quantitative reachability game (see [BBDG13] and Chapter 5). To our knowledge, the existence of a secure equilibrium in multiplayer quantitative reachability games is still an open problem.

Jointly with Thomas Brihaye and Sven Schewe, we consider more general objectives: the quantitative objectives of the players are expressed through a cost function for each player. Each cost function assigns, for every play of the game, a value that represents the cost that is incurred for a player by this play. Cost functions allow to express classical quantitative objectives such as quantitative reachability or mean-payoff objectives. In this framework, all players are supposed to be rational: they want to minimise their own cost.

Our results are twofold. Firstly, we prove the existence of simple Nash equilibria for a large class of cost games that includes quantitative reachability and mean-payoff objectives (see [BDS13] and Chapter 4). This result then gives a quantitative counterpart to a result of E. Grädel and M. Ummels [GU08] about qualitative games. Secondly, we study the complexity of these Nash equilibria in terms of the memory needed in the strategies of the individual players in these Nash equilibria. More precisely, we ensure existence of Nash equilibria whose strategies only requires a number of memory states that is linear in the size of the game for a wide class of cost games, including games with quantitative reachability and mean-payoff objectives.

Some of these results are summarised in Table 1.1. In this table, 'NE' (resp. 'SE', 'SPE', 'SPSE') means 'Nash (resp. secure, subgame perfect, subgame perfect secure) equilibrium'.

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 [CH06], mean-payoff parity games [CHJ05], games with prioritised requirements [AKW08], request-response games where the waiting times between the requests and the responses are minimised

Table 1.1: Summary of the main results

NE	• Existence of a finite-memory NE in multiplayer quantitative reachability games [BBD10, BBD12]. • Existence of a simple NE in a large class of multiplayer cost games [BDS13].
SE	• Existence of a finite-memory SE in two-player quantitative reachability games [BBD10, BBD12]. • Algorithm to decide in ExpSpace the existence of a SE in multiplayer quantitative reachability games [BBDG12, BBDG13].
SPE	• Existence of a SPE in multiplayer quantitative reachability games [BBDG12, BBDG13] *(non-constructive proof)*.
SPSE	• Existence of a SPSE in two-player quantitative reachability games [BBDG12, BBDG13] *(non-constructive proof)*.

[HTW08, Zim09], games whose winning conditions are expressed via quantitative languages [BCHJ09], and more recently, cost-parity and cost-Streett games [FZ12].

Other work concerns *qualitative non-zero-sum* games. In [CHJ04] where the notion of secure equilibrium has been introduced, it is proved that a unique maximal payoff profile of secure equilibria always exists for two-player non-zero-sum games with regular objectives. In [GU08], general criteria ensuring existence of Nash equilibria and subgame perfect equilibria (resp. secure equilibria) are provided for multiplayer (resp. two-player) games, as well as complexity results. The complexity of Nash equilibria in multiplayer concurrent games with Büchi objectives has been discussed in [BBMU11]. [BBM10b] studies the existence of Nash equilibria for timed games with qualitative reachability objectives

Finally, beyond the classical literature about game theory (see for instance [FT91, FL08]), there is a series of recent results on the combination of *non-zero-sum* aspects with *quantitative objectives*. In [BG09], the authors study games played on graphs with terminal vertices where quantitative payoffs are assigned to the players. In [KLŠT12], the authors provide an algorithm to decide the existence of Nash equilibria for concurrent priced games with quantitative reachability objectives. In [PS09], the authors prove existence of a Nash equilibrium in Muller games on finite graphs where players have a preference ordering on the sets of the Muller table. In [FKMY+10] (resp. [PS11]), it is shown that every multiplayer sequential game has a subgame-perfect ε-equilibrium for every $\varepsilon > 0$ if the payoff functions of the players are bounded and lower-semicontinuous (resp. upper-semicontinuous).

Outline. This document is divided into two parts. Part I gives a brief overview of necessary definitions and some results about games played on graphs (Chapter 2), and then states the main contributions of this thesis (Chapter 3). Chapter 2 presents the context of this thesis in order to better understand where our contributions lie.

Part II details the results we obtained about multiplayer cost games, as well as their complete proofs and additional related results. They are grouped according to the kind of equilibrium that is considered: Chapter 4 concerns Nash equilibrium, Chapter 5 concerns secure equilibrium, Chapter 6 concerns subgame perfect equilibrium, and Chapter 7 concerns subgame perfect secure equilibrium.

Finally, Chapter 8 lists some open problems, as well as other possible research directions.

Part I

About General Games Played on Graphs

Chapter 2

Background

This chapter presents the context of this thesis.

The games we study in this document are *games played on (directed) graphs.* Intuitively, a token is moved by the players from vertex to vertex, along the edges. In our case, each vertex of the graph is controlled by one and only one player, meaning that whenever the token reaches a vertex, the player who controls this vertex moves the token along an outgoing edge to a successor vertex (*turn-based game*). A *play* of the game is then a path through the graph constructed by the moves of the token. In some sense, the graph determines the rules of the game and the possible *interactions* between the players. To make the game complete, we need to know what are the *goals* of the players. We can distinguish between *qualitative* and *quantitative* objectives. A qualitative objective for a player means that he wants to guarantee that some property holds. For example, a player *wins* a play if it reaches a certain vertex, otherwise, he *loses* this play. A quantitative objective for a player signifies that he aims at optimising (maximising or minimising) a certain value. A player may, for instance, wish to *minimise* the number of edges until the token reaches a certain vertex.

To summarise, a *game*, in this document, is composed of an *arena* (roughly, a graph) and an *objective* for each player.

Once a game has been fixed, it is natural to wonder how the game will be played. Depending of the kind of the players' rational behaviour [1], we can predict what might be a potential result of the game, which is formalised through a particular *solution concept*. In this document, we will study different solution concepts, according to the sort of games that is considered.

The next sections describe the notion of arena and different kinds of objectives.

2.1 Arena

This section defines what is an arena and gives some related vocabulary words. Notations and definitions are inspired from [GU08, GTW02, BBDG13, BDS13].

Definition 2.1.1. Given a finite set Π of *players*, an (finite) *arena* is a tuple $\mathcal{A}_\Pi = (V, (V_i)_{i\in\Pi}, E)$ where

- $G = (V, E)$ is a (finite) *directed graph* with vertices V and edges $E \subseteq V \times V$ such that for all $v \in V$, there exists $v' \in V$ with $(v, v') \in E$ (i.e., each vertex has at least one outgoing edge), and
- $(V_i)_{i\in\Pi}$ is a partition of V such that V_i is the set of vertices controlled by player i.

When the set of players is clear from the context, we simply write $\mathcal{A}$ for $\mathcal{A}_\Pi$. In the sequel, an arena is always supposed to be finite, unless it is written explicitly that it can be infinite, or when we consider the infinite unravelling of the graph (see Section 4.1.3). Given an arena, the set V of vertices is partitioned between the players, then we only consider *turn-based games*. Usually, we call the players player 1, player 2,. . .

Example 2.1.2. Let us consider the following arena $\mathcal{A} = (V, (V_i)_{i\in\Pi}, E)$ for a set $\Pi = \{1, 2\}$ of two players. The directed graph $G = (V, E)$ has four vertices and six edges: $V = \{A, B, C, D\}$ and $E = \{(A, B), (B, A), (B, C), (C, B), (A, D), (D, B)\}$. The set of vertices controlled by player 1 (resp. player 2) is $V_1 = \{A, C, D\}$ (resp. $V_2 = \{B\}$). This arena is

1. For example, a player may only care about his own objective, or he may want to harm the other players if he does not harm himself by doing so.

depicted in Figure 2.1. The vertices of player 1 (resp. player 2) are represented by circles (resp. squares). We will keep this convention throughout the document.

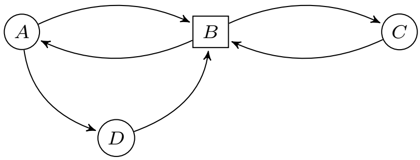


Figure 2.1: An arena.

Intuitively, a game played on an arena $\mathcal{A} = (V, (V_i)_{i\in\Pi}, E)$ proceeds as follows. First, a token is placed on some initial vertex v_0. Whenever a token is on a vertex $v \in V_i$ controlled by player $i \in \Pi$, player i chooses one of the outgoing edges $(v, v') \in E$ and moves the token along this edge to v'. This way, the players together determine a infinite path in $\mathcal{A}$ starting in v_0.

We here give most of the notations that will be used in the sequel of the document.

A finite (resp. infinite) path through the graph $G = (V, E)$ is called a *history* (resp. a *play*). Let ϵ denote the empty history, and Hist (resp. Plays) denote the set of all histories (resp. plays) in $\mathcal{A}$. For $i \in \Pi$, the set Hist_i is the set of all histories $h \in \mathsf{Hist} \setminus \{\epsilon\}$ whose last vertex belongs to V_i. We write $h = h_0 \dots h_k$, where $h_0, \dots, h_k \in V$ $(k \in \mathbb{N})$, for a non-empty history h, and similarly, $\rho = \rho_0\rho_1 \dots$, where $\rho_0, \rho_1, \dots \in V$, for a play ρ.

A *prefix* (resp. *proper prefix*) α of a non-empty history $h = h_0 \dots h_k$ is a finite sequence $h_0 \dots h_l$, with $l \leq k$ (resp. $l < k$), denoted by $\alpha \leq h$ (resp. $\alpha < h$). We analogously consider a prefix α of a play ρ, denoted by $\alpha < \rho$. The *length* $|h|$ of h is the number k of its *edges*[2]. Note that the length is *not* defined as the number of vertices. Given a play $\rho = \rho_0\rho_1 \dots$, we denote by $\rho_{\leq l}$ the prefix of ρ of length l (for some $l \in \mathbb{N}$), i.e. $\rho_{\leq l} = \rho_0\rho_1 \dots \rho_l$. Similarly, $\rho_{<l} = \rho_0\rho_1 \dots \rho_{l-1}$.

2. As a convention, we let $|\epsilon| = -1$.

Given a non-empty history $h = h_0 \dots h_k$ and a vertex v such that $(h_k, v) \in E$, we denote by hv the history $h_0 \dots h_k v$. Moreover, given a history $h = h_0 \dots h_k$ and a play $\rho = \rho_0 \rho_1 \dots$ such that $(h_k, \rho_0) \in E$, we denote by $h\rho$ the play $h_0 \dots h_k \rho_0 \rho_1 \dots$. The function First (resp. Last) returns, for a non-empty history $h = h_0 \dots h_k$, the first vertex h_0 (resp. the last vertex h_k) of h. The function First naturally extends to plays.

We say that a play $\rho = \rho_0 \rho_1 \dots$ *visits* a set $S \subseteq V$ (resp. a vertex $v \in V$) if there exists $l \in \mathbb{N}$ such that ρ_l is in S (resp. $\rho_l = v$). The same terminology also stands for a history h. More precisely, we say that ρ visits a set S at (resp. before) *depth* $d \in \mathbb{N}$ if ρ_d is in S (resp. if there exists $l \leq d$ such that ρ_l is in S).

A *strategy*[3] of player $i \in \Pi$ in $\mathcal{A}$ is a function $\sigma : \mathsf{Hist}_i \to V$ assigning to each non-empty history h that ends in a vertex controlled by player i ($\mathsf{Last}(h) \in V_i$), a successor $v = \sigma(h)$ of $\mathsf{Last}(h)$. That is, $\big(\mathsf{Last}(h), \sigma(h)\big) \in E$. We say that a play $\rho = \rho_0 \rho_1 \dots$ of $\mathcal{A}$ 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 any history h. A *strategy profile* of $\mathcal{A}$ is a tuple $(\sigma_i)_{i \in \Pi}$ of strategies, where σ_i refers to a strategy for player i. Given an initial vertex v, a strategy profile determines a unique play starting in v that is consistent with all strategies σ_i. This play is called the *outcome* of $(\sigma_i)_{i \in \Pi}$ from v, and is denoted by $\langle (\sigma_i)_{i \in \Pi} \rangle_v$. We write σ_{-j} for $(\sigma_i)_{i \in \Pi \setminus \{j\}}$, the set of strategies σ_i for all the players except for player j. We say that a player *deviates* from a strategy (resp. from a play) if he does not carefully follow this strategy (resp. this play). More formally, for a strategy profile $(\sigma_i)_{i \in \Pi}$ with outcome ρ, we say that *player j deviates from ρ* if there exists a strategy σ'_j of player j and a prefix h of ρ consistent with σ'_j, such that $h \in \mathsf{Hist}_j$ and $\sigma'_j(h) \neq \sigma_j(h)$.

A *finite strategy automaton* [Umm05] for player $i \in \Pi$ over an arena $\mathcal{A}$ is a Mealy automaton $\mathcal{M}_i = (M, m_0, V, \delta, \nu)$ where:

- M is a non-empty, finite set of memory states,
- $m_0 \in M$ is the initial memory state,
- $\delta : M \times V \to M$ is the memory update function,
- $\nu : M \times V_i \to V$ is the transition choice function, which satisfies

3. In this document, a strategy is always a pure one.

$(v, \nu(m, v)) \in E$ for all $m \in M$ and $v \in V_i$.

We can extend the memory update function δ to a function $\delta^* : M \times \mathsf{Hist} \to M$ defined by $\delta^*(m, \epsilon) = m$ and $\delta^*(m, hv) = \delta(\delta^*(m, h), v)$ for all $m \in M$ and $hv \in \mathsf{Hist}$. The strategy $\sigma_{\mathcal{M}_i}$ computed by a finite strategy automaton $\mathcal{M}_i$ is defined by $\sigma_{\mathcal{M}_i}(hv) = \nu(\delta^*(m_0, h), v)$ for all $hv \in \mathsf{Hist}$ such that $v \in V_i$. We say that σ is a *finite-memory strategy* if there exists [4] a finite strategy automaton $\mathcal{M}$ such that $\sigma = \sigma_{\mathcal{M}}$. Moreover, we say that $\sigma = \sigma_{\mathcal{M}}$ has a memory of size at most $|M|$, where $|M|$ is the number of states of $\mathcal{M}$. In particular, if $|M| = 1$, we say that σ is a *positional* (or *memoryless*) *strategy* (the current vertex of the play determines the choice of the next vertex, that is, $\sigma : V_i \to V$). We call $(\sigma_i)_{i \in \Pi}$ a strategy profile with memory m if for all $i \in \Pi$, the strategy σ_i has a memory of size at most m. In an equivalent way, σ is a finite-memory strategy if the equivalence relation $\approx_\sigma$ on Hist defined by $h \approx_\sigma h'$ if h, h' end in the same vertex, and $\sigma(h\delta) = \sigma(h'\delta)$ for all histories $h\delta, h'\delta \in \mathsf{Hist}_i$, has finite index.

A strategy profile $(\sigma_i)_{i \in \Pi}$ is called *positional* or *finite-memory* if each σ_i is a positional or a finite-memory strategy, respectively.

Sometimes, it is useful to specify an initial vertex $v_0 \in V$ for an arena $\mathcal{A}$. We then call the pair $(\mathcal{A}, v_0)$ an *initialised arena*. A history (resp. a play) of $(\mathcal{A}, v_0)$ is a history (resp. a play) of $\mathcal{A}$ starting in v_0.

Example 2.1.3. Let us come back to the arena $\mathcal{A}$ of Example 2.1.2 (on page 18). For instance, ABC and $(BC)^\omega$ are respectively a history and a play of this arena. [5] On the contrary, AC (resp. $(BD)^\omega$) is not one of its histories (resp. plays).

The function $\sigma_1 : V_1 \to V$ defined [6] by $\sigma_1(A) = B$ is a positional strategy of player 1. In particular, the plays $(AB)^\omega$ and $(BC)^\omega$ are consistent with σ_1.

Let us now define the following finite-memory strategy of player 2:

---

4. Note that there exist several finite strategy automata such that $\sigma = \sigma_{\mathcal{M}}$.

5. The notation $(BC)^\omega$ corresponds to the infinite word $BCBCBC\ldots$ on the alphabet V. In the sequel, we will keep using classical notations for (sets of) (in)finite words [PP04] in order to write (sets of) paths in a graph.

6. Notice that vertices C and D have only one successor, so player 1 has no choice to make in these vertices.

$\sigma_2(B) = C$ and $\sigma_2(hB) = A$ for all histories hB, with $h \neq \epsilon$. The outcome of the strategy profile (σ_1, σ_2) in the initialised arena $(\mathcal{A}, B)$ is the play $BC(BA)^\omega$.

We can construct a finite strategy automaton $\mathcal{M}_{\sigma_2}$ that computes the strategy σ_2. The set M of memory states is $M = \{m_0, m_1\}$, and the initial state is m_0. The memory update function $\delta : M \times V \to M$ is defined as $\delta(m_0, v) = m_1$ and $\delta(m_1, v) = m_1$ for all $v \in V$, and the transition choice function $\nu : M \times V_2 \to V$ is given by $\nu(m_0, B) = C$ and $\nu(m_1, B) = A$. The automaton $\mathcal{M}_{\sigma_2}$ is depicted in Figure 2.2: a label $-/v'$ on an edge (m, m') means that $\delta(m, v) = m'$ for all $v \in V$, and $\nu(m, B) = v'$.

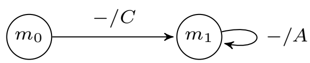


Figure 2.2: The finite strategy automaton $\mathcal{M}_{\sigma_2}$.

Sections 2.2 and 2.3 give some definitions and results for different kinds of games (depending on the objectives of the players). The list of results is not exhaustive but contains those that are useful for the comprehension of the rest of the document. For the interested reader, the references given in each section contain more information.

2.2 Qualitative Objectives

In this section, we focus on games where the players have qualitative objectives. This means that each player wants to guarantee that some property holds along a play. In Section 2.2.1, we study zero-sum games, while in Section 2.2.2, we explore non-zero-sum games.

2.2.1 Zero-sum Games

This section, inspired from [GTW02], deals with *qualitative zero-sum games*. These games are *two-player* games where the two players have *antagonistic objectives*. In other words, one player plays in such a way as to satisfy some property, while the other player plays in order to prevent this to happen. More formally, the objective of player 1 is represented as a set of plays in the arena of the game. A play is then won by player 1 (resp. player 2) if it belongs (resp. does not belong) to this set.

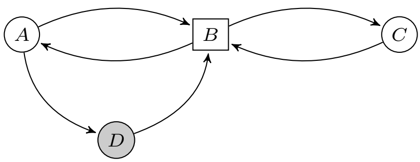
Definition 2.2.1. Given a two-player arena $\mathcal{A} = (V, (V_1, V_2), E)$, a *qualitative zero-sum game* is a pair $\mathcal{G} = (\mathcal{A}, \mathsf{Win})$, where $\mathsf{Win} \subseteq V^\omega$ is the *winning condition* for player 1.

In this definition, we implicitly assume that the set of players is $\{1, 2\}$.

A play ρ is *won* by player 1 if and only if ρ belongs to Win. Player 2 wins ρ if and only if player 1 does not win ρ, meaning that $V^\omega \setminus \mathsf{Win}$ (also denoted by Win^c) is the winning condition for player 2.

Given an initialised arena $(\mathcal{A}, v_0)$, we denote by $(\mathcal{G}, v_0)$ the game $\mathcal{G}$ played on $(\mathcal{A}, v_0)$, and we call it an *initialised game*. The term "initialised game" and the notation $(\mathcal{G}, v_0)$ will be used for any kind of game $\mathcal{G}$ that we study in this document.

Example 2.2.2. Consider the two-player arena $\mathcal{A}$ of Example 2.1.2 (on page 18), which is depicted below. We remind that the vertices of player 1 (resp. player 2) are represented by circles (resp. squares).



Let $\mathcal{G} = (\mathcal{A}, \mathsf{Win})$ be the qualitative zero-sum game played on the arena $\mathcal{A}$, and where the winning condition of player 1 is given by $\mathsf{Win} = \{\rho \in V^\omega \mid \exists i \in \mathbb{N} \;\; \rho_i = D\}$. In this game, player 1 wants a play to visit vertex D at least once (*reachability winning condition*), and player 2 wants to avoid this. For instance, player 1 wins the play $(ADB)^\omega$, whereas player 2 wins the play $(BC)^\omega$.

Given a play ρ, we can define the gain [7] of player 1 for this play as:

$$\mathsf{Gain}_1(\rho) = \begin{cases} 1 & \text{if } \rho \in \mathsf{Win}, \\ -1 & \text{otherwise}, \end{cases} \tag{2.1}$$

and the gain of player 2 as $\mathsf{Gain}_2(\rho) = -\mathsf{Gain}_1(\rho)$. For any play ρ, it holds that $\mathsf{Gain}_1(\rho) + \mathsf{Gain}_2(\rho) = 0$, that is why the game is called *zero-sum*.

Winning strategy

In qualitative zero-sum games, we assume that each player is rational in the sense that he wants to win. As the objectives of the players are completely antagonistic, it is natural to wonder whether a player can play in such a way that he wins for sure, however the other player plays. Such a behaviour is formalised through the solution concept of *winning strategy* defined below.

Definition 2.2.3. Given a qualitative zero-sum game $\mathcal{G}$, a strategy σ of a player in $\mathcal{G}$ is a *winning strategy* for this player from a vertex v_0 if all plays starting in v_0 and consistent with σ are won by this player.

We say that a player *wins* an initialised game $(\mathcal{G}, v_0)$ if he has a winning strategy from v_0. Given a game $\mathcal{G}$, we define the *winning region* for player $i \in \{1, 2\}$, denoted W_i, as the set of all vertices v such that player i wins the game $(\mathcal{G}, v)$. Clearly, since the game is zero-sum, at most one player wins the initialised game, and then $W_1 \cap W_2 = \emptyset$.

Example 2.2.4. Let us come back to the qualitative zero-sum game $\mathcal{G}$ of Example 2.2.2 (on page 23), where player 1 wants to reach vertex D. From vertex B, player 1 has no winning strategy, but player 2 does have a winning strategy: if he always chooses the edge (B, C), then every play consistent with this strategy never visits vertex D. In other words, player 2 wins the game $(\mathcal{G}, B)$. One can easily show that the winning regions for both players are $W_1 = \{A, D\}$ and $W_2 = \{B, C\}$.

7. In the literature, the gain of a player is also known as his *payoff*. We here choose the word "gain" as opposed to the word "cost", introduced later in the document. Using gains (resp. costs) implies that each player aims at maximising his (resp. minimising his).

In this example, one can notice that the winning regions for both players form a partition of V. Whenever a game has this property, we say that it is *determined*.

Definition 2.2.5. A qualitative zero-sum game $\mathcal{G} = (\mathcal{A}, \mathsf{Win})$ is *determined* if the winning regions for player 1 and player 2 form a partition of the set V of vertices of $\mathcal{G}$. That is, $V = W_1 \cup W_2$.

Martin showed [Mar75] that if the winning condition is a Borel subset [8] of V^ω (*Borel winning condition*), then the game is determined.

Theorem 2.2.6 ([Mar75])**.** *Every qualitative zero-sum game with a Borel winning condition is determined.*

The following definitions give some examples of particular Borel winning conditions [Tho95].

Definition 2.2.7. A *reachability winning condition* given by a goal set $\mathsf{R} \subseteq V$ is the set of plays $\rho = \rho_0\rho_1 \dots$ such that there exists $i \in \mathbb{N}$ with $\rho_i \in \mathsf{R}$.

A qualitative zero-sum game with a reachability winning condition is also called a (zero-sum qualitative) *reachability game* and is denoted by $\mathcal{G} = (\mathcal{A}, \mathsf{R})$.

Definition 2.2.8. A *safety winning condition* given by a goal set $\mathsf{S} \subseteq V$ is the set of plays $\rho = \rho_0\rho_1 \dots$ such that for all $i \in \mathbb{N}$, $\rho_i \in \mathsf{S}$.

A qualitative zero-sum game with a safety winning condition is also called a (zero-sum qualitative) *safety game* and is denoted by $\mathcal{G} = (\mathcal{A}, \mathsf{S})$.

Definition 2.2.9. A *reachability under safety winning condition* given by two sets $\mathsf{R}, \mathsf{S} \subseteq V$ is the set of plays $\rho = \rho_0\rho_1 \dots$ such that there exists $i \in \mathbb{N}$ with $\rho_i \in \mathsf{R}$, and for all $i \in \mathbb{N}$, $\rho_i \in \mathsf{S}$.

A qualitative zero-sum game with a reachability under safety winning condition is also called a (zero-sum qualitative) *reachability under safety game* and is denoted by $\mathcal{G} = (\mathcal{A}, \mathsf{R}, \mathsf{S})$.

8. We assume that the reader is familiar with the product topology on V^ω and the notion of Borel sets (see for instance [PP04]).

Reachability games (resp. safety games) are special cases of reachability under safety games where $\mathsf{S} = V$ (resp. $\mathsf{R} = V$).

Definition 2.2.10. A *Büchi winning condition* given by a goal set $\mathsf{F} \subseteq V$ is the set of plays $\rho = \rho_0\rho_1 \dots$ such that for all $i \in \mathbb{N}$, there exists $j \in \mathbb{N}$ with $j \geq i$ and $\rho_j \in \mathsf{F}$.

A qualitative zero-sum game with a Büchi winning condition is also called a (zero-sum qualitative) *Büchi game* and is denoted by $\mathcal{G} = (\mathcal{A}, \mathsf{F})$.

Definition 2.2.11. A *weak parity winning condition* given by a colouring function $c : V \to \mathbb{N}$ is the set of plays $\rho = \rho_0\rho_1 \dots$ such that the minimal colour occurring in the sequence $c(\rho_0)c(\rho_1)\dots$ is even.

A qualitative zero-sum game with a weak parity winning condition is also called a (zero-sum qualitative) *weak parity game* and is denoted by $\mathcal{G} = (\mathcal{A}, c)$.

Remark 2.2.12. The reachability under safety condition can be encoded with a weak parity condition by defining the colouring function c as follows: for any $v \in V$, $c(v) = 3$ if $v \notin \mathsf{S}$, $c(v) = 2$ if $v \in \mathsf{R}$ and $c(v) = 1$ otherwise.

Definition 2.2.13. A *parity winning condition* given by a colouring function $c : V \to \mathbb{N}$ is the set of plays $\rho = \rho_0\rho_1 \dots$ such that the minimal colour occurring infinitely often in the sequence $c(\rho_0)c(\rho_1)\dots$ is even.

A qualitative zero-sum game with a parity winning condition is also called a (zero-sum qualitative) *parity game* and is denoted by $\mathcal{G} = (\mathcal{A}, c)$.

Definition 2.2.14. A *Muller winning condition* given by a family of goal sets $\mathsf{F}_1, \dots, \mathsf{F}_m \subseteq V$ is the set of plays ρ such that the set of vertices appearing infinitely often in ρ is exactly F_k for a certain $k \in \{1, \dots, m\}$.

A qualitative zero-sum game with a Muller winning condition is also called a (zero-sum qualitative) *Muller game* and is denoted by $\mathcal{G} = (\mathcal{A}, \mathsf{F}_1, \dots, \mathsf{F}_m)$.

For all these winning conditions except the last one, the player who wins the game has a *memoryless* winning strategy (*memoryless determinacy*).

Theorem 2.2.15 ([GTW02, Tho08]). *Every qualitative zero-sum game with a reachability/safety/reachability under safety/Büchi/(weak) parity winning condition is determined. Moreover, for every vertex v, one of the two players has a memoryless winning strategy from v.*

For the Muller case however, memory is necessary, but one can show that the player who wins the game has a *finite-memory* winning strategy.

Theorem 2.2.16 ([Tho08]). *Every qualitative zero-sum game with a Muller winning condition is determined. Moreover, for every vertex v, one of the two players has a finite-memory winning strategy from v.*

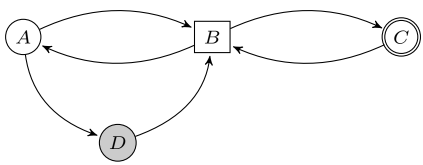
2.2.2 Non-zero-sum Games

This section, based on [GU08], studies a natural extension of zero-sum games: *non-zero-sum games*. In such games, there may be more than two players, each one with his own objective, and several players may win the same play, or no player may win a play (the objectives are not necessarily antagonistic to each other).

Definition 2.2.17. Given a finite set Π of players and an arena $\mathcal{A} = (V, (V_i)_{i\in\Pi}, E)$, a *qualitative non-zero-sum game* is a tuple $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{Win}_i)_{i\in\Pi})$, where $\mathsf{Win}_i \subseteq V^\omega$ is the *winning condition* for player i, for all $i \in \Pi$.

For all $i \in \Pi$, a play ρ is *won* by player i if and only if ρ belongs to Win_i. According to this definition, qualitative zero-sum games are then particular cases of two-player qualitative non-zero-sum games.

Example 2.2.18. Consider the two-player arena $\mathcal{A}$ of Example 2.1.2 (on page 18), which is depicted below.



Let $\mathcal{G} = (\{1, 2\}, \mathcal{A}, (\mathsf{Win}_1, \mathsf{Win}_2))$ be the qualitative non-zero-sum game played on the arena $\mathcal{A}$, and where player 1 has a reachability winning condition for the goal set $\mathsf{R} = \{D\}$ (as in Example 2.2.2), and player 2 has a Büchi winning condition for the goal set $\mathsf{F} = \{C\}$ (he wants a play to visit infinitely often vertex C). In particular, player 1 wins the play $(ADB)^\omega$, player 2 wins the play $(BC)^\omega$, both players win the play $AD(BC)^\omega$, and none of them wins the play $(BA)^\omega$.

In zero-sum games, the *rational behaviour* of a player is to play according to a winning strategy, because the players have conflicting objectives. In non-zero-sum games, it is usually not the case, so we rather look for a "contract" that makes everyone "satisfied" (in the sense that no one wants to break the contract if the others follow it), that is, we search for something like an *equilibrium*.

Nash Equilibrium

A Nash equilibrium is a solution concept that captures the idea that each player is *selfish* (he only cares about his own objective) and *rational* (he prefers winning to losing). Roughly, a Nash equilibrium is a strategy profile in which each player has chosen the *best* strategy for him, taking into account the strategies of the other players. This famous notion was introduced by J. F. Nash [Nas50].

Definition 2.2.19. Given a qualitative non-zero-sum game $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{Win}_i)_{i \in \Pi})$ and an initial vertex $v_0 \in V$, a strategy profile $(\sigma_i)_{i \in \Pi}$ of $\mathcal{G}$ is a *Nash equilibrium* of $(\mathcal{G}, v_0)$ if, for every player $j \in \Pi$ and for every strategy σ'_j of player j, we have that

$$\rho' \in \mathsf{Win}_j \quad \Rightarrow \quad \rho \in \mathsf{Win}_j \, ,$$

where $\rho = \langle (\sigma_i)_{i \in \Pi} \rangle_{v_0}$ and $\rho' = \langle \sigma'_j, \sigma_{-j} \rangle_{v_0}$.

A Nash equilibrium $(\sigma_i)_{i \in \Pi}$ thus ensures that any deviation of a player will not be *better* for him (i.e., a deviation can not make him win if he was losing the play $\langle (\sigma_i)_{i \in \Pi} \rangle_{v_0}$).

To see that from the point of view of gains, we define, like in Equation (2.1), the gain of player i for a play ρ as:

$$\mathsf{Gain}_i(\rho) = \begin{cases} 1 & \text{if } \rho \in \mathsf{Win}_i, \\ -1 & \text{otherwise.} \end{cases} \tag{2.2}$$

The *gain profile* of this play is then $(\mathsf{Gain}_i(\rho))_{i\in\Pi}$ and is denoted by $\mathsf{Gain}(\rho)$. We can now rephrase Definition 2.2.19 by using these gain functions.

Definition 2.2.20. Given a qualitative non-zero-sum game $(\mathcal{G}, v_0)$, a strategy profile $(\sigma_i)_{i\in\Pi}$ of $\mathcal{G}$ is a *Nash equilibrium* of $(\mathcal{G}, v_0)$ if, for every player $j \in \Pi$ and for every strategy σ'_j of player j, we have that

$$\mathsf{Gain}_j(\rho') \leq \mathsf{Gain}_j(\rho)\,,$$

where $\rho = \langle(\sigma_i)_{i\in\Pi}\rangle_{v_0}$ and $\rho' = \langle\sigma'_j, \sigma_{-j}\rangle_{v_0}$.

In a Nash equilibrium, each player is *happy* because none of his deviations can give him a greater gain. A Nash equilibrium can thus be seen as a *contract* between all players: if a player deviates from his contract strategy, then he can not get a strictly greater gain provided that the other players follow the contract.

Definition 2.2.21. Given a strategy profile $(\sigma_i)_{i\in\Pi}$, a strategy σ'_j of player j is called a *profitable deviation* for player j w.r.t. $(\sigma_i)_{i\in\Pi}$ in $(\mathcal{G}, v_0)$ if $\mathsf{Gain}_j(\langle(\sigma_i)_{i\in\Pi}\rangle_{v_0}) < \mathsf{Gain}_j(\langle\sigma'_j, \sigma_{-j}\rangle_{v_0})$.

As a consequence, a strategy profile $(\sigma_i)_{i\in\Pi}$ is a Nash equilibrium if none of the players has a profitable deviation.

Example 2.2.22. Let us come back to the two-player qualitative non-zero-sum game $\mathcal{G}$ described in Example 2.2.18 (on page 27), where player 1 has a reachability winning condition for the goal set $\mathsf{R} = \{D\}$, and player 2 has a Büchi winning condition for the goal set $\mathsf{F} = \{C\}$.

We define the following positional strategies for player 1 and player 2 respectively: $\sigma_1(A) = B$ and $\sigma_2(B) = C$. In $(\mathcal{G}, B)$, the outcome of (σ_1, σ_2) is the play $(BC)^\omega$, which is winning for player 2 and losing for player 1. This strategy profile is in fact a Nash equilibrium in $(\mathcal{G}, B)$ because player 2 wins this play (so he can not make a better choice), and player 1 has no possibility to win while deviating since this play never visits vertex A.

Let us now define two other strategies: a positional strategy σ_1' for player 1 such that $\sigma_1'(A) = D$, and a finite-memory strategy σ_2' for player 2 such that $\sigma_2'(B) = A$ and $\sigma_2'(hB) = C$ for all histories hB, with $h \neq \epsilon$. The strategy profile (σ_1', σ_2') leads to the outcome $BAD(BC)^\omega$ in $(\mathcal{G}, B)$. As this play is won by both players, (σ_1', σ_2') is also a Nash equilibrium in $(\mathcal{G}, B)$.

Remark that if we consider another positional strategy of player 2 defined by $\sigma_2''(B) = A$, then (σ_1', σ_2'') is not a Nash equilibrium in $(\mathcal{G}, B)$. Indeed, σ_2 is a profitable deviation for player 2 w.r.t. (σ_1', σ_2''), since he wins the play $\langle \sigma_1', \sigma_2 \rangle_B = (BC)^\omega$ and loses the play $\langle \sigma_1', \sigma_2'' \rangle_B = (BAD)^\omega$.

In [CMJ04], it has been shown that there exists a Nash equilibrium in every initialised multiplayer game with Borel winning conditions. Note that this result has been generalised in [GU08].

Theorem 2.2.23 ([CMJ04])**.** *Every initialised qualitative non-zero-sum game with Borel winning conditions has a Nash equilibrium.*

Subgame Perfect Equilibrium

Let us first motivate the introduction of a more refined notion of equilibrium than the Nash equilibrium. For this purpose, we consider once again the Nash equilibrium (σ_1, σ_2) of Example 2.2.22. Assume that player 2 changes his mind at the first step and decides to choose the edge (B, A) (instead of (B, C)). Then, according to σ_1 (which belongs to the Nash equilibrium), player 1 chooses the edge (A, B), whereas it would be more rational for him to choose the edge (A, D) in order to win the play. This shows that Nash equilibria do not take into account the sequential nature of games played on graphs: they are not anymore robust as soon as one (or several) player does not play according to the equilibrium for some finite amount of time. The notion of subgame perfect equilibrium avoids these non-rational behaviours.

This notion is a stronger solution concept than the Nash equilibrium. To be a subgame perfect equilibrium, a strategy profile has to be a Nash equilibrium not only from the initial vertex, but also after every possible history of the game, i.e. in every *subgame*. The notion of subgame perfect equilibrium was introduced by R. Selten [Sel65].

Before presenting the formal definition, let us first give some notations. Given a qualitative non-zero-sum game $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{Win}_i)_{i\in\Pi})$ and a history h of $\mathcal{G}$, we denote by $\mathcal{G}|_h$ the game $\mathcal{G}|_h = (\Pi, \mathcal{A}, (\mathsf{Win}_i|_h)_{i\in\Pi})$ where $\mathsf{Win}_i|_h = \{\rho \in V^\omega \mid h\rho \in \mathsf{Win}_i\}$, and we say that $\mathcal{G}|_h$ is a *subgame* of $\mathcal{G}$. For an initialised game $(\mathcal{G}, v_0)$ and a history hv of $(\mathcal{G}, v_0)$ (with $v \in V$), the initialised game $(\mathcal{G}|_h, v)$ is called the *subgame of* $(\mathcal{G}, v_0)$ *with history* hv. Given a strategy σ_i for player i in $\mathcal{G}$, we define the strategy [9] $\sigma_i|_h$ in $(\mathcal{G}|_h, v)$ as $\sigma_i|_h(h') = \sigma_i(hh')$ for all non-empty histories h' of $\mathcal{G}$ such that $\mathsf{First}(h') = v$ and $\mathsf{Last}(h') \in V_i$. We usually write $\sigma_{-j}|_h$ for $(\sigma_i|_h)_{i\in\Pi\setminus\{j\}}$.

Then, we say that $(\sigma_i|_h)_{i\in\Pi}$ is a Nash equilibrium in the subgame $(\mathcal{G}|_h, v)$ if, for every player $j \in \Pi$ and every strategy σ'_j of player j, we have that $\mathsf{Gain}_j(h\langle\sigma'_j|_h, \sigma_{-j}|_h\rangle_v) \leq \mathsf{Gain}_j(h\langle(\sigma_i|_h)_{i\in\Pi}\rangle_v)$. [10]

Definition 2.2.24. Given a qualitative non-zero-sum game $(\mathcal{G}, v_0)$, a strategy profile $(\sigma_i)_{i\in\Pi}$ of $\mathcal{G}$ is a *subgame perfect equilibrium* of $(\mathcal{G}, v_0)$ if $(\sigma_i|_h)_{i\in\Pi}$ is a Nash equilibrium in $(\mathcal{G}|_h, v)$, for every history hv of $(\mathcal{G}, v_0)$, with $v \in V$.

In particular, a subgame perfect equilibrium is also a Nash equilibrium.

Example 2.2.25. Let us come back to the two-player qualitative non-zero-sum game $\mathcal{G}$ described in Example 2.2.18 (on page 27), where player 1 has a reachability winning condition for the goal set $\mathsf{R} = \{D\}$, and player 2 has a Büchi winning condition for the goal set $\mathsf{F} = \{C\}$.

---

9. Notice that $\sigma_i|_h$ is only defined for histories beginning with hv, but it is not a problem since when we consider plays consistent with such a strategy, it always happens in the subgame $(\mathcal{G}|_h, v)$.

10. Recall that the notation $h\langle(\sigma_i|_h)_{i\in\Pi}\rangle_v$ represents the play of $(\mathcal{G}, v_0)$ with prefix h that is consistent with $(\sigma_i|_h)_{i\in\Pi}$ from v.

We showed that the strategy profile (σ_1, σ_2) of Example 2.2.22, where player 1 (resp. player 2) always chooses the edge (A, B) (resp. (B, C)), is a Nash equilibrium in $(\mathcal{G}, B)$. But it is not a subgame perfect equilibrium in this game, because $(\sigma_1|_h, \sigma_2|_h)$ is not a Nash equilibrium in the subgame $(\mathcal{G}|_h, v)$ for $h = B$ and $v = A$. Indeed, the strategy $\sigma_1'|_h$ defined in Example 2.2.22, which always chooses the edge (A, D), is a profitable deviation for player 1 w.r.t. $(\sigma_1|_h, \sigma_2|_h)$ in $(\mathcal{G}|_h, v)$.

One can be convinced that the strategy profile (σ_1', σ_2) is a subgame perfect equilibrium in $(\mathcal{G}, B)$, since after any possible history of the game, the choices made by the players are always rational compared to their objectives.

It has been proved in [Umm06] that there exists a subgame perfect equilibrium in every initialised multiplayer game with Borel winning conditions. Note that this result has been generalised in [GU08].

Theorem 2.2.26 ([Umm06])**.** *Every initialised qualitative non-zero-sum game with Borel winning conditions has a subgame perfect equilibrium.*

Secure Equilibrium

A secure equilibrium is another refinement of the concept of Nash equilibrium. While in a Nash equilibrium, each player only cares about his own gain, in a secure equilibrium, each player cares about his own gain, as well as the other players' gains (but in a negative way). First, each player aims at maximising his own gain, and then, he aims at minimising the other players' gains. This kind of behaviour naturally appears when verifying systems with multiple components, which motivated the definition of secure equilibria.

The concept of secure equilibrium was introduced by K. Chatterjee, T. A. Henzinger and M. Jurdzinski [CHJ04], and adapts the famous Nash equilibrium to the context of assume-guarantee synthesis [CH07, CR10].

We first define the notion of secure equilibria in the two-player case, we will consider the multiplayer case further.

Definition 2.2.27. Given a two-player qualitative non-zero-sum game $\mathcal{G} = (\{1, 2\}, \mathcal{A}, (\mathsf{Win}_1, \mathsf{Win}_2))$ and an initial vertex $v_0 \in V$, a strategy profile (σ_1, σ_2) of $\mathcal{G}$ is called *secure* in $(\mathcal{G}, v_0)$ if, for every strategy σ_1' of player 1, we have that

$$\left(\rho \in \mathsf{Win}_1 \;\Rightarrow\; \rho' \in \mathsf{Win}_1\right) \quad \Rightarrow \quad \left(\rho \in \mathsf{Win}_2 \;\Rightarrow\; \rho' \in \mathsf{Win}_2\right),$$

where $\rho = \langle \sigma_1, \sigma_2 \rangle_{v_0}$ and $\rho' = \langle \sigma_1', \sigma_2 \rangle_{v_0}$; and symmetrically for every strategy σ_2' of player 2.

A secure strategy profile (σ_1, σ_2) thus ensures that any deviation of player 1 that is *not bad* for him (i.e., a deviation that does not make him lose if he was winning the play $\langle \sigma_1, \sigma_2 \rangle_{v_0}$) will not be bad for player 2, and symmetrically for any deviation of player 2.

We can rephrase Definition 2.2.27 by using the gain functions of Equation (2.2).

Definition 2.2.28. Given a two-player qualitative non-zero-sum game $(\mathcal{G}, v_0)$, a strategy profile (σ_1, σ_2) of $\mathcal{G}$ is called *secure* in $(\mathcal{G}, v_0)$ if, for every strategy σ_1' of player 1, we have that

$$\mathsf{Gain}_1(\rho) \leq \mathsf{Gain}_1(\rho') \quad \Rightarrow \quad \mathsf{Gain}_2(\rho) \leq \mathsf{Gain}_2(\rho'),$$

where $\rho = \langle \sigma_1, \sigma_2 \rangle_{v_0}$ and $\rho' = \langle \sigma_1', \sigma_2 \rangle_{v_0}$; and symmetrically for every strategy σ_2' of player 2.

In a secure strategy profile, any deviation of player 1 that does not decrease his gain will not decrease the gain of player 2. A secure profile can thus be seen as a *contract* between the two players which strengthens *cooperation*: if a player chooses another strategy that is not harmful to himself, then this cannot harm the other player if the latter follows the contract.

Definition 2.2.29. Given a two-player qualitative non-zero-sum game $(\mathcal{G}, v_0)$, a strategy profile of $\mathcal{G}$ is a *secure equilibrium* of $(\mathcal{G}, v_0)$ if it is a Nash equilibrium and it is secure in $(\mathcal{G}, v_0)$.

In other words, secure equilibria are those Nash equilibria which reinforce cooperation between the two players.

In [CHJ04], an equivalent characterisation is given for secure equilibria in two-player games. To this end, two binary relations $\prec_1$ and $\prec_2$ are defined on gain profiles:

$$(x_1, x_2) \prec_1 (y_1, y_2) \quad \text{iff} \quad (x_1 < y_1) \vee (x_1 = y_1 \wedge x_2 > y_2)\,; \tag{2.3}$$

symmetrically, $(x_1, x_2) \prec_2 (y_1, y_2)$ iff $(x_2 < y_2) \vee (x_2 = y_2 \wedge x_1 > y_1)$, for any gain profiles (x_1, x_2) and (y_1, y_2). The relation $\prec_1$ means that player 1 prefers a gain profile that gives him a greater gain, and if two gain profiles give him the same gain, then he prefers the gain profiles in which player 2's gain is lower. So, for player 1, we have that $(0, 1) \prec_1 (0, 0) \prec_1 (1, 1) \prec_1 (1, 0)$. Note that the relations $\prec_1$ and $\prec_2$ are transitive.

Definition 2.2.30 ([CHJ04])**.** Given a two-player qualitative non-zero-sum game $(\mathcal{G}, v_0)$, a strategy profile (σ_1, σ_2) of $\mathcal{G}$ is a *secure equilibrium* of $(\mathcal{G}, v_0)$ iff there does not exist a strategy σ_1' of player 1 such that:

$$\mathsf{Gain}(\rho) \prec_1 \mathsf{Gain}(\rho')\,,$$

where $\rho = \langle \sigma_1, \sigma_2 \rangle_{v_0}$ and $\rho' = \langle \sigma_1', \sigma_2 \rangle_{v_0}$; and symmetrically for player 2.

In other words, player 1 (resp. player 2) has no incentive to deviate w.r.t. the relation $\prec_1$ (resp. $\prec_2$). Remember that $\mathsf{Gain}(\rho)$ represents the gain profile of the play ρ.

Example 2.2.31. Let us come back to the two-player qualitative non-zero-sum game $\mathcal{G}$ described in Example 2.2.18 (on page 27), where player 1 has a reachability winning condition for the goal set $\mathsf{R} = \{D\}$, and player 2 has a Büchi winning condition for the goal set $\mathsf{F} = \{C\}$.

We show that the Nash equilibrium (σ_1, σ_2) of Example 2.2.22 (on page 29), where player 1 (resp. player 2) always chooses the edge (A, B) (resp. (B, C)), is also a secure equilibrium in $(\mathcal{G}, B)$. Indeed, this profile is secure as, on one hand, player 1 loses the outcome $(BC)^\omega$, then player 2 can not hope to lower player 1's gain, and on the other hand, the outcome never visits a vertex controlled by player 1, so he has no chance to lower player 2's gain.

As a counterexample, we consider the Nash equilibrium (σ_1', σ_2') of Example 2.2.22, where σ_1' always chooses the edge (A, D) and σ_2' chooses the edge (B, A) if the play has just started in B, otherwise it always chooses the edge (B, C). This strategy profile is not a secure equilibrium in $(\mathcal{G}, B)$ since the strategy σ_2 enables player 2 to get the same gain while decreasing player 1's gain. More precisely, we have that: $\mathsf{Gain}(\langle\sigma'_1, \sigma'_2\rangle_B) = (1,1) \prec_2 (0,1) = \mathsf{Gain}(\langle\sigma'_1, \sigma_2\rangle_B)$, as $\langle\sigma'_1, \sigma'_2\rangle_B = BAD(BC)^\omega$ and $\langle\sigma'_1, \sigma_2\rangle_B = (BC)^\omega$.

In [CHJ04], it is not only shown that there exists a secure equilibrium in two-player qualitative non-zero-sum game with Borel winning conditions, but also a *unique maximal secure equilibrium gain profile.* A maximal secure equilibrium gain profile is a gain profile $(g_i)_{i\in\Pi}$ such that there exists a secure equilibrium with gain profile [11] $(g_i)_{i\in\Pi}$ and for every secure equilibrium with gain profile $(g'_i)_{i\in\Pi}$, it holds that $g'_i \leq g_i$, for all $i \in \Pi$.

Theorem 2.2.32 ([CHJ04])**.** *Every initialised two-player qualitative non-zero-sum game with Borel winning conditions has a unique maximal secure equilibrium gain profile.*

The proof of this result relies on a partition of the set V of vertices of the graph into four sets of vertices from which the players have particular interesting strategies.

Theorem 2.2.32 is generalised in [GU08] for *well-behaved* winning conditions. Let $\mathcal{G} = (\mathcal{A}, (\mathsf{Win}_1, \mathsf{Win}_2))$ be a two-player qualitative non-zero-sum game. We say that the pair $(\mathsf{Win}_1, \mathsf{Win}_2)$ of winning conditions is *determined* if any *two-player zero-sum* game $\mathcal{G}' = (\mathcal{A}, \mathsf{Win})$, such that the winning condition Win is a Boolean combination of Win_1 and Win_2, is determined.

Theorem 2.2.33 ([GU08])**.** *Let $(\mathcal{G}, v_0)$ be a two-player qualitative non-zero-sum game. If the pair $(\mathsf{Win}_1, \mathsf{Win}_2)$ of winning conditions of $\mathcal{G}$ is determined, then there exists a unique maximal secure equilibrium gain profile in $(\mathcal{G}, v_0)$.*

Note that, in fact, it suffices that the two-player zero-sum games $\mathcal{G}' = (\mathcal{A}, \mathsf{Win}_1)$ and $\mathcal{G}'' = (\mathcal{A}, \mathsf{Win}_2^c)$ are determined.

11. The gain profile of a strategy profile is the gain profile of its outcome.

As regards the multiplayer case, two definitions of a secure equilibrium are given in [CHJ04]. It is said there that they are equivalent but they are not. The first one is quite strong: there exists a simple three-player game with no "secure equilibrium" according to this definition (see Example 2.2.34). Let us explain the idea. A strategy profile $(\sigma_i)_{i\in\Pi}$ is said to be "secure" in a game $(\mathcal{G}, v_0)$ if, for every player $j \in \Pi$ and for every strategy σ'_j of player j,

$$\mathsf{Gain}_j(\rho) \leq \mathsf{Gain}_j(\rho') \quad \Rightarrow \quad \big(\forall i \neq j \;\; \mathsf{Gain}_i(\rho) \leq \mathsf{Gain}_i(\rho')\big)\,, \tag{2.4}$$

where $\rho = \langle(\sigma_i)_{i\in\Pi}\rangle_{v_0}$ and $\rho' = \langle\sigma'_j, \sigma_{-j}\rangle_{v_0}$. Then, a strategy profile is a "secure equilibrium" if it is both a Nash equilibrium and secure. It means that if a player deviates from a "secure equilibrium" while keeping the same gain, then *all* the other players must have a greater (or the same) gain than in the equilibrium. We show in the following example that there exists a game with no such equilibrium.

Example 2.2.34. Let us consider the three-player qualitative non-zero-sum game $\mathcal{G} = (\{1,2,3\}, \mathcal{A}, (\mathsf{Win}_1, \mathsf{Win}_2, \mathsf{Win}_3))$, where the arena $\mathcal{A}$ is depicted in Figure 2.3, and player 1 (resp. player 2, player 3) has reachability winning condition for the goal set $\mathsf{R}_1 = \{B, C\}$ (resp. $\mathsf{R}_2 = \{B\}$, $\mathsf{R}_3 = \{C\}$). Only player 1 (who controls the circle vertices) plays in this game.

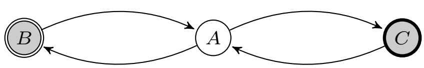


Figure 2.3: Arena of the three-player game $\mathcal{G}$.

Let σ that always chooses the edge (A, B). Then, the outcome is won by players 1 and 2, and lost by player 3. This strategy is not "secure". To see that, let us consider the strategy σ' that always chooses the edge (A, C). Its outcome is won by players 1 and 3, and lost by player 2. So, player 1's gain remains the same, but player 2's gain strictly decreases with the strategy σ' of player 1, which contradicts the definition of a "secure" strategy. With a similar argument, one can show that the strategy σ' is not "secure" either.

Let σ'' be a strategy that chooses the edges (A, B) and (A, C), each one at least once. The outcome of such a strategy is won by the three players, but the strategy that always chooses the edge (A, B) strictly decreases player 3's gain. As a consequence, such a strategy is not "secure", and then, this game has no "secure equilibrium".

Let us now give the second definition of secure equilibria that is proposed in [CHJ04] for multiplayer games. For this purpose, we need to associate a binary relation $\prec_j$ on gain profiles with each player j. Given two gain profiles $(x_i)_{i\in\Pi}$ and $(y_i)_{i\in\Pi}$:

$$(x_i)_{i\in\Pi} \prec_j (y_i)_{i\in\Pi} \quad \text{iff} \quad \big(x_j < y_j\big) \vee \big(x_j = y_j \wedge (\forall i \neq j \; x_i \geq y_i) \wedge (\exists i \neq j \; x_i > y_i)\big). \tag{2.5}$$

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 gain profile to another one either if he has a strictly greater gain, or if he keeps the same gain, the other players have a lower gain, and at least one has a strictly lower gain. For example, we have that $(0, 1, 0) \prec_1 (1, 1, 1) \prec_1 (1, 1, 0) \prec_1 (1, 0, 0)$. One can show that each relation $\prec_j$ is transitive. Moreover, the relation $\prec_j$ given here exactly corresponds to the definitions of $\prec_1$ and $\prec_2$ in the two-player case (see Equation (2.3)).

Definition 2.2.35. Given a multiplayer qualitative non-zero-sum game $(\mathcal{G}, v_0)$, a strategy profile (σ_1, σ_2) of $\mathcal{G}$ is a *secure equilibrium* of $(\mathcal{G}, v_0)$ if for every player $j \in \Pi$, there does not exist any strategy σ'_j of player j such that:

$$\mathsf{Gain}(\rho) \prec_j \mathsf{Gain}(\rho'),$$

where $\rho = \langle (\sigma_i)_{i\in\Pi} \rangle_{v_0}$ and $\rho' = \langle \sigma'_j, \sigma_{-j} \rangle_{v_0}$.

In other words, player j has no incentive to deviate w.r.t. relation $\prec_j$.

Definition 2.2.36. Given a strategy profile $(\sigma_i)_{i\in\Pi}$, a strategy σ'_j of player j is called a *$\prec_j$-profitable deviation* for player j w.r.t. $(\sigma_i)_{i\in\Pi}$ in $(\mathcal{G}, v_0)$ if $\mathsf{Gain}(\langle (\sigma_i)_{i\in\Pi} \rangle_{v_0}) \prec_j \mathsf{Gain}(\langle \sigma'_j, \sigma_{-j} \rangle_{v_0})$.

As a consequence, a strategy profile $(\sigma_i)_{i\in\Pi}$ is a secure equilibrium if no player j has a $\prec_j$-profitable deviation.

Unfortunately, Theorem 2.2.32 (about the existence of a unique maximal secure equilibrium gain profile) does not hold in the multiplayer case. Let us show this with the following example.

Example 2.2.37 ([CHJ04]). Let us come back to the three-player game $\mathcal{G}$ of Example 2.2.34 (on page 36) and consider the strategies σ, σ' and σ'' defined in that example. The strategy σ'' is not a secure equilibrium because σ is a $\prec_1$-profitable deviation for player 1 (since $(1,1,1) \prec_1 (1,1,0)$). On the other hand, one can prove that σ and σ' are secure equilibria. Their gain profiles are $(1,1,0)$ and $(1,0,1)$, which are both maximal, but incomparable. Then, there does not exist a *unique* maximal secure equilibrium gain profile in this game.

To our knowledge, the existence of secure equilibria in *multiplayer* qualitative non-zero-games is still an open problem.

Let us now give an equivalent characterisation of a secure equilibrium. For this, we first give the definition of a secure strategy profile.

Definition 2.2.38. Given a multiplayer qualitative non-zero-sum game $(\mathcal{G}, v_0)$, a strategy profile $(\sigma_i)_{i\in\Pi}$ of $\mathcal{G}$ is called *secure* in $(\mathcal{G}, v_0)$ if, for every player $j \in \Pi$ and every strategy σ'_j of player j, we have that

$$\mathsf{Gain}_j(\rho) \leq \mathsf{Gain}_j(\rho') \quad \Rightarrow \quad \Big(\big(\forall i \neq j \;\; \mathsf{Gain}_i(\rho) \leq \mathsf{Gain}_i(\rho')\big) \\ \vee \;\; \big(\exists i \neq j \;\; \mathsf{Gain}_i(\rho) < \mathsf{Gain}_i(\rho')\big) \Big)\,,$$

where $\rho = \langle(\sigma_i)_{i\in\Pi}\rangle_{v_0}$ and $\rho' = \langle\sigma'_j, \sigma_{-j}\rangle_{v_0}$.

Like in the two-player case, a secure profile ensures that any deviation of a player that does not put him at a disadvantage cannot put the other players at a disadvantage either, if they follow the contract.

Definition 2.2.39. Given a multiplayer qualitative non-zero-sum game $(\mathcal{G}, v_0)$, a strategy profile of $\mathcal{G}$ is a *secure equilibrium* of $(\mathcal{G}, v_0)$ if it is a Nash equilibrium and it is secure in $(\mathcal{G}, v_0)$.

Complexity Results

In this section, we give a non-exhaustive list of some complexity results, which are the most related to our work. We call NE (resp. SPE) the following decision problems.

> Given a qualitative non-zero-sum game with *reachability* winning conditions and some thresholds $(x_i)_{i\in\Pi}, (y_i)_{i\in\Pi} \in \{0,1\}^{|\Pi|}$, decide whether the game has a Nash (resp. subgame perfect) equilibrium with gain profile $(g_i)_{i\in\Pi}$ such that $x_i \leq g_i \leq y_i$, for all $i \in \Pi$.[12]

By adapting the proof of [Umm08, Theorem 8], one can show that this problem is in NP for Nash equilibria.

Theorem 2.2.40 ([Umm08])**.** *NE is in* NP*.*

Theorem 2.2.41 ([CMJ04, Umm05])**.** *NE (and SPE) is* NP-hard*, even with the threshold* $(y_i)_{i\in\Pi} = (1,\ldots,1)$*.*

Theorem 2.2.42 ([GU08])**.** *The problem of deciding whether, in an initialised two-player qualitative game with* parity *winning conditions, there exists a secure equilibrium with gain profile* $(0,0)$ *(resp.* $(1,0)$ *or* $(0,1)$*, resp.* $(1,1)$*) is in* UP ∩ co-UP *(resp. is* co-NP-complete*, resp. is in* NP*). If the number of priorities is bounded, these four problems are in* P*.*

2.3 Quantitative Objectives

In this section, we focus on games where the players have quantitative objectives. This means that each player aims at optimising (maximising or minimising) a certain value along a play. Section 2.3.1 is about Min-Max cost games, while Section 2.3.2 refers to multiplayer cost games.

2.3.1 Min-Max Cost Games

This section, mainly inspired from [Tri09], considers *Min-Max cost games*. These games generalise qualitative zero-sum games to the quantitative framework. In Min-Max cost games, there are two players: player Min, who wants to *minimise* the cost he pays for plays, and player Max, who wants to *maximise* the gain he gets for plays.

12. Note that, thanks to Theorem 2.2.26, deciding the existence of a subgame perfect equilibrium in an initialised qualitative non-zero-sum game with Borel winning conditions is trivial, since the answer is always ’yes’.

Definition 2.3.1. Given a two-player arena $\mathcal{A} = (V, (V_{\text{Min}}, V_{\text{Max}}), E)$, a *Min-Max cost game* is a triple $\mathcal{G} = (\mathcal{A}, \mathsf{Cost}_{\text{Min}}, \mathsf{Gain}_{\text{Max}})$, where $\mathsf{Cost}_{\text{Min}} : \mathsf{Plays} \to \mathbb{R} \cup \{+\infty, -\infty\}$ is the *cost function* of player Min, and $\mathsf{Gain}_{\text{Max}} : \mathsf{Plays} \to \mathbb{R} \cup \{+\infty, -\infty\}$ is the *gain function* of player Max, such that $\mathsf{Cost}_{\text{Min}}(\rho) \geq \mathsf{Gain}_{\text{Max}}(\rho)$ for all $\rho \in \mathsf{Plays}$.

In this definition, we implicitly assume that the set of players is $\{\text{Min}, \text{Max}\}$.

For every play ρ, the value $\mathsf{Cost}_{\text{Min}}(\rho)$ represents the amount that player Min *loses* for this play, and $\mathsf{Gain}_{\text{Max}}(\rho)$ represents the amount that player Max *wins* for this play. In such a game, player Min wants to *minimise* his cost, while player Max wants to *maximise* his gain.

Let us stress that, according to this definition, a Min-Max cost game is *zero-sum* if $\mathsf{Cost}_{\text{Min}} = \mathsf{Gain}_{\text{Max}}$, but this might not always be the case[13]. We also point out that Definition 2.3.1 allows to take quite unrelated functions $\mathsf{Cost}_{\text{Min}}$ and $\mathsf{Gain}_{\text{Max}}$, but usually they are similar (see Definition 2.3.3). In the sequel, we denote by Σ_{Min} (resp. Σ_{Max}) the set of strategies of player Min (resp. Max) in a Min-Max cost game.

Example 2.3.2. We describe the following Min-Max cost game $\mathcal{G} = (\mathcal{A}, \mathsf{Cost}_{\text{Min}}, \mathsf{Gain}_{\text{Max}})$. Its arena $\mathcal{A}$ is the one of Example 2.1.2 (on page 18), but we enrich the graph with prices on the edges in order to define the cost function $\mathsf{Cost}_{\text{Min}}$. The price function $\phi : E \to \{1, 2, 3\}$, which assigns a price to each edge of the graph, is as follows: $\phi(A, B) = \phi(B, A) = \phi(B, C) = 1$, $\phi(A, D) = 2$ and $\phi(C, B) = \phi(D, B) = 3$ (see Figure 2.4). We remind that vertices controlled by player Min (resp. player Max) are represented by circles (resp. squares).

The cost function $\mathsf{Cost}_{\text{Min}}$ of player Min is defined as follows:

$$\mathsf{Cost}_{\text{Min}}(\rho) = \begin{cases} \sum_{i=1}^{n} \phi(\rho_{i-1}, \rho_i) & \text{if } n \text{ is the } \textit{least} \text{ index s.t. } \rho_n = C, \\ +\infty & \text{otherwise,} \end{cases}$$

for every play $\rho = \rho_0 \rho_1 \ldots$ of $\mathcal{G}$. This cost function expresses a *quantitative reachability objective*: player Min wants to reach vertex C (shaded vertex) while minimising the sum of prices up to this vertex. For player

13. For an example, see the average-price game in Remark 2.3.4.

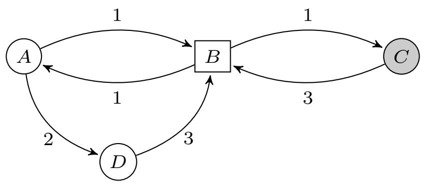


Figure 2.4: Arena $\mathcal{A}$ with prices on edges.

Max, we set $\mathsf{Gain}_{\mathrm{Max}} = \mathsf{Cost}_{\mathrm{Min}}$, which means that this player wants to avoid vertex C or maximise the sum of the prices until reaching it (*quantitative safety objective*).

For example, the play $(AB)^\omega$ of this game leads to an infinite cost/gain for player Min/Max, whereas the play $AD(BC)^\omega$ induces a cost/gain of 6 for player Min/Max.

We here give four well-known kinds of Min-Max cost games. For each sort of game, the cost and gain functions are defined from a price function $\phi : E \to \mathbb{R}$ (and a reward function $\vartheta : E \to \mathbb{R}$ in the last case), which labels the edges of the arena with prices (and rewards). For a play ρ, we use the following notations: $\phi(\rho_{\leq n}) = \sum_{i=1}^{n} \phi(\rho_{i-1}, \rho_i)$ and $\vartheta(\rho_{\leq n}) = \sum_{i=1}^{n} \vartheta(\rho_{i-1}, \rho_i)$.

Definition 2.3.3 ([Tri09])**.** Given an arena $\mathcal{A} = (V, (V_{\mathrm{Min}}, V_{\mathrm{Max}}), E)$, a price function $\phi : E \to \mathbb{R}$ that assigns a price to each edge, a diverging [14] reward function $\vartheta : E \to \mathbb{R}$ that assigns a reward to each edge, and a play $\rho = \rho_0\rho_1 \ldots$ in $\mathcal{A}$, we define the following Min-Max cost games:

1. a *reachability-price game* is a Min-Max cost game $\mathcal{G} = (\mathcal{A}, \mathrm{RP}_{\mathrm{Min}}, \mathrm{RP}_{\mathrm{Max}})$ together with a given goal set $\mathsf{R} \subseteq V$, where

$$\mathrm{RP}_{\mathrm{Min}}(\rho) = \begin{cases} \phi(\rho_{\leq n}) & \text{if } n \text{ is the \textit{least} index s.t. } \rho_n \in \mathsf{R}, \\ +\infty & \text{otherwise,} \end{cases}$$

and $\mathrm{RP}_{\mathrm{Min}} = \mathrm{RP}_{\mathrm{Max}}$;

14. For all plays $\rho = \rho_0\rho_1 \ldots$ in $\mathcal{A}$, it holds that $\lim_{n\to\infty} |\sum_{i=1}^{n} \vartheta(\rho_{i-1}, \rho_i)| = +\infty$. This is equivalent to requiring that every cycle has a positive sum of rewards.

2. a *discounted-price game* is a Min-Max cost game $\mathcal{G} = (\mathcal{A}, \mathrm{DP}_{\mathrm{Min}}(\lambda), \mathrm{DP}_{\mathrm{Max}}(\lambda))$ together with a given discount factor $\lambda \in]0,1[$, where

$$\mathrm{DP}_{\mathrm{Min}}(\lambda)(\rho) = (1-\lambda) \cdot \sum_{i=1}^{+\infty} \lambda^{i-1} \phi(\rho_{i-1}, \rho_i)\,,$$

and $\mathrm{DP}_{\mathrm{Min}}(\lambda) = \mathrm{DP}_{\mathrm{Max}}(\lambda)$;

3. an *average-price game*[15] is a Min-Max cost game $\mathcal{G} = (\mathcal{A}, \mathrm{AP}_{\mathrm{Min}}, \mathrm{AP}_{\mathrm{Max}})$, where

$$\mathrm{AP}_{\mathrm{Min}}(\rho) = \limsup_{n\to+\infty} \frac{\phi(\rho_{\leq n})}{n}\,, \qquad \text{and} \quad \mathrm{AP}_{\mathrm{Max}}(\rho) = \liminf_{n\to+\infty} \frac{\phi(\rho_{\leq n})}{n}\,;$$

4. a *price-per-reward-average game* is a Min-Max cost game $\mathcal{G} = (\mathcal{A}, \mathrm{PRAvg}_{\mathrm{Min}}, \mathrm{PRAvg}_{\mathrm{Max}})$, where

$$\mathrm{PRAvg}_{\mathrm{Min}}(\rho) = \limsup_{n\to+\infty} \frac{\phi(\rho_{\leq n})}{\vartheta(\rho_{\leq n})}\,, \qquad \text{and} \quad \mathrm{PRAvg}_{\mathrm{Max}}(\rho) = \liminf_{n\to+\infty} \frac{\phi(\rho_{\leq n})}{\vartheta(\rho_{\leq n})}\,.$$

An average-price game is then a particular case of a price-per-reward-average game. Let us observe that the Min-Max cost game $\mathcal{G}$ of Example 2.3.2 is a reachability-price game for the goal set $\mathsf{R} = \{C\}$.

Remark 2.3.4. Reachability-price and discounted-price games are zero-sum games, whereas average-price and price-per-reward-average games are not. For example, let us consider the average-price game $\mathcal{G} = (\mathcal{A}, \mathrm{AP}_{\mathrm{Min}}, \mathrm{AP}_{\mathrm{Max}})$ depicted in Figure 2.5. The vertices of the arena are A and B, and are both controlled by player Min. The number 0 or 1 associated to each edge corresponds with the price of this edge ($\phi(A,B) = \phi(B,B) = 1$ and the price of the other edges is zero).

Let ρ be the play $ABAB^2A^2B^4A^4 \dots B^{2^n}A^{2^n} \dots$, where A^i means the concatenation of i A. Then, the sequence of prices appearing along ρ is

15. When the cost function of a player is $\mathrm{AP}_{\mathrm{Min}}$, we say that he has a *mean-payoff* objective.

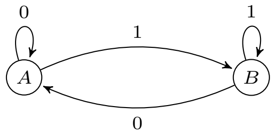


Figure 2.5: Average-price game $\mathcal{G}$.

$101^2 0^2 1^4 0^4 \dots 1^{2^n} 0^{2^n} \dots$, and so we get: $\mathrm{AP}_{\mathrm{Min}}(\rho) = \frac{2}{3}$ and $\mathrm{AP}_{\mathrm{Max}}(\rho) = \frac{1}{2}$. As these costs are not equal, the average-price game $\mathcal{G}$ depicted in Figure 2.5 is not a zero-sum game. As a consequence, we can conclude that average-price and price-per-reward-average games are non-zero-sum games.

Value and optimal strategies

Let us remind that in qualitative zero-sum games, we wonder if a player can *ensure* to win, no matter how the other player plays. In Min-Max cost games, we look for some value that a player can *guarantee* to have, however the other player plays.

Given a Min-Max cost game and a vertex v of the graph, we call the *upper value* the lowest cost that player Min can guarantee to pay from v, and the *lower value* the greatest gain that player Max can guarantee to get from v.

Definition 2.3.5. Given a Min-Max cost game $\mathcal{G}$, we define, for every vertex $v \in V$, the *upper value* $\mathsf{Val}^*(v)$ as:

$$\mathsf{Val}^*(v) = \inf_{\sigma_1 \in \Sigma_{\mathrm{Min}}} \sup_{\sigma_2 \in \Sigma_{\mathrm{Max}}} \mathsf{Cost}_{\mathrm{Min}}(\langle \sigma_1, \sigma_2 \rangle_v)\,,$$

and the *lower value* $\mathsf{Val}_*(v)$ as:

$$\mathsf{Val}_*(v) = \sup_{\sigma_2 \in \Sigma_{\mathrm{Max}}} \inf_{\sigma_1 \in \Sigma_{\mathrm{Min}}} \mathsf{Gain}_{\mathrm{Max}}(\langle \sigma_1, \sigma_2 \rangle_v)\,.$$

Note that for every vertex v, it always holds that $\mathsf{Val}^*(v) \geq \mathsf{Val}_*(v)$. When these values are equal for all v, we say that the game is *determined*.

Definition 2.3.6. A Min-Max cost game $\mathcal{G}$ is *determined* if, for every $v \in V$, we have $\mathsf{Val}^*(v) = \mathsf{Val}_*(v)$. We also say that the game $\mathcal{G}$ has a *value* from v, and we write $\mathsf{Val}(v) = \mathsf{Val}^*(v) = \mathsf{Val}_*(v)$.

In a game $\mathcal{G}$, an *ε-optimal strategy* for player Min ensures that for all v, his cost will not exceed $\mathsf{Val}^*(v) + \varepsilon$ in $(\mathcal{G}, v)$, against any strategy of player Max. Similarly, an ε-optimal strategy for player Max ensures that for all v, his gain will not fall below $\mathsf{Val}_*(v) - \varepsilon$ in $(\mathcal{G}, v)$, against any strategy of player Min.

Definition 2.3.7. Given a Min-Max cost game $\mathcal{G}$ and $\varepsilon \geq 0$, we say that $\sigma_1^\star \in \Sigma_{\mathrm{Min}}$ is an *ε-optimal strategy* for player Min if, for every $v \in V$, we have that

$$\sup_{\sigma_2 \in \Sigma_{\mathrm{Max}}} \mathsf{Cost}_{\mathrm{Min}}(\langle \sigma_1^\star, \sigma_2 \rangle_v) \leq \mathsf{Val}^*(v) + \varepsilon\,.$$

Similarly, $\sigma_2^\star \in \Sigma_{\mathrm{Max}}$ is an *ε-optimal strategy* for player Max if, for every $v \in V$, we have that

$$\inf_{\sigma_1 \in \Sigma_{\mathrm{Min}}} \mathsf{Gain}_{\mathrm{Max}}(\langle \sigma_1, \sigma_2^\star \rangle_v) \geq \mathsf{Val}_*(v) - \varepsilon\,.$$

In particular, if the game is determined, $\sigma_1^\star \in \Sigma_{\mathrm{Min}}$ and $\sigma_2^\star \in \Sigma_{\mathrm{Max}}$ are 0-optimal strategies, also called *optimal strategies*, for the respective players if, for every $v \in V$, we have that

$$\inf_{\sigma_1 \in \Sigma_{\mathrm{Min}}} \mathsf{Gain}_{\mathrm{Max}}(\langle \sigma_1, \sigma_2^\star \rangle_v) = \mathsf{Val}(v) = \sup_{\sigma_2 \in \Sigma_{\mathrm{Max}}} \mathsf{Cost}_{\mathrm{Min}}(\langle \sigma_1^\star, \sigma_2 \rangle_v)\,.$$

In other words, if player Min plays according to an optimal strategy from v, then he loses at most $\mathsf{Val}(v)$. On the other hand, if player Max plays according to an optimal strategy from v, then he wins at least $\mathsf{Val}(v)$.

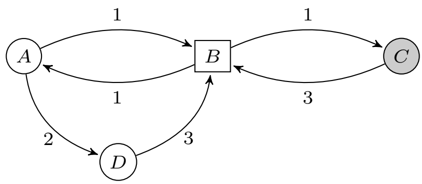
Remark 2.3.8. Given a game $\mathcal{G}$ and $\varepsilon \geq 0$, if for all $v \in V$, there exists a strategy σ_1^v for player Min such that

$$\sup_{\sigma_2 \in \Sigma_{\mathrm{Max}}} \mathsf{Cost}_{\mathrm{Min}}(\langle \sigma_1^v, \sigma_2 \rangle_v) \leq \mathsf{Val}^*(v) + \varepsilon\,, \tag{2.6}$$

then player Min has an ε-optimal strategy $\sigma_1^\star$ in $\mathcal{G}$. Indeed, we can define $\sigma_1^\star(h) := \sigma_1^v(h)$ such that $v = \mathsf{First}(h)$. Then, any play starting in v and consistent with $\sigma_1^\star$ will be consistent with σ_1^v and satisfy Equation (2.6).

Also note that the converse is true. A similar remark can be done for ε-optimal strategies of player Max.

Example 2.3.9. Let us come back to the zero-sum Min-Max cost game $\mathcal{G}$ described in Example 2.3.2 (on page 40), whose arena is depicted below, and where player Min has a quantitative reachability objective for the goal set $\mathsf{R} = \{C\}$.



Let us show that this game is determined. From vertices A, B and D, the lowest cost (resp. greatest gain) that player Min (resp. player Max) can guarantee to have is $+\infty$. Then, $\mathsf{Val}(A) = \mathsf{Val}(B) = \mathsf{Val}(D) = +\infty$. From vertex C, no matter how both players play, their cost/gain is 0, and so, $\mathsf{Val}(C) = 0$.

Let us define the following two positional strategies $\sigma_1^\star$ and $\sigma_2^\star$ for player Min and player Max respectively: $\sigma_1^\star(A) = B$ and $\sigma_2^\star(B) = A$. These strategies are optimal for the respective players: from every vertex v of the game, $\sigma_1^\star$ ensures that player Min's cost is at most $\mathsf{Val}(v)$, and $\sigma_2^\star$ ensures that player Max's gain is at least $\mathsf{Val}(v)$.

The following theorem is a well-known result about the particular cost games described in Definition 2.3.3.

Theorem 2.3.10 ([FV97, Tri09])**.** *Reachability-price games, discounted-price games, average-price games, and price-per-reward games are determined and have* positional *optimal strategies for both players.*

The following result is more general, and states the determinacy of zero-sum Min-Max cost games with bounded, Borel measurable cost functions.

Theorem 2.3.11 ([Mer86])**.** *Every zero-sum Min-Max cost game, where* Cost_{Min} *is a bounded, Borel measurable function, is determined, and both players have* ε*-optimal strategies, for any* $\varepsilon > 0$*.*

This is a classical and well-known result, but we have not found any detailed proof in the literature, so we give one here.

Proof. Let $\mathcal{G} = (\mathcal{A}, \mathsf{Cost}_{\text{Min}}, \mathsf{Gain}_{\text{Max}})$ be a zero-sum Min-Max cost game where $\mathsf{Cost}_{\text{Min}}$ is a bounded, Borel measurable function (and $\mathsf{Gain}_{\text{Max}} = \mathsf{Cost}_{\text{Min}}$). For any real number $r \in \mathbb{R}$, we define the winning condition $\mathsf{Win}_r = \{\rho \in \mathsf{Plays} \mid \mathsf{Cost}_{\text{Min}}(\rho) \leq r\}$, and we consider the qualitative zero-sum game $\mathcal{G}_r = (\mathcal{A}, \mathsf{Win}_r)$. Since $\mathsf{Cost}_{\text{Min}}$ is Borel measurable, $\mathsf{Win}_r = (\mathsf{Cost}_{\text{Min}})^{-1}(]-\infty, r])$ is a Borel set, and so we can apply Martin's result (see Theorem 2.2.6) to the game $\mathcal{G}_r$. Then, we know that one of the players can win from every vertex v. If we see the game as a Min-Max cost game, it implies that it is determined, and $\mathsf{Val}(v) = 0$ or 1 for every vertex v.

Let us fix $\varepsilon > 0$ and some vertex $v \in V$. Let r^v be the infimum of all r such that player 1 (player Min in $\mathcal{G}$) wins the game $\mathcal{G}_r = (\mathcal{A}, \mathsf{Win}_r)$ from v. Notice that r^v is finite as $\mathsf{Cost}_{\text{Min}}$ is bounded, and consequently, there exists $c \in \mathbb{R}$ such that $\mathsf{Win}_r = \mathsf{Plays}$ for all $r \geq c$, and $\mathsf{Win}_r = \emptyset$ for all $r \leq -c$. It is also easy to see that for all $r_1 \leq r_2$, if player 1 wins $(\mathcal{G}_{r_1}, v)$, then he also wins $(\mathcal{G}_{r_2}, v)$ (as $\mathsf{Win}_{r_1} \subseteq \mathsf{Win}_{r_2}$), and by contrapositive and determinacy, if player 2 wins $(\mathcal{G}_{r_2}, v)$, then he also wins $(\mathcal{G}_{r_1}, v)$. Let us show that $\mathsf{Val}(v) = r^v$.

By definition of r^v, player 1 (player Min in $\mathcal{G}$) has a winning strategy σ_1^v from v in the game $\mathcal{G}_{r^v+\varepsilon} = (\mathcal{A}, \mathsf{Win}_{r^v+\varepsilon})$. Thus, it holds that $\sup_{\sigma_2 \in \Sigma_{\text{Max}}} \mathsf{Cost}_{\text{Min}}(\langle \sigma_1^v, \sigma_2 \rangle_v) \leq r^v + \varepsilon$, meaning that player Min can guarantee that his cost is at most $r^v + \varepsilon$.

Similarly, we have that $r^v - \varepsilon < r^v$, and so, player 2 (player Max in $\mathcal{G}$) has a winning strategy σ_2^v from v in the game $\mathcal{G}_{r^v-\varepsilon} = (\mathcal{A}, \mathsf{Win}_{r^v-\varepsilon})$. Then we have that $\inf_{\sigma_1 \in \Sigma_{\text{Min}}} \mathsf{Cost}_{\text{Min}}(\langle \sigma_1, \sigma_2^v \rangle_v) \geq r^v - \varepsilon$, meaning that player Max can guarantee that his gain is at least $r^v - \varepsilon$ (recall that $\mathsf{Gain}_{\text{Max}} = \mathsf{Cost}_{\text{Min}}$).

As a consequence, for all $\varepsilon > 0$, $\mathsf{Val}^*(v) \leq r^v + \varepsilon$ and $\mathsf{Val}_*(v) \geq r^v - \varepsilon$, which implies that $\mathsf{Val}^*(v) = \mathsf{Val}_*(v) = \mathsf{Val}(v) = r^v$. By Remark 2.3.8, player Min and player Max have ε-optimal strategies in $\mathcal{G}$ for all $\varepsilon > 0$, and $\mathsf{Val}(v) = r^v$ for all $v \in V$. □

Let us show with the following example that there exist zero-sum Min-Max cost games where the players have ε-optimal strategies for all $\varepsilon > 0$, but they do not have optimal strategies.

Example 2.3.12. Let $\mathcal{G}$ be the zero-sum Min-Max cost game whose arena is depicted in Figure 2.6, and where the cost function $\mathsf{Cost}_{\mathrm{Min}}$ is defined by $\mathsf{Cost}_{\mathrm{Min}}(A^n B^\omega) = -1 + \frac{1}{n}$ for $n \in \mathbb{N}_0$ and $\mathsf{Cost}_{\mathrm{Min}}(A^\omega) = 0$. One can be convinced that $\mathsf{Val}(A) = -1$. For all $\varepsilon > 0$, player Min has an ε-optimal strategy to ensure to pay at most $-1+\varepsilon$ from A, but he has no optimal strategy from A to guarantee that his cost will not exceed -1.

Figure 2.6: A Min-Max cost game without optimal strategies.

Remark 2.3.13. In a zero-sum Min-Max cost game where $\mathsf{Cost}_{\mathrm{Min}}$ is bounded, Borel measurable, and also *continuous*[16], then both players have *optimal* strategies.

2.3.2 Multiplayer Cost Games

This section, mainly inspired from [BDS13], introduces the games on which we will focus in the sequel of this document: *multiplayer cost games*. These games generalise the other kinds of games we have defined before. They are played by any number of players, and each player wants to minimise the cost he pays for plays. In particular, they are *quantitative non-zero-sum games*.

Definition 2.3.14. Given a finite set Π of players and an arena $\mathcal{A} = (V, (V_i)_{i\in\Pi}, E)$, a *multiplayer cost game* is a tuple $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{Cost}_i)_{i\in\Pi})$, where $\mathsf{Cost}_i : \mathsf{Plays} \to \mathbb{R} \cup \{+\infty, -\infty\}$ is the *cost function* of player i, for all $i \in \Pi$.

16. Notice that the function $\mathsf{Cost}_{\mathrm{Min}}$ of Example 2.3.12 is not continuous.

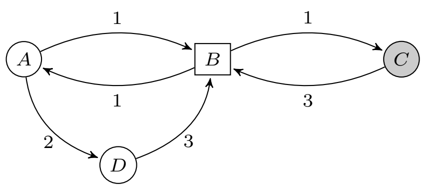
For every play ρ of the game, the value $\mathsf{Cost}_i(\rho)$ represents the amount that player i *loses* for this play. The *cost profile* of this play is then $(\mathsf{Cost}_i(\rho))_{i\in\Pi}$ and is denoted by $\mathsf{Cost}(\rho)$.

For the sake of simplicity, we assume that each player has a cost function that he wants to minimise. But note that minimising cost or maximising gain are essentially equivalent, as maximising the gain for player i can be modelled by using Cost_i to be minus this gain and then minimising the cost. Then, a Min-Max cost game is a particular case of a two-player cost game.

Notice that the cost functions of a cost game are defined on the set Plays, and not on V^ω, as they often depend on prices put on the edges of the game graph.

Let us insist on the fact that the players of a multiplayer cost game may have completely different cost functions, as in the following example.

Example 2.3.15. Let $\mathcal{G} = (\{1,2\}, \mathcal{A}, (\mathsf{Cost}_1, \mathsf{Cost}_2))$ be the two-player cost game whose arena $\mathcal{A}$ is depicted below (see Example 2.3.2 on page 40 for more details), and such that player 1 has the same cost function as player Min in that example, that is, $\mathsf{Cost}_1 = \mathrm{RP}_{\mathrm{Min}}$ for the goal set $\mathsf{R} = \{C\}$. As for player 2, we set $\mathsf{Cost}_2 = \mathrm{AP}_{\mathrm{Min}}$, for the same[17] price function ϕ defined in that example. Then, for player 2, his cost of a play is the long-run average of the prices that appear along this play. Recall that the cost functions $\mathrm{RP}_{\mathrm{Min}}$ and $\mathrm{AP}_{\mathrm{Min}}$ are specified in Definition 2.3.3.



An example of a play in $\mathcal{G}$ can be given by $\rho = (AB)^\omega$, leading to the costs $\mathsf{Cost}_1(\rho) = +\infty$ and $\mathsf{Cost}_2(\rho) = 1$. In the same way, the play $\rho' = A(BC)^\omega$ induces the following costs: $\mathsf{Cost}_1(\rho') = 2$ and $\mathsf{Cost}_2(\rho') = 2$.

17. Note that we could have defined a different price function for each player. In this case, the edges of the graph would have been labelled by couples of numbers.

Nash Equilibrium

In this section, we define the concept of *Nash equilibrium* in the context of cost games.

Definition 2.3.16. Given a multiplayer cost game $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{Cost}_i)_{i\in\Pi})$ and an initial vertex $v_0 \in V$, a strategy profile $(\sigma_i)_{i\in\Pi}$ of $\mathcal{G}$ is a *Nash equilibrium* of $(\mathcal{G}, v_0)$ if, for every player $j \in \Pi$ and for every strategy σ'_j of player j, we have that:

$$\mathsf{Cost}_j(\rho) \leq \mathsf{Cost}_j(\rho') \, ,$$

where $\rho = \langle(\sigma_i)_{i\in\Pi}\rangle_{v_0}$ and $\rho' = \langle\sigma'_j, (\sigma_i)_{i\in\Pi\setminus\{j\}}\rangle_{v_0}$.

Notice that this definition exactly corresponds to Definition 2.2.20 in the qualitative framework, except that the inequality is here reversed as we consider costs to minimise instead of gains to maximise.

Definition 2.3.17. Given a strategy profile $(\sigma_i)_{i\in\Pi}$, a strategy σ'_j of player j is called a *profitable deviation* for player j w.r.t. $(\sigma_i)_{i\in\Pi}$ in $(\mathcal{G}, v_0)$ if $\mathsf{Cost}_j(\langle(\sigma_i)_{i\in\Pi}\rangle_{v_0}) > \mathsf{Cost}_j(\langle\sigma'_j, \sigma_{-j}\rangle_{v_0})$.

As in the qualitative case, a strategy profile $(\sigma_i)_{i\in\Pi}$ is a Nash equilibrium if none of the players has a profitable deviation.

Example 2.3.18. Let us come back to the two-player cost game $\mathcal{G}$ described in Example 2.3.15 (on page 48), where $\mathsf{Cost}_1 = \mathrm{RP}_{\mathrm{Min}}$ for the goal set $\mathsf{R} = \{C\}$ (quantitative reachability objective), and $\mathsf{Cost}_2 = \mathrm{AP}_{\mathrm{Min}}$ (mean-payoff objective). We consider the positional strategies σ_1 and σ_2 of player 1 and player 2 respectively, defined by $\sigma_1(A) = B$ and $\sigma_2(B) = A$. If we fix the initial vertex A, the outcome of the strategy profile (σ_1, σ_2) is the play $\rho = (AB)^\omega$, with cost profile $(+\infty, 1)$. One can show that (σ_1, σ_2) is in fact a Nash equilibrium in $(\mathcal{G}, A)$: player 2 gets the least cost he can expect in this game, and player 1 has no incentive to choose the edge (A, D) (it does not allow the play to pass through vertex C).

We now consider the positional strategy profile (σ'_1, σ'_2) with $\sigma'_1(A) = B$ and $\sigma'_2(B) = C$. Its outcome is the play $\rho' = A(BC)^\omega$. However, this strategy profile is not a Nash equilibrium in $(\mathcal{G}, A)$, because player 2 can strictly lower his cost by always choosing the edge (B, A) instead of (B, C), thus lowering his cost from 2 to 1. In other words, the strategy σ_2 is a profitable deviation for player 2 w.r.t. (σ_1', σ_2').

It is important to notice that there exist simple cost games with no Nash equilibrium, as it is shown in the following example.

Example 2.3.19. Let $(\mathcal{G}, A)$ be the one-player cost game[18] whose arena is depicted in Figure 2.7, and where the cost function Cost_1 is defined by $\mathsf{Cost}_1(A^n B^\omega) = -1 + \frac{1}{n}$ for $n \in \mathbb{N}_0$ and $\mathsf{Cost}_1(A^\omega) = 0$. One can be convinced that there is no Nash equilibrium in $(\mathcal{G}, A)$. Indeed, looping infinitely often in A is not a Nash equilibrium, and for the other strategy profiles, looping one more time in A is always a profitable deviation.

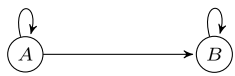


Figure 2.7: A one-player cost game without Nash equilibrium.

Subgame Perfect Equilibrium

In this section, we define the concept of *subgame perfect equilibrium* in the context of cost games, and give some existence results about it.

Let us first generalise the notion of *subgame* in this framework. Given a multiplayer cost game $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{Cost}_i)_{i \in \Pi})$ and a history h of $\mathcal{G}$, we denote by $\mathcal{G}|_h$ the game $\mathcal{G}|_h = (\Pi, \mathcal{A}, (\mathsf{Cost}_i|_h)_{i \in \Pi})$ where $\mathsf{Cost}_i|_h(\rho) = \mathsf{Cost}_i(h\rho)$ for any play ρ such that $(\mathsf{Last}(h), \mathsf{First}(\rho)) \in E$,[19] and we say that $\mathcal{G}|_h$ is a *subgame* of $\mathcal{G}$. We keep the same notations for initialised subgames and particular strategies $\sigma_i|_h$ as in Section 2.2.2 (just before Definition 2.2.24).

Then, given a history hv of a cost game $(\mathcal{G}, v_0)$ (with $v \in V$), we say that $(\sigma_i|_h)_{i \in \Pi}$ is a Nash equilibrium in the subgame $(\mathcal{G}|_h, v)$ if, for

18. This game is in the same vein as the Min-Max cost game of Example 2.3.12.
19. Notice that $\mathsf{Cost}_i|_h$ is not defined on all of Plays, but this is not a problem since it is always applied on plays beginning with a successor of $\mathsf{Last}(h)$.

every player $j \in \Pi$ and every strategy σ'_j of player j, we have that $\mathsf{Cost}_j|_h(\langle \sigma'_j|_h, \sigma_{-j}|_h\rangle_v) \geq \mathsf{Cost}_j|_h(\langle (\sigma_i|_h)_{i\in\Pi}\rangle_v)$, or in an equivalent way, $\mathsf{Cost}_j(h\langle \sigma'_j|_h, \sigma_{-j}|_h\rangle_v) \geq \mathsf{Cost}_j(h\langle (\sigma_i|_h)_{i\in\Pi}\rangle_v)$.

Exactly as in the qualitative case, a strategy profile is a subgame perfect equilibrium in a game if it is a Nash equilibrium in every subgame. In particular, a subgame perfect equilibrium is also a Nash equilibrium.

Definition 2.3.20. Given a multiplayer cost game $(\mathcal{G}, v_0)$, a strategy profile $(\sigma_i)_{i\in\Pi}$ of $\mathcal{G}$ is a *subgame perfect equilibrium* of $(\mathcal{G}, v_0)$ if $(\sigma_i|_h)_{i\in\Pi}$ is a Nash equilibrium in $(\mathcal{G}|_h, v)$, for every history hv of $(\mathcal{G}, v_0)$, with $v \in V$.

Let us illustrate this on an example.

Example 2.3.21. Let us come back to the two-player cost game $\mathcal{G}$ described in Example 2.3.15 (on page 48), where $\mathsf{Cost}_1 = \mathrm{RP}_{\mathrm{Min}}$ for the goal set $\mathsf{R} = \{C\}$, and $\mathsf{Cost}_2 = \mathrm{AP}_{\mathrm{Min}}$. The Nash equilibrium (σ_1, σ_2) of Example 2.3.18 (on page 49) is also a subgame perfect equilibrium in $(\mathcal{G}, A)$.

The first result we present is the existence of a subgame perfect equilibrium in multiplayer cost games played on *finite trees*. This directly follows from the classical Kuhn's theorem [Kuh53] (see below). A *preference relation* is a total, reflexive and transitive binary relation.

Theorem 2.3.22 ([Kuh53])**.** *Given a multiplayer cost game $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{Cost}_i)_{i\in\Pi})$ whose graph is a* finite tree[20]*, and a preference relation $\precsim_i$ on cost profiles*[21] *for each player $i \in \Pi$, there exists a strategy profile $(\sigma_i)_{i\in\Pi}$ such that for every history hv in $\mathcal{G}$, with $v \in V$, for every player $j \in \Pi$, and every strategy σ'_j of player j, we have*

$$\mathsf{Cost}(\rho') \precsim_j \mathsf{Cost}(\rho)$$

where $\rho = h\langle (\sigma_i|_h)_{i\in\Pi}\rangle_v$ and $\rho' = h\langle \sigma'_j|_h, \sigma_{-j}|_h\rangle_v$.

20. When a game is played on a tree, we always assume that the game is initialised, and the initial vertex is the root of the tree.

21. Given two cost profiles x and y, the notation $x \precsim_j y$ means that player j prefers y to x, or the two cost profiles are equivalent for him.

Let $\precsim_j$ be the binary relation on cost profiles defined by $(x_i)_{i\in\Pi} \precsim_j (y_i)_{i\in\Pi}$ iff $x_j \geq y_j$. It is clearly a preference relation which captures the concept of Nash equilibrium (see Definition 2.3.16). We thus have the following corollary.

Corollary 2.3.23. *In every multiplayer cost game whose graph is a* finite tree, *there exists a subgame perfect equilibrium.*

In particular, it implies that there exists a Nash equilibrium in every multiplayer cost game whose graph is a finite tree.

The following result states the existence of a subgame perfect equilibrium (and thus, a Nash equilibrium) in a certain class of cost games.

Theorem 2.3.24 ([FL83, Har85])**.** *Given a multiplayer cost game* $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{Cost}_i)_{i\in\Pi})$ *and an initial vertex* $v_0 \in V$*, if the cost functions* $(\mathsf{Cost}_i)_{i\in\Pi}$ *are continuous and real-valued*[22]*, then there exists a subgame perfect equilibrium in* $(\mathcal{G}, v_0)$*.*

Notice that if Cost_i is a continuous real-valued function, then Cost_i is bounded, as Plays is compact[23].

Remark 2.3.25. Theorem 2.3.24 does not apply on games with mean-payoff objectives, as the cost function $\mathrm{AP}_{\mathrm{Min}}$ (see Definition 2.3.3) is not continuous. Indeed, let us consider the one-player cost game whose arena is depicted in Figure 2.8, and where the cost function Cost_1 is $\mathrm{AP}_{\mathrm{Min}}$ for the price function ϕ defined by $\phi(A, A) = \phi(A, B) = 0$ and $\phi(B, B) = 1$.

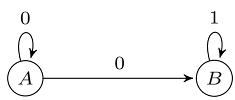


Figure 2.8: The cost function $\mathrm{AP}_{\mathrm{Min}}$ is not continuous.

The sequence of plays $(A^n B^\omega)_{n\in\mathbb{N}_0}$ converges to the play A^ω. For all $n \in \mathbb{N}_0$, we have that $\mathrm{AP}_{\mathrm{Min}}(A^n B^\omega) = 1$, but $\mathrm{AP}_{\mathrm{Min}}(A^\omega) = 0$.

22. A *real-valued* cost function assigns a real number to every play.
23. Plays is compact since it is a closed set in V^ω, which is compact.

Chapter 3

Contributions

The aim of this chapter is to summarise, in a few pages, the main contributions of this thesis, and to give the history of the principal results. The sequel of this document is devoted to the details and the proofs of these results.

In this thesis, we focus on *existence results* for several kinds of equilibrium in *multiplayer cost games*. Many parameters appear when studying cost games: the graph can be enriched with *prices on edges* or not; there can be *two or more players*; the *objectives* of the players can be various and also complicated; different kinds of rational behaviour can be considered for the players, leading to *different concepts of equilibria*;... We have worked in a gradual manner: we fixed some of these parameters and we studied the existence of the chosen notion of equilibrium in the chosen class of games, as well as the complexity of the equilibria in terms of the *memory* needed in the strategies of the individual players in these equilibria. Then, we proceeded to more or less restrictive parameters, like more general objectives or more refined notions of equilibrium. Note that most of our results hold for any number of players.

As Nash equilibria do not always exist in cost games (see Example 2.3.19), we first focused on quantitative reachability objectives, and we studied *quantitative reachability games* [BBD10, BBD12, BBDG12, BBDG13]. These games are cost games where each player has a goal set of vertices of the graph, and aims at reaching his own goal set as soon as possible. In other words, he wants to minimise the number of edges until the play reaches his goal set for the first time (see Definition 4.1.1 for more details).

Once the objectives were fixed, we studied the existence of different notions of equilibria. We proved in [BBD10, BBD12] several results about the famous *Nash equilibrium*, and the more recent concept of *secure equilibrium*. To our knowledge, it was the first time that the notion of secure equilibrium was considered in the quantitative framework. The formal definition can be found in Section 5.1, and naturally extends the definition given in the qualitative framework ([CHJ04], or see Definition 2.2.35 or 2.2.39).

The first problem we considered is the following one.

Problem 1. Does there exist a Nash equilibrium (resp. a secure equilibrium) in every initialised quantitative reachability game?

We provided the following positive answers.

- In every initialised multiplayer quantitative reachability game, there exists a finite-memory Nash equilibrium (Theorem 4.1.5).
- In every initialised two-player quantitative reachability game, there exists a finite-memory secure equilibrium (Theorem 5.2.1).
- In every initialised multiplayer quantitative reachability game, one can decide whether there exists a secure equilibrium in ExpSpace (Theorem 5.2.7).
- In every initialised multiplayer quantitative reachability game, one can decide in ExpSpace whether there exists a secure equilibrium such that the players' costs are below some thresholds (Proposition 5.2.22).

We also somewhat extended our existence result of Nash equilibria in two directions. The first one concerns *quantitative reachability/safety games* (Definition 4.2.1). These are cost games where some players have quantitative reachability objectives, whereas others have quantitative *safety* objectives (they want to avoid some *bad set* of vertices or, if impossible, delay its visit as long as possible).

In another direction, we enriched the graph with tuples of prices (one price for each player) on the edges of the graph, and studied *quantitative reachability games with tuples of prices on edges* (Definition 4.3.1). In such games, the edges of the graph are labelled with tuples of positive prices (one price for each player), and every player has a reachability objective, but in this framework, we do not only count the number of edges to reach the goal of a player, but we sum up his prices along the path until his goal is reached.

We positively answered Problem 1 for Nash equilibria in these two kinds of cost games [BBD12].

- In any initialised multiplayer quantitative reachability/safety game, there exists a finite-memory Nash equilibrium (Theorem 4.2.2).
- In any initialised multiplayer quantitative reachability game with tuples of prices on edges, there exists a finite-memory Nash equilibrium (Theorem 4.3.2).

The second problem asks, given a Nash (resp. secure) equilibrium, whether there exists a *finite-memory* Nash (resp. secure) equilibrium with the same *type*, meaning that the sets of players who reach their goal set along the outcome of the initial equilibrium and along the outcome of the finite-memory one are the same.

Problem 2. Given a Nash equilibrium (resp. a secure equilibrium) in an initialised quantitative reachability game, does there exist a finite-memory Nash equilibrium (resp. secure equilibrium) with the same type?

The answers to this problem for Nash and secure equilibria are both positive in multiplayer quantitative reachability games.

- Given a Nash equilibrium in an initialised multiplayer quantitative reachability game, there exists a finite-memory Nash equilibrium of the same type (Theorem 4.1.12).
- Given a secure equilibrium in an initialised multiplayer quantitative reachability game, there exists a finite-memory secure equilibrium of the same type (Theorem 5.2.8).

The first result comes from [BBD10, BBD12] and the second one comes from [BBDG12, BBDG13].

Then, we considered more refined notions of equilibria, like *subgame perfect equilibria*. And we introduced [BBDG12, BBDG13] a new and even stronger solution concept with the notion of *subgame perfect secure equilibrium* (see Section 7.1 for the definition), which gathers both the sequential nature of subgame perfect equilibria and the verification-oriented aspects of secure equilibria.

The third problem concerns the existence of subgame perfect (secure) equilibria in quantitative reachability games.

Problem 3. Does there exist a subgame perfect equilibrium (resp. a subgame perfect secure equilibrium) in every initialised quantitative reachability game?

We gave the two following positive answers [BBDG12, BBDG13].

- In every initialised multiplayer quantitative reachability game, there exists a subgame perfect equilibrium (Theorem 6.1.1).
- In every initialised two-player quantitative reachability game, there exists a subgame perfect secure equilibrium (Theorem 7.3.1).

In order to prove these two results, we had to use proof techniques (namely, topology) completely different from the ones given in [BBD10, BBD12] for the existence of Nash and secure equilibria.

The first result extends to quantitative reachability games with tuples of prices on edges.

- In every initialised multiplayer quantitative reachability game with tuples of prices on edges, there exists a subgame perfect equilibrium (Theorem 6.2.1).

Once multiplayer quantitative reachability games have been quite well understood, we wanted to extend our results to other classes of cost games. We studied in [BDS13] quantitative objectives expressed through a cost function for each player. Each cost function assigns, for every play of the game, a value that represents the cost that is incurred for a player by this play. In this framework, we allow tuples of prices and rewards on the edges of the graph.

Problems 4 and 5 extend the question of existence of (finite-memory) Nash equilibria to general multiplayer cost games.

Problem 4. Does there exist a Nash equilibrium in every initialised multiplayer cost game?

Problem 5. Does there exist a finite-memory Nash equilibrium in every initialised multiplayer cost game?

If we make no restriction on the cost games, the answer to Problem 4 (and thus to Problem 5) is negative (see again Example 2.3.19). That is why we identified [BDS13] large classes of cost games for which the answers to Problems 4 and 5 are positive. These classes of cost games include, in particular, quantitative reachability objectives, mean-payoff objectives, discounted objectives, and many other ones. This result then gave a quantitative counterpart to a result of E. Grädel and M. Ummels [GU08] about qualitative games. Moreover, it generalised some results obtained in [BBD10, BBD12] about quantitative reachability/safety objectives to a wider class of objectives.

- In every initialised multiplayer cost game where each cost function is prefix-linear and positionally coalition-determined, there exists a Nash equilibrium with memory (at most) linear in the size of the game (Proposition 4.4.6).
- In every initialised multiplayer cost game where each cost function is prefix-independent and coalition-determined, there exists a Nash equilibrium (Proposition 4.4.11).
- In every initialised multiplayer cost game where each cost function is prefix-independent and finite-memory coalition-determined, there exists a finite-memory Nash equilibrium (Proposition 4.4.12).

These results are partly established in [BDS13]. The definitions concerning the hypotheses required for the cost functions of the game can be found in Section 4.4.1.

We deduced from the first result that simple Nash equilibria exists in cost games where the players have quantitative reachability objectives or mean-payoff objectives (Corollary 4.4.15). We also got a more general result, which states that if each cost function satisfies either the hypotheses of the first result, or the ones of the third result, then there exists a finite-memory Nash equilibrium in the cost game (Theorem 4.4.14).

The general philosophy of the proofs is as follows: we derive existence of Nash equilibria in multiplayer non-zero-sum quantitative games (and characterisation of their complexity) through determinacy results (and characterisation of the optimal strategies) of several well-chosen two-player quantitative games obtained from the initial multiplayer game.

All these results are covered in Part II, and are grouped according to the kind of equilibrium that is considered. They were not a direct consequence of the existing results in the qualitative framework, we had to develop new proof techniques.

On another hand, we have established some collaborations with Hugo Gimbert (CNRS researcher in LaBRI-Bordeaux) and Sven Schewe (lecturer at the University of Liverpool), which led to several publications. We are currently working with János Flesch, Jeroen Kuipers, Gijs Schoenmakers, and Koos Vrieze (from Maastricht university).

Part II

Focus on Multiplayer Cost Games

Chapter 4

Nash Equilibrium

In this chapter, based on [BBD10, BBD12, BDS13], we study the notion of *Nash equilibrium* in multiplayer cost games. Let us remind that this concept is defined in Section 2.3.2 (see Definition 2.3.16), and the notations about arenas, plays, histories and strategies can be found in Section 2.1.

In Sections 4.1, 4.2 and 4.3, we mainly focus on cost games where all players have quantitative reachability (or safety) objectives. Then, in Section 4.4, we turn to larger classes of cost games, where existence of Nash equilibria can be proved.

4.1 Quantitative Reachability Objectives

In this section, we define quantitative reachability games, which are particular cost games, and give several results concerning these games and Nash equilibria.

4.1.1 Definition

Quantitative reachability games are cost games where all players have *quantitative reachability objectives*. It means that, given a certain set R_i of vertices, each player i wants to reach one of these vertices as soon as possible (that is, while minimising the number of edges to visit this set for the first time).

Definition 4.1.1. A *multiplayer quantitative reachability game* is a multiplayer cost game $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{Cost}_i)_{i\in\Pi})$ such that for any $i \in \Pi$, $\mathsf{Cost}_i = \mathrm{RP}_{\mathrm{Min}}$ for a given goal set $\mathsf{R}_i \subseteq V$, and the shared price function $\phi : E \to \mathbb{R}$ with $\phi(e) = 1$ for all $e \in E$.

Abusively, such a game is denoted by $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{R}_i)_{i\in\Pi})$.

Thus, in a quantitative reachability game $\mathcal{G}$, the cost of player i of a play $\rho = \rho_0\rho_1 \dots$ is [1]:

$$\mathsf{Cost}_i(\rho) = \begin{cases} l & \text{if } l \text{ is the } \textit{least} \text{ index such that } \rho_l \in \mathsf{R}_i, \\ +\infty & \text{otherwise.} \end{cases} \tag{4.1}$$

Let us introduce two notations that will be useful in the sequel. For any play ρ, we denote by $\mathsf{Visit}(\rho)$ the set of players $i \in \Pi$ such that ρ visits R_i. The set $\mathsf{Visit}(h)$ for a history h is defined similarly.

The *type* of a strategy profile $(\sigma_i)_{i\in\Pi}$ in a quantitative reachability game $(\mathcal{G}, v_0)$ is the set of players $j \in \Pi$ such that the outcome of $(\sigma_i)_{i\in\Pi}$ in $(\mathcal{G}, v_0)$ visits R_j, and is denoted by $\mathsf{Type}(v_0, (\sigma_i)_{i\in\Pi})$. In other words, $\mathsf{Type}(v_0, (\sigma_i)_{i\in\Pi}) = \mathsf{Visit}(\langle(\sigma_i)_{i\in\Pi}\rangle_{v_0})$. When the context is clear, we simply write $\mathsf{Type}((\sigma_i)_{i\in\Pi})$.

Example 4.1.2. Let $\mathcal{G} = (\{1, 2\}, \mathcal{A}, (\mathsf{R}_1, \mathsf{R}_2))$ be the two-player quantitative reachability game whose arena $\mathcal{A}$ is depicted in Figure 4.1 and where $\mathsf{R}_1 = \{C\}$ and $\mathsf{R}_2 = \{D\}$. We remind that the vertices of player 1 (resp. player 2) are represented by circles (resp. squares). Notice that the vertex of R_1 (resp. R_2) is shaded (resp. doubly circled) in the figure. [2]

Let us define the following memoryless strategies σ_1 and σ_2 for player 1 and player 2 respectively, as: $\sigma_1(A) = D$ and $\sigma_2(B) = C$. [3] The outcome of (σ_1, σ_2) in $(\mathcal{G}, A)$ is the play $(AD)^\omega$, with cost profile $(+\infty, 1)$, since the play $(AD)^\omega$ does not visit R_1 and visits R_2 within one edge.

The strategy profile (σ_1, σ_2) is not a Nash equilibrium in $(\mathcal{G}, A)$, since the memoryless strategy σ'_1 defined by $\sigma'_1(A) = B$ is a profitable deviation

1. See Definition 2.3.3.
2. We will keep this convention in the sequel of the document.
3. Notice that player 1 has no choice to make in vertices C and D.

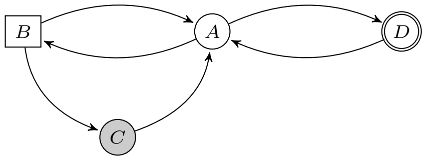


Figure 4.1: A two-player quantitative reachability game.

for player 1. Indeed, the outcome of (σ'_1, σ_2) in $(\mathcal{G}, A)$ is the play $(ABC)^\omega$, and $\mathsf{Cost}_1((ABC)^\omega) = 2 < +\infty$.

On the opposite side, one can show that (σ'_1, σ_2) is a Nash equilibrium in $(\mathcal{G}, A)$.

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

$$\tau_1(hA) = \begin{cases} D & \text{if } h = \epsilon \\ B & \text{if } h \neq \epsilon \end{cases} \quad ; \quad \tau_2(h'B) = \begin{cases} C & \text{if } h' \text{ visits } D \\ A & \text{otherwise} \end{cases}$$

for all histories hA and $h'B$. The outcome of (τ_1, τ_2) in $(\mathcal{G}, A)$ is the play $AD(ABC)^\omega$, with cost profile $(4, 1)$. We claim that the strategy profile (τ_1, τ_2) is a Nash equilibrium in $(\mathcal{G}, A)$. For player 2, it is clearly impossible to pay a cost less than 1 in this game. Moreover, player 1 has no incentive to deviate, because if he chooses the edge (A, B) at the first step, then player 2 chooses the edge (B, A) according to τ_2. And so, the cost of player 1 will be greater than 4.

We will see in Section 4.1.4 that there always exists a finite-memory Nash equilibrium in an initialised multiplayer quantitative reachability game.

4.1.2 Qualitative Games vs Quantitative Games

We show in this section that Problems 1 and 2 can not be reduced to problems on qualitative games in the following sense. Existence of Nash equilibria in *qualitative* non-zero-sum games where each player has a reachability objective (given by Theorem 2.2.23) does *not* directly imply existence of Nash equilibria in quantitative reachability games.

Given a quantitative reachability game $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{R}_i)_{i\in\Pi})$, one can naturally define a *qualitative* version of $\mathcal{G}$, denoted by $\overline{\mathcal{G}}$ and given by the qualitative non-zero-sum game $\overline{\mathcal{G}} = (\Pi, \mathcal{A}, (\mathsf{Win}_i)_{i\in\Pi})$, where $\mathsf{Win}_i = \{\rho \in V^\omega \mid \exists n \in \mathbb{N},\ \rho_n \in \mathsf{R}_i\}$, for all $i \in \Pi$.

By Theorem 2.2.23, we know that there exists a Nash equilibrium in $\overline{\mathcal{G}}$. Nevertheless, the next example illustrates that lifting Nash equilibria in $\overline{\mathcal{G}}$ to Nash equilibria in $\mathcal{G}$ does not work. That is why we developed new ideas to solve Problem 1.

Example 4.1.3. Let us now consider the two-player quantitative reachability game $\mathcal{G} = (\{1, 2\}, \mathcal{A}, (\mathsf{R}_1, \mathsf{R}_2))$ whose arena $\mathcal{A}$ is depicted in Figure 4.2, and such that $\mathsf{R}_1 = \{B, E\}$ and $\mathsf{R}_2 = \{C\}$. Notice that only player 1 effectively plays in this game.

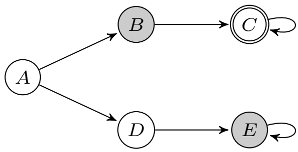


Figure 4.2: A two-player quantitative reachability game.

We are going to exhibit a Nash equilibrium σ_1 in the corresponding qualitative game $(\overline{\mathcal{G}}, A)$ that can not be lifted to a Nash equilibrium in the quantitative game $(\mathcal{G}, A)$. The positional strategy σ_1 of player 1 is defined by $\sigma_1(A) = D$. It is a Nash equilibrium in $(\overline{\mathcal{G}}, A)$, with outcome ADE^ω that player 1 wins and player 2 loses. However, σ_1 is not a Nash equilibrium in $(\mathcal{G}, A)$. Indeed, choosing the edge (A, B) provides a smaller cost to player 1: $\mathsf{Cost}_1(ABC^\omega) < \mathsf{Cost}_1(ADE^\omega)$. Notice that in this example, there is no Nash equilibrium in $(\mathcal{G}, A)$ where only player 1 reaches his goal set.

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

Proposition 4.1.4. *Any Nash equilibrium in a multiplayer quantitative reachability game $(\mathcal{G}, v_0)$ is also a Nash equilibrium in the corresponding qualitative game $(\overline{\mathcal{G}}, v_0)$.*

Proof. Let $(\sigma_i)_{i\in\Pi}$ be a Nash equilibrium in a multiplayer quantitative reachability game $(\mathcal{G}, v_0)$. For a contradiction, let us assume that in $(\overline{\mathcal{G}}, v_0)$, player j has a profitable deviation σ'_j w.r.t. $(\sigma_i)_{i\in\Pi}$. This is only possible if $\langle(\sigma_i)_{i\in\Pi}\rangle_{v_0}$ is lost by player j and $\langle\sigma'_j, \sigma_{-j}\rangle_{v_0}$ is won by player j. Thus when playing σ'_j against σ_{-j}, player j manages to reach R_j. Clearly enough, σ'_j would also be a profitable deviation w.r.t. $(\sigma_i)_{i\in\Pi}$ in $(\mathcal{G}, v_0)$, contradicting the hypothesis. □

4.1.3 Unravelling of a Graph

Let $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{R}_i)_{i\in\Pi})$ be a multiplayer quantitative reachability game played on an arena $\mathcal{A} = (V, (V_i)_{i\in\Pi}, E)$. In order to prove most of the results concerning quantitative reachability games (even for the other concepts of equilibria studied in this document), it will often be useful to *unravel* the graph $G = (V, E)$ from an 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 of histories of $\mathcal{A}$ starting in v_0, the initial vertex is the history v_0, and an edge of T is a pair (h, hv) of histories starting in v_0 such that $(\mathsf{Last}(h), v) \in E$. A history h is a vertex of player i in T if $h \in \mathsf{Hist}_i$, and h belongs to the goal set of player i if $\mathsf{Last}(h) \in \mathsf{R}_i$.

We denote by $\mathcal{T}$ the corresponding quantitative reachability game played on this new arena[4]. This game $\mathcal{T}$ played on the unravelling T of G from v_0 is equivalent to the game $(\mathcal{G}, v_0)$ played on the graph G in the following sense. A play $(\rho_0)(\rho_0\rho_1)(\rho_0\rho_1\rho_2)\ldots$ in $\mathcal{T}$ induces a unique play $\rho = \rho_0\rho_1\rho_2\ldots$ in $(\mathcal{G}, v_0)$, and conversely. Thus, we denote a play in $\mathcal{T}$ by the respective play in $(\mathcal{G}, v_0)$. The bijection between plays of $(\mathcal{G}, v_0)$ and plays of $\mathcal{T}$ allows us to use the same cost functions $(\mathsf{Cost}_i)_{i\in\Pi}$, and to transform easily strategies in $\mathcal{G}$ to strategies in $\mathcal{T}$ (and conversely).

4. Notice that the graph T is infinite, but comes from a finite graph.

For practical reasons, we often use equivalently $\mathcal{T}$ in our proofs instead of $(\mathcal{G}, v_0)$, and the equilibria defined in $\mathcal{T}$ are obviously equilibria in $(\mathcal{G}, v_0)$. Moreover, figures given in proofs to help the understanding roughly represent the unravelling T of G and plays in game $\mathcal{T}$.

We also sometimes need to consider the tree T limited to a certain depth $d \in \mathbb{N}$: we denote by $\mathsf{Trunc}_d(T)$ the *truncated tree of T of depth d* and $\mathsf{Trunc}_d(\mathcal{T})$ the *finite* game played on $\mathsf{Trunc}_d(T)$. More precisely, the set of vertices of $\mathsf{Trunc}_d(T)$ is the set of histories starting in v_0 of length $\leq d$; the edges of $\mathsf{Trunc}_d(T)$ 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 in $\mathsf{Trunc}_d(\mathcal{T})$ corresponds to a history of $(\mathcal{G}, v_0)$ of length *equal to d*. The notions of cost and strategy are defined exactly like in the game $\mathcal{T}$, but limited to the depth d. For instance, a player pays an infinite cost for a play ρ in $\mathsf{Trunc}_d(\mathcal{T})$ if his goal set is not visited by ρ. Remark that $\mathcal{T}$ and $\mathsf{Trunc}_d(\mathcal{T})$ are always supposed to be initialised in v_0.

The reason why we consider the games $\mathsf{Trunc}_d(\mathcal{T})$ for certain depths d is that Corollary 2.3.23 (Kuhn's theorem) applies in these games since they are played on finite trees.

For an example of a game $(\mathcal{G}, v_0)$ and an associated truncated tree $\mathsf{Trunc}_d(T)$, see Figures 4.4 and 4.5 on page 76.

4.1.4 Results

In this section, we present and prove the two main results about Nash equilibria in multiplayer quantitative reachability games. On one hand, we state that there always exists a *finite-memory Nash equilibrium* in such games (Theorem 4.1.5). On the other hand, we show that, given a Nash equilibrium, there exists a *finite-memory* Nash equilibrium *of the same type* (Theorem 4.1.12).

Existence of a Finite-Memory Nash Equilibrium

In this section, we show the existence of a finite-memory Nash equilibrium in multiplayer quantitative reachability games, and then positively solve Problem 1 for Nash equilibria.

Theorem 4.1.5. *In every initialised multiplayer quantitative reachability game, there exists a finite-memory Nash equilibrium.*

The proof of this theorem is based on the following ideas. Given a quantitative reachability game $(\mathcal{G}, v_0)$ played on a finite graph G, we unravel the graph from v_0, as in Section 4.1.3, to get an equivalent game $\mathcal{T}$ played on the infinite tree T. By Kuhn's theorem (Corollary 2.3.23), there exists a Nash equilibrium in the game $\mathsf{Trunc}_d(\mathcal{T})$ played on the finite tree $\mathsf{Trunc}_d(T)$, for any depth d. By choosing an adequate depth d, Proposition 4.1.6 enables to extend this Nash equilibrium to a Nash equilibrium in $\mathcal{T}$, and thus in $\mathcal{G}$. Let us detail these ideas.

Proposition 4.1.6 states that it is possible to extend a Nash equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$ to a Nash equilibrium in the game $\mathcal{T}$, if depth d is equal to $(|\Pi|+1)\cdot 2\cdot |V|$. We then obtain Theorem 4.1.5 as a consequence of Corollary 2.3.23 and Proposition 4.1.6.

Proposition 4.1.6. *Let $(\mathcal{G}, v_0)$ be a multiplayer quantitative reachability game and $\mathcal{T}$ be the corresponding game played on the unravelling of G from v_0. If there exists a Nash equilibrium in the game $\mathsf{Trunc}_d(\mathcal{T})$ where $d = (|\Pi|+1)\cdot 2\cdot |V|$, then there exists a finite-memory Nash equilibrium in the game $\mathcal{T}$.*

The proof of Proposition 4.1.6 roughly works as follows. Let $(\sigma_i)_{i\in\Pi}$ be a Nash equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$. A well-chosen prefix $\alpha\beta$, where β can be repeated (as 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 to punish him in a way that this deviation is not profitable for him. These ideas are detailed in Lemmas 4.1.7 and 4.1.8. One can see Lemma 4.1.7 as a technical result used to prove Lemma 4.1.8, which is the main ingredient to show Proposition 4.1.6. The proof of Lemma 4.1.7 relies on the memoryless determinacy of zero-sum qualitative reachability games (Theorem 2.2.15). More precisely, given a multiplayer quantitative reachability game $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{R}_i)_{i\in\Pi})$, we consider the (zero-sum qualitative) reachability game $\mathcal{G}^j = (\mathcal{A}^j, \mathsf{R}_j)$, where $\mathcal{A}^j = (V, (V_j, V\setminus V_j), E)$, and player j plays in order to reach his goal set R_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 $V \setminus V_j$.

Lemma 4.1.7. *Let $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{R}_i)_{i\in\Pi})$ be a multiplayer quantitative reachability game, and $\mathcal{T}$ be the corresponding game played on the unravelling of G from a vertex v_0. For any depth $d \in \mathbb{N}$, let $(\sigma_i)_{i\in\Pi}$ be a Nash equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$, and ρ the (finite) outcome of $(\sigma_i)_{i\in\Pi}$. Assume that ρ has a prefix $\alpha\beta\gamma$, where $\alpha, \beta, \gamma \in V^+$, such that*

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

for some $l \geq 1$.
Let $j \in \Pi$ be such that α does not visit R_j, and let us consider the zero-sum qualitative reachability game $\mathcal{G}^j = (\mathcal{A}^j, \mathsf{R}_j)$. Then for all histories hv of $(\mathcal{G}, v_0)$ (with $v \in V$) consistent with σ_{-j} and such that $|hv| \leq |\alpha\beta|$, the coalition of the players $i \neq j$ wins the game $\mathcal{G}^j$ from v.

Condition $\mathsf{Visit}(\alpha) = \mathsf{Visit}(\alpha\beta\gamma)$ means that if R_i is visited by $\alpha\beta\gamma$ for any $i \in \Pi$, then it has already been visited by α. Condition $\mathsf{Last}(\alpha) = \mathsf{Last}(\alpha\beta)$ means that β can be repeated (as a cycle). The play ρ of Lemma 4.1.7 is illustrated in Figure 4.3 (in the proof of Proposition 4.1.6, on page 75).

Lemma 4.1.7 says in particular that the players $i \neq j$ can play together to prevent player j from reaching his goal set R_j, from any vertex of the history $\alpha\beta$ (as $\alpha\beta$ is consistent with σ_{-j}).

Proof of Lemma 4.1.7. Let us assume that the hypotheses of the lemma are fulfilled. By contradiction suppose that player j wins the game $\mathcal{G}^j$ from v. By Theorem 2.2.15, player j has a memoryless winning strategy μ_j^v which enables him to reach his goal set R_j within at most $|V| - 1$ edges from v. We show that μ_j^v leads to a profitable deviation for player j w.r.t. $(\sigma_i)_{i\in\Pi}$ in the game $\mathsf{Trunc}_d(\mathcal{T})$, which is impossible by hypothesis.

Let ρ' be a play in $\mathsf{Trunc}_d(\mathcal{T})$ such that hv is a prefix of ρ', and from hv, player j plays according to the strategy μ_j^v and each other player $i \neq j$ continues to play according to σ_i. As the play ρ' is consistent with the memoryless winning strategy μ_j^v from hv, it visits R_j, and so,

$$\begin{aligned}\mathsf{Cost}_j(\rho') &\leq |hv| + |V| \\ &\leq |\alpha\beta| + |V| \\ &\leq l \cdot |V| + |V| = (l+1) \cdot |V| = |\alpha\beta\gamma| \,.\end{aligned}$$

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

$$\mathsf{Cost}_j(\rho') < \mathsf{Cost}_j(\rho) = +\infty.$$

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

$$\mathsf{Cost}_j(\rho') < \mathsf{Cost}_j(\rho)$$

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

Since ρ' is consistent with σ_{-j}, the strategy of player j induced by the play ρ' is a profitable deviation for player j w.r.t. $(\sigma_i)_{i\in\Pi}$, which is a contradiction. □

Now that we have proved Lemma 4.1.7, we use it in order to obtain Lemma 4.1.8, which states that one can define a Nash equilibrium $(\tau_i)_{i\in\Pi}$ in the game $\mathcal{T}$, based on the Nash equilibrium $(\sigma_i)_{i\in\Pi}$ given in some game $\mathsf{Trunc}_d(\mathcal{T})$.

Lemma 4.1.8. *Let $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{R}_i)_{i\in\Pi})$ be a multiplayer quantitative reachability game, and $\mathcal{T}$ be the corresponding game played on the unravelling of G from a vertex v_0. For any depth $d \in \mathbb{N}$, let $(\sigma_i)_{i\in\Pi}$ be a Nash equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$, and $\alpha\beta\gamma$ be a prefix of $\rho = \langle(\sigma_i)_{i\in\Pi}\rangle_{v_0}$ as defined in Lemma 4.1.7. 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 $\mathsf{Type}((\tau_i)_{i\in\Pi}) = \mathsf{Visit}(\alpha)$.*

Proof. Let us assume that the hypotheses of the lemma are fulfilled, and let us set $\Pi = \{1, \ldots, n\}$. 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 $i \in \Pi$ so that

$$\begin{array}{lll} \forall i \leq k & \mathsf{Cost}_i(\alpha\beta^\omega) < +\infty & (\alpha \text{ visits } \mathsf{R}_i) \\ \forall i > k & \mathsf{Cost}_i(\alpha\beta^\omega) = +\infty & (\alpha \text{ does not visit } \mathsf{R}_i) \end{array}$$

where $0 \leq k \leq n$. In the second case, notice that ρ could visit R_i (but after the prefix $\alpha\beta\gamma$).

The Nash equilibrium $(\tau_i)_{i\in\Pi}$ required by Lemma 4.1.8 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 $\mathsf{Trunc}_d(T)$, then the coalition plays according to σ_{-j} in this tree. It prevents player j from reaching his goal set R_j faster than in $\alpha\beta^\omega$. And if $j > k$, the coalition wins the zero-sum qualitative reachability game $\mathcal{G}^j$ from the vertex v where player j has deviated from $\alpha\beta^\omega$ (see Lemma 4.1.7), then the coalition plays according to its memoryless winning strategy $\mu^v_{\mathtt{C}j}$ (given by Theorem 2.2.15), so that player j does not reach his goal set at all. We denote by $\mu^v_{i,j}$ the strategy of player $i \neq j$ derived from $\mu^v_{\mathtt{C}j}$.

We begin by defining a punishment function $P : \mathsf{Hist} \to \Pi \cup \{\perp\}$, 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 v_0, we define $P(v_0) = \perp$, and for every history $hv \in \mathsf{Hist}$ ($v \in V$) starting in v_0, we let:

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

The Nash equilibrium $(\tau_i)_{i\in\Pi}$ is then defined as follows: for all $i \in \Pi$, let h be a history of $(\mathcal{G}, v_0)$ ending in a vertex of V_i,

$$\tau_i(h) := \begin{cases} v & \text{if } P(h) = \bot \ (h < \alpha\beta^\omega), \text{ s.t. } \ hv < \alpha\beta^\omega, \\ \sigma_i(h) & \text{if } P(h) \neq \bot, i, \ P(h) \leq k \text{ and } |h| < d, \\ \mu^v_{i,P(h)}(v'') & \text{if } P(h) \neq \bot, i \text{ and } P(h) > k, \\ & \quad \text{s.t. } \big(h = h'vv'h''v'' \ (v, v', v'' \in V), \\ & \quad \ P(h'v) = \bot, \text{ and } P(h'vv') = P(h)\big), \\ \textit{arbitrary} & \text{otherwise,} \end{cases} \tag{4.2}$$

where *arbitrary* means that the next vertex is chosen arbitrarily (in a memoryless way). Notice that in the third case, the strategy $\mu^v_{i,P(h)}$ is well-defined since we can apply Lemma 4.1.7 ($P(h) > k$) when considering the history $\bar{h}v$ such that $\bar{h}v \leq h'v$ (see h' above) and $|\bar{h}v| \leq |\alpha\beta|$. Clearly, the outcome of $(\tau_i)_{i\in\Pi}$ is the play $\alpha\beta^\omega$, and $\mathsf{Type}((\tau_i)_{i\in\Pi})$ is equal to $\mathsf{Visit}(\alpha)$ ($= \mathsf{Visit}(\alpha\beta)$).

We first show that the strategy profile $(\tau_i)_{i\in\Pi}$ 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. $(\tau_i)_{i\in\Pi}$ in $\mathcal{T}$. We distinguish the following two cases:

(*i*) $j \leq k$ ($\mathsf{Cost}_j(\alpha\beta^\omega) < +\infty$, α visits R_j).

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 us set $\pi = \langle(\tau_i)_{i\in\Pi}\rangle_{v_0}$ and $\pi' = \langle\tau'_j, \tau_{-j}\rangle_{v_0}$. Then

$$\mathsf{Cost}_j(\pi') < \mathsf{Cost}_j(\pi).$$

On the other hand we know that

$$\mathsf{Cost}_j(\pi) = \mathsf{Cost}_j(\rho) \leq |\alpha|.$$

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

$$\mathsf{Cost}_j(\rho') = \mathsf{Cost}_j(\pi') < \mathsf{Cost}_j(\rho).$$

By Equation (4.2), the play ρ' is consistent with $(\sigma_i)_{i\in\Pi\setminus\{j\}}$, and so, the strategy τ'_j restricted to the tree $\mathsf{Trunc}_d(T)$ is a profitable deviation for player j w.r.t. $(\sigma_i)_{i\in\Pi}$ in the game $\mathsf{Trunc}_d(\mathcal{T})$. This contradicts the fact that $(\sigma_i)_{i\in\Pi}$ is a Nash equilibrium in this game.

(ii) $j > k$ ($\mathsf{Cost}_j(\alpha\beta^\omega) = +\infty$, $\alpha\beta^\omega$ does not visit R_j).

If player j deviates from $\alpha\beta^\omega$ (with the strategy τ'_j), then by Equation (4.2), the other players combine against him and play according to a memoryless strategy $\mu^v_{\mathsf{C}j}$ given by Lemma 4.1.7. This strategy of the coalition keeps the play $\langle\tau'_j, \tau_{-j}\rangle_{v_0}$ away from the set R_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 2.1), we have to prove that each relation $\approx_{\tau_i}$ on Hist has finite index (recall that $h \approx_{\tau_i} h'$ if h, h' end in the same vertex, and $\tau_i(h\delta) = \tau_i(h'\delta)$ for all histories $h\delta, h'\delta \in \mathsf{Hist}_i$). In this aim, we define for each player i an equivalence relation $\sim_{\tau_i}$ with finite index such that

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

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 has to be 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 remembers the first player who has deviated from $\alpha\beta^\omega$. The definition of $\sim_P$ is as follows:

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

The relation $\sim_P$ is an equivalence relation on Hist with finite index, and $h \sim_P h'$ implies that $P(h) = P(h')$.

We now turn to the definition of $\sim_{\tau_i}$. It is based on the definition of τ_i (given in Equation (4.2)) and $\sim_P$. To get an equivalence with finite index, we proceed as follows. Recall that each strategy $\mu^v_{i,P(h)}$ 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 $|h| < d$. 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 $h, h' \in \mathsf{Hist}$, we define $\sim_{\tau_i}$ in the following way:

$$\begin{aligned} h \sim_{\tau_i} h' \quad \text{if} \quad \Big(h \sim_P h' \quad \wedge \quad & \mathsf{Last}(h) = \mathsf{Last}(h') \quad \wedge \\ & \Big(P(h) = \bot \quad \vee \quad P(h) = i \\ & \vee \big(P(h) \neq \bot, i \ \wedge \ P(h) > k\big) \\ & \vee \big(P(h) \neq \bot, i \ \wedge \ P(h) \leq k \ \wedge \ |h|, |h'| \geq d\big)\Big)\Big). \end{aligned}$$

Notice that this relation satisfies

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

and has finite index. Moreover, if $h \sim_{\tau_i} h'$, then $h\delta \sim_{\tau_i} h'\delta$ for all histories $h\delta, h'\delta \in \mathsf{Hist}_i$, and so $h \approx_{\tau_i} h'$, and the relation $\approx_{\tau_i}$ has finite index. □

Remark 4.1.9. The Nash equilibrium given in the proof of Lemma 4.1.8 needs a memory (at most) exponential[5] in the size of the game.

We can now proceed to the proof of Proposition 4.1.6, which states that if there exists a Nash equilibrium in the game $\mathsf{Trunc}_d(\mathcal{T})$ where $d = (|\Pi| + 1) \cdot 2 \cdot |V|$, then there exists a finite-memory Nash equilibrium in the game $\mathcal{T}$.

Proof of Proposition 4.1.6. Let $(\mathcal{G}, v_0)$ be a multiplayer quantitative reachability game and $\mathcal{T}$ be the corresponding game played on the unravelling of G from v_0. Assume that there exists a Nash equilibrium $(\sigma_i)_{i\in\Pi}$

5. Note that Corollary 4.4.15 improves this result: it states, in particular, that in every initialised multiplayer quantitative reachability game, there exists a Nash equilibrium with memory (at most) linear in the size of the game.

in the game $\mathsf{Trunc}_d(\mathcal{T})$, where $d = (|\Pi|+1)\cdot 2\cdot|V|$. We set $\Pi = \{1,\dots,n\}$ and $\rho = \langle(\sigma_i)_{i\in\Pi}\rangle_{v_0}$.

To be able to use Lemma 4.1.8, we consider the prefix $\mathfrak{pq}$ of ρ of minimal length such that

$$\exists\, l \geq 1 \qquad \begin{aligned} &|\mathfrak{p}| = (l-1)\cdot|V| \\ &|\mathfrak{pq}| = (l+1)\cdot|V| \\ &\mathsf{Visit}(\mathfrak{p}) = \mathsf{Visit}(\mathfrak{pq})\,. \end{aligned} \tag{4.3}$$

It means that $|\mathfrak{q}| = 2\cdot|V|-1$ and no new goal set is visited by $\mathfrak{q}$. Let us show that such a prefix $\mathfrak{pq}$ exists. In the worst case, the play ρ visits the goal set of a new player in each prefix of length $i\cdot 2\cdot|V|$, for $1 \leq i \leq n$, i.e. $|\mathfrak{p}| = n\cdot 2\cdot|V|$. But the length d of ρ is equal to $(n+1)\cdot 2\cdot|V|$ by hypothesis. As a consequence, such a prefix $\mathfrak{pq}$ exists. Moreover, the following statements are true.

- $l \leq 2\cdot n+1$.
- If $\mathsf{Visit}(\mathfrak{p}) \subsetneq \mathsf{Visit}(\rho)$, then $l < 2\cdot n+1$.

The first statement results from the fact that $|\mathfrak{p}| \leq n\cdot 2\cdot|V|$ (see above). For the second statement, suppose that there exists $i \in \mathsf{Visit}(\rho)\setminus\mathsf{Visit}(\mathfrak{p})$, then ρ visit R_i after the prefix $\mathfrak{pq}$ by Equation (4.3). And so, it can not be the case that $l = 2\cdot n+1$.

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

$$\mathfrak{pq} = \alpha\beta\gamma \quad \text{with} \quad \begin{aligned} &\alpha,\beta,\gamma \in V^+ \\ &\mathsf{Last}(\alpha) = \mathsf{Last}(\alpha\beta) \\ &|\alpha| \geq (l-1)\cdot|V| \\ &|\alpha\beta| \leq l\cdot|V|\,. \end{aligned}$$

In particular, $|\mathfrak{p}| \leq |\alpha|$ (see Figure 4.3). Moreover, $\mathsf{Visit}(\alpha) = \mathsf{Visit}(\alpha\beta\gamma)$, and $|\alpha\beta\gamma| = (l+1)\cdot|V|$.

As the hypotheses of Lemma 4.1.8 are verified, we can apply it in this context to get a finite-memory Nash equilibrium $(\tau_i)_{i\in\Pi}$ in the game $\mathcal{T}$ with $\mathsf{Type}((\tau_i)_{i\in\Pi}) = \mathsf{Visit}(\alpha)$. □

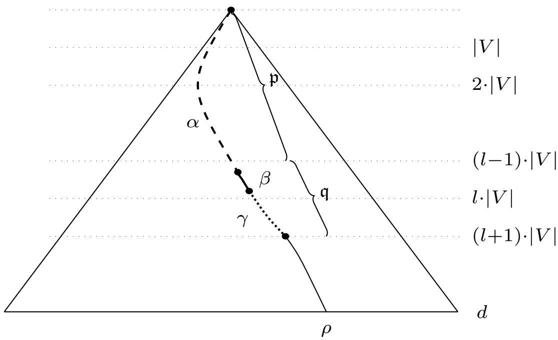


Figure 4.3: Slicing of the play ρ in the tree $\mathsf{Trunc}_d(T)$.

Remark 4.1.10. From the proof of Proposition 4.1.6, we can construct a Nash equilibrium such that each player pays either an infinite cost, or a cost bounded by $|\Pi| \cdot 2 \cdot |V|$.

Proposition 4.1.6 asserts that given a game $\mathcal{G}$ and the game $\mathsf{Trunc}_d(\mathcal{T})$ 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 $\mathsf{Trunc}_d(\mathcal{T})$ to a Nash equilibrium $(\tau_i)_{i\in\Pi}$ of $\mathcal{G}$. The proof of Proposition 4.1.6 states that the type of $(\tau_i)_{i\in\Pi}$ is equal to $\mathsf{Visit}(\alpha)$, which might be different from the type of $(\sigma_i)_{i\in\Pi}$. We here give an example that shows that it is in fact impossible to preserve the type of the lifted Nash equilibrium $(\sigma_i)_{i\in\Pi}$.

Example 4.1.11. Let us consider the two-player quantitative reachability game $\mathcal{G} = (\{1,2\}, \mathcal{A}, (\mathsf{R}_1, \mathsf{R}_2))$, whose arena is depicted in Figure 4.4, and where $V_1 = \{A, C, E\}$, $V_2 = \{B, D\}$, $\mathsf{R}_1 = \{C\}$ and $\mathsf{R}_2 = \{E\}$. One can show that $(\mathcal{G}, A)$ admits only Nash equilibria of type $\{2\}$ or $\emptyset$. Indeed, on one hand, there is no play of $\mathcal{G}$ where both goal sets are visited, and on the other hand, any strategy profile such that its outcome visits R_1 (i.e., is of the form A^+BC^ω) is not a Nash equilibrium, because choosing the edge (B, D) instead of (B, C) is clearly a profitable deviation for player 2.

We will now see that for each $d \geq 2$, the game played on $\mathsf{Trunc}_d(T)$ admits a Nash equilibrium of type $\{1\}$. From the above discussion, this equilibrium can not be lifted to a Nash equilibrium of the same type in $\mathcal{G}$. A truncated tree $\mathsf{Trunc}_d(T)$ is depicted in Figure 4.5. 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 $\mathsf{Trunc}_d(\mathcal{T})$ of type $\{1\}$. Following the lines of the proof of Proposition 4.1.6, we see that this Nash equilibrium is lifted to a Nash equilibrium of $\mathcal{G}$ with outcome A^ω and type $\emptyset$.

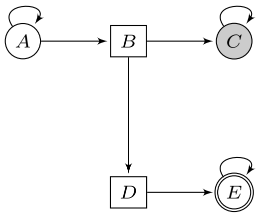


Figure 4.4: Game $(\mathcal{G}, A)$.

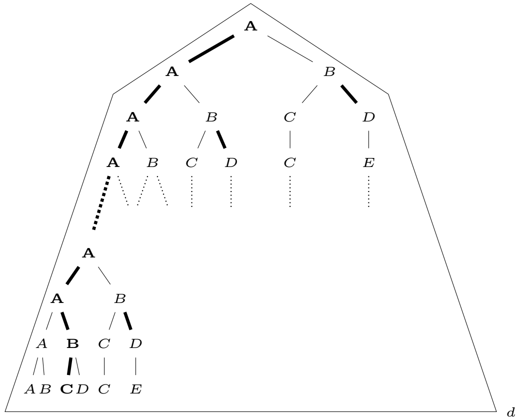


Figure 4.5: The truncated tree $\mathsf{Trunc}_d(T)$.

Finite-Memory Nash Equilibria Preserving Types

In this section, 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 4.1.12. *Given a Nash equilibrium in an initialised multiplayer quantitative reachability game, there exists a finite-memory Nash equilibrium of the same type.*

The proof is based on two steps. Given a Nash equilibrium $(\sigma_i)_{i\in\Pi}$ in a quantitative reachability game $(\mathcal{G}, v_0)$, 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 $\langle(\tau_i)_{i\in\Pi}\rangle_{v_0}$ is of the form $\alpha\beta^\omega$ with $\mathsf{Visit}(\alpha) = \mathsf{Type}((\sigma_i)_{i\in\Pi})$.

This is possible thanks to Lemmas 4.1.14 and 4.1.15, by first eliminating *unnecessary cycles* (see the formal definition below) from the play $\langle(\sigma_i)_{i\in\Pi}\rangle_{v_0}$, and then locating a prefix $\alpha\beta$ such that β can be infinitely repeated after α.

Definition 4.1.13. Given a play $\rho = \alpha\beta\tilde{\rho}$ in a quantitative reachability game $(\mathcal{G}, v_0)$ such that

$$\begin{aligned}
&\alpha, \beta \in V^+, \ \tilde{\rho} \in V^\omega, \\
&\mathsf{Last}(\alpha) = \mathsf{Last}(\alpha\beta), \\
&\mathsf{Visit}(\alpha) = \mathsf{Visit}(\alpha\beta), \text{ and} \\
&\mathsf{Visit}(\alpha) \neq \mathsf{Visit}(\rho),
\end{aligned}$$

if $v = \mathsf{Last}(\alpha)$, then the cycle $v\beta$ is called an *unnecessary cycle*.

In other words, the cycle $\mathsf{Last}(\alpha)\beta$ does not visit a new goal set, but the play ρ visits a new one after this cycle, which means, in some sense, that this cycle is “unnecessary”.

The second step of the proof of Theorem 4.1.12 transforms the Nash equilibrium $(\tau_i)_{i\in\Pi}$ given by the first step into a finite-memory one by means of Lemma 4.1.8. For that purpose, we consider the strategy profile $(\tau_i)_{i\in\Pi}$ limited to the unravelling T of G truncated at a well-chosen depth.

The next lemma indicates how to eliminate a cycle from the outcome of a Nash equilibrium.

Lemma 4.1.14. *Let $(\sigma_i)_{i\in\Pi}$ be a strategy profile in a multiplayer quantitative reachability game $(\mathcal{G}, v_0)$, and $\rho = \langle(\sigma_i)_{i\in\Pi}\rangle_{v_0}$ its outcome. Suppose that $\rho = \mathfrak{p}\mathfrak{q}\tilde{\rho}$, where $\mathfrak{p}, \mathfrak{q} \in V^+$ and $\tilde{\rho} \in V^\omega$, such that*

$$\begin{aligned}
\mathsf{Visit}(\mathfrak{p}) &= \mathsf{Visit}(\mathfrak{p}\mathfrak{q}) \\
\mathsf{Last}(\mathfrak{p}) &= \mathsf{Last}(\mathfrak{p}\mathfrak{q}).
\end{aligned}$$

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

$$\tau_i(h) = \begin{cases} \sigma_i(h) & \text{if } \mathfrak{p} \not\leq h, \\ \sigma_i(\mathfrak{p}\mathfrak{q}\delta) & \text{if } h = \mathfrak{p}\delta \text{ and } \delta \in V^* \end{cases}$$

for all $i \in \Pi$ *and* $h \in \mathsf{Hist}_i$*. We get the outcome* $\langle(\tau_i)_{i\in\Pi}\rangle_{v_0} = \mathfrak{p}\tilde{\rho}$*.*
If there exists a profitable deviation τ'_j *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. Assume that the hypotheses of the lemma are fulfilled. We keep the same notations, and we write $\pi = \langle(\tau_i)_{i\in\Pi}\rangle_{v_0}$.

We observe that as $\rho = \mathfrak{p}\mathfrak{q}\tilde{\rho}$, we have $\pi = \mathfrak{p}\tilde{\rho}$ (see Figures 4.6 and 4.7). It follows that

$$\forall i \in \Pi, \quad \mathsf{Cost}_i(\pi) \leq \mathsf{Cost}_i(\rho). \tag{4.4}$$

More precisely,

- if $\mathsf{Cost}_i(\rho) = +\infty$, then $\mathsf{Cost}_i(\pi) = +\infty$;
- if $\mathsf{Cost}_i(\rho) < +\infty$, then
  - $\mathsf{Cost}_i(\pi) = \mathsf{Cost}_i(\rho)$, if $i \in \mathsf{Visit}(\mathfrak{p})$;
  - $\mathsf{Cost}_i(\pi) = \mathsf{Cost}_i(\rho) - (|\mathfrak{q}| + 1)$, if $i \notin \mathsf{Visit}(\mathfrak{p})$.

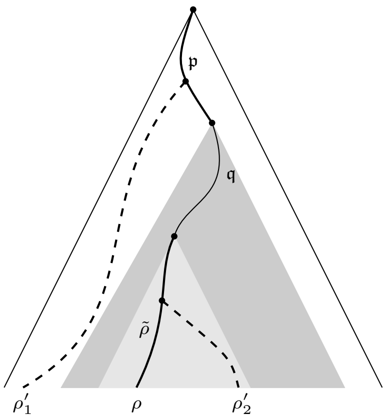


Figure 4.6: Play ρ and deviations.

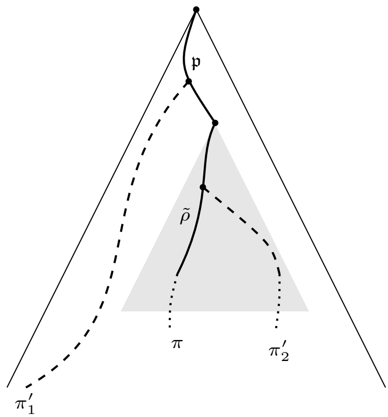


Figure 4.7: Play π and deviations.

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 (τ'_j, τ_{-j}) from v_0. Then,

$$\mathsf{Cost}_j(\pi') < \mathsf{Cost}_j(\pi).$$

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

(i) The history $\mathfrak{p}$ is not a prefix of π' (like for the play π'_1 in Figure 4.7).

We define $\sigma'_j := \tau'_j$ and we denote by ρ' the outcome of (σ'_j, σ_{-j}) from v_0. 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 4.6). Thus,

$$\mathsf{Cost}_j(\rho') = \mathsf{Cost}_j(\pi') < \mathsf{Cost}_j(\pi) \leq \mathsf{Cost}_j(\rho)$$

by Equation (4.4), which implies that σ'_j is a profitable deviation of player j w.r.t. $(\sigma_i)_{i\in\Pi}$.

(ii) The history $\mathfrak{p}$ is a prefix of π' (like for the play π'_2 in Figure 4.7).

We define for all histories $h \in \mathsf{Hist}_j$:

$$\sigma'_j(h) := \begin{cases} \sigma_j(h) & \text{if } \mathfrak{pq} \not\leq h, \\ \tau'_j(\mathfrak{p}\delta) & \text{if } h = \mathfrak{pq}\delta \text{ for } \delta \in V^*. \end{cases}$$

Let us set $\rho' = \langle \sigma'_j, \sigma_{-j}\rangle_{v_0}$. As player j deviates after $\mathfrak{p}$ with the strategy τ'_j, one can prove that

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

by definition of $(\tau_i)_{i\in\Pi}$ (see the play ρ'_2 in Figure 4.6). Moreover, as $\mathsf{Cost}_j(\pi') < \mathsf{Cost}_j(\pi)$, it means that $j \in \mathsf{Visit}(\pi')$ but $j \notin \mathsf{Visit}(\mathfrak{p})$ (as $\mathfrak{p}$ is also a prefix of π). Since $\mathsf{Visit}(\mathfrak{p}) = \mathsf{Visit}(\mathfrak{pq})$, we have

$$\mathsf{Cost}_j(\pi') + (|\mathfrak{q}| + 1) = \mathsf{Cost}_j(\rho').$$

Then, it holds that
– either $\mathsf{Cost}_j(\rho) = \mathsf{Cost}_j(\pi) = +\infty$, and so, $\mathsf{Cost}_j(\rho') < \mathsf{Cost}_j(\rho)$,
– or $\mathsf{Cost}_j(\rho) = \mathsf{Cost}_j(\pi) + (|\mathfrak{q}| + 1)$, and so, $\mathsf{Cost}_j(\rho') < \mathsf{Cost}_j(\rho)$.
In both cases, it proves that the strategy σ'_j is a profitable deviation for player j w.r.t. $(\sigma_i)_{i\in\Pi}$. □

While Lemma 4.1.14 deals with elimination of unnecessary cycles, Lemma 4.1.15 deals with repetition of a useful cycle.

Lemma 4.1.15. *Let $(\sigma_i)_{i\in\Pi}$ be a strategy profile in a multiplayer quantitative reachability game $(\mathcal{G}, v_0)$, and $\rho = \langle(\sigma_i)_{i\in\Pi}\rangle_{v_0}$ its outcome. We assume that $\rho = \mathfrak{p}\mathfrak{q}\tilde{\rho}$, where $\mathfrak{p}, \mathfrak{q} \in V^+$ and $\tilde{\rho} \in V^\omega$, such that*

$$\mathsf{Visit}(\mathfrak{p}) = \mathsf{Visit}(\rho)$$
$$\mathsf{Last}(\mathfrak{p}) = \mathsf{Last}(\mathfrak{p}\mathfrak{q}).$$

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

$$\tau_i(h) = \begin{cases} \sigma_i(h) & \text{if } \mathfrak{p} \not\leq h, \\ \sigma_i(\mathfrak{p}\delta) & \text{if } h = \mathfrak{p}\mathfrak{q}^k\delta,\ k \in \mathbb{N},\ \delta \in V^* \text{ and } \mathfrak{q} \not\leq \delta \end{cases}$$

for all $i \in \Pi$ and $h \in \mathsf{Hist}_i$. We get the outcome $\langle(\tau_i)_{i\in\Pi}\rangle_{v_0} = \mathfrak{p}\mathfrak{q}^\omega$.
If there exists a profitable deviation τ'_j 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 us assume that the hypotheses of the lemma are satisfied. We keep the same notations, and we write $\pi = \langle(\tau_i)_{i\in\Pi}\rangle_{v_0}$. We have that $\mathsf{Cost}_i(\rho) = \mathsf{Cost}_i(\pi)$ for all $i \in \Pi$, since $\mathsf{Visit}(\mathfrak{p}) = \mathsf{Visit}(\rho)$. One can prove that $\pi = \mathfrak{p}\mathfrak{q}^\omega$ (see Figures 4.8 and 4.9).

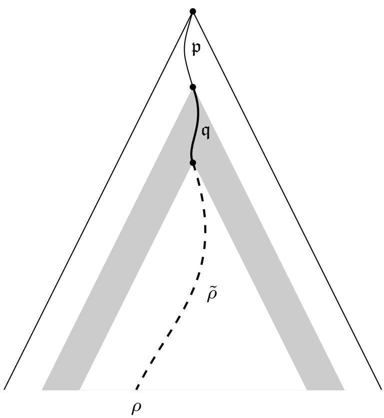


Figure 4.8: Play ρ and its prefix $\mathfrak{p}\mathfrak{q}$.

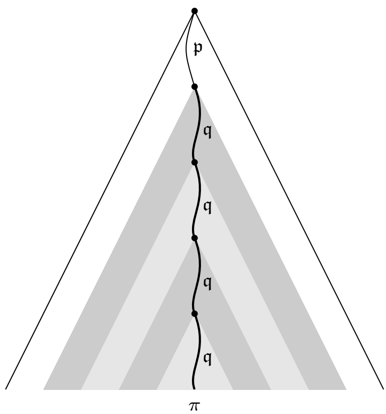


Figure 4.9: Play $\pi = \mathfrak{p}\mathfrak{q}^\omega$.

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 (τ'_j, τ_{-j}) from v_0. We show how to define a profitable deviation σ'_j for player j w.r.t. $(\sigma_i)_{i\in\Pi}$ from the deviation τ'_j. We distinguish the following two cases:

(i) The history $\mathfrak{pq}$ is not a prefix of π'.

We define $\sigma'_j := \tau'_j$. As in the first case of the proof of Lemma 4.1.14, we have $\mathsf{Cost}_j(\rho') < \mathsf{Cost}_j(\rho)$, which implies that σ'_j is a profitable deviation of player j w.r.t. $(\sigma_i)_{i\in\Pi}$.

(ii) The history $\mathfrak{pq}^k$ is a prefix of π', for a certain $k \geq 1$, with k being maximal.

We define for all histories $h \in \mathsf{Hist}_j$:

$$\sigma'_j(h) := \begin{cases} \sigma_j(h) & \text{if } \mathfrak{p} \not\leq h, \\ \tau'_j(\mathfrak{pq}^k\delta) & \text{if } h = \mathfrak{p}\delta \text{ for } \delta \in V^*. \end{cases}$$

Note that $\mathfrak{pq}^k\delta$ is indeed a history of the game as $\mathsf{Last}(\mathfrak{p}) = \mathsf{Last}(\mathfrak{pq})$. One can prove that

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

And then, from the point of view of costs, we have

$$\mathsf{Cost}_j(\rho') < \mathsf{Cost}_j(\pi') < y_j = x_j,$$

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

The next proposition achieves the first step of the proof of Theorem 4.1.12 as mentioned at the beginning of Section 4.1.4. 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 4.1.14 and 4.1.15.

Proposition 4.1.16. *Let $(\sigma_i)_{i\in\Pi}$ be a Nash equilibrium in a multiplayer quantitative reachability game $(\mathcal{G}, v_0)$. Then there exists a Nash equilibrium $(\tau_i)_{i\in\Pi}$ with the same type and such that $\langle(\tau_i)_{i\in\Pi}\rangle_{v_0} = \alpha\beta^\omega$, where $\mathsf{Visit}(\alpha) = \mathsf{Type}((\sigma_i)_{i\in\Pi})$ and $|\alpha\beta| < (|\Pi| + 1)\cdot|V|$.*

Proof. Let $(\sigma_i)_{i\in\Pi}$ be a Nash equilibrium in $(\mathcal{G}, v_0)$ and let ρ be its outcome in this game. Let us set $\Pi = \{1, \dots, n\}$, and assume, without loss of generality, that the players are ordered in the following way:

$$\begin{aligned}
&\mathsf{Cost}_1(\rho) \le \dots \le \mathsf{Cost}_k(\rho) < +\infty \\
&\mathsf{Cost}_{k+1}(\rho) = \dots = \mathsf{Cost}_n(\rho) = +\infty
\end{aligned}$$

for a certain k such that $0 \le k \le n$. We consider two cases:

(*i*) $\mathsf{Cost}_1(\rho) \ge |V|$.

Then, there exists a prefix $\mathfrak{pq}$ of ρ, with $\mathfrak{p}, \mathfrak{q} \in V^+$, such that

$$\begin{aligned}
&|\mathfrak{pq}| < \mathsf{Cost}_1(\rho) \\
&\mathsf{Visit}(\mathfrak{p}) = \mathsf{Visit}(\mathfrak{pq}) = \emptyset \\
&\mathsf{Last}(\mathfrak{p}) = \mathsf{Last}(\mathfrak{pq}).
\end{aligned}$$

We define the strategy profile $(\tau_i)_{i\in\Pi}$ as proposed in Lemma 4.1.14. By this lemma, it is actually a Nash equilibrium in $(\mathcal{G}, v_0)$. If we write $\pi = \langle(\tau_i)_{i\in\Pi}\rangle_{v_0}$, we have

$$\rho = \mathfrak{pq}\tilde{\rho} \quad \text{and} \quad \pi = \mathfrak{p}\tilde{\rho} \quad \text{for a certain } \tilde{\rho} \in V^\omega.$$

It follows that

$$\begin{aligned}
&\forall\, i \le k \qquad && \mathsf{Cost}_i(\pi) < \mathsf{Cost}_i(\rho), \text{ and} \\
&\forall\, i > k \qquad && \mathsf{Cost}_i(\pi) = \mathsf{Cost}_i(\rho) = +\infty.
\end{aligned}$$

(*ii*) $(\mathsf{Cost}_{l+1}(\rho) - \mathsf{Cost}_l(\rho)) \ge |V|$ for $1 \le l \le k-1$.

Then, there exists a prefix $\mathfrak{pq}$ of ρ, with $\mathfrak{p}, \mathfrak{q} \in V^+$, such that

$$\begin{aligned}
&\mathsf{Cost}_l(\rho) < |\mathfrak{pq}| < \mathsf{Cost}_{l+1}(\rho) \\
&\mathsf{Visit}(\mathfrak{p}) = \mathsf{Visit}(\mathfrak{pq}) = \{1, \dots, l\} \\
&\mathsf{Last}(\mathfrak{p}) = \mathsf{Last}(\mathfrak{pq}).
\end{aligned}$$

We define the strategy profile $(\tau_i)_{i\in\Pi}$ given in Lemma 4.1.14. It is then a Nash equilibrium in $(\mathcal{G}, v_0)$, and for $\pi = \langle(\tau_i)_{i\in\Pi}\rangle_{v_0}$, we have

$$\rho = \mathfrak{pq}\tilde{\rho} \quad \text{and} \quad \pi = \mathfrak{p}\tilde{\rho} \quad \text{for a certain } \tilde{\rho} \in V^\omega.$$

Hence, it holds that

$$\begin{aligned}
\mathsf{Cost}_i(\pi) &= \mathsf{Cost}_i(\rho) && \text{for } i \leq l\,; \\
\mathsf{Cost}_i(\pi) &< \mathsf{Cost}_i(\rho) && \text{for } l < i \leq k\,; \\
\mathsf{Cost}_i(\pi) &= \mathsf{Cost}_i(\rho) = +\infty && \text{for } k < i \leq n\,.
\end{aligned}$$

By applying finitely many times the two previous cases, we can assume, without loss of generality, that $(\sigma_i)_{i\in\Pi}$ is a Nash equilibrium such that

$$\begin{aligned}
\mathsf{Cost}_i(\rho) &< i \cdot |V| && \text{for } i \leq k\,; \\
\mathsf{Cost}_i(\rho) &= +\infty && \text{for } i > k\,.
\end{aligned}$$

Let us go further. We can write $\rho = \alpha\beta\tilde{\rho}$ such that $\alpha, \beta \in V^+$, $\tilde{\rho} \in V^\omega$, and

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

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

If we define the strategy profile $(\tau_i)_{i\in\Pi}$ like in Lemma 4.1.15, we get a Nash equilibrium in $(\mathcal{G}, v_0)$ with outcome $\alpha\beta^\omega$ and the same type as $(\sigma_i)_{i\in\Pi}$. □

We are now ready to prove Theorem 4.1.12.

Proof of Theorem 4.1.12. Let $(\mathcal{G}, v_0)$ be a multiplayer quantitative reachability game, and let $(\sigma_i)_{i\in\Pi}$ be a Nash equilibrium in this game. The first step consists in constructing a Nash equilibrium as in Proposition 4.1.16. Let us denote it again by $(\sigma_i)_{i\in\Pi}$, and set $\rho := \langle(\sigma_i)_{i\in\Pi}\rangle_{v_0}$. By this proposition, we have that $\rho = \alpha\beta^\omega$, $\mathsf{Visit}(\alpha) = \mathsf{Type}((\sigma_i)_{i\in\Pi})$, and $|\alpha\beta| < (|\Pi| + 1) \cdot |V|$. The strategy profile $(\sigma_i)_{i\in\Pi}$ is also a Nash equilibrium in the corresponding game $\mathcal{T}$ played on the unravelling T of G from v_0.

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

We know that $|\alpha\beta| < (|\Pi| + 1) \cdot |V|$, and so, we set $\gamma \in V^+$ such that $\alpha\beta\gamma$ is a prefix of ρ and $|\alpha\beta\gamma| = (|\Pi|+2)\cdot|V|$. Furthermore, we have $\mathsf{Last}(\alpha) = \mathsf{Last}(\alpha\beta)$ and $\mathsf{Visit}(\alpha) = \mathsf{Visit}(\alpha\beta\gamma)$ (since $\mathsf{Visit}(\alpha) = \mathsf{Type}(\rho)$). Then, this prefix $\alpha\beta\gamma$ satisfies the properties described in Lemma 4.1.7 (with $l = |\Pi|+1$). By Lemma 4.1.8, we conclude that there exists a finite-memory Nash equilibrium $(\tau_i)_{i\in\Pi}$ such that $\mathsf{Type}((\tau_i)_{i\in\Pi}) = \mathsf{Visit}(\alpha)$, that is, with the same type as the initial Nash equilibrium $(\sigma_i)_{i\in\Pi}$. $\square$

4.2 Quantitative Reachability or Safety Objectives

In this section, we extend our result of existence of finite-memory Nash equilibria in quantitative reachability games (Theorem 4.1.5) to *multiplayer quantitative reachability/safety games.*

These games are cost games where some players have *quantitative reachability objectives*, whereas others have *quantitative safety objectives.* As previously, the players with reachability objectives want to reach their goal set as soon as possible. The players with safety objectives want to avoid[6] their *bad set* or, if impossible, delay its visit as long as possible. Let us make that precise through the following definition.

Definition 4.2.1. A *multiplayer quantitative reachability/safety game* is a multiplayer cost game $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{Cost}_i)_{i\in\Pi})$ such that the set Π of players into Π_r and Π_s, the sets of players with reachability and safety objectives respectively, and for all $i \in \Pi_r$, $\mathsf{Cost}_i = \mathrm{RP}_{\mathrm{Min}}$ for a given goal set $\mathsf{R}_i \subseteq V$, and the shared price function $\phi : E \to \mathbb{R}$ with $\phi(e) = 1$ for all $e \in E$, while for all $i \in \Pi_s$, Cost_i is defined, for a given bad set $\mathsf{S}_i \subseteq V$,

6. Note that in Definition 2.2.8, a play satisfies a safety winning condition if it stays forever in a given goal set $\mathsf{S} \subseteq V$, which is equivalent to avoid the bad set $V \setminus \mathsf{S}$.

as:

$$\mathsf{Cost}_i(\rho) = \begin{cases} -l & \text{if } l \text{ is the } \textit{least} \text{ index such that } \rho_l \in \mathsf{S}_i, \\ -\infty & \text{otherwise,} \end{cases}$$

for any play ρ.
Abusively, such a game is written $\mathcal{G} = (\Pi, \Pi_r, \Pi_s, \mathcal{A}, (\mathsf{R}_i)_{i\in\Pi_r}, (\mathsf{S}_i)_{i\in\Pi_s})$.

As before, the aim of each player i is to *minimise* his cost, i.e. reach his goal set R_i as soon as possible for $i \in \Pi_r$, or delay the visit of S_i as long as possible for $i \in \Pi_s$. The main result of this section is the following theorem which solves Problem 1 for Nash equilibria in this framework.

Theorem 4.2.2. *In every initialised multiplayer quantitative reachability/safety game, there exists a finite-memory Nash equilibrium.*

In order to prove Theorem 4.2.2, we have to review the results of Section 4.1.4 for quantitative reachability games. In the context of quantitative reachability/safety games, the notation $\mathsf{Visit}(\rho)$ refers to the set of players $i \in \Pi_r$ and $j \in \Pi_s$ such that ρ visits R_i and S_j. Let us first notice that Lemma 4.1.7 remains true in this framework when player j belongs to Π_r. As a reminder, this lemma roughly says that, given a Nash equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$, if its outcome has a prefix that fulfils some conditions, then the coalition of the players $i \neq j$ wins the game $\mathcal{G}^j$ from any vertex of this prefix.

Furthermore, Lemma 4.1.8 remains true, however we have to slightly adapt its proof. Let us remind this lemma and prove it again in the current context.

Lemma 4.2.3. *Let $\mathcal{G} = (\Pi, \Pi_r, \Pi_s, \mathcal{A}, (\mathsf{R}_i)_{i\in\Pi_r}, (\mathsf{S}_i)_{i\in\Pi_s})$ be a multiplayer quantitative reachability/safety game, and $\mathcal{T}$ be the corresponding game played on the unravelling of G from a vertex v_0. For any depth $d \in \mathbb{N}$, let $(\sigma_i)_{i\in\Pi}$ be a Nash equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$, and $\alpha\beta\gamma$ be* *a prefix of* $\rho = \langle(\sigma_i)_{i\in\Pi}\rangle_{v_0}$, *such that* $\alpha, \beta, \gamma \in V^+$, *and*

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

for some $l \geq 1$.
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* $\mathsf{Type}((\tau_i)_{i\in\Pi}) = \mathsf{Visit}(\alpha)$.

Proof. As in the proof of Lemma 4.1.8, we denote by $\mu^v_{\mathtt{C}j}$ the memoryless winning strategy of the coalition $\Pi \setminus \{j\}$ given by Lemma 4.1.7 (when its hypotheses are satisfied) and Theorem 2.2.15, and we write $\mu^v_{i,j}$ for the strategy of player $i \neq j$ derived from $\mu^v_{\mathtt{C}j}$.

Moreover, we denote by Π^f_r (resp. Π^f_s) the subset of players $i \in \Pi_r$ (resp. $i \in \Pi_s$) such that α visits R_i (resp. S_i) and by Π^∞_r (resp. Π^∞_s) the set $\Pi_r \setminus \Pi^f_r$ (resp. $\Pi_s \setminus \Pi^f_s$).

The punishment function P is defined exactly as in the proof of Lemma 4.1.8. We remind that for v_0, we define $P(v_0) = \bot$, and for every history $hv \in \mathsf{Hist}$ ($v \in V$) starting in v_0 , we let:

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

The difference with the proof of Lemma 4.1.8 arises in the definition of the Nash equilibrium. A Nash equilibrium in this context needs to incorporate an adequate punishment for the players with safety objectives. More precisely, in order to dissuade a player $j \in \Pi^f_s$ from deviating, the other players punish him by playing the strategies $(\sigma_i)_{i\in\Pi\setminus\{j\}}$ in $\mathsf{Trunc}_d(T)$. Notice that a player $j \in \Pi^\infty_s$ has no incentive to deviate.

Formally, we define the Nash equilibrium $(\tau_i)_{i\in\Pi}$ as follows. For $h \in \mathsf{Hist}_i$,

$$\tau_i(h) := \begin{cases} v & \text{if } P(h) = \bot \ (h < \alpha\beta^\omega)\text{; such that } hv < \alpha\beta^\omega, \\ \sigma_i(h) & \text{if } P(h) \neq \bot, i,\ P(h) \in \Pi_r^f \cup \Pi_s^f \text{ and } |h| < d, \\ \mu^v_{i,P(h)}(v'') & \text{if } P(h) \neq \bot, i \text{ and } P(h) \in \Pi_r^\infty, \\ & \quad \text{s.t. } \big(h = h'vv'h''v'' \ (v, v', v'' \in V), \\ & \quad\ P(h'v) = \bot, \text{ and } P(h'vv') = P(h)\big), \\ \textit{arbitrary} & \text{otherwise,} \end{cases}$$

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 $\mathsf{Type}((\tau_i)_{i\in\Pi})$ is equal to $\mathsf{Visit}(\alpha)$ $(= \mathsf{Visit}(\alpha\beta))$.

It remains to prove that $(\tau_i)_{i\in\Pi}$ is a finite-memory Nash equilibrium in the game $\mathcal{T}$. In order to do so, we prove that none of the players has a profitable deviation. For players with reachability objectives, the arguments are exactly the same as the ones provided in the proof of Lemma 4.1.8. Let us now consider players with safety objectives. In the case where $j \in \Pi_s^\infty$, player j has clearly no incentive to deviate. In the case where $j \in \Pi_s^f$, player j has no incentive to deviate after the prefix α if he wants to decrease his cost (recall that α visits S_j). Thus, we assume that the strategy τ'_j causes a deviation from a vertex visited in α. By definition of $(\tau_i)_{i\in\Pi}$, the other players first play according to σ_{-j} in $\mathsf{Trunc}_d(\mathcal{T})$, 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 us set $\pi = \langle(\tau_i)_{i\in\Pi}\rangle_{v_0}$ and $\pi' = \langle\tau'_j, \tau_{-j}\rangle_{v_0}$. Then,

$$\mathsf{Cost}_j(\pi') < \mathsf{Cost}_j(\pi).$$

On the other hand, we know that

$$\mathsf{Cost}_j(\pi) = \mathsf{Cost}_j(\rho) \leq |\alpha|.$$

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

$$\mathsf{Cost}_j(\rho') \leq \mathsf{Cost}_j(\pi') < \mathsf{Cost}_j(\rho).$$

Notice that we do not necessarily have that $\mathsf{Cost}_j(\rho') = \mathsf{Cost}_j(\pi')$ (as in the proof of Lemma 4.1.8) since the bad set S_j can be visited by π' and not by ρ' (if it is visited after depth d). As the play ρ' is consistent with the strategies σ_{-j} by definition of $(\tau_i)_{i\in\Pi}$, the strategy τ'_j restricted to the tree $\mathsf{Trunc}_d(T)$ is a profitable deviation for player j w.r.t. $(\sigma_i)_{i\in\Pi}$ in the game $\mathsf{Trunc}_d(\mathcal{T})$. This is impossible. Moreover, as done in the proof of Lemma 4.1.8, $(\tau_i)_{i\in\Pi}$ is a finite-memory strategy profile. □

Thanks to Lemma 4.2.3, Proposition 4.1.6 ensures that the Nash equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$ provided by Kuhn's theorem (Corollary 2.3.23) can be lifted to $\mathcal{T}$. This proves Theorem 4.2.2.

4.3 General Quantitative Reachability Objectives

In this section, we come back to a pure reachability framework and we extend our model in the following way: we assume that edges are labelled with tuples of *positive* prices (one price for each player). Here we do not only count the number of edges to reach the goal of a player, but we sum up his prices along the path until his goal is reached. His aim is still to minimise his global cost for a play. We generalise Definition 4.1.1, and extend Theorem 4.1.5 to these games.

Definition 4.3.1. A *multiplayer quantitative reachability game with tuples of prices on edges* is a multiplayer cost game $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{Cost}_i)_{i\in\Pi})$ such that for any $i \in \Pi$, $\mathsf{Cost}_i = \mathrm{RP}_{\mathrm{Min}}$ for a given goal set $\mathsf{R}_i \subseteq V$, and a given price function $\phi_i : E \to \mathbb{R}^{>0}$.

Abusively, such a game is denoted by $\mathcal{G} = (\Pi, \mathcal{A}, (\phi_i)_{i\in\Pi}, (\mathsf{R}_i)_{i\in\Pi})$.

Then, in such a game, the cost function Cost_i of player i is defined, for any play $\rho = \rho_0\rho_1\ldots$, as[7]:

$$\mathsf{Cost}_i(\rho) = \begin{cases} \sum_{j=1}^{n} \phi_i(\rho_{j-1}, \rho_j) & \text{if } n \text{ is the } \textit{least} \text{ index s.t. } \rho_n \in \mathsf{R}_i, \\ +\infty & \text{otherwise.} \end{cases}$$

7. See Definition 2.3.3.

We also positively solve Problem 1 for Nash equilibria in this context.

Theorem 4.3.2. *In every initialised multiplayer quantitative reachability game with tuples of prices on edges, there exists a finite-memory Nash equilibrium.*

To prove Theorem 4.3.2, we follow the same scheme as in Section 4.1.4 for quantitative reachability games. In particular, we rely on Kuhn's theorem (Corollary 2.3.23) and need to prove a counterpart of Lemma 4.1.7, Lemma 4.1.8 and Proposition 4.1.6 in this framework.

Let us first introduce some notations that will be useful in this context. We define $\mathsf{c_{min}} := \min_{i\in\Pi}\min_{e\in E}\phi_i(e)$, $\mathsf{c_{max}} := \max_{i\in\Pi}\max_{e\in E}\phi_i(e)$ (resp. the minimal and maximal price appearing in the graph), and $\mathsf{K} := \left\lceil \frac{\mathsf{c_{max}}}{\mathsf{c_{min}}} \right\rceil$. It is clear that $\mathsf{c_{min}}, \mathsf{c_{max}} > 0$ (since $\phi_i : E \to \mathbb{R}^{>0}$ for all $i \in \Pi$), and $\mathsf{K} \geq 1$. Furthermore, for any play $\rho = \rho_0\rho_1\ldots$ of $\mathcal{G}$ and any player $i \in \Pi$, we define $\mathsf{Index}_i(\rho)$ as the least index l such that $\rho_l \in \mathsf{R}_i$ if it exists, or -1 if not [8].

The counterpart of Lemma 4.1.7 is the following one, taking into account the constant K defined before (see the length of $\alpha\beta\gamma$).

Lemma 4.3.3. *Let $\mathcal{G} = (\Pi, \mathcal{A}, (\phi_i)_{i\in\Pi}, (\mathsf{R}_i)_{i\in\Pi})$ be a multiplayer quantitative reachability game with tuples of prices on edges, and $\mathcal{T}$ be the corresponding game played on the unravelling of G from a vertex v_0. For any depth $d \in \mathbb{N}$, let $(\sigma_i)_{i\in\Pi}$ be a Nash equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$, and ρ the (finite) outcome of $(\sigma_i)_{i\in\Pi}$. Assume that ρ has a prefix $\alpha\beta\gamma$, where $\alpha, \beta, \gamma \in V^+$, such that*

$$\mathsf{Visit}(\alpha) = \mathsf{Visit}(\alpha\beta\gamma)$$
$$\mathsf{Last}(\alpha) = \mathsf{Last}(\alpha\beta)$$
$$|\alpha\beta| \leq l \cdot |V|$$
$$|\alpha\beta\gamma| = (l + \mathsf{K}) \cdot |V|$$

for some $l \geq 1$.
Let $j \in \Pi$ be such that α does not visit R_j, and let us consider the zero-sum qualitative reachability game $\mathcal{G}^j = (\mathcal{A}^j, \mathsf{R}_j)$. Then for all histories hv

8. We are conscious that it is counter-intuitive to use the particular value -1, but it is helpful in the proofs of Section 5.2.

of $(\mathcal{G}, v_0)$ *(with* $v \in V$*) consistent with* σ_{-j} *and such that* $|hv| \leq |\alpha\beta|$*, the coalition of the players* $i \neq j$ *wins the game* $\mathcal{G}_j$ *from* v*.*

Sketch of proof. As for the proof of Lemma 4.1.7, we proceed by contradiction, and define a play ρ' in the very same way. It follows that

$$\begin{aligned}
\mathsf{Index}_j(\rho') &\leq |hv| + |V| && \text{(by Theorem 2.2.15)}\\
&\leq (l+1)\cdot|V| && \text{(by hypothesis)}\\
&\leq (l+\mathsf{K})\cdot|V| && \text{(as } \mathsf{K} \geq 1)\\
&\leq d && \text{(as } \alpha\beta\gamma \leq \rho).
\end{aligned}$$

Then, $\mathsf{Cost}_j(\rho') < +\infty$. If $\mathsf{Cost}_j(\rho) = +\infty$, then $\mathsf{Cost}_j(\rho') < \mathsf{Cost}_j(\rho)$.

If $\mathsf{Cost}_j(\rho) < +\infty$, we write, as before, $\phi_j(hv)$ for the sum of the prices of player j along the prefix hv. We have the following inequalities (see Figure 4.10):

$$\begin{aligned}
\mathsf{Cost}_j(\rho') &\leq \phi_j(hv) + \mathsf{c_{max}}\cdot|V|\\
\mathsf{Cost}_j(\rho) &> \phi_j(hv) + \mathsf{c_{min}}\cdot\mathsf{K}\cdot|V| && \text{(as } \mathsf{Index}_j(\rho) > (l+\mathsf{K})\cdot|V|)\\
&\geq \phi_j(hv) + \mathsf{c_{min}}\cdot\frac{\mathsf{c_{max}}}{\mathsf{c_{min}}}\cdot|V| && \text{(by definition of } \mathsf{K})\\
&= \phi_j(hv) + \mathsf{c_{max}}\cdot|V|\,.
\end{aligned}$$

Then, we have $\mathsf{Cost}_j(\rho') < \mathsf{Cost}_j(\rho)$.

In both cases, since ρ' is consistent with σ_{-j}, the strategy of player j induced by the play ρ' is a profitable deviation for player j w.r.t. $(\sigma_i)_{i\in\Pi}$. This contradicts the fact that $(\sigma_i)_{i\in\Pi}$ is a Nash equilibrium in the game $\mathsf{Trunc}_d(\mathcal{T})$. ☐

The following lemma is the counterpart of Lemma 4.1.8.

Lemma 4.3.4. *Let* $\mathcal{G} = (\Pi, \mathcal{A}, (\phi_i)_{i\in\Pi}, (\mathsf{R}_i)_{i\in\Pi})$ *be a multiplayer quantitative reachability game with tuples of prices on edges, and* $\mathcal{T}$ *be the corresponding game played on the unravelling of* G *from a vertex* v_0*. For any depth* $d \in \mathbb{N}$*, let* $(\sigma_i)_{i\in\Pi}$ *be a Nash equilibrium in* $\mathsf{Trunc}_d(\mathcal{T})$*, and* $\alpha\beta\gamma$ *be a prefix of* $\rho = \langle(\sigma_i)_{i\in\Pi}\rangle_{v_0}$ *as defined in Lemma 4.3.3 where* $|\alpha\beta\gamma| = (l+\mathsf{K})\cdot|V|$ *for some* $l \geq 1$ *such that* $l \leq \frac{d}{|V|\cdot\mathsf{K}}$*. 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* $\mathsf{Type}((\tau_i)_{i\in\Pi}) = \mathsf{Visit}(\alpha)$*.*

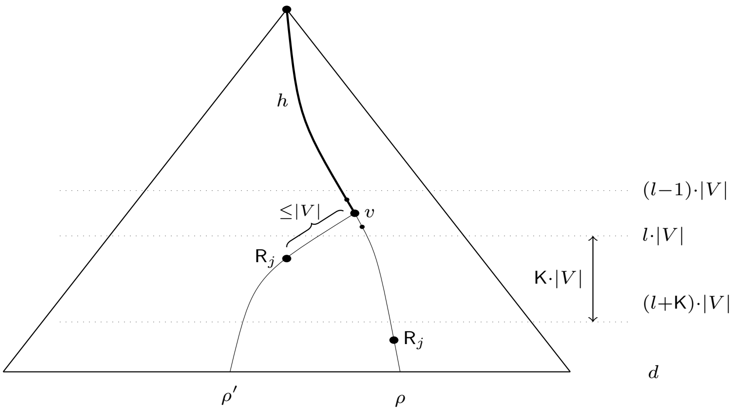


Figure 4.10: Plays ρ and ρ' with their common prefix hv.

Proof. We prove this result in the very same way as Lemma 4.1.8. We define $(\tau_i)_{i\in\Pi}$ exactly as in the proof of this lemma, and we show that it is a Nash equilibrium in $\mathcal{T}$. The only difference lies in the case[9] $j \leq k$. We suppose that τ'_j is a profitable deviation for player j w.r.t. $(\tau_i)_{i\in\Pi}$ in the game $\mathcal{T}$. So we have $\mathsf{Cost}_j(\pi') < \mathsf{Cost}_j(\pi)$, where $\pi = \langle(\tau_i)_{i\in\Pi}\rangle_{v_0}$ and $\pi' = \langle\tau'_j, \tau_{-j}\rangle_{v_0}$. As $\mathsf{Index}_j(\pi) \leq |\alpha|$, we know that $\mathsf{Cost}_j(\pi) \leq |\alpha| \cdot \mathsf{c_{max}}$. It follows that $\mathsf{Cost}_j(\pi') < |\alpha| \cdot \mathsf{c_{max}}$ and

$$\begin{aligned}\mathsf{Index}_j(\pi') &< |\alpha| \cdot \frac{\mathsf{c_{max}}}{\mathsf{c_{min}}} \\ &\leq l \cdot |V| \cdot \mathsf{K} \\ &\leq d && \text{(by hypothesis).}\end{aligned}$$

The first inequality can be justified as follows. For a contradiction, let us assume that $\mathsf{Index}_j(\pi') \geq |\alpha| \cdot \frac{\mathsf{c_{max}}}{\mathsf{c_{min}}}$. It follows that $\mathsf{Cost}_j(\pi') \geq \mathsf{c_{min}} \cdot |\alpha| \cdot \frac{\mathsf{c_{max}}}{\mathsf{c_{min}}}$, this contradicts the fact that $\mathsf{Cost}_j(\pi') < |\alpha| \cdot \mathsf{c_{max}}$.

As in the proof of Lemma 4.1.8, we limit the play π' in $\mathcal{T}$ to its prefix of length d and get a profitable deviation for player j w.r.t. $(\sigma_i)_{i\in\Pi}$ in the

9. Indeed, when $j > k$, i.e. when player j has not reached his goal set along $\alpha\beta^\omega$, the coalition punishes him in the exact same way as Lemma 4.1.8 by preventing him from visiting his goal set.

game $\mathsf{Trunc}_d(\mathcal{T})$, contradicting the fact that $(\sigma_i)_{i\in\Pi}$ is a Nash equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$.

Moreover, as done in the proof of Lemma 4.1.8, $(\tau_i)_{i\in\Pi}$ is a finite-memory strategy profile. □

As a consequence of the two previous lemmas, Proposition 4.1.6 remains true in this context, we only have to adjust the depth d of the finite tree.

Proposition 4.3.5. *Let $(\mathcal{G}, v_0)$ be a multiplayer quantitative reachability game with tuples of prices on edges, and $\mathcal{T}$ be the corresponding game played on the unravelling of G from v_0. If there exists a Nash equilibrium in the game $\mathsf{Trunc}_d(\mathcal{T})$ where $d = \max\{(|\Pi|+1)\cdot(\mathsf{K}+1)\cdot|V|, (|\Pi|\cdot(\mathsf{K}+1)+1)\cdot|V|\cdot\mathsf{K}\}$, then there exists a finite-memory Nash equilibrium in the game $\mathcal{T}$.*

Proof. The proof is similar to the proof of Proposition 4.1.6. Let $(\sigma_i)_{i\in\Pi}$ be a Nash equilibrium in the game $\mathsf{Trunc}_d(\mathcal{T})$ and ρ its outcome. We consider the prefix $\mathfrak{pq}$ of ρ of minimal length such that

$$\begin{aligned} \exists\, l \geq 1 \qquad & |\mathfrak{p}| = (l-1)\cdot|V| \\ & |\mathfrak{pq}| = (l+\mathsf{K})\cdot|V| \\ & \mathsf{Visit}(\mathfrak{p}) = \mathsf{Visit}(\mathfrak{pq}). \end{aligned}$$

In the worst case, the play ρ visits the goal set of a new player in each prefix of length $i\cdot(\mathsf{K}+1)\cdot|V|$, for $1 \leq i \leq |\Pi|$, i.e. $|\mathfrak{p}| = |\Pi|\cdot(\mathsf{K}+1)\cdot|V|$. So we know that $l \leq |\Pi|\cdot(\mathsf{K}+1)+1$ and $\mathfrak{pq}$ exists as a prefix of ρ, because the length d of ρ is greater or equal to $(|\Pi|+1)\cdot(\mathsf{K}+1)\cdot|V|$ by hypothesis.

Given the length of $\mathfrak{q}$ ($\mathsf{K} \geq 1$), one vertex of V is visited at least twice by $\mathfrak{q}$. More precisely, we can write

$$\begin{aligned} \mathfrak{pq} = \alpha\beta\gamma \quad \text{with} \quad & \mathsf{Last}(\alpha) = \mathsf{Last}(\alpha\beta) \\ & |\alpha| \geq (l-1)\cdot|V| \\ & |\alpha\beta| \leq l\cdot|V|. \end{aligned}$$

We have $\mathsf{Visit}(\alpha) = \mathsf{Visit}(\alpha\beta\gamma)$, and $|\alpha\beta\gamma| = (l+\mathsf{K})\cdot|V|$.

Moreover, the following inequality holds:

$$d \geq (|\Pi| \cdot (\mathsf{K}+1)+1) \cdot |V| \cdot \mathsf{K} \geq l \cdot |V| \cdot \mathsf{K} \quad \text{and so,} \quad l \leq \frac{d}{|V| \cdot \mathsf{K}}.$$

Then, we can apply Lemma 4.3.4 and get a finite-memory Nash equilibrium $(\tau_i)_{i\in\Pi}$ in the game $\mathcal{T}$ such that $\mathsf{Type}((\tau_i)_{i\in\Pi}) = \mathsf{Visit}(\alpha)$. □

Thanks to Kuhn's theorem (Corollary 2.3.23) and Proposition 4.3.5, one can easily deduce Theorem 4.3.2.

Remark 4.3.6. Let us comment on the depth d that is chosen in Proposition 4.3.5. It is defined as the maximum between $d_1 := (|\Pi|+1)\cdot(\mathsf{K}+1)\cdot|V|$ and $d_2 := (|\Pi| \cdot (\mathsf{K}+1)+1) \cdot |V| \cdot \mathsf{K}$. One can easily prove that $d_1 < d_2$ if and only if $\mathsf{K}^2 > \frac{|\Pi|+1}{|\Pi|}$.

We now investigate an alternative method to handle *simple price functions*. More precisely, we only consider price functions $(\phi_i)_{i\in\Pi}$ such that for all $i, j \in \Pi$, we have that $\phi_i = \phi_j$ and $\phi_i : E \to \mathbb{N}_0$. In other words, it means that there is a unique non-zero natural price on every edge. Later on we are going to compare the depths of the finite trees obtained by the two methods.

In the case of these simple price functions, we can directly deduce Theorem 4.3.2 by replacing any edge of price c by a path of length c composed of c new edges (of price 1) and then applying Theorem 4.1.5 on this new game. If we write $\mathcal{A}' = (V', (V_i)_{i\in\Pi}, E')$ the new arena obtained by adding new vertices and edges when necessary, it holds that:

$$\begin{aligned} |V'| &\leq |V| + (\mathsf{c_{max}} - 1) \cdot |E| \\ &\leq |V| + (\mathsf{c_{max}} - 1) \cdot |V|^2, \text{ and} \\ |E'| &\leq \mathsf{c_{max}} \cdot |E| \,. \end{aligned}$$

If we apply Proposition 4.1.6, the depth d' of the finite tree that is considered satisfies:

$$\begin{aligned} d' &= (|\Pi|+1) \cdot 2 \cdot |V'| \\ &\leq (|\Pi|+1) \cdot 2 \cdot (|V| + (\mathsf{c_{max}} - 1) \cdot |E|) \\ &\leq (|\Pi|+1) \cdot 2 \cdot \left(|V| + (\mathsf{c_{max}} - 1) \cdot |V|^2\right) \,. \end{aligned}$$

Whereas if we apply Proposition 4.3.5 directly on the initial game $\mathcal{G}$, we have the following equality:

$$d = \max\{(|\Pi| + 1) \cdot (\mathsf{K} + 1) \cdot |V|, (|\Pi| \cdot (\mathsf{K} + 1) + 1) \cdot |V| \cdot \mathsf{K}\} .$$

Let us first notice that if all the edges of $\mathcal{G}$ are labelled with the same price (i.e., $\mathsf{c}_{\mathsf{max}} = \mathsf{c}_{\mathsf{min}}$ and $\mathsf{K} = 1$), then

$$\begin{array}{lcl} d' & = & (|\Pi| + 1) \cdot 2 \cdot (|V| + (\mathsf{c}_{\mathsf{max}} - 1) \cdot |E|), \text{ and} \\ d & = & (|\Pi| + 1) \cdot 2 \cdot |V| . \end{array}$$

And so,

$$\begin{array}{lll} \text{if } \mathsf{c}_{\mathsf{max}} = \mathsf{c}_{\mathsf{min}} = 1, & \text{then} & d' = d = (|\Pi| + 1) \cdot 2 \cdot |V|, \text{ and} \\ \text{if } \mathsf{c}_{\mathsf{max}} = \mathsf{c}_{\mathsf{min}} > 1, & \text{then} & d' > d . \end{array}$$

When $\mathsf{K} > 1$, the comparison between d and d' depends on the values of many parameters of the game. For example, if the graph of the game has five vertices, three edges of price 1 and one edge of price 100, then it is more interesting to study the game played on $\mathcal{A}'$ and use techniques of the proof of Theorem 4.1.5 to construct the Nash equilibrium, because in this case, $d' = (\Pi + 1) \cdot 2 \cdot 104$ and $d = (|\Pi| \cdot 101 + 1) \cdot 5 \cdot 101$, and so $d >> d'$.

Remark 4.3.7. The problem of deciding, given an initialised multiplayer quantitative reachability game with tuples of prices on edges, and a tuple of thresholds $(t_i)_{i \in \Pi} \in (\mathbb{R} \cup \{+\infty\})^{|\Pi|}$, whether there exists a Nash equilibrium with cost profile at most $(t_i)_{i \in \Pi}$, is NP-complete (see [KLŠT12]). Let us notice that in [KLŠT12], null prices on the edges are allowed. However, even for multiplayer quantitative reachability games (without tuples of prices on edges), this decision problem is NP-hard (it can be deduced from the proof of [Umm05, Proposition 6.29], see also [CMJ04]).

4.4 Various Objectives

In this section, we show the existence of Nash equilibria in a large class of multiplayer cost games. Section 4.4.1 presents and proves the results we obtained, and Section 4.4.2 exhibits some particular classes of cost games in which our results apply.

4.4.1 Results

We here define a large class of cost games for which Problems 4 and 5 can be answered positively (see Theorem 4.4.14). For the sake of clarity, we first show three propositions (Propositions 4.4.6, 4.4.11 and 4.4.12), which state the existence of Nash equilibria in different subclasses of cost games.

The general philosophy of our work is as follows: we try to derive existence of simple Nash equilibria in multiplayer cost games (and characterisation of their complexity) through determinacy results (and characterisation of the optimal strategies) of several well-chosen Min-Max cost games derived from the multiplayer game. These ideas were already successfully exploited in the qualitative framework [GU08], and in the case of limit-average objectives [TR98].

To describe the interesting class of cost games where Nash equilibria exist, we need the concepts of *prefix-independent*, *prefix-linear* and *coalition-determined* cost functions.

Definition 4.4.1. Given a multiplayer cost game $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{Cost}_i)_{i\in\Pi})$, a cost function Cost_i is *prefix-independent* in $\mathcal{G}$ if, for every vertex $v \in V$, every history $hv \in \mathsf{Hist}$, and every play $\rho \in \mathsf{Plays}$ with $\mathsf{First}(\rho) = v$, we have:

$$\mathsf{Cost}_i(h\rho) = \mathsf{Cost}_i(\rho)\,.$$

Note that usually the definition of prefix-independent does not depend on the selected game.

For example, the cost functions $\mathrm{AP}_{\mathrm{Min}}$ and $\mathrm{PRAvg}_{\mathrm{Min}}$ of Definition 2.3.3 are prefix-independent in any cost game (see the proof of Corollary 4.4.15). As a counterexample, let us consider the cost function $\mathrm{RP}_{\mathrm{Min}}$ in the two-player cost game of Example 2.3.15 on page 48, with goal set $\mathsf{R} = \{C\}$. While choosing $v = B$, $h = A$ and $\rho = (BC)^\omega$, we have that $\mathrm{RP}_{\mathrm{Min}}(h\rho) = \mathrm{RP}_{\mathrm{Min}}(A(BC)^\omega) = 2$, whereas $\mathrm{RP}_{\mathrm{Min}}(\rho) = \mathrm{RP}_{\mathrm{Min}}((BC)^\omega) = 1$.

Definition 4.4.2. Given a multiplayer cost game $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{Cost}_i)_{i\in\Pi})$, a cost function Cost_i is *prefix-linear* in $\mathcal{G}$ if, for every vertex $v \in V$ and history $hv \in \mathsf{Hist}$, there exists $a \in \mathbb{R}$ and $b \in \mathbb{R}^+$ such that, for every play $\rho \in \mathsf{Plays}$ with $\mathsf{First}(\rho) = v$, we have:

$$\mathsf{Cost}_i(h\rho) = a + b \cdot \mathsf{Cost}_i(\rho)\,.$$

In particular, prefix-independent cost functions are special cases of prefix-linear cost functions (take $a = 0$ and $b = 1$ for all histories hv). Let us show an example of a one-player cost game where the cost function of the player is not prefix-linear.

Example 4.4.3. Multiplayer cost games allow to encode energy games. In our framework, where each player aims at minimising his cost, an energy objective [BFL+08] (with threshold $T \in \mathbb{R}$ and price function $\phi : E \to \mathbb{R}$) could be encoded as follows:

$$\mathsf{Cost}(\rho) = \begin{cases} \sup_{n \geq 0} \phi(\rho_{\leq n}) & \text{if } \sup_{n \geq 0} \phi(\rho_{\leq n}) \leq T \\ +\infty & \text{otherwise,} \end{cases}$$

with $\phi(\rho_{\leq n}) = \sum_{i=1}^{n} \phi(\rho_{i-1}, \rho_i)$.

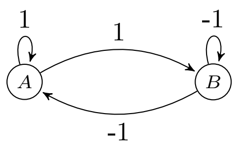


Figure 4.11: A game where the cost function which is not prefix-linear.

Let us consider the one-player cost game with an energy objective played on the arena depicted in Figure 4.11. We show that the cost function Cost, with threshold $T = 2$ and price function as in the figure, is not prefix-linear in this game. For this, we exhibit a history $hv \in \mathsf{Hist}$ such that for all a, $b \in \mathbb{R}$ there exists a play $\rho \in \mathsf{Plays}$ with $\mathsf{First}(\rho) = v$, such that $\mathsf{Cost}(h\rho) \neq a + b \cdot \mathsf{Cost}(\rho)$. We in fact give a play ρ independent of a and b. Let hv be the history AA and ρ be the play $A(AB)^\omega$. We have that $\mathsf{Cost}(\rho) = 2$ and $\mathsf{Cost}(h\rho) = \mathsf{Cost}(AA(AB)^\omega) = +\infty$, since $\sup_{n \geq 0} \phi((h\rho)_{\leq n}) = 3$, which is above the threshold $T = 2$. It is thus impossible to find a, $b \in \mathbb{R}$ such that:

$$+\infty = \mathsf{Cost}(h\rho) = a + b \cdot \mathsf{Cost}(\rho) = a + b \cdot 2.$$

Notice that if a could be infinite, there would still be a problem. Indeed, let us fix $a = +\infty$ and $\rho = (AB)^\omega$. Then, we have that $\mathsf{Cost}(\rho) = 1$, and $\mathsf{Cost}(h\rho) = \mathsf{Cost}(A(AB)^\omega) = 2 \neq +\infty$.

Let us now define the concept of *coalition-determined* cost function.

Definition 4.4.4. Given an arena $\mathcal{A} = (V, (V_i)_{i\in\Pi}, E)$ and a multiplayer cost game $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{Cost}_i)_{i\in\Pi})$, a cost function Cost_i is *(positionally/finite-memory) coalition-determined* in $\mathcal{G}$ if there exists a gain function $\mathsf{Gain}^i_{\mathrm{Max}} : \mathsf{Plays} \to \mathbb{R} \cup \{+\infty, -\infty\}$ such that the Min-Max cost game[10] $\mathcal{G}^i = (\mathcal{A}^i, \mathsf{Cost}_i, \mathsf{Gain}^i_{\mathrm{Max}})$, where $\mathcal{A}^i = (V, (V_i, V \setminus V_i), E)$ and player i (player Min) plays against the coalition $\Pi \setminus \{i\}$ (player Max), is determined and has (positional/finite-memory) optimal strategies for both players.
That is: $\exists\, \sigma_i^\star \in \Sigma_{\mathrm{Min}}$, $\exists\, \sigma_{\mathtt{C}i}^\star \in \Sigma_{\mathrm{Max}}$ (both positional/finite-memory) such that $\forall v \in V$,

$$\inf_{\sigma_i \in \Sigma_{\mathrm{Min}}} \mathsf{Gain}^i_{\mathrm{Max}}(\langle \sigma_i, \sigma_{\mathtt{C}i}^\star \rangle_v) = \mathsf{Val}^i(v) = \sup_{\sigma_{\mathtt{C}i} \in \Sigma_{\mathrm{Max}}} \mathsf{Cost}_i(\langle \sigma_i^\star, \sigma_{\mathtt{C}i} \rangle_v)\,.$$

Given $i \in \Pi$, note that $\mathcal{G}^i$ does not depend on the cost functions Cost_j, with $j \neq i$.

Example 4.4.5. Let us consider the two-player cost game $\mathcal{G}$ (the same as in Example 2.3.15 on page 48), whose arena $\mathcal{A}$ is depicted in Figure 4.12, and where player 1 has a quantitative reachability objective ($\mathsf{Cost}_1 = \mathrm{RP}_{\mathrm{Min}}$ for the goal set $\mathsf{R} = \{C\}$) and player 2 has a mean-payoff objective ($\mathsf{Cost}_2 = \mathrm{AP}_{\mathrm{Min}}$). We show that both cost functions[11] are positionally coalition-determined in this game.

Let us set $\mathsf{Gain}^1_{\mathrm{Max}} = \mathsf{Cost}_1$ and study the Min-Max cost game $\mathcal{G}^1 = (\mathcal{A}^1, \mathsf{Cost}_1, \mathsf{Gain}^1_{\mathrm{Max}})$, where player Min (resp. Max) is player 1 (resp. 2) and wants to minimise Cost_1 (resp. maximise $\mathsf{Gain}^1_{\mathrm{Max}}$). This game (exactly the same as the Min-Max cost game of Example 2.3.2, on page 40) is positionally determined by Theorem 2.3.10. In Example 2.3.9 (on page 45), we proved that the positional strategies $\sigma_1^\star$ and $\sigma_{\mathtt{C}1}^\star$ of player 1 and player 2 respectively, defined as: $\sigma_1^\star(A) = B$ and $\sigma_{\mathtt{C}1}^\star(B) = A$, are

10. Remember that in a Min-Max cost game, it must hold that $\mathsf{Cost}_i \geq \mathsf{Gain}^i_{\mathrm{Max}}$.
11. See Definition 2.3.3 for the description of the $\mathrm{RP}_{\mathrm{Min}}$ and $\mathrm{AP}_{\mathrm{Min}}$ functions.

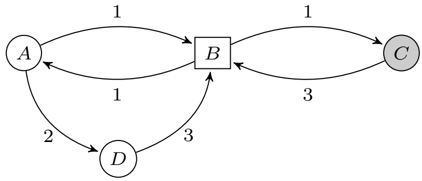


Figure 4.12: A two-player cost game.

optimal in $\mathcal{G}^1$. From A, their outcome is $\langle(\sigma_1^\star, \sigma_{\mathtt{C}1}^\star)\rangle_A = (AB)^\omega$, and $\mathsf{Cost}_1((AB)^\omega) = \mathsf{Gain}^1_{\mathrm{Max}}((AB)^\omega) = +\infty$. Then, Cost_1 is positionally coalition-determined in $\mathcal{G}$.

Note that the positional strategy $\tilde{\sigma}_1^\star$ defined by $\tilde{\sigma}_1^\star(A) = D$ is also optimal (for player 1) in $\mathcal{G}^1$. With this strategy, we have that $\langle(\tilde{\sigma}_1^\star, \sigma_{\mathtt{C}1}^\star)\rangle_A = (ADB)^\omega$, and $\mathsf{Cost}_1((ADB)^\omega) = \mathsf{Gain}^1_{\mathrm{Max}}((ADB)^\omega) = +\infty$.

We now examine the Min-Max cost game $\mathcal{G}^2 = (\mathcal{A}^2, \mathsf{Cost}_2, \mathsf{Gain}^2_{\mathrm{Max}})$, where $\mathsf{Gain}^2_{\mathrm{Max}} = \mathrm{AP}_{\mathrm{Max}}$. In this game, player Min (resp. Max) is player 2 (resp. 1) and wants to minimise Cost_2 (resp. maximise $\mathsf{Gain}^2_{\mathrm{Max}}$). This game is also positionally determined (Theorem 2.3.10). Let $\sigma_2^\star$ and $\sigma_{\mathtt{C}2}^\star$ be the positional strategies of player 2 and player 1 respectively, defined as follows: $\sigma_2^\star(B) = C$ and $\sigma_{\mathtt{C}2}^\star(A) = D$. From A, their outcome is $\langle(\sigma_2^\star, \sigma_{\mathtt{C}2}^\star)\rangle_A = AD(BC)^\omega$, and $\mathsf{Cost}_2(AD(BC)^\omega) = \mathsf{Gain}^2_{\mathrm{Max}}(AD(BC)^\omega) = 2$. We claim that $\sigma_2^\star$ and $\sigma_{\mathtt{C}2}^\star$ are the only positional optimal strategies in $\mathcal{G}^2$. In particular, Cost_2 is positionally coalition-determined in $\mathcal{G}$.

Proposition 4.4.6 positively answers Problem 5 for cost games with prefix-linear, *positionally*[12] coalition-determined cost functions.

Proposition 4.4.6. *In every initialised multiplayer cost game where each cost function is prefix-linear and* positionally *coalition-determined, there exists a Nash equilibrium with memory (at most)* $|V| + |\Pi|$.

Proof. Let $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{Cost}_i)_{i\in\Pi})$ be a multiplayer cost game where each cost function Cost_i is prefix-linear and positionally coalition-determined,

12. The proof of [BDS13, Theorem 10] only works if the cost functions of the game are *positionally* coalition-determined.

and let $v_0 \in V$ be an initial vertex. Since the cost functions are positionally coalition-determined in $\mathcal{G}$, we know that, for every $i \in \Pi$, there exists a gain function $\mathsf{Gain}^i_{\mathrm{Max}}$ such that the Min-Max cost game $\mathcal{G}^i = (\mathcal{A}^i, \mathsf{Cost}_i, \mathsf{Gain}^i_{\mathrm{Max}})$ is determined, and there exist positional optimal strategies $\sigma^\star_i$ and $\sigma^\star_{\mathtt{C}i}$ for player i and the coalition $\Pi \backslash \{i\}$ respectively. In particular, for $j \neq i$, we denote by $\sigma^\star_{j,i}$ the strategy of player $j \neq i$ derived from the strategy $\sigma^\star_{\mathtt{C}i}$ of the coalition $\Pi \setminus \{i\}$.

The idea is to define a Nash equilibrium in $\mathcal{G}$ as follows: each player i plays according to his strategy $\sigma^\star_i$, and punishes the first player $j \neq i$ who deviates from his strategy $\sigma^\star_j$, by playing according to $\sigma^\star_{i,j}$ (the strategy of player i derived from $\sigma^\star_{\mathtt{C}j}$ in the game $\mathcal{G}^j$).

Formally, we consider the outcome of the optimal strategies $(\sigma^\star_i)_{i\in\Pi}$ from v_0, and set $\rho := \langle(\sigma^\star_i)_{i\in\Pi}\rangle_{v_0}$. We need to specify a punishment function $P : \mathsf{Hist} \to \Pi \cup \{\bot\}$ that detects who is the first player to deviate from the play ρ, i.e. who has to be punished. For the initial vertex v_0, we define $P(v_0) = \bot$ (meaning that nobody has deviated from ρ) and for every history $hv \in \mathsf{Hist}$ starting in v_0 ($v \in V$), we let:

$$P(hv) := \begin{cases} \bot & \text{if } P(h) = \bot \text{ and } hv \text{ is a prefix of } \rho, \\ i & \text{if } P(h) = \bot,\ hv \text{ is not a prefix of } \rho, \text{ and } h \in \mathsf{Hist}_i, \\ P(h) & \text{otherwise } (P(h) \neq \bot). \end{cases}$$

Then the definition of the Nash equilibrium $(\tau_i)_{i\in\Pi}$ in $(\mathcal{G}, v_0)$ is as follows. For all $i \in \Pi$ and $hv \in \mathsf{Hist}_i$ ($v \in V_i$),

$$\tau_i(hv) := \begin{cases} \sigma^\star_i(v) & \text{if } P(hv) = \bot \text{ or } i, \\ \sigma^\star_{i,P(hv)}(v) & \text{otherwise.} \end{cases}$$

We remind that the strategies $\sigma^\star_i$ and $\sigma^\star_{\mathtt{C}i}$ are positional, for all $i \in \Pi$. Clearly, the outcome of $(\tau_i)_{i\in\Pi}$ is the play ρ $(= \langle(\sigma^\star_i)_{i\in\Pi}\rangle_{v_0})$.

Now we show that the strategy profile $(\tau_i)_{i\in\Pi}$ is a Nash equilibrium in $(\mathcal{G}, v_0)$. As a contradiction, let us assume that there exists a profitable deviation τ'_j for some player $j \in \Pi$. We denote by $\rho' := \langle\tau'_j, \tau_{-j}\rangle_{v_0}$ the outcome where player j plays according to his profitable deviation τ'_j and the players of the coalition $\Pi \setminus \{j\}$ keep their strategies $(\tau_i)_{i\in\Pi\setminus\{j\}}$. Since

τ'_j is a profitable deviation for player j w.r.t. $(\tau_i)_{i\in\Pi}$, we have that:

$$\mathsf{Cost}_j(\rho') < \mathsf{Cost}_j(\rho). \tag{4.5}$$

Let hv (with $v \in V$) be the longest common prefix of ρ and ρ'. This prefix exists and is finite as both plays ρ and ρ' start from vertex v_0 and $\rho \neq \rho'$ (remark that h could be empty). By definition of the strategy profile $(\tau_i)_{i\in\Pi}$, we can write in the case of the outcome ρ that $\rho = h\langle(\sigma^\star_i)_{i\in\Pi}\rangle_v$. Note that the strategies $\sigma^\star_i$ do not depend on h in the notation $h\langle(\sigma^\star_i)_{i\in\Pi}\rangle_v$, since they are positional. In the case of the outcome ρ', player j does not follow his strategy $\sigma^\star_j$ any more from vertex v, and so, the coalition $\Pi \setminus \{j\}$ punishes him by playing according to the strategy $\sigma^\star_{\mathtt{C}j}$ after history hv, and so $\rho' = h\langle\tau'_j|_h, \sigma^\star_{\mathtt{C}j}\rangle_v$ (see Figure 4.13).[13] Here, $\tau'_j|_h$ depends on the history h (as we do not know if τ'_j is positional or not), whereas $\sigma^\star_{\mathtt{C}j}$ does not (as it is a positional strategy).

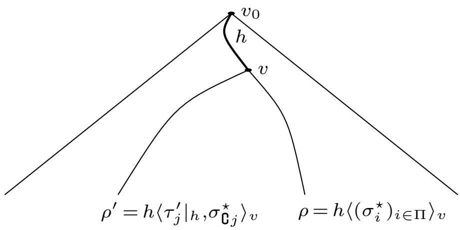


Figure 4.13: Sketch of the tree representing the unravelling of $(\mathcal{G}, v_0)$.

Since $\sigma^\star_{\mathtt{C}j}$ is an optimal strategy for the coalition $\Pi \setminus \{j\}$ in the determined Min-Max cost game $\mathcal{G}^j$, we have:

$$\begin{aligned} \mathsf{Val}^j(v) &= \inf_{\sigma_j\in\Sigma_{\mathrm{Min}}} \mathsf{Gain}^j_{\mathrm{Max}}(\langle\sigma_j, \sigma^\star_{\mathtt{C}j}\rangle_v) \\ &\leq \mathsf{Gain}^j_{\mathrm{Max}}(\langle\tau'_j|_h, \sigma^\star_{\mathtt{C}j}\rangle_v) \\ &\leq \mathsf{Cost}_j(\langle\tau'_j|_h, \sigma^\star_{\mathtt{C}j}\rangle_v)\,. \end{aligned} \tag{4.6}$$

The last inequality comes from the fact that $\mathcal{G}^j$ is a Min-Max cost game.

13. Recall that $\tau'_j|_h$ is defined by $\tau'_j|_h(h') = \tau'_j(hh')$ for all non-empty histories h' of $\mathcal{G}$ such that $(\mathsf{Last}(h), \mathsf{First}(h')) \in E$ and $\mathsf{Last}(h') \in V_j$.

Moreover, Cost_j is prefix-linear in $\mathcal{G}$, and then, when considering the history hv, there exist $a \in \mathbb{R}$ and $b \in \mathbb{R}^+$ such that

$$\mathsf{Cost}_j(\rho') = \mathsf{Cost}_j(h\langle \tau'_j|_h, \sigma^\star_{\mathtt{C}j}\rangle_v) = a + b \cdot \mathsf{Cost}_j(\langle \tau'_j|_h, \sigma^\star_{\mathtt{C}j}\rangle_v)\,. \tag{4.7}$$

As $b \geq 0$, Equations (4.6) and (4.7) imply:

$$\mathsf{Cost}_j(\rho') \geq a + b \cdot \mathsf{Val}^j(v)\,. \tag{4.8}$$

Since h is also a prefix of ρ, we have:

$$\mathsf{Cost}_j(\rho) = \mathsf{Cost}_j(h\langle(\sigma^\star_i)_{i\in\Pi}\rangle_v) = a + b \cdot \mathsf{Cost}_j(\langle(\sigma^\star_i)_{i\in\Pi}\rangle_v)\,. \tag{4.9}$$

Furthermore, as $\sigma^\star_j$ is an optimal strategy for player j in the Min-Max cost game $\mathcal{G}^j$, it follows that:

$$\begin{aligned} \mathsf{Val}^j(v) &= \sup_{\sigma_{\mathtt{C}j} \in \Sigma_{\mathrm{Max}}} \mathsf{Cost}_j(\langle \sigma^\star_j, \sigma_{\mathtt{C}j}\rangle_v) \\ &\geq \mathsf{Cost}_j(\langle(\sigma^\star_i)_{i\in\Pi}\rangle_v)\,. \end{aligned} \tag{4.10}$$

Then, Equations (4.9) and (4.10) imply:

$$\mathsf{Cost}_j(\rho) \leq a + b \cdot \mathsf{Val}^j(v)\,. \tag{4.11}$$

Finally, Equations (4.8) and (4.11) lead to the following inequality:

$$\mathsf{Cost}_j(\rho) \leq a + b \cdot \mathsf{Val}^j(v) \leq \mathsf{Cost}_j(\rho')\,,$$

which contradicts Equation (4.5). The strategy profile $(\tau_i)_{i\in\Pi}$ is then a Nash equilibrium in the game $\mathcal{G}$.

Now we show that $(\tau_i)_{i\in\Pi}$ is a strategy profile with memory (at most) $|V|+|\Pi|$. For this purpose, we define a finite strategy automaton for each player that remembers the play ρ and who has to be punished. As the play ρ is the outcome of the positional strategy profile $(\sigma^\star_i)_{i\in\Pi}$, we can write $\rho := v_0 \dots v_{k-1}(v_k \dots v_n)^\omega$ where $0 \leq k \leq n \leq |V|$, $v_l \in V$ for all $0 \leq l \leq n$ and these vertices are all different. For any $i \in \Pi$, let $\mathcal{M}_i = (M, m_0, V, \delta, \nu)$ be the strategy automaton of player i, where:

- $M = \{v_0v_0, v_0v_1, \dots, v_{n-1}v_n, v_nv_k\} \cup \Pi \setminus \{i\}$.
As we want to be sure that the play ρ is followed by all players, we need to memorise which movement (edge) has to be chosen at each step of ρ. This is the role of $\{v_0v_0, v_0v_1, \dots, v_{n-1}v_n, v_nv_k\}$. But in case a player deviates from ρ, we only have to remember this player during the rest of the play (no matter if another player later deviates from ρ). This is the role of $\Pi \setminus \{i\}$.
- $m_0 = v_0v_0$ (this memory state means that the play has not begun yet).
- $\delta : M \times V \to M$ is defined in this way: given $m \in M$ and $v \in V$,

$$\delta(m,v) := \begin{cases} j & \text{if } m = j \in \Pi \text{ or } (m = u_1u_2, \text{ with } u_1, u_2 \in V, \\ & \quad v \neq u_2 \text{ and } u_1 \in V_j), \\ v_lv_{l+1} & \text{if } m = uv_l \text{ for a certain } l \in \{0, \dots, n-1\}, \\ & \quad u \in V, \text{ and } v = v_l, \\ v_nv_k & \text{otherwise } (m = uv_n \text{ and } v = v_n). \end{cases}$$

Intuitively, m represents either a player to punish, or the edge that should, if following ρ, have been chosen at the last step of the current stage of the play, and v is the real last vertex of the current stage of the play.
Notice that in this definition of δ, j is different from i because if player i follows the strategy computed by this strategy automaton, one can be convinced that he does not deviate from the play ρ.
- $\nu : M \times V_i \to V$ is defined in this way: given $m \in M$ and $v \in V_i$,

$$\nu(m,v) := \begin{cases} \sigma_i^\star(v) & \text{if } m = u_1u_2 \text{ with } u_1, u_2 \in V \text{ and } v = u_2, \\ \sigma_{i,j}^\star(v) & \text{if } m = j \in \Pi \text{ or } (m = u_1u_2, \text{ with } u_1, u_2 \in V, \\ & \quad v \neq u_2 \text{ and } u_1 \in V_j). \end{cases}$$

The idea is to play according to $\sigma_i^\star$ if everybody follows the play ρ, and switch to $\sigma_{i,j}^\star$ if player j is the first player who has deviated from ρ.

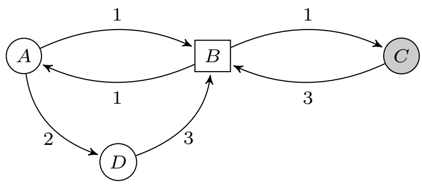
Obviously, the strategy $\sigma_{\mathcal{M}_i}$ computed by the strategy automaton $\mathcal{M}_i$ exactly corresponds to the strategy τ_i of the Nash equilibrium. And so, we can conclude that each strategy τ_i requires a memory of size at most $|M| \leq |\Pi| + |V|$. □

Applications of Proposition 4.4.6 to specific classes of cost games are provided in Section 4.4.2.

Remark 4.4.7. If, in the definition of prefix-linear (Definition 4.4.2), a can be a real number *or* $+\infty$, and b is a *strictly positive* real number, and if we assume that the cost functions are bounded from below, then the proof of Proposition 4.4.6 still holds.

Let us also notice that the proof of Proposition 4.4.6 remains valid for cost functions $\mathsf{Cost}_i : \mathsf{Plays} \to K$, where $\langle K, +, \cdot, 0, 1, \leq \rangle$ is an ordered field. This allows for instance to consider non-standard real costs and enjoy infinitesimals to model the costs of a player. In particular, this might be a road to explore in order to prove existence of secure equilibria in cost games.

Example 4.4.8. Let us come back to the two-player cost game $\mathcal{G}$ of Example 4.4.5 (on page 97), whose arena $\mathcal{A}$ is depicted below, and where player 1 has a quantitative reachability objective ($\mathsf{Cost}_1 = \mathrm{RP}_{\mathrm{Min}}$ for the goal set $\mathsf{R} = \{C\}$) and player 2 has a mean-payoff objective ($\mathsf{Cost}_2 = \mathrm{AP}_{\mathrm{Min}}$).



One can show that both cost functions are prefix-linear (see the proof of Corollary 4.4.15). Since we saw in Example 4.4.5 (on page 97) that these cost functions are also positionally coalition-determined, we can apply the construction in the proof of Proposition 4.4.6 to get a Nash equilibrium in $(\mathcal{G}, A)$. The construction from this proof results in two different Nash equilibria, depending on the selection of the strategies $\sigma_1^\star/\tilde{\sigma}_1^\star$, $\sigma_{\mathtt{C}1}^\star$, $\sigma_2^\star$ and $\sigma_{\mathtt{C}2}^\star$ as defined in Example 4.4.5.

The first Nash equilibrium (τ_1, τ_2) with outcome $\rho = \langle \sigma_1^\star, \sigma_2^\star \rangle_A = A(BC)^\omega$ is given, for all histories hA and $h'B$, by:

$$\tau_1(hA) = \begin{cases} B & \text{if } P(hA) = \bot, 1 \\ D & \text{otherwise} \end{cases} \quad ; \quad \tau_2(h'B) := \begin{cases} C & \text{if } P(h'B) = \bot, 2 \\ A & \text{otherwise} \end{cases}$$

where the punishment function P is defined as in the proof of Proposition 4.4.6 and depends on the play ρ. The cost for this finite-memory Nash equilibrium in $(\mathcal{G}, A)$ is $\mathsf{Cost}_1(\rho) = 2 = \mathsf{Cost}_2(\rho)$.

The strategy $\tilde{\tau}_1$ of the second Nash equilibrium $(\tilde{\tau}_1, \tau_2)$ in $(\mathcal{G}, A)$ with outcome $\tilde{\rho} = \langle \tilde{\sigma}_1^\star, \sigma_2^\star \rangle_A = AD(BC)^\omega$ is given by $\tilde{\tau}_1(hA) := D$ for all histories hA. The costs for this finite-memory Nash equilibrium in $(\mathcal{G}, A)$ are $\mathsf{Cost}_1(\tilde{\rho}) = 6$ and $\mathsf{Cost}_2(\tilde{\rho}) = 2$, respectively.

Note that there is no positional Nash equilibrium with outcome ρ or $\tilde{\rho}$ in $(\mathcal{G}, A)$. The only positional Nash equilibrium in this game is the one with outcome $(AB)^\omega$ and cost profile $(+\infty, 1)$.

Thanks to the proof of Proposition 4.4.6, we can construct a finite strategy automaton $\mathcal{M}_{\tau_1}$ that computes the strategy τ_1 of player 1, for instance. The set M of memory states is $M = \{AA, AB, BC, CB\} \cup \{2\}$ since $\rho = A(BC)^\omega$, and the initial state is $m_0 = AA$. The memory update function $\delta : M \times V \to M$ and the transition choice function $\nu : M \times V_1 \to V$ are depicted in Figure 4.14: a label v/v' on an edge (m_1, m_2) means that $\delta(m_1, v) = m_2$, and $\nu(m_1, v) = v'$ if $v \in V_1$. If $v \notin V_1$, we indicate that ν does not return any advice by a '$-$', and label the edge with $v/-$.

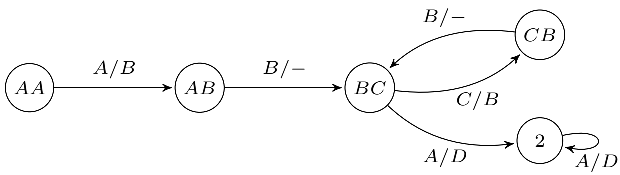


Figure 4.14: The finite strategy automaton $\mathcal{M}_{\tau_1}$.

When the optimal strategies in the Min-Max cost games $\mathcal{G}^i$ are not positional, the proof of Proposition 4.4.6 does not hold anymore. Indeed, in that proof, we wrote $\rho = h\langle(\sigma_i^\star)_{i\in\Pi}\rangle_v$, but if the strategies $(\sigma_i^\star)_{i\in\Pi}$ use memory, we have to write $\rho = h\langle(\sigma_i^\star|_h)_{i\in\Pi}\rangle_v$, since playing from v as it was the initial vertex, and playing from v while taking into account the initial history h might be done in different ways (as the $\sigma_i^\star$'s are not positional). Moreover, playing according to $\sigma_i^\star|_h$ from v might not ensure that player i loses at most $\mathsf{Val}^i(v)$ in the game $\mathcal{G}^i$. In fact, an optimal strategy $\sigma_i^\star$ might be "inconsistent" after certain histories, as it can been in the following example.

Example 4.4.9. Let us consider the Min-Max cost game $\mathcal{G} = (\mathcal{A}, \mathsf{Cost}_{\mathrm{Min}}, \mathsf{Gain}_{\mathrm{Max}})$ whose arena $\mathcal{A}$ is depicted in Figure 4.15. We set $\mathsf{Gain}_{\mathrm{Max}} = \mathsf{Cost}_{\mathrm{Min}}$, and for any play ρ, $\mathsf{Cost}_{\mathrm{Min}}(\rho) = 1$ if ρ visits vertex B or E, and $\mathsf{Cost}_{\mathrm{Min}}(\rho) = 0$ otherwise (ρ visits vertex D).

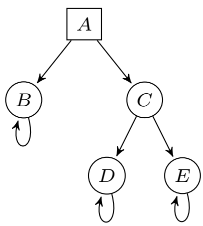


Figure 4.15: Min-Max cost game $\mathcal{G}$.

The finite-memory strategy $\sigma_1^\star$ defined by $\sigma_1^\star(C) = D$ and $\sigma_1^\star(hC) = E$ for all histories hC with $h \neq \epsilon$, is optimal for player Min in $\mathcal{G}$. But if he plays according to $\sigma_1^\star|_A$ from C, he pays a cost of 1, losing strictly more than $\mathsf{Val}(C) = 0$.

In the case where the optimal strategies are not positional, we also want to compare the cost $\mathsf{Cost}_j(\rho)$ to the value $\mathsf{Val}^j(v)$ as in the proof of Proposition 4.4.6 (see Equation (4.11)). But Example 4.4.9 shows that sometimes the optimal strategies might be inadequate. That is why we prove the existence of an "intelligent" optimal strategy for player Min in Min-Max cost games with a prefix-independent cost function $\mathsf{Cost}_{\mathrm{Min}}$. What we mean by an "intelligent" optimal strategy is that player Min's cost of the outcome of such a strategy and any strategy of player Max, is smaller or equal to the value of any vertex visited by this outcome. This property is formalised in the following lemma. Notice that we relax the hypothesis of *positional* optimal strategies, but we require the stronger condition about *prefix-independent* (instead of prefix-linear) cost functions.

Lemma 4.4.10. *Let $\mathcal{G} = (\mathcal{A}, \mathsf{Cost}_{Min}, \mathsf{Gain}_{Max})$ be a determined Min-Max cost game such that the cost function Cost_{Min} is prefix-independent. If there exists an optimal strategy $\sigma_1^\star$ for player Min, then there exists an optimal strategy $\tau_1^\star$ for player Min such that for all vertices $v_0 \in V$ and all strategies σ_2 of player Max, we have that*

$$\mathsf{Cost}_{Min}(\rho) \leq \min\{\mathsf{Val}(u) \mid u \textit{ is a vertex of } \rho\}\,, \textit{ where } \rho = \langle \tau_1^\star, \sigma_2 \rangle_{v_0}.$$

Note that the strategy $\sigma_1^\star$ of Example 4.4.9 does not satisfy the property stated for $\tau_1^\star$ in the above lemma. Indeed, with $v_0 := A$ and the strategy σ_2 of player Max defined as $\sigma_2(A) := C$, we have that $\mathsf{Cost}_{\mathrm{Min}}(\langle \sigma_1^\star, \sigma_2 \rangle_{v_0}) = \mathsf{Cost}_{\mathrm{Min}}(ACE) = 1 > 0 = \mathsf{Val}(C)$.

Proof of Lemma 4.4.10. For all histories $h \in \mathsf{Hist}_1$, we define

$$\tau_1^\star(h) := \sigma_1^\star(vh'') \;\text{ s.t. } \begin{cases} h = h'vh'' \text{ (with } v \in V), \\ \mathsf{Val}(v) = \min\{\mathsf{Val}(u) \mid u \text{ is a vertex of } h\}, \\ |h'| \text{ is minimal w.r.t. the above property}^{14}. \end{cases}$$

We first show that $\tau_1^\star$ satisfies the property stated in Lemma 4.4.10. Given $v_0 \in V$ and $\sigma_2 \in \Sigma_{\mathrm{Max}}$, let ρ be the play $\langle \tau_1^\star, \sigma_2 \rangle_{v_0}$. We assert that there exists a prefix $h'v$ of ρ (with $v \in V$) such that $\rho = h' \langle \sigma_1^\star, \sigma_2|_{h'} \rangle_v$. Roughly, it means that after $h'v$, player Min plays according to $\sigma_1^\star$ as if v was the initial vertex (and forgets h').

Indeed, let $h'v$ be the smallest prefix of ρ such that

$$\mathsf{Val}(v) = \min\{\mathsf{Val}(u) \mid u \text{ is a vertex of } \rho\}\,. \tag{4.12}$$

Note that $\min\{\mathsf{Val}(u) \mid u \text{ is a vertex of } \rho\}$ exists since $|V|$ is finite. We now prove that $\rho = h' \langle \sigma_1^\star, \sigma_2|_{h'} \rangle_v$. Let h'' be a history such that $h'vh''$ is

14. It means that for all prefixes $\bar{h}\bar{v}$ (with $\bar{v} \in V$) of the history h such that $\mathsf{Val}(\bar{v}) = \min\{\mathsf{Val}(u) \mid u \text{ is a vertex of } h\}$, we have that $|h'| \leq |\bar{h}|$.

a prefix of ρ, and $h'vh''$ is in Hist_1. We have

$$\tau_1^\star(h'vh'') = \sigma_1^\star(vh'') \quad \text{by the definitions of } h'v \text{ and } \tau_1^\star.$$

Then, it holds that $\rho = h'\langle \sigma_1^\star, \sigma_2|_{h'}\rangle_v$.

As $\mathsf{Cost}_{\mathrm{Min}}$ is prefix-independent and $\sigma_1^\star$ is optimal, it follows that

$$\begin{aligned}\mathsf{Cost}_{\mathrm{Min}}(\rho) &= \mathsf{Cost}_{\mathrm{Min}}(h'\langle \sigma_1^\star, \sigma_2|_{h'}\rangle_v) \\ &= \mathsf{Cost}_{\mathrm{Min}}(\langle \sigma_1^\star, \sigma_2|_{h'}\rangle_v) \\ &\leq \sup_{\sigma_2' \in \Sigma_{\mathrm{Max}}} \mathsf{Cost}_{\mathrm{Min}}(\langle \sigma_1^\star, \sigma_2'\rangle_v) = \mathsf{Val}(v)\,.\end{aligned}$$

$\mathsf{Val}(v) = \min\{\mathsf{Val}(u) \mid u \text{ is a vertex of } \rho\}$ (Equation (4.12)) implies that

$$\mathsf{Cost}_{\mathrm{Min}}(\rho) \leq \min\{\mathsf{Val}(u) \mid u \text{ is a vertex of } \rho\}\,,$$

which proves the property stated in Lemma 4.4.10.

Furthermore, since v_0 is a vertex of ρ, we have that

$$\mathsf{Cost}_{\mathrm{Min}}(\langle \tau_1^\star, \sigma_2\rangle_{v_0}) = \mathsf{Cost}_{\mathrm{Min}}(\rho) \leq \mathsf{Val}(v_0)$$

for any strategy $\sigma_2 \in \Sigma_{\mathrm{Max}}$, which confirms that the strategy $\tau_1^\star$ is optimal for player Min. □

Lemma 4.4.10 enables to prove the existence of a Nash equilibrium in multiplayer cost games where the cost functions are prefix-independent and coalition-determined. Note that here we do not make any restriction on the memory needed by the optimal strategies in the Min-Max cost games $\mathcal{G}^i$, contrary to Proposition 4.4.6 where the cost functions are required to be positionally coalition-determined. Nevertheless, we here ask for prefix-independent cost functions (stronger condition than prefix-linear cost functions). Proposition 4.4.11 then positively answers Problem 4 for cost games with prefix-independent, coalition-determined cost functions.

Proposition 4.4.11. *In every initialised multiplayer cost game where each cost function is prefix-independent and coalition-determined, there exists a Nash equilibrium.*

Proof. Let $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{Cost}_i)_{i\in\Pi})$ be a multiplayer cost game such that each cost function Cost_i is prefix-independent and coalition-determined, and let $v_0 \in V$ be an initial vertex. By hypothesis, we know that, for every $i \in \Pi$, there exists a gain function $\mathsf{Gain}^i_{\mathrm{Max}}$ such that the Min-Max cost game $\mathcal{G}^i = (\mathcal{A}^i, \mathsf{Cost}_i, \mathsf{Gain}^i_{\mathrm{Max}})$ is determined and there exist optimal strategies $\sigma^\star_i$ and $\sigma^\star_{\complement i}$ for player i and the coalition $\Pi \setminus \{i\}$ respectively. By Lemma 4.4.10, we can assume that $\sigma^\star_i$ verifies: for all strategies $\sigma_{\complement i}$ of the coalition $\Pi \setminus \{i\}$,

$$\mathsf{Cost}_i(\langle \sigma^\star_i, \sigma_{\complement i}\rangle_{v_0}) \leq \min\{\mathsf{Val}^i(u) \mid u \text{ is a vertex of } \langle \sigma^\star_i, \sigma_{\complement i}\rangle_{v_0}\}\,. \quad (4.13)$$

For $j \neq i$, we denote by $\sigma^\star_{j,i}$ the strategy of player j derived from the strategy $\sigma^\star_{\complement i}$ of the coalition $\Pi \setminus \{i\}$.

The idea is to define the Nash equilibrium in $\mathcal{G}$ in a similar way to the proof of Proposition 4.4.6. Each player i plays according to his strategy $\sigma^\star_i$, and punishes the first player $j \neq i$ who deviates from his strategy $\sigma^\star_j$, by playing according to $\sigma^\star_{i,j}$ from the first vertex where player j does not follow $\sigma^\star_j$, while forgetting the initial history until there.

We set $\rho := \langle(\sigma^\star_i)_{i\in\Pi}\rangle_{v_0}$, and consider the same punishment function $P : \mathsf{Hist} \to \Pi \cup \{\perp\}$ as the one defined in the proof of Proposition 4.4.6. As a reminder, it detects who is the first player to deviate from the play ρ, i.e. who has to be punished. Then the definition of the Nash equilibrium $(\tau_i)_{i\in\Pi}$ in $\mathcal{G}$ is as follows. For all $i \in \Pi$ and $h \in \mathsf{Hist}_i$,

$$\tau_i(h) := \begin{cases} \sigma^\star_i(h) & \text{if } P(h) = \perp \text{ or } i, \\ \sigma^\star_{i,P(h)}(vv'h'') & \text{if } P(h) \neq \perp, i, \text{ s.t. } h = h'vv'h'' \ (v, v' \in V), \\ & \quad P(h'vv') = P(h), \text{ and } |h'| \text{ is minimal} \\ & \quad \text{w.r.t. this property.} \end{cases}$$

Note that if $P(h) \neq \perp$, then h has at least two vertices since $P(v_0) = \perp$ by definition of P. The decomposition $h = h'vv'h''$ as described in the second part of the definition of τ_i means that v is the vertex where player $P(h)$ has not chosen the outgoing edge that follows the play ρ. At this very moment, the other players start to punish player $P(h)$, while forgetting the past history. Clearly, the outcome of $(\tau_i)_{i\in\Pi}$ from v_0 is the play ρ $(= \langle(\sigma^\star_i)_{i\in\Pi}\rangle_{v_0})$.

We prove that the strategy profile $(\tau_i)_{i\in\Pi}$ is a Nash equilibrium in $\mathcal{G}$. As a contradiction, assume that there exists a profitable deviation τ'_j for some player $j \in \Pi$. We denote by $\rho' := \langle \tau'_j, \tau_{-j}\rangle_{v_0}$ the outcome where player j plays according to his profitable deviation τ'_j and the players of the coalition $\Pi \setminus \{j\}$ keep their strategies $(\tau_i)_{i\in\Pi\setminus\{j\}}$. Since τ'_j is a profitable deviation for player j w.r.t. $(\tau_i)_{i\in\Pi}$, we have that:

$$\mathsf{Cost}_j(\rho') < \mathsf{Cost}_j(\rho). \tag{4.14}$$

Let hv (with $v \in V$) be the longest common prefix of ρ and ρ'. This prefix exists and is finite as both plays ρ and ρ' start from vertex v_0 and $\rho \neq \rho'$ (remark that h could be empty). It follows that $\rho' = h\langle \tau'_j|_h, \tau_{-j}|_h\rangle_v$. By definition of τ_{-j}, we have that $\rho' = h\langle \tau'_j|_h, \sigma^\star_{\mathtt{C}j}\rangle_v$. That is, the coalition $\Pi \setminus \{j\}$ starts playing according to $\sigma^\star_{\mathtt{C}j}$ from v as if it was the initial vertex. Indeed, let $v'h'$ be a history such that $hvv'h'$ is a prefix of ρ' and $hvv'h'$ is in Hist_i ($v' \in V$), for $i \neq j$. Notice that $v \in V_j$ as player j is the only player who deviates from ρ. We then have:

$$\tau_i(hvv'h') = \sigma^\star_{i,j}(vv'h')$$

by definition of τ_i (as $P(hvv') = j$ and $|h|$ is minimal w.r.t. this property).

Since $\sigma^\star_{\mathtt{C}j}$ is an optimal strategy for the coalition $\Pi \setminus \{j\}$ in the determined Min-Max cost game $\mathcal{G}^j$, we can deduce that:

$$\begin{aligned}
\mathsf{Val}^j(v) &= \inf_{\sigma_j \in \Sigma_{\mathrm{Min}}} \mathsf{Gain}^j_{\mathrm{Max}}(\langle \sigma_j, \sigma^\star_{\mathtt{C}j}\rangle_v) \\
&\leq \mathsf{Gain}^j_{\mathrm{Max}}(\langle \tau'_j|_h, \sigma^\star_{\mathtt{C}j}\rangle_v) \\
&\leq \mathsf{Cost}_j(\langle \tau'_j|_h, \sigma^\star_{\mathtt{C}j}\rangle_v)\,.
\end{aligned}$$

The last inequality comes from the fact that $\mathcal{G}^j$ is a Min-Max cost game. As Cost_j is prefix-independent, it follows that:

$$\mathsf{Val}^j(v) \leq \mathsf{Cost}_j(h\langle \tau'_j|_h, \sigma^\star_{\mathtt{C}j}\rangle_v) = \mathsf{Cost}_j(\rho')\,. \tag{4.15}$$

Furthermore, by Equation (4.13), we have that:

$$\begin{aligned}
\mathsf{Cost}_j(\rho) &= \mathsf{Cost}_j(\langle(\sigma^\star_i)_{i\in\Pi}\rangle_{v_0}) \\
&\leq \min\{\mathsf{Val}^j(u) \mid u \text{ is a vertex of } \langle(\sigma^\star_i)_{i\in\Pi}\rangle_{v_0}\} \\
&\leq \mathsf{Val}^j(v) \qquad (\text{since } v \text{ is a vertex of } \langle(\sigma^\star_i)_{i\in\Pi}\rangle_{v_0}) \\
&\leq \mathsf{Cost}_j(\rho') \qquad (\text{by Equation (4.15)})\,,
\end{aligned}$$

which contradicts Equation (4.14). The strategy profile $(\tau_i)_{i\in\Pi}$ is then a Nash equilibrium in the game $(\mathcal{G}, v_0)$. □

When the cost functions are prefix-independent and *finite-memory* coalition-determined in a cost game, we can prove that there exists a *finite-memory* Nash equilibrium, and then positively answer Problem 5 for such games.

Proposition 4.4.12. *In every initialised multiplayer cost game where all cost functions are prefix-independent and* finite-memory *coalition-determined, there exists a finite-memory Nash equilibrium.*

The proof of this theorem relies on the construction of the Nash equilibrium provided in the proof of Proposition 4.4.11.

Sketch of proof of Proposition 4.4.12. Let $(\mathcal{G}, v_0)$ be an initialised multiplayer cost game with prefix-independent and *finite-memory* coalition-determined cost functions. The proof follows the same philosophy concerning memory as the proof of Proposition 4.4.6. We consider the Nash equilibrium $(\tau_i)_{i\in\Pi}$ defined in the proof of Proposition 4.4.11, whose outcome is $\rho := \langle(\sigma_i^\star)_{i\in\Pi}\rangle_{v_0}$. We keep the same notations as in that proof. We remind that for all $i \in \Pi$, the strategy τ_i depends on the strategies $\sigma_i^\star$ and $\sigma_{i,j}^\star$ for $j \in \Pi \setminus \{i\}$. As the cost functions are finite-memory coalition-determined in $\mathcal{G}$ by hypothesis, these strategies are assumed to be finite-memory. Given $i \in \Pi$ (and $j \in \Pi \setminus \{i\}$), we denote by $\mathcal{M}^{\sigma_i^\star}$ (resp. $\mathcal{M}^{\sigma_{i,j}^\star}$) a finite strategy automaton for the strategy $\sigma_i^\star$ (resp. $\sigma_{i,j}^\star$).

As in the proof of Proposition 4.4.6, each player needs to remember both the play ρ and who has to be punished. But here the play ρ is not anymore the outcome of a positional strategy profile: each $\sigma_i^\star$ is a finite-memory strategy. Nevertheless, in some sense, we can see the $\sigma_i^\star$'s as positional strategies played on the product graph $G \times \prod_{i\in\Pi} \mathcal{M}^{\sigma_i^\star}$, where G is the graph of the arena of $\mathcal{G}$. This allows us to write $\rho := v_0 \dots v_{k-1}(v_k \dots v_n)^\omega$ where[15] $0 \le k \le n \le |V| \cdot \prod_{i\in\Pi} |\mathcal{M}^{\sigma_i^\star}|$, $v_l \in V$ for all $0 \le l \le n$. Like in the proof of Proposition 4.4.6, we can now define, for any $i \in \Pi$, $\mathcal{M}^{\tau_i}$, a finite strategy automaton for

15. $|\mathcal{M}|$ denotes the number of states of the automaton A.

τ_i. In order to build explicitly $\mathcal{M}^{\tau_i}$, we need to take into account, on one hand, the path ρ, and on the other hand, the memory of the punishing strategies $\sigma^\star_{i,j}$. This enables to bound the size of $\mathcal{M}^{\tau_i}$ by $|V| \cdot \prod_{i\in\Pi} |\mathcal{M}^{\sigma^\star_i}| + \sum_{i\in\Pi} \sum_{j\in\Pi\setminus\{i\}} |\mathcal{M}^{\sigma^\star_{i,j}}|$. □

Remark 4.4.13. As we have seen before, the outcomes ρ of the Nash equilibria $(\tau_i)_{i\in\Pi}$ given in the proofs of Propositions 4.4.6, 4.4.11 and 4.4.12 are consistent with the optimal strategies $(\sigma^\star_i)_{i\in\Pi}$ from v_0. Then, it holds that $\mathsf{Cost}_i(\rho) \leq \mathsf{Val}^i(v_0)$, for all $i \in \Pi$. That is, each player i pays at most $\mathsf{Val}^i(v_0)$ for the outcomes ρ.

An interesting problem could be to find an algorithm to decide, given a cost game and a tuple of thresholds $(t_i)_{i\in\Pi} \in (\mathbb{R} \cup \{-\infty, +\infty\})^{|\Pi|}$, whether there exists a *simple* Nash equilibrium with cost profile $(c_i)_{i\in\Pi}$ such that for all $i \in \Pi$, $c_i \leq t_i$.

In the same idea as Propositions 4.4.6 and 4.4.12, we can state a more general result, which answers Problem 4 (and Problem 5) for a large class of cost games.

Theorem 4.4.14. *In every initialised multiplayer cost game where each cost function is either prefix-linear and positionally coalition-determined, or prefix-independent and (finite-memory) coalition-determined, there exists a (finite-memory) Nash equilibrium.*

In order to show this theorem, we can reuse the proofs of Propositions 4.4.6, 4.4.11 and 4.4.12, depending on the cost function of the considered player. But we must be cautious in the proof of Proposition 4.4.6: when we write ρ with the prefix h, we must now write $\rho = h\langle\sigma^\star_j, (\sigma^\star_i|_h)_{i\in\Pi\setminus\{j\}}\rangle_v$, because here the strategies $\sigma^\star_i$ $(i \neq j)$ may need memory. Then, Equation (4.9) becomes

$$\mathsf{Cost}_j(h\langle\sigma^\star_j, (\sigma^\star_i|_h)_{i\in\Pi\setminus\{j\}}\rangle_v) = a + b \cdot \mathsf{Cost}_j(\langle\sigma^\star_j, (\sigma^\star_i|_h)_{i\in\Pi\setminus\{j\}}\rangle_v),$$

and Equation (4.10) becomes

$$\mathsf{Val}^j(v) = \sup_{\sigma_{\mathtt{C}_j}\in\Sigma_{\mathrm{Max}}} \mathsf{Cost}_j(\langle\sigma^\star_j, \sigma_{\mathtt{C}_j}\rangle_v) \geq \mathsf{Cost}_j(\langle\sigma^\star_j, (\sigma^\star_i|_h)_{i\in\Pi\setminus\{j\}}\rangle_v).$$

Then, the rest of the proof remains the same.

4.4.2 Applications

In this section, we exhibit several classes of *classical objectives* that can be encoded in our general setting. The list we propose is far from being exhaustive.

Qualitative Objectives

A qualitative objective can naturally be encoded via a cost function $\mathsf{Cost}_i : \mathsf{Plays} \to \{-1, 1\}$, where -1 (resp. 1) means that the play is won (resp. lost) by player i.[16] If we assume that the objective is Borel, then the cost function is coalition-determined, as a consequence of the Borel determinacy theorem (see Theorem 2.2.6). By applying Proposition 4.4.11, we obtain the existence of a Nash equilibrium for qualitative games with prefix-independent[17] Borel objectives. Let us notice that this result is already present in [GU08].

When considering more specific subclasses of qualitative games with prefix-independent objectives enjoying a positional determinacy result, such as parity games (see Theorem 2.2.15), we can apply Proposition 4.4.6 and ensure existence of a Nash equilibrium whose memory is (at most) linear.

Furthermore, one can prove that the cost function encoding a qualitative reachability or safety objective is prefix-linear in any cost game. As reachability games and safety games are positionally determined (see Theorem 2.2.15), Proposition 4.4.6 also applies to cost games with such objectives.

Classical Quantitative Objectives

If we consider the particular cost functions of Definition 2.3.3, Theorem 2.3.10 implies that the cost functions $\mathrm{RP}_{\mathrm{Min}}$, $\mathrm{DP}_{\mathrm{Min}}$, $\mathrm{AP}_{\mathrm{Min}}$ and $\mathrm{PRAvg}_{\mathrm{Min}}$ are positionally coalition-determined in any cost game. Then,

16. Note that we here minimise costs, that is why the cost function is defined in the opposite way to the gain function of Equation (2.2) on page 29.

17. A qualitative objective $\mathsf{Win} \subseteq V^\omega$ is prefix-independent if and only if for every path $\rho = \rho_0\rho_1 \ldots \in V^\omega$, we have that $\rho \in \mathsf{Win}$ iff for every $n \in \mathbb{N}$, $\rho_n\rho_{n+1} \ldots \in \mathsf{Win}$ (same idea as in the quantitative case, see Definition 4.4.1).

if we show that they are also prefix-linear, Corollary 4.4.15 will follow from Proposition 4.4.6.

Corollary 4.4.15. *In every initialised multiplayer cost game where each cost function is RP_{Min}, DP_{Min}, AP_{Min} or $PRAvg_{Min}$, there exists a Nash equilibrium with memory (at most) $|V| + |\Pi|$.*

In particular, this result applies in multiplayer quantitative reachability games, as they are multiplayer cost games where each cost function is $\mathrm{RP}_{\mathrm{Min}}$ for some goal sets $(\mathsf{R}_i)_{i\in\Pi}$. Note that the existence of finite-memory Nash equilibria in such games has already been established in Theorem 4.1.5. Nevertheless, the Nash equilibrium given for its proof needs a memory (at most) exponential in the size of the cost game (see Remark 4.1.9). Thus, Corollary 4.4.15 significantly improves the complexity of the strategies constructed in the case of multiplayer quantitative reachability games.

Proof of Corollary 4.4.15. Let $\mathcal{G}$ be a a multiplayer cost game where each cost function is $\mathrm{RP}_{\mathrm{Min}}$, $\mathrm{DP}_{\mathrm{Min}}$, $\mathrm{AP}_{\mathrm{Min}}$ or $\mathrm{PRAvg}_{\mathrm{Min}}$. Let us show that these cost functions are prefix-linear in $\mathcal{G}$. Given $j \in \Pi$, $v \in V$ and $hv \in \mathsf{Hist}$, we consider the four possible cases for Cost_j. For the sake of simplicity, we write $hv := h_0 \ldots h_k$ with $k \in \mathbb{N}$, $h_k = v$ and $h_l \in V$ for $l = 0, \ldots, k$. Let $\phi : E \to \mathbb{R}$ be a price function and $\vartheta : E \to \mathbb{R}$ be a diverging reward function. To avoid heavy notations, we do not explicitly show the dependency between j and ϕ, ϑ, R in the first case or λ in the second case. That is, we write ϕ, ϑ, R and λ for ϕ_j, ϑ_j, R_j and λ_j, respectively.

1. Case $\mathsf{Cost}_j = \mathrm{RP}_{\mathrm{Min}}$ for a given goal set $\mathsf{R} \subseteq V$:

   Let us distinguish two situations. If there exists $l \in \{0, \ldots, k\}$ such that $h_l \in \mathsf{R}$, then we set $a := \sum_{i=1}^{n} \phi(h_{i-1}, h_i) \in \mathbb{R}$ and $b := 0 \in \mathbb{R}^+$, where n is the least index such that $h_n \in \mathsf{R}$. Let ρ be a play with $\mathsf{First}(\rho) = v$, then it implies that $\mathrm{RP}_{\mathrm{Min}}(h\rho) = \sum_{i=1}^{n} \phi(h_{i-1}, h_i) = a + b \cdot \mathrm{RP}_{\mathrm{Min}}(\rho)$ (with the convention that $0 \cdot +\infty = 0$).

   If there does not exist $l \in \{0, \ldots, k\}$ such that $h_l \in \mathsf{R}$, then we set $a := \sum_{i=1}^{k} \phi(h_{i-1}, h_i) \in \mathbb{R}$ and $b := 1 \in \mathbb{R}^+$. Let $\rho = \rho_0\rho_1 \ldots$

be a play such that $\mathsf{First}(\rho) = v$. If $\mathrm{RP}_{\mathrm{Min}}(\rho)$ is infinite, then $\mathrm{RP}_{\mathrm{Min}}(h\rho) = +\infty = a + b \cdot \mathrm{RP}_{\mathrm{Min}}(\rho)$. Otherwise, if n is the least index in $\mathbb{N}$ such that $\rho_n \in \mathsf{R}$, then we have that:

$$\begin{aligned} \mathrm{RP}_{\mathrm{Min}}(h\rho) &= \sum_{i=1}^{k} \phi(h_{i-1}, h_i) + \sum_{i=1}^{n} \phi(\rho_{i-1}, \rho_i) \\ &= a + b \cdot \mathrm{RP}_{\mathrm{Min}}(\rho). \end{aligned}$$

2. Case $\mathsf{Cost}_j = \mathrm{DP}_{\mathrm{Min}}(\lambda)$ for a given discount factor $\lambda \in \,]0,1[$:
We set $a := (1-\lambda)\sum_{i=1}^{k} \lambda^{i-1}\phi(h_{i-1}, h_i) \in \mathbb{R}$ and $b := \lambda^k \in \mathbb{R}^+$. Given a play $\rho = \rho_0\rho_1\ldots$ such that $\mathsf{First}(\rho) = v$ and $\eta := h\rho \in \mathsf{Plays}$ (with $\eta = \eta_0\eta_1\ldots$), we have that:

$$\begin{aligned} \mathrm{DP}_{\mathrm{Min}}(\lambda)(h\rho) &= \mathrm{DP}_{\mathrm{Min}}(\lambda)(\eta) \\ &= (1-\lambda)\sum_{i=1}^{+\infty} \lambda^{i-1}\phi(\eta_{i-1}, \eta_i) \\ &= (1-\lambda)\sum_{i=1}^{k} \lambda^{i-1}\phi(\eta_{i-1}, \eta_i) \\ &\qquad + (1-\lambda)\sum_{i=k+1}^{+\infty} \lambda^{i-1}\phi(\eta_{i-1}, \eta_i) \\ &= (1-\lambda)\sum_{i=1}^{k} \lambda^{i-1}\phi(h_{i-1}, h_i) \\ &\qquad + \lambda^k(1-\lambda)\sum_{i=1}^{+\infty} \lambda^{i-1}\phi(\rho_{i-1}, \rho_i) \\ &= a + b \cdot \mathrm{DP}_{\mathrm{Min}}(\lambda)(\rho)\,. \end{aligned}$$

3. Case $\mathsf{Cost}_j = \mathrm{AP}_{\mathrm{Min}}$:
We set $a := 0 \in \mathbb{R}$ and $b := 1 \in \mathbb{R}^+$. Given $\rho \in \mathsf{Plays}$ such that $\mathsf{First}(\rho) = v$ and $\eta := h\rho \in \mathsf{Plays}$ (with $\eta = \eta_0\eta_1\ldots$), we show that:

$$\mathrm{AP}_{\mathrm{Min}}(h\rho) \;=\; \mathrm{AP}_{\mathrm{Min}}(\eta) \;=\; \mathrm{AP}_{\mathrm{Min}}(\rho)\,.$$

If $\mathrm{AP}_{\mathrm{Min}}(\eta) = \mathrm{AP}_{\mathrm{Min}}(\rho) = +\infty$ or $-\infty$, the desired result obviously holds. Otherwise, let us set $x_n := \frac{1}{n}\sum_{i=1}^{n} \phi(\eta_{i-1}, \eta_i)$ and

$y_n := \frac{1}{n}\sum_{i=1}^{n}\phi(\rho_{i-1},\rho_i)$, for all $n \in \mathbb{N}_0$. By properties of the limit superior and definition of the $\mathrm{AP}_{\mathrm{Min}}$ function, it holds that:

$$\limsup_{n\to+\infty}(x_n - y_n) \geq \mathrm{AP}_{\mathrm{Min}}(\eta) - \mathrm{AP}_{\mathrm{Min}}(\rho) \geq \liminf_{n\to+\infty}(x_n - y_n)\,.$$

It remains to prove that the sequence $(x_n - y_n)_{n\in\mathbb{N}}$ converges to 0. For all $n > k$, we have that:

$$\begin{aligned} |x_n - y_n| &= \left|\frac{1}{n}\cdot\left(\sum_{i=1}^{n}\phi(\eta_{i-1},\eta_i) - \sum_{i=k+1}^{k+n}\phi(\eta_{i-1},\eta_i)\right)\right| \\ &= \frac{1}{n}\cdot\left|\sum_{i=1}^{k}\phi(\eta_{i-1},\eta_i) - \sum_{i=n+1}^{n+k}\phi(\eta_{i-1},\eta_i)\right|. \end{aligned}$$

As the absolute value is bounded independently of n (let us remind that E is finite), we can conclude that $(x_n - y_n)_{n\in\mathbb{N}}$ converges to 0, and so $\mathrm{AP}_{\mathrm{Min}}(\eta) = \mathrm{AP}_{\mathrm{Min}}(\rho)$.

4. Case $\mathsf{Cost}_j = \mathrm{PRAvg}_{\mathrm{Min}}$:

We set $a := 0 \in \mathbb{R}$ and $b := 1 \in \mathbb{R}^+$. Given $\rho \in \mathsf{Plays}$ such that $\mathsf{First}(\rho) = v$ and $\eta := h\rho \in \mathsf{Plays}$ (with $\eta = \eta_0\eta_1\ldots$), we show that:

$$\mathrm{PRAvg}_{\mathrm{Min}}(h\rho) \;=\; \mathrm{PRAvg}_{\mathrm{Min}}(\eta) \;=\; \mathrm{PRAvg}_{\mathrm{Min}}(\rho)\,.$$

Thanks to several properties of lim sup, we have that:

$$\begin{aligned}
\mathrm{PRAvg}_{\mathrm{Min}}(\rho) &= \limsup_{n\to+\infty} \frac{\sum_{i=1}^{n}\phi(\rho_{i-1},\rho_i)}{\sum_{i=1}^{n}\vartheta(\rho_{i-1},\rho_i)} \\
&= \limsup_{n\to+\infty} \frac{\sum_{i=1}^{n}\phi(\eta_{k+i-1},\eta_{k+i})}{\sum_{i=1}^{n}\vartheta(\eta_{k+i-1},\eta_{k+i})} \\
&= \limsup_{n\to+\infty} \frac{\sum_{i=1}^{n+k}\phi(\eta_{i-1},\eta_i) - \sum_{i=1}^{k}\phi(\eta_{i-1},\eta_i)}{\sum_{i=1}^{n+k}\vartheta(\eta_{i-1},\eta_i) - \sum_{i=1}^{k}\vartheta(\eta_{i-1},\eta_i)} \\
&= \limsup_{n\to+\infty} \frac{\sum_{i=1}^{n+k}\phi(\eta_{i-1},\eta_i)}{\sum_{i=1}^{n+k}\vartheta(\eta_{i-1},\eta_i) - \sum_{i=1}^{k}\vartheta(\eta_{i-1},\eta_i)} \\
&= \limsup_{n\to+\infty} \frac{\sum_{i=1}^{n+k}\phi(\eta_{i-1},\eta_i)}{\sum_{i=1}^{n+k}\vartheta(\eta_{i-1},\eta_i)} \cdot \left(1 - \frac{\sum_{i=1}^{k}\vartheta(\eta_{i-1},\eta_i)}{\sum_{i=1}^{n+k}\vartheta(\eta_{i-1},\eta_i)}\right)^{-1} \\
&= \limsup_{n\to+\infty} \frac{\sum_{i=1}^{n+k}\phi(\eta_{i-1},\eta_i)}{\sum_{i=1}^{n+k}\vartheta(\eta_{i-1},\eta_i)} \\
&= \limsup_{n\to+\infty} \frac{\sum_{i=1}^{n}\phi(\eta_{i-1},\eta_i)}{\sum_{i=1}^{n}\vartheta(\eta_{i-1},\eta_i)} \\
&= \mathrm{PRAvg}_{\mathrm{Min}}(\eta) = \mathrm{PRAvg}_{\mathrm{Min}}(h\rho)\,.
\end{aligned}$$

Line 4 comes from the fact that the reward function ϑ is diverging, and from the following property: if $\lim_{n\to+\infty} b_n = b \in \mathbb{R}$, then $\limsup_{n\to+\infty}(a_n+b_n) = (\limsup_{n\to+\infty} a_n)+b$. Line 5 is implied by this property: if $\lim_{n\to+\infty} b_n = b > 0$, then $\limsup_{n\to+\infty}(a_n \cdot b_n) = (\limsup_{n\to+\infty} a_n) \cdot b$.

Note that, if the history h is empty, then $k = 0$ and, in all cases, a is equal to 0 and b to 1. This actually implies that $\mathsf{Cost}_i(h\rho) = \mathsf{Cost}_i(\rho)$ holds.

Moreover, the cost functions $\mathrm{RP}_{\mathrm{Min}}$, $\mathrm{DP}_{\mathrm{Min}}$, $\mathrm{AP}_{\mathrm{Min}}$ and $\mathrm{PRAvg}_{\mathrm{Min}}$ are positionally coalition-determined in $\mathcal{G}$. Indeed, given a player $i \in \Pi$, if $\mathsf{Cost}_i = \mathrm{RP}_{\mathrm{Min}}$, then we take $\mathsf{Gain}^i_{\mathrm{Max}} = \mathrm{RP}_{\mathrm{Max}}$. We do the same for the other cases by defining the gain function $\mathsf{Gain}^i_{\mathrm{Max}}$ for the coalition as the counterpart of Cost_i in Definition 2.3.3. By Theorem 2.3.10, the Min-Max cost game $\mathcal{G}^i = (\mathcal{A}^i, \mathsf{Cost}_i, \mathsf{Gain}^i_{\mathrm{Max}})$ is determined and has positional optimal strategies. Then, Proposition 4.4.6 applies and concludes the proof. □

Combining Qualitative and Quantitative Objectives

Multiplayer cost games also allow to encode games combining both qualitative and quantitative objectives in one objective, such as *mean-payoff parity games* [CHJ05]. In our framework, where each player aims at minimising his cost, the mean-payoff parity objective could be encoded as follows: $\mathsf{Cost}_i(\rho) = \mathrm{AP}_{\mathrm{Min}}(\rho)$ if the parity condition is satisfied, $+\infty$ otherwise.

The determinacy of mean-payoff parity games, together with the existence of optimal strategies (that could require infinite memory), have been proved in [CHJ05]. This result implies that a cost function encoding a mean-payoff parity objective is coalition-determined. Moreover, such objective is prefix-independent. By applying Proposition 4.4.11, we obtain the existence of a Nash equilibrium for multiplayer cost games with mean-payoff parity objectives. As far as we know, this is the first result about the existence of a Nash equilibrium in cost games with mean-payoff parity objectives.

Remark 4.4.16. Let us emphasise that Theorem 4.4.14 applies to cost games where the players have different kinds of cost functions (as in Example 2.3.15 on page 48). In particular, one player could have a qualitative Büchi objective, a second player a price-per-reward-average objective, a third player a mean-payoff parity objective,...

Chapter 5

Secure Equilibrium

In this chapter, based on [BBD12, BBDG13], we first extend the concept of *secure equilibrium* to the quantitative framework (in Section 5.1), and then we study this notion in quantitative reachability games (in Section 5.2).

5.1 Definition

The concept of secure equilibrium has already been defined for qualitative non-zero-sum games in Section 2.2.2. We here naturally extend [1] this notion to the quantitative framework. We first begin with the definition of a secure strategy profile, which generalises Definition 2.2.28.

Definition 5.1.1. Given a multiplayer cost game $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{Cost}_i)_{i\in\Pi})$ and an initial vertex $v_0 \in V$, a strategy profile $(\sigma_i)_{i\in\Pi}$ of $\mathcal{G}$ is called *secure* in $(\mathcal{G}, v_0)$ if, for every player $j \in \Pi$ and every strategy σ'_j of player j, we have that

$$\begin{aligned}\mathsf{Cost}_j(\rho') \leq \mathsf{Cost}_j(\rho) \quad \Rightarrow \quad & \Big((\forall i \neq j \ \ \mathsf{Cost}_i(\rho') \leq \mathsf{Cost}_i(\rho)) \\ & \vee \ (\exists i \neq j \ \ \mathsf{Cost}_i(\rho') < \mathsf{Cost}_i(\rho)) \Big),\end{aligned}$$

1. In fact, the inequalities comparing the gains of the players in Definitions 2.2.35 and 2.2.38 are just reversed, as we here consider cost functions to minimise.

where $\rho = \langle (\sigma_i)_{i\in\Pi}\rangle_{v_0}$ and $\rho' = \langle \sigma'_j, \sigma_{-j}\rangle_{v_0}$.

Exactly like in the qualitative case, a secure profile ensures that any deviation of a player that does not put him at a disadvantage cannot put the other players at a disadvantage either, if they follow the contract. Then, the definition of secure equilibrium follows.

Definition 5.1.2. Given a multiplayer cost game $(\mathcal{G}, v_0)$, a strategy profile of $\mathcal{G}$ is a *secure equilibrium* of $(\mathcal{G}, v_0)$ if it is a Nash equilibrium and it is secure in $(\mathcal{G}, v_0)$.

We also give an equivalent characterisation for secure equilibria, based on binary relations $(\prec_i)_{i\in\Pi}$ on cost profiles. These relations are defined in the same idea as in the qualitative case (see Equation (2.5)). Given $j \in \Pi$ and two cost profiles $(x_i)_{i\in\Pi}$, $(y_i)_{i\in\Pi}$:

$$(x_i)_{i\in\Pi} \prec_j (y_i)_{i\in\Pi} \quad \text{iff} \quad \big(x_j > y_j\big) \vee \big(x_j = y_j \wedge (\forall i \neq j\ x_i \leq y_i) \wedge (\exists i \neq j\ x_i < y_i)\big). \tag{5.1}$$

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 cost profile to another one either if he has a strictly lower cost, or if he keeps the same cost, the other players have a greater cost, and at least one has a strictly greater cost.

Definition 5.1.3. Given a multiplayer cost game $(\mathcal{G}, v_0)$, a strategy profile $(\sigma_i)_{i\in\Pi}$ of $\mathcal{G}$ is a *secure equilibrium* of $(\mathcal{G}, v_0)$ if, for every player $j \in \Pi$, there does not exist any strategy σ'_j of player j such that:

$$\mathsf{Cost}(\rho) \prec_j \mathsf{Cost}(\rho'),$$

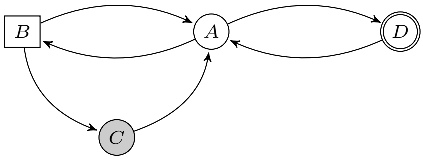
where $\rho = \langle (\sigma_i)_{i\in\Pi}\rangle_{v_0}$ and $\rho' = \langle \sigma'_j, \sigma_{-j}\rangle_{v_0}$.

In other words, player j has no incentive to deviate w.r.t. relation $\prec_j$.

Definition 5.1.4. Given a strategy profile $(\sigma_i)_{i\in\Pi}$, a strategy σ'_j of player j is called a *$\prec_j$-profitable deviation* for player j w.r.t. $(\sigma_i)_{i\in\Pi}$ in $(\mathcal{G}, v_0)$ if $\mathsf{Cost}(\langle (\sigma_i)_{i\in\Pi}\rangle_{v_0}) \prec_j \mathsf{Cost}(\langle \sigma'_j, \sigma_{-j}\rangle_{v_0})$.

Then, a strategy profile $(\sigma_i)_{i\in\Pi}$ is a secure equilibrium if no player j has a $\prec_j$-profitable deviation.

Example 5.1.5. Let $\mathcal{G} = (\{1,2\}, \mathcal{A}, (\mathsf{R}_1, \mathsf{R}_2))$ be the two-player quantitative reachability game whose arena $\mathcal{A}$ is depicted below and where $\mathsf{R}_1 = \{C\}$ and $\mathsf{R}_2 = \{D\}$ (same game as in Example 4.1.2 on page 62).



The memoryless strategy profile (σ'_1, σ_2), such that $\sigma'_1(A) = B$ and $\sigma_2(B) = C$, is a Nash equilibrium in $(\mathcal{G}, A)$, with outcome $(ABC)^\omega$ and cost profile $(2, +\infty)$ (see Example 4.1.2). However, it is *not* a secure equilibrium in $(\mathcal{G}, A)$. Indeed, player 2 has a $\prec_2$-profitable deviation: for instance, the memoryless strategy σ'_2 of player 2 defined by $\sigma'_2(B) = A$. Indeed, the outcome of (σ'_1, σ'_2) in $(\mathcal{G}, A)$ is the play $(AB)^\omega$, with cost profile $(+\infty, +\infty)$, and $(2, +\infty) \prec_2 (+\infty, +\infty)$.

In fact, any secure equilibrium in $(\mathcal{G}, A)$ must be of type $\emptyset$. For example, the memoryless strategy profile (σ'_1, σ'_2) is a secure equilibrium in this game.

5.2 Quantitative Reachability Objectives

In this section, we study the existence of secure equilibria in quantitative reachability games. In Section 5.2.1, we show, in particular, that there always exists a secure equilibrium in the two-player case. To our knowledge, it is still an open problem in the multiplayer case. Nevertheless, in Section 5.2.2, we prove that one can decide in $\mathsf{ExpSpace}$ whether there exists a secure equilibrium in an initialised multiplayer quantitative reachability game.

We remind that the definition of quantitative reachability games and some related notations can be found in Section 4.1.1.

5.2.1 Two-Player Case

In Section 4.1.4, we positively solved Problems 1 and 2 for Nash equilibria in multiplayer quantitative reachability games. We here solve these two problems for secure equilibria, but in the two-player case only. The main results are stated in Theorems 5.2.1 and 5.2.2 below. In this section, we exclusively consider two-player games.

Theorem 5.2.1. *In every initialised two-player quantitative reachability game, there exists a finite-memory secure equilibrium.*

Theorem 5.2.2. *Given a secure equilibrium in an initialised two-player quantitative reachability game, there exists a finite-memory secure equilibrium of the same type.*

Note that Theorem 5.2.2 is generalised to the multiplayer case in the next section (see Theorem 5.2.8).

The proof of Theorem 5.2.1 is based on the same ideas as for the proof of Theorem 4.1.5 (existence of a Nash equilibrium in quantitative reachability games). Given a two-player game $(\mathcal{G}, v_0)$ played on a finite graph G, we unravel the graph from v_0, as in Section 4.1.3, to get an equivalent game $\mathcal{T}$ played on the infinite tree T. We first show that while choosing suitable preference relations, Kuhn's theorem (Theorem 2.3.22) implies the existence of a secure equilibrium in the game $\mathsf{Trunc}_d(\mathcal{T})$ played on the finite tree $\mathsf{Trunc}_d(T)$, for any depth d. By choosing an adequate depth d, Proposition 5.2.4 enables to extend this secure equilibrium to a secure equilibrium in $\mathcal{T}$, and thus in $\mathcal{G}$.

To be able to apply Kuhn's theorem, we define the following preference relation $\precsim_j$ for player $j = 1, 2$, based on the relation $\prec_j$ of Equation (5.1). Given two cost profiles (x_1, x_2) and (y_1, y_2):

$$(x_1, x_2) \precsim_j (y_1, y_2) \text{ iff } \big((x_1, x_2) \prec_j (y_1, y_2) \vee (x_1 = y_1 \wedge x_2 = y_2)\big). \quad (5.2)$$

The relations $\precsim_1$ and $\precsim_2$ are clearly preference relations[2]. One can also be easily convinced that a strategy profile (σ_1, σ_2) is a *secure equilibrium* in $(\mathcal{G}, v_0)$ if for all strategies σ_1' of player 1, $\mathsf{Cost}(\rho') \precsim_1 \mathsf{Cost}(\rho)$, where

2. Remark that $\precsim_j$ is a kind of lexicographic order on $(\mathbb{N} \cup \{+\infty\}) \times (\mathbb{N} \cup \{+\infty\})$.

$\rho = \langle \sigma_1, \sigma_2 \rangle_{v_0}$ and $\rho' = \langle \sigma_1', \sigma_2 \rangle_{v_0}$, and symmetrically for all strategies σ_2' of player 2.

Note that we will see in Section 7.2 that these statements do not hold anymore in the multiplayer case.

Since $\precsim_1$ and $\precsim_2$ are preference relations in two-player games, we get the next corollary by Kuhn's theorem (Theorem 2.3.22).

Corollary 5.2.3. *In every two-player cost game whose graph is a* finite tree*, there exists a secure equilibrium.*

Now that we can guarantee the existence of a secure equilibrium in a two-player quantitative reachability game played on a finite tree $\mathsf{Trunc}_d(T)$, it remains to show how to lift it to the game played on the infinite tree T. The next proposition states that it is possible to extend a secure equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$ to a secure equilibrium in the game $\mathcal{T}$ with the same type, if the depth d is greater or equal to $(|\Pi|+1)\cdot 2\cdot |V| = 6\cdot|V|$ (since we consider two-player games). It also says that we can construct a secure equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$ from a secure equilibrium in $\mathcal{T}$, while keeping the same type. Recall that the type of a strategy profile in an initialised game is the set of players such that its outcome visits their goal sets.

Proposition 5.2.4. *Let $(\mathcal{G}, v_0)$ be a two-player quantitative reachability game and $\mathcal{T}$ be the corresponding game played on the unravelling of G from v_0.*

(i) If there exists a secure equilibrium of a certain type in the game $\mathcal{T}$, then there exists a secure equilibrium of the same type in the game $\mathsf{Trunc}_d(\mathcal{T})$, for some depth $d \geq 6 \cdot |V|$.

(ii) If there exists a secure equilibrium of a certain type in the game $\mathsf{Trunc}_d(\mathcal{T})$, where $d \geq 6 \cdot |V|$, then there exists a finite-memory secure equilibrium of the same type in the game $\mathcal{T}$.

To prove Proposition 5.2.4, we need the following technical lemma whose hypotheses are the same as in Lemma 4.1.7. As a reminder, this lemma roughly says that, given a Nash equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$, if its outcome has a prefix that fulfils some conditions, then the coalition of the players $i \neq j$ can play together to prevent player j from reaching his goal set R_j, from any vertex of this prefix.

Lemma 5.2.5. *Let $(\mathcal{G}, v_0)$ be a two-player quantitative reachability game, and $\mathcal{T}$ be the corresponding game played on the unravelling of G from v_0. For any depth $d \in \mathbb{N}$, let (σ_1, σ_2) be a secure equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$, and ρ the (finite) outcome of (σ_1, σ_2). Assume that ρ has a prefix $\alpha\beta\gamma$, where $\alpha, \beta, \gamma \in V^+$, such that*

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

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

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

In particular, Lemma 5.2.5 implies that if α visits none of the goal sets, then ρ visits either both goal sets or none. Notice that in the case of Nash equilibria, we can have situations contradicting Lemma 5.2.5, and in particular the previous situation, as it can be seen in Example 4.1.11 (on page 75).

Proof of Lemma 5.2.5. Let us suppose that the hypotheses of the lemma are fulfilled, and, by contradiction, we assume that $\big(\mathsf{Visit}(\alpha) \neq \emptyset \ \vee\ \mathsf{Visit}(\rho) \neq \{1,2\}\big) \ \wedge\ \mathsf{Visit}(\alpha) \neq \mathsf{Visit}(\rho)$. The last conjunct implies that $2 \in \mathsf{Visit}(\rho) \setminus \mathsf{Visit}(\alpha)$ or $1 \in \mathsf{Visit}(\rho) \setminus \mathsf{Visit}(\alpha)$. We consider the first case (the other one is symmetric). Then, by the first conjunct, we have that $1 \in \mathsf{Visit}(\alpha)$ or $1 \notin \mathsf{Visit}(\rho)$ (otherwise, $\mathsf{Visit}(\alpha) = \emptyset$ and $\mathsf{Visit}(\rho) = \{1,2\}$).

As α does not visit R_2, by Lemma 4.1.7, player 1 wins the game $\mathcal{G}^2$ from $v := \mathsf{Last}(\alpha)$, that is, has a memoryless winning strategy $\mu^v_{1,2}$ from this vertex, which prevents player 2 from reaching his goal set R_2. Then, if player 1 plays according to σ_1 until depth $|\alpha|$, and then switches to $\mu^v_{1,2}$ from $v = \mathsf{Last}(\alpha)$, this strategy is a $\prec_1$-profitable deviation for player 1 w.r.t. (σ_1, σ_2). Indeed, if $1 \in \mathsf{Visit}(\alpha)$, player 1 manages to increase player 2's cost while keeping his own cost. On the other hand, if $1 \notin \mathsf{Visit}(\rho)$, either player 1 succeeds in reaching his goal set while deviating (i.e. strictly decreases his cost), or he does not reach it (then gets the same cost as in ρ) but succeeds in increasing player 2's cost. We get thus a contradiction with the fact that (σ_1, σ_2) is a secure equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$. □

We can now give the proof of Proposition 5.2.4. The idea for showing case (i) is to look at the outcome π of the secure equilibrium in $\mathcal{T}$, and consider the minimal depth d needed to visit all the goal sets of the players in $\mathsf{Visit}(\pi)$. Then, the secure equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$ is defined exactly as the secure equilibrium of $\mathcal{T}$.

The proof of case (ii) works pretty much as the one of Proposition 4.1.6 (whereas the latter proposition does not preserve the type of the Nash equilibrium). Thanks to Lemma 5.2.5, the proof reduces into only two cases depending on when the goal sets are visited along the outcome of the secure equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$. In the most interesting case, a well-chosen prefix $\alpha\beta$, where β can be repeated (as a cycle), is first extracted from this outcome, and the outcome of the required secure equilibrium in $\mathcal{T}$ 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 the punishment is here more involved than in the proof of Proposition 4.1.6.

Before entering the details, let us remind the following notation: for any play $\rho = \rho_0\rho_1 \ldots$ of $\mathcal{G}$ and any player $i \in \{1, 2\}$, we write $\mathsf{Index}_i(\rho)$ for the least index l such that $\rho_l \in \mathsf{R}_i$ if it exists, or -1 if not [3].

Proof of Proposition 5.2.4. First, let us begin with the proof of (i). Suppose that there exists a secure equilibrium (τ_1, τ_2) in $\mathcal{T}$. We denote its outcome by π. Let us set $d := \max\{6 \cdot |V|, \mathsf{Index}_1(\pi), \mathsf{Index}_2(\pi)\}$, and define (σ_1, σ_2) as the strategy profile in $\mathsf{Trunc}_d(\mathcal{T})$ corresponding to the strategies (τ_1, τ_2) restricted to the finite tree $\mathsf{Trunc}_d(T)$. Clearly, the outcome ρ of (σ_1, σ_2) is a prefix of π, and $\mathsf{Visit}(\rho) = \mathsf{Visit}(\pi)$, so (σ_1, σ_2) and

3. We are conscious that it is counter-intuitive to use the particular value -1, but it is helpful in the proofs.

(τ_1, τ_2) are of the same type. It remains to show that (σ_1, σ_2) is a secure equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$.

Assume by contradiction that player 1 has a $\prec_1$-profitable deviation σ_1' w.r.t. (σ_1, σ_2) (the case of player 2 is symmetric). We write ρ' for the outcome of (σ_1', σ_2) in $\mathsf{Trunc}_d(\mathcal{T})$. There are two cases to consider: either player 1 manages to decrease his cost in ρ' w.r.t. ρ, or he pays the same cost as in ρ but he is able to increase the cost of player 2 in ρ' w.r.t. ρ. In both cases, if player 1 plays according to σ_1' in $\mathcal{T}$ until depth d and then arbitrarily, one can easily be convinced that we get a $\prec_1$-profitable deviation[4] w.r.t. (τ_1, τ_2) in $\mathcal{T}$. This leads to a contradiction.

Now let us proceed to the proof of (ii). Let (σ_1, σ_2) be a secure equilibrium in the game $\mathsf{Trunc}_d(\mathcal{T})$, where $d \geq 6 \cdot |V|$, and ρ be its outcome. We define the prefixes $\mathfrak{pq}$ and $\alpha\beta\gamma$ as in the proof of Proposition 4.1.6 on page 73 (see Figure 4.3 on page 75).

By Lemma 5.2.5, there are only two cases to consider:

(a) $\mathsf{Visit}(\alpha) = \emptyset$ and $\mathsf{Visit}(\rho) = \{1, 2\}$;

(b) $\mathsf{Visit}(\alpha) = \mathsf{Visit}(\rho)$.

We define a different secure equilibrium in $\mathcal{T}$ according to the case.

Let us start with case (a): $\mathsf{Visit}(\alpha) = \emptyset$ and $\mathsf{Visit}(\rho) = \{1, 2\}$. We define the following strategy profile: for every history $h \in \mathsf{Hist}_i$,

$$\tau_i(h) := \begin{cases} \sigma_i(h) & \text{if } |h| < \max\{\mathsf{Index}_1(\rho), \mathsf{Index}_2(\rho)\}, \\ \textit{arbitrary} & \text{otherwise.} \end{cases}$$

where $i = 1, 2$, and *arbitrary* means that the next vertex is chosen arbitrarily, but in a memoryless way. Note that the outcome of (τ_1, τ_2) is of the form $\alpha'(\beta')^\omega$ where $\mathsf{Visit}(\alpha') = \mathsf{Visit}(\rho) = \{1, 2\}$ and β' is a cycle. So, (τ_1, τ_2) has the same type as (σ_1, σ_2). It remains to prove that (τ_1, τ_2) is a finite-memory secure equilibrium in $\mathcal{T}$.

Assume by contradiction that player 1 has a $\prec_1$-profitable deviation τ_1' w.r.t. (τ_1, τ_2) in $\mathcal{T}$ (the case for player 2 is symmetric). The strategy σ_1' equal to τ_1' in $\mathsf{Trunc}_d(\mathcal{T})$ is clearly a $\prec_1$-profitable deviation

4. Notice that in the second case, when ρ and ρ' do not visit R_1 in $\mathsf{Trunc}_d(\mathcal{T})$, player 1 may reach his goal set in $\mathcal{T}$ when deviating in this way, and this would be profitable for him in this game.

w.r.t. (σ_1, σ_2), which is a contradiction with the fact that (σ_1, σ_2) is a secure equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$. Indeed, we have that $\mathsf{Type}(\tau_1, \tau_2) = \mathsf{Type}(\sigma_1, \sigma_2) = \{1, 2\}$, and the outcomes of the strategy profiles (τ_1, τ_2) and (σ_1, σ_2), as well as the outcomes of (τ_1', τ_2) and (σ_1', σ_2), coincide until a depth $\geq \max\{\mathsf{Index}_1(\rho), \mathsf{Index}_2(\rho)\}$. So, if player 1 strictly lowers his cost thanks to τ_1' in $\mathcal{T}$, then he also strictly lowers his cost thanks to σ_1' in $\mathsf{Trunc}_d(\mathcal{T})$. In the same way, if player 1 gets the same cost but strictly increases player 2's cost while deviating according to τ_1' in $\mathcal{T}$, then deviating according to σ_1' in $\mathsf{Trunc}_d(\mathcal{T})$ induces the same effect on the players' costs in this game.

Moreover, (τ_1, τ_2) is obviously a finite-memory strategy profile (for more details, see the proof of Lemma 4.1.8).

Now we consider case (b): $\mathsf{Visit}(\alpha) = \mathsf{Visit}(\rho)$. Like in the proof of Lemma 4.1.8, 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 following: player 2 (resp. 1) plays according to $\alpha\beta^\omega$ and punishes player 1 (resp. 2) if he deviates from $\alpha\beta^\omega$, in the following way. Suppose that player 1 deviates (the case for player 2 is similar). Then player 2 plays according to σ_2 until depth $|\alpha|$, and after that, he plays arbitrarily if α visits R_1, otherwise he plays according to a memoryless strategy given by Lemma 4.1.7 that prevents player 1 from reaching his goal set.

In order to describe the secure equilibrium in $\mathcal{T}$, we define the same punishment function P as in the proof of Lemma 4.1.8: for v_0, we define $P(v_0) = \perp$, and for every history $hv \in \mathsf{Hist}$ ($v \in V$) starting in v_0, we let:

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

The definition of the secure equilibrium (τ_1, τ_2) in $\mathcal{T}$ is as follows: for every history h of $(\mathcal{G}, v_0)$ ending in a vertex of V_i,

$$\tau_i(h) := \begin{cases} v & \text{if } P(h) = \bot \ (h < \alpha\beta^\omega)\text{; such that } hv < \alpha\beta^\omega, \\ \sigma_i(h) & \text{if } P(h) \neq \bot, i, \text{ and } |h| \leq |\alpha|, \\ \mu^v_{i,P(h)}(v') & \text{if } P(h) \neq \bot, i,\ |h| > |\alpha| \text{ and } P(h) \notin \mathsf{Visit}(\alpha), \\ & \qquad \text{s.t. } h = h'vh''v' \ (v, v' \in V) \text{ and } |h'| = |\alpha|, \\ \textit{arbitrary} & \text{otherwise,} \end{cases}$$

where $i = 1, 2$, and *arbitrary* means that the next vertex is chosen arbitrarily (in a memoryless way). Notice that in the third case, the strategy $\mu^v_{i,P(h)}$ is the memoryless winning strategy of player i given by Lemma 4.1.7 ($P(h) \notin \mathsf{Visit}(\alpha)$) when considering the history $h'v$ (see h' above).

Clearly, the outcome of (τ_1, τ_2) is the play $\alpha\beta^\omega$, and the type of (τ_1, τ_2) is equal to $\mathsf{Visit}(\alpha) = \mathsf{Visit}(\rho)$, the type of (σ_1, σ_2). Moreover, as done in the proof of Lemma 4.1.8, (τ_1, τ_2) is a finite-memory strategy profile.

Remark that the definition of the strategy profile (τ_1, τ_2) is a little different from the one in the proof of Lemma 4.1.8 because here, if player 1 deviates (for example), then player 2 has to prevent him from reaching his goal set R_1 (faster), or having the same cost but succeeding in increasing player 2's cost.

It remains to show that (τ_1, τ_2) is a secure equilibrium in the game $\mathcal{T}$. Assume by contradiction 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 $\mathsf{Trunc}_d(\mathcal{T})$ as follows: player 1 plays according to the strategy τ'_1 restricted to $\mathsf{Trunc}_d(\mathcal{T})$ (denoted by σ'_1) and player 2 plays according to σ_2. Thus the play ρ' coincide with the play $\pi' = \langle \tau'_1, \tau_2 \rangle_{v_0}$ at least until depth $|\alpha|$ (by definition of τ_2); it can differ afterwards. We let:

$$\begin{array}{rclll} \rho & = & \langle \sigma_1, \sigma_2 \rangle_{v_0} & \text{of cost profile} & (x_1, x_2) \\ \rho' & = & \langle \sigma'_1, \sigma_2 \rangle_{v_0} & \text{of cost profile} & (x'_1, x'_2) \\ \pi & = & \langle \tau_1, \tau_2 \rangle_{v_0} & \text{of cost profile} & (y_1, y_2) \\ \pi' & = & \langle \tau'_1, \tau_2 \rangle_{v_0} & \text{of cost profile} & (y'_1, y'_2). \end{array}$$

The situation is depicted in Figure 5.1.

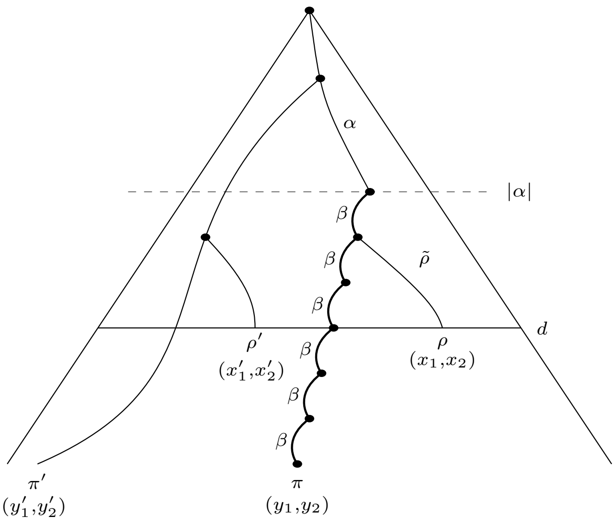


Figure 5.1: Plays ρ and π, and their respective deviations ρ' and π'.

By contradiction, we assumed that τ'_1 is a $\prec_1$-profitable deviation for player 1 w.r.t. (τ_1, τ_2), i.e. $(y_1, y_2) \prec_1 (y'_1, y'_2)$. Now 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 $\mathsf{Trunc}_d(\mathcal{T})$. 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.

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

As $\pi = \alpha\beta^\omega$, it means that α visits R_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 $x'_1 < x_1$, and $(x_1, x_2) \prec_1 (x'_1, x'_2)$.

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

If $y'_1 \leq |\alpha|$, we have $x'_1 = y'_1$ (by the same argument as before). As $\mathsf{Visit}(\alpha) = \mathsf{Visit}(\rho)$, we have $x_1 = y_1 = +\infty$ and $x'_1 < x_1$ (and so $(x_1, x_2) \prec_1 (x'_1, x'_2)$).

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 $\mu^v_{2,1}$ (as $y_1 = +\infty$). By Lemma 4.1.7, the play π' can not visit R_1 after a depth $> |\alpha|$.

(3) $y_1 = y'_1$ 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 cost of player 2 is increased. In order to ensure that σ'_1 is a $\prec_1$-profitable deviation, it remains to show that either player 1 keeps the same cost, or he decreases his cost.

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 cost in ρ and ρ'. And so, $(x_1, x_2) \prec_1 (x'_1, x'_2)$.

On the contrary, if $y'_1 = y_1 = +\infty$, it follows that $x_1 = +\infty$ (as $\mathsf{Visit}(\alpha) = \mathsf{Visit}(\rho)$). And so, we have that $x'_1 < +\infty = x_1$, or $x'_1 = x_1$. But in both cases, it holds that $(x_1, x_2) \prec_1 (x'_1, x'_2)$.

In conclusion, we constructed a $\prec_1$-profitable deviation σ'_1 w.r.t. (σ_1, σ_2) in $\mathsf{Trunc}_d(\mathcal{T})$, and then we get a contradiction. □

Remark 5.2.6. Let us notice that in case (i) of Proposition 5.2.4, the proof remains valid if we take $d = \max\{0, \mathsf{Index}_1(\pi), \mathsf{Index}_2(\pi)\}$. Thus, in the statement of case (i), the constraint $d \geq 6 \cdot |V|$ can be replaced by $d \in \mathbb{N}$.

We can now proceed to the proof of Theorem 5.2.1, which states the existence of a finite-memory secure equilibrium in every two-player quantitative reachability game.

Proof of Theorem 5.2.1. Let $(\mathcal{G}, v_0)$ be a two-player quantitative reachability game and $\mathcal{T}$ be the corresponding game played on the unravelling of G from v_0. Let us set $d := 6 \cdot |V|$ and apply Corollary 5.2.3 on the game $\mathsf{Trunc}_d(\mathcal{T})$. Then we get a secure equilibrium in this game. By Proposition 5.2.4, there exists a finite-memory secure equilibrium in $\mathcal{G}$. □

Finally, we prove Theorem 5.2.2, which asserts that, if there exists a secure equilibrium in an initialised two-player quantitative reachability game, then there exists a finite-memory secure equilibrium of the same type.

Proof of Theorem 5.2.2. Let $(\mathcal{G}, v_0)$ be a two-player quantitative reachability game and $\mathcal{T}$ be the corresponding game played on the unravelling of G from v_0. Let (τ_1, τ_2) be a secure equilibrium in $(\mathcal{G}, v_0)$. By the first part of Proposition 5.2.4, there exists a secure equilibrium of the same type in the game $\mathsf{Trunc}_d(\mathcal{T})$, for a certain depth $d \geq 6 \cdot |V|$. If we apply the second part of Proposition 5.2.4, we get a finite-memory secure equilibrium of the same type as (τ_1, τ_2) in $(\mathcal{G}, v_0)$. □

The proof of Theorem 5.2.2 is based on Proposition 5.2.4 which, roughly speaking, ensures that every secure equilibrium of $\mathsf{Trunc}_d(\mathcal{T})$ can be lifted to a secure equilibrium *of the same type* in $\mathcal{T}$, and conversely. Notice that Proposition 5.2.4 has no counterpart for Nash equilibria, since we can not guarantee that the type can be preserved, as it can be seen from Example 4.1.11 (on page 75). This approach makes the proof of Theorem 5.2.2 rather different than the proof of Theorem 4.1.12 (which states that, given a Nash equilibrium in an initialised multiplayer quantitative reachability game, there exists a finite-memory Nash equilibrium of the same type).

Notice that Proposition 5.2.4 stands for two-player games because its proof uses Lemma 5.2.5 that has only been proved in the two-player case.

5.2.2 Multiplayer Case

In this section, we study Problems 1 and 2 in the context of secure equilibria and multiplayer games. Both problems have been positively solved in Section 5.2.1 for two-player games only. To the best of our knowledge, the existence of secure equilibria in the multiplayer framework is still an open problem. We here provide an algorithm that *decides* the existence of a secure equilibrium in multiplayer quantitative reachability games.

Theorem 5.2.7. *In every initialised multiplayer quantitative reachability game, one can decide whether there exists a secure equilibrium in* ExpSpace.

We also show that if there exists a secure equilibrium, then there exists one that is finite-memory and has the same type.

Theorem 5.2.8. *If there exists a secure equilibrium in an initialised multiplayer quantitative reachability game, then there exists a finite-memory secure equilibrium of the same type.*

This theorem gives a counterpart to Theorem 4.1.12, which states the same result but for Nash equilibria.

The proof of Theorem 5.2.7 is inspired from ideas developed in Sections 4.1.4 and 5.2.1. The key point is to show that the existence of a secure equilibrium in a quantitative reachability game $(\mathcal{G}, v_0)$ is equivalent to the existence of a secure equilibrium (with two additional properties) in the finite game $\mathsf{Trunc}_d(\mathcal{T})$ for a well-chosen depth d. The existence of the latter equilibrium is decidable. Notice that we can show [5] that a secure equilibrium always exists in $\mathsf{Trunc}_d(\mathcal{T})$; however, we do not know if there always exists a secure equilibrium with the two required additional properties in $\mathsf{Trunc}_d(\mathcal{T})$.

Let us formally introduce these two properties. The first one requires that the secure equilibrium is *goal-optimised*, meaning that all the goal sets visited along its outcome are visited for the first time *before a certain given depth*. For any game $\mathcal{G}$ played on a graph with $|V|$ vertices by $|\Pi|$ players, we fix the following constant: $d_{goal}(\mathcal{G}) := 2 \cdot |\Pi| \cdot |V|$.

Definition 5.2.9. Given a quantitative reachability game $(\mathcal{G}, v_0)$ and a strategy profile $(\sigma_i)_{i\in\Pi}$ in $\mathcal{G}$, with outcome ρ, we say that $(\sigma_i)_{i\in\Pi}$ is

5. See Section 7.2 and Corollary 7.2.2 for the existence of a stronger solution concept in multiplayer cost games played on finite trees.

goal-optimised if and only if for all $i \in \Pi$ such that $\mathsf{Cost}_i(\rho) < +\infty$, we have that $\mathsf{Cost}_i(\rho) < d_{goal}(\mathcal{G})$.

The second property asks for a secure equilibrium that is *deviation-optimised*, meaning that whenever a player deviates from its outcome, he realises *within a certain given number of steps* that his deviation is not profitable for him.

Definition 5.2.10. Given a quantitative reachability game $(\mathcal{G}, v_0)$ and a secure equilibrium $(\sigma_i)_{i\in\Pi}$ in $\mathcal{G}$, with outcome ρ, we say that $(\sigma_i)_{i\in\Pi}$ is *deviation-optimised* if and only if for every player $j \in \Pi$ and every strategy σ'_j of player j, we have that

$$\mathsf{Cost}(\rho_{<d_{dev}}) \not\prec_j \mathsf{Cost}(\rho'_{<d_{dev}}),$$

where $d_{dev} = \max\{\mathsf{Cost}_i(\rho) \mid \mathsf{Cost}_i(\rho) < +\infty\} + |V|$ and $\rho' = \langle \sigma'_j, \sigma_{-j} \rangle_{v_0}$.

Remark that Definitions 5.2.9 and 5.2.10 extend to games $\mathsf{Trunc}_d(\mathcal{T})$ where $d \geq d_{goal}(\mathcal{G})$.

We can now state the key proposition for proving Theorems 5.2.7 and 5.2.8.

Proposition 5.2.11. *Let $(\mathcal{G}, v_0)$ be a multiplayer quantitative reachability game and $\mathcal{T}$ be the corresponding game played on the unravelling of G from v_0. We set $d := d_{goal}(\mathcal{G}) + 3 \cdot |V|$.*

1. *If there exists a goal-optimised and deviation-optimised secure equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$, then there exists a finite-memory secure equilibrium in $(\mathcal{G}, v_0)$.*
2. *If there exists a secure equilibrium in $(\mathcal{G}, v_0)$, then there exists a secure equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$ that is goal-optimised and deviation-optimised.*

At this stage, it is difficult to give some intuition about the choice of the values $d_{goal}(\mathcal{G})$, d_{dev} and $d = d_{goal}(\mathcal{G}) + 3 \cdot |V|$. These values are linked to the proofs contained in this section.

The proof of Proposition 5.2.11 is long and technical, it will be done thereafter. We first prove Theorems 5.2.7 and 5.2.8 thanks to this proposition.

Proof of Theorem 5.2.7. Proposition 5.2.11 implies that there exists a secure equilibrium in $(\mathcal{G}, v_0)$ iff there exists a goal-optimised and deviation-optimised secure equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$, with $d = d_{goal}(\mathcal{G}) + 3 \cdot |V|$. The latter property is decidable in NExpSpace (in $|V|$ and $|\Pi|$). Indeed, the size of $\mathsf{Trunc}_d(\mathcal{T})$ is exponential in the size of $\mathcal{G}$. Guessing a strategy profile $(\sigma_i)_{i\in\Pi}$ in this tree also needs an exponential size (in the size of $\mathcal{G}$). Then we can test in exponential size (in the size of $\mathcal{G}$) whether $(\sigma_i)_{i\in\Pi}$ is a goal-optimised and deviation-optimised secure equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$. By Savitch's theorem, deciding the existence of a secure equilibria is thus in ExpSpace. □

Proof of Theorem 5.2.8. This theorem is a direct consequence of Proposition 5.2.11. Indeed, consider a secure equilibrium in a game $(\mathcal{G}, v_0)$. We first apply Proposition 5.2.11 (Part (ii)) to this strategy profile to get a goal-optimised and deviation-optimised secure equilibrium $(\sigma_i)_{i\in\Pi}$ in $\mathsf{Trunc}_d(\mathcal{T})$, for $d = d_{goal}(\mathcal{G}) + 3 \cdot |V|$. Then, we apply Proposition 5.2.11 (Part (i)) to the equilibrium $(\sigma_i)_{i\in\Pi}$, to get a finite-memory secure equilibrium back in $(\mathcal{G}, v_0)$. Moreover, Remarks 5.2.13 and 5.2.21 imply that this secure equilibrium has the same type as the initial one. □

The next two sections are devoted to the proof of the two parts of Proposition 5.2.11.

Part (i) of Proposition 5.2.11

This section is devoted to the proof of Proposition 5.2.11, Part (i). We begin with a useful characterisation of a deviation-optimised secure equilibrium.

Lemma 5.2.12. *With the previous notations of Definition 5.2.10, a secure equilibrium $(\sigma_i)_{i\in\Pi}$ is deviation-optimised if and only if for every player $j \in \Pi$ and every strategy σ'_j of player j, if*

1. $\mathsf{Cost}_j(\rho) = \mathsf{Cost}_j(\rho')$,
2. $\forall\, i \in \Pi$ *such that* $\mathsf{Cost}_i(\rho) < +\infty$, *we have that* $\mathsf{Cost}_i(\rho) \le \mathsf{Cost}_i(\rho')$,
3. $\exists\, i \in \Pi$ $\mathsf{Cost}_i(\rho) < \mathsf{Cost}_i(\rho')$,

then there exists $l \in \Pi$ *such that* $\mathsf{Cost}_l(\rho) = +\infty$ *and* $\mathsf{Cost}_l(\rho') < d_{dev}$.

Proof. Let us first assume that $(\sigma_i)_{i\in\Pi}$ is a deviation-optimised secure equilibrium whose outcome is denoted by ρ. Given any player $j \in \Pi$, let σ'_j be a strategy fulfilling the hypotheses of the lemma and ρ' the outcome given by $\langle \sigma'_j, \sigma_{-j}\rangle_{v_0}$. Let us denote respectively by $(x_i)_{i\in\Pi}$ and $(y_i)_{i\in\Pi}$ the cost profiles of the histories $\rho_{<d_{dev}}$ and $\rho'_{<d_{dev}}$.

Notice that by definition of d_{dev}, $\mathsf{Cost}_i(\rho) = x_i$ for all $i \in \Pi$. For ρ', we have $\mathsf{Cost}_i(\rho') = y_i$ provided $\mathsf{Cost}_i(\rho') < d_{dev}$. Otherwise, it may happen that $y_i = +\infty$ and $\mathsf{Cost}_i(\rho') < +\infty$. So, it holds that $\mathsf{Cost}_i(\rho') \leq y_i$ for all $i \in \Pi$. These observations will be often used in the sequel of the proof.

By hypothesis, we know that $\mathsf{Cost}(\rho_{<d_{dev}}) \not\prec_j \mathsf{Cost}(\rho'_{<d_{dev}})$, which means:

$$\big(x_j \leq y_j\big) \wedge \big(x_j \neq y_j \vee (\exists i \in \Pi\ x_i > y_i) \vee (\forall i \in \Pi\ x_i \geq y_i)\big)\,. \tag{5.3}$$

By hypothesis (i), it holds that $x_j = y_j$. By hypothesis (iii), we cannot have: $\forall i \in \Pi, x_i \geq y_i$. Therefore to satisfy Equation (5.3), there must exist a player i such that $x_i > y_i$. If $\mathsf{Cost}_i(\rho) < +\infty$, then by definition of d_{dev}, $\mathsf{Cost}_i(\rho) = x_i > y_i = \mathsf{Cost}_i(\rho')$ in contradiction with hypothesis (ii). Therefore $\mathsf{Cost}_i(\rho) = +\infty$. From $x_i > y_i$, it follows that $\mathsf{Cost}_i(\rho') < d_{dev}$, which concludes the first implication of the proof.

For the converse, let us now assume that $(\sigma_i)_{i\in\Pi}$ is a secure equilibrium that fulfils the property stated in Lemma 5.2.12. We will prove that it is deviation-optimised, that is, for any player $j \in \Pi$, and any deviation σ'_j of player j, we have that $\mathsf{Cost}(\rho_{<d_{dev}}) \not\prec_j \mathsf{Cost}(\rho'_{<d_{dev}})$, with $\rho = \langle(\sigma_i)_{i\in\Pi}\rangle_{v_0}$ and $\rho' = \langle \sigma'_j, \sigma_{-j}\rangle_{v_0}$. By denoting respectively by $(x_i)_{i\in\Pi}$ and $(y_i)_{i\in\Pi}$ the cost profiles of $\rho_{<d_{dev}}$ and $\rho'_{<d_{dev}}$, it is equivalent to prove Equation (5.3).

Since $(\sigma_i)_{i\in\Pi}$ is a secure equilibrium, we know that σ'_j is not a $\prec_j$-profitable deviation. In particular, player j can not strictly decrease his cost along ρ', and thus $x_j \leq y_j$. It remains to prove that the second conjunct of Equation (5.3) is true. For this, we first show that as soon as one of the hypotheses among (i), (ii) or (iii) is not fulfilled, this conjunct is satisfied.

- If $\mathsf{Cost}_j(\rho) < \mathsf{Cost}_j(\rho')$, by choice of d_{dev}, we also have that $x_j < y_j$. Moreover, the case $\mathsf{Cost}_j(\rho) > \mathsf{Cost}_j(\rho')$ is not possible as $(\sigma_i)_{i\in\Pi}$

is a secure equilibrium.

- If there exists $i \in \Pi$ such that $\mathsf{Cost}_i(\rho) < +\infty$ and $\mathsf{Cost}_i(\rho) > \mathsf{Cost}_i(\rho')$, then $x_i > y_i$.
- If for all $i \in \Pi$, $\mathsf{Cost}_i(\rho) \geq \mathsf{Cost}_i(\rho')$, we also have that $x_i \geq y_i$, for all $i \in \Pi$.

Thus the remaining deviations to consider fulfil hypotheses (i), (ii) and (iii). In this case, there exists $l \in \Pi$ such that $\mathsf{Cost}_l(\rho) = +\infty$ and $\mathsf{Cost}_l(\rho') < d_{dev}$. In particular we have that $x_l > y_l$, and the second conjunct of Equation (5.3) is true. □

The ideas of the proof for Part (i) of Proposition 5.2.11 are as follows. Suppose that there exists a goal-optimised and deviation-optimised secure equilibrium $(\sigma_i)_{i\in\Pi}$ in $\mathsf{Trunc}_d(\mathcal{T})$, for $d = d_{goal}(\mathcal{G}) + 3 \cdot |V|$. In order to get from $(\sigma_i)_{i\in\Pi}$ a finite-memory secure equilibrium in $(\mathcal{G}, v_0)$, we use a similar construction as in the proof of Proposition 5.2.4, where it is shown, in the context of two-player games, how to extend a secure equilibrium in a finite truncation of $(\mathcal{G}, v_0)$ to a secure equilibrium in $(\mathcal{G}, v_0)$. The rough idea is as follows. Due to the hypotheses, the outcome π of $(\sigma_i)_{i\in\Pi}$ has a prefix $\alpha\beta$ such that all goal sets visited by π are already visited by α, and such that β can be repeated (as a cycle). The required secure equilibrium is specified such that its outcome is equal to $\alpha\beta^\omega$ and any deviating player is punished by the coalition of the other players in a way that this deviation is not profitable for him. This secure equilibrium can be constructed in a way to be finite-memory.

Proof of Proposition 5.2.11, Part (i). Let $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{R}_i)_{i\in\Pi})$ be a multiplayer quantitative game, and let us set $\Pi = \{1, \ldots, n\}$. Let $(\tau_i)_{i\in\Pi}$ be a goal-optimised and deviation-optimised secure equilibrium in the game $\mathsf{Trunc}_d(\mathcal{T})$ and π its outcome. Since $|\pi| = d_{goal}(\mathcal{G}) + 3 \cdot |V|$, we can write

$$\begin{aligned}\pi = \alpha\beta\gamma \quad \text{with} \quad & \alpha, \beta \in V^+, \gamma \in V^* \\ & \mathsf{Last}(\alpha) = \mathsf{Last}(\alpha\beta) \\ & |\alpha| \geq d_{goal}(\mathcal{G}) + |V| \\ & |\alpha\beta| \leq d_{goal}(\mathcal{G}) + 2 \cdot |V| \,.\end{aligned}$$

We have that $\mathsf{Visit}(\alpha) = \mathsf{Visit}(\alpha\beta\gamma)$ (no new goal set is visited after α) because $|\alpha| \geq d_{goal}(\mathcal{G})$ and $(\tau_i)_{i\in\Pi}$ is goal-optimised. This enables us to use Lemma 4.1.7 as follows. Let $j \in \Pi$ be such that α does not visit R_j, and suppose that player j deviates from the history α. This lemma states that for all histories hv consistent with τ_{-j} and such that $|hv| \leq |\alpha\beta|$, then the coalition formed by all the players $i \in \Pi \setminus \{j\}$ can play to prevent player j from reaching his goal set R_j from vertex v. It means that this coalition has a memoryless winning strategy $\mu^v_{\mathtt{C}j}$ from vertex v in the zero-sum qualitative reachability game $\mathcal{G}^j = (\mathcal{A}^j, \mathsf{R}_j)$, where $\mathcal{A}^j = (V, (V_j, V \setminus V_j), E)$, and player j aims at reaching R_j while the coalition $\Pi \setminus \{j\}$ wants to prevent this (see Theorem 2.2.15). For each player $i \neq j$, let $\mu^v_{i,j}$ be the memoryless strategy of player i in $\mathcal{G}$ induced by $\mu^v_{\mathtt{C}j}$.

We define a finite-memory secure equilibrium in the game $\mathcal{T}$ using the same idea as in the proof of Proposition 5.2.4. The idea is to specify the required secure equilibrium as follows: each player i plays according to $\alpha\beta^\omega$ (which is the outcome of this equilibrium) and punishes player $j \neq i$ if he deviates from $\alpha\beta^\omega$, by playing according to τ_i until depth $|\alpha|$, and after that, by playing arbitrarily if α visits R_j, and according to $\mu^v_{i,j}$ otherwise (where v is the vertex visited at depth $|\alpha|$ when deviating).

Formally we first need to specify a punishment function P. For the initial vertex v_0, we define $P(v_0) = \bot$ and for all histories $hv \in \mathsf{Hist}$ such that $h \in \mathsf{Hist}_i$, we let:

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

Then the definition of the secure equilibrium $(\sigma_i)_{i\in\Pi}$ in $\mathcal{T}$ is as follows.

For all $i \in \Pi$ and $h \in \mathsf{Hist}_i$,

$$\sigma_i(h) := \begin{cases} v & \text{if } P(h) = \bot\ (h < \alpha\beta^\omega)\text{; such that } hv < \alpha\beta^\omega, \\ \tau_i(h) & \text{if } P(h) \neq \bot, i \text{ and } |h| \leq |\alpha|, \\ \mu^v_{i,P(h)}(v') & \text{if } P(h) \neq \bot, i,\ |h| > |\alpha|, \text{ and } P(h) \notin \mathsf{Visit}(\alpha); \\ & \quad \text{s.t. } h = h'vh''v'\ (v, v' \in V),\ |h'| = |\alpha|, \\ & \quad \text{and } h'v \text{ consistent with } \tau_{-j}, \\ \textit{arbitrary} & \text{otherwise,} \end{cases}$$

where *arbitrary* means that the next vertex is chosen arbitrarily (in a memoryless way). Clearly the outcome of $(\sigma_i)_{i\in\Pi}$ is the play $\alpha\beta^\omega$.

Let us show that $(\sigma_i)_{i\in\Pi}$ is a secure equilibrium in the game $\mathcal{T}$. Assume by contradiction that there exists a $\prec_j$-profitable deviation σ'_j for player j w.r.t. $(\sigma_i)_{i\in\Pi}$ in $\mathcal{T}$. Let τ'_j be the strategy σ'_j restricted to $\mathsf{Trunc}_d(\mathcal{T})$. We are going to show that τ'_j is a $\prec_j$-profitable deviation for player j w.r.t. $(\tau_i)_{i\in\Pi}$ in $\mathsf{Trunc}_d(\mathcal{T})$, which is impossible by hypothesis. Here are some useful notations:

$$\begin{array}{lcll} \pi & = & \langle(\tau_i)_{i\in\Pi}\rangle_{v_0} & \text{of cost profile} \quad (x_1,\ldots,x_n) \\ \pi' & = & \langle\tau'_j, \tau_{-j}\rangle_{v_0} & \text{of cost profile} \quad (x'_1,\ldots,x'_n) \\ \rho & = & \langle(\sigma_i)_{i\in\Pi}\rangle_{v_0} & \text{of cost profile} \quad (y_1,\ldots,y_n) \\ \rho' & = & \langle\sigma'_j, \sigma_{-j}\rangle_{v_0} & \text{of cost profile} \quad (y'_1,\ldots,y'_n). \end{array}$$

Notice that the play π' coincide with the play ρ' at least until depth $|\alpha|$ (by definition of τ'_j and σ_{-j}); they can differ afterwards. Clearly π and ρ coincide at least until depth $|\alpha\beta|$. The situation is depicted in Figure 5.2.

As σ'_j is a $\prec_j$-profitable deviation for player j w.r.t. $(\sigma_i)_{i\in\Pi}$, we have that

$$(y_1,\ldots,y_n) \prec_j (y'_1,\ldots,y'_n). \tag{5.4}$$

Let us show that τ'_j is a $\prec_j$-profitable deviation for player j w.r.t. $(\tau_i)_{i\in\Pi}$, i.e.,

$$(x_1,\ldots,x_n) \prec_j (x'_1,\ldots,x'_n).$$

By Equation (5.4), one of the next three cases stands.

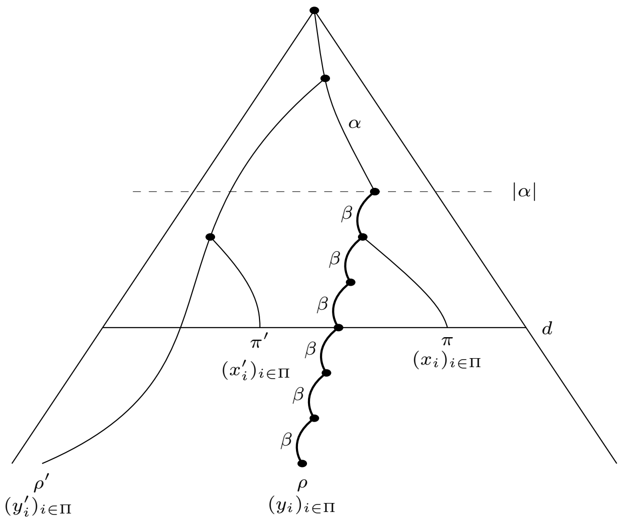


Figure 5.2: Plays π and ρ, and their respective deviations π' and ρ'.

(1) $y'_j < y_j < +\infty$.

As $\rho = \alpha\beta^\omega$ and $\mathsf{Visit}(\alpha) = \mathsf{Visit}(\alpha\beta\gamma)$, it means that α visits R_j, and then $y_j = x_j$. Since $y'_j < |\alpha|$, we also have $x'_j = y'_j$ (as π' and ρ' coincide until depth $|\alpha|$). Therefore $x'_j < x_j$, and $(x_1, \ldots, x_n) \prec_j (x'_1, \ldots, x'_n)$.

(2) $y'_j < y_j = +\infty$.

If $y'_j \leq |\alpha|$, we have again $x'_j = y'_j$. Since $\mathsf{Visit}(\alpha) = \mathsf{Visit}(\pi)$, it follows that $x_j = y_j = +\infty$. Thus $x'_j < x_j$, and so $(x_1, \ldots, x_n) \prec_j (x'_1, \ldots, x'_n)$.

We show that the case $y'_j > |\alpha|$ is impossible. By definition of σ_{-j}, the play ρ' is consistent with τ_{-j} until depth $|\alpha|$, and then with $\mu^v_{\mathtt{C}j}$ from $\rho'_{|\alpha|}$ (as $y_j = +\infty$). The play ρ' cannot visit R_j after a depth $> |\alpha|$ by definition of $\mu^v_{\mathtt{C}j}$.

(3) $y_j = y'_j$, $\forall i \in \Pi\ y_i \leq y'_i$ and $\exists\, i \in \Pi\ y_i < y'_i$.

The fact that $y_j = y'_j$ implies $y_j = x_j \geq x'_j$ (as $\mathsf{Visit}(\alpha) = \mathsf{Visit}(\pi)$). If $x'_j < x_j$, then $(x_1, \ldots, x_n) \prec_j (x'_1, \ldots, x'_n)$.

We show that the case $x'_j = x_j$ is impossible. We can show that for all $i \in \Pi$ such that $x_i < +\infty$, we have $x_i \leq x'_i$, and that there exists $i \in \Pi$ such that $x_i < x'_i$. Since $(\tau_i)_{i\in\Pi}$ is deviation-optimised, Lemma 5.2.12 implies that there exists some $l \in \Pi$ such that $x_l = +\infty$, and $x'_l < d_{dev} = \max\{x_i \mid x_i < +\infty\} + |V|$. As $(\tau_i)_{i\in\Pi}$ is also goal-optimised, we have that $d_{dev} \leq d_{goal}(\mathcal{G}) + |V| \leq |\alpha|$. As ρ' is consistent with τ_{-j} until depth $|\alpha|$, it follows that $y'_l = x'_l < y_l = x_l = +\infty$. Thus case (3) is impossible.

Therefore, each case is either impossible or shows that $(x_i)_{i\in\Pi} \prec_j (x'_i)_{i\in\Pi}$. This is in contradiction with $(\tau_i)_{i\in\Pi}$ being a secure equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$, and therefore, $(\sigma_i)_{i\in\Pi}$ is a secure equilibrium in $\mathcal{T}$, thus in $(\mathcal{G}, v_0)$.

It remains to show that $(\sigma_i)_{i\in\Pi}$ is a finite-memory strategy profile. This proof is very similar to the proof of Proposition 5.2.4 and thus is not given in details. Roughly speaking, a finite amount of memory is enough to produce the outcome $\alpha\beta^\omega$; outside of this outcome it is enough to remember how $(\sigma_i)_{i\in\Pi}$ is defined for histories up to length $|\alpha|$ (after depth $|\alpha|$, memoryless strategies are used). □

Remark 5.2.13. This proof shows in fact a little stronger result: if there exists a goal-optimised and deviation-optimised secure equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$, then there exists a finite-memory secure equilibrium in $(\mathcal{G}, v_0)$ *with the same cost profile.*

Part (ii) of Proposition 5.2.11

Part (ii) of Proposition 5.2.11 states that if there exists a secure equilibrium in a quantitative reachability game $(\mathcal{G}, v_0)$, then there exists a goal-optimised and deviation-optimised secure equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$, for $d = d_{goal}(\mathcal{G}) + 3 \cdot |V|$. The proof needs several steps. Suppose that there exists a secure equilibrium $(\sigma_i)_{i\in\Pi}$ in $(\mathcal{G}, v_0)$. The first step consists in transforming $(\sigma_i)_{i\in\Pi}$ into a goal-optimised and deviation-optimised secure equilibrium in $(\mathcal{G}, v_0)$ (Proposition 5.2.14); the second step in showing that its restriction to $\mathsf{Trunc}_d(\mathcal{T})$ with $d = d_{goal}(\mathcal{G}) + 3 \cdot |V|$ is still a goal-optimised and deviation-optimised secure equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$.

Proposition 5.2.14. *If there exists a secure equilibrium in an initialised multiplayer quantitative reachability game, then there exists one which is goal-optimised and deviation-optimised.*

To get a goal-optimised equilibrium, the idea is to eliminate some unnecessary cycles (see Definition 4.1.13). Such an idea has already been developed in Lemma 4.1.14 for Nash equilibria. Unfortunately, this lemma cannot be applied for secure equilibria (as shown in Example 5.2.15). Adapting it to the context of secure equilibria is not trivial, the underlying constructions are more involved: we need to modify the strategies of the coalition against a deviating player. By using specific punishing strategies for the coalitions, we are then able to get a goal-optimised equilibrium that is also deviation-optimised, due to the particular form of these strategies.

Example 5.2.15. Let us consider the three-player quantitative reachability game whose arena is depicted in Figure 5.3, where the initial vertex is A, $V_1 = \{A, C, D\}$, $V_2 = \{B\}$, $V_3 = \emptyset$, $\mathsf{R}_1 = \mathsf{R}_2 = \{A\}$ and $\mathsf{R}_3 = \{D\}$. The number 4 labelling the edge (A, D) is a shortcut to indicate that there are in fact four consecutive edges from A to D (through three intermediate vertices), and similarly for the number 2 on the edge (B, D). The strategy profile (σ_1, σ_2) defined [6] below is a secure equilibrium from A whose outcome is $ABCBD^\omega$ and cost profile is $(0, 0, 5)$:

$$\sigma_1(h) = \begin{cases} B & \text{if } h = A \text{ or } ABC \\ D & \text{otherwise} \end{cases} \quad ; \quad \sigma_2(h') = \begin{cases} C & \text{if } h' = AB \\ D & \text{otherwise} \end{cases}$$

for every history h (resp. h') ending in V_1 (resp. V_2).

Let us notice that the cycle BCB of the outcome $ABCBD^\omega$ is an unnecessary cycle (see Definition 4.1.13). If we modify (σ_1, σ_2) in order to remove this cycle, as done in Lemma 4.1.14 for Nash equilibria, the resulting strategy profile is a Nash equilibrium with outcome ABD^ω and cost profile $(0, 0, 3)$, however it is no longer a secure equilibrium. Indeed, player 1 has a $\prec_1$-profitable deviation by choosing the edge (A, D) instead of (A, B), which leads to a cost of 4 for player 3 (instead of 3). In the

6. We do not have to choose a strategy for player 3 since $V_3 = \emptyset$.

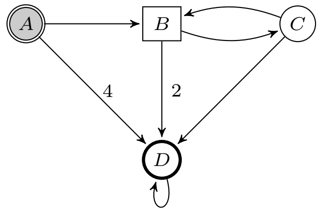


Figure 5.3: A three-player game with $\mathsf{R}_1 = \mathsf{R}_2 = \{A\}$ and $\mathsf{R}_3 = \{D\}$.

sequel, we show how to modify the approach of Lemma 4.1.14 in a way to keep the property of secure equilibrium.

In order to prove Proposition 5.2.14, we need three lemmas: Lemmas 5.2.17, 5.2.18 and 5.2.19. Given a secure equilibrium, Lemma 5.2.17 describes some particular memoryless strategies for the coalition when a player deviates. Lemma 5.2.18 (counterpart of Lemma 4.1.14 for secure equilibria) states that we can remove a cycle from the outcome of a secure equilibrium, but the strategies have to be somewhat modified with these specific coalition strategies. This lemma is used in the proof of Proposition 5.2.14 to get a goal-optimised secure equilibrium. Lemma 5.2.19 states that we can also get a deviation-optimised secure equilibrium.

Memoryless coalition strategies. Given a secure equilibrium in a quantitative reachability game $(\mathcal{G}, v_0)$, we here prove the existence of interesting memoryless strategies for the coalition against a deviating player.

Let us first introduce the definition of a *j-promising history* for some deviating player j. Intuitively player j deviates from a strategy profile $(\sigma_i)_{i\in\Pi}$ and constructs a history h consistent with σ_{-j}. This history h is called j-promising w.r.t. $(\sigma_i)_{i\in\Pi}$ if player j does not know yet if this deviation will be *$\prec_j$-profitable* for him w.r.t. $(\sigma_i)_{i\in\Pi}$, but he can still *hope* that it will be, without knowing what he will play after h.

Definition 5.2.16. Let $(\sigma_i)_{i\in\Pi}$ be a strategy profile in a quantitative reachability game $(\mathcal{G}, v_0)$, with cost profile $(x_i)_{i\in\Pi}$. Let us assume that

$\Pi = \{1, \dots, n\}$ and

$$x_1 \leq \dots \leq x_k < x_{k+1} \leq \dots \leq x_n$$

where $0 \leq k < n$. Let h be a history of $(\mathcal{G}, v_0)$ such that $x_k \leq |h| < x_{k+1}$.

For any player $j \in \Pi$, we say that h is *j-promising* w.r.t. $(\sigma_i)_{i\in\Pi}$ if h is consistent with σ_{-j} and if

- in the case where $x_{k+1} < +\infty$:
  - if $j \leq k$, we have that $\mathsf{Cost}_j(h) = x_j$ and $\forall\, i \in \Pi\ \mathsf{Cost}_i(h) \geq x_i$,
  - if $j > k$, we have that $\mathsf{Cost}_j(h) = +\infty$;
- in the case where $x_{k+1} = +\infty$:
  $\mathsf{Cost}_j(h) = x_j$, $\forall\, i \in \Pi\ \mathsf{Cost}_i(h) \geq x_i$ and $\exists\, i \in \Pi\ \mathsf{Cost}_i(h) > x_i$.

In the case where $x_{k+1} < +\infty$ and $j \leq k$, *along* h, player j has been able to get the same cost as along ρ ($\mathsf{Cost}_j(h) = x_j$) and to not decrease the cost of the other players ($\mathsf{Cost}_i(h) \geq x_i$). *After* h, he hopes to be able to play such that the resulting deviation $h\rho'$ will satisfy $(x_i)_{i\in\Pi} \prec_j \mathsf{Cost}(h\rho')$. In the case where $j > k$, player j has not visited his goal set along h, so he does not know yet if his deviation will be $\prec_j$-profitable for him. However he hopes to visit it early enough after h along $h\rho'$, such that $\mathsf{Cost}_j(h\rho') < x_j$, or to get the same cost while increasing the cost of the other players in a way that $(x_i)_{i\in\Pi} \prec_j \mathsf{Cost}(h\rho')$.

In the case where $x_{k+1} = +\infty$, the history $\rho_{\leq |h|}$ has visited all the goal sets R_i such that $\mathsf{Cost}_i(\rho) < +\infty$. Thus player j could have a $\prec_j$-profitable deviation $h\rho'$ if he can avoid visiting the goal sets R_i, where $i \geq k+1$ ($i \neq j$).

Given a j-promising history h of player j, the next lemma describes the existence of interesting memoryless strategies of the coalition $\Pi \setminus \{j\}$ from $\mathsf{Last}(h)$. For that purpose, the lemma considers particular zero-sum qualitative reachability under safety games $\mathcal{G}^{-j} = (\mathcal{A}^{-j}, \mathsf{R}^{-j}, \mathsf{S}^{-j})$, where $\mathcal{A}^{-j} = (V, (V \setminus V_j, V_j), E)$, and the coalition $\Pi \setminus \{j\}$ aims at reaching R^{-j} while staying in S^{-j}, whereas player j wants to prevent this.

Lemma 5.2.17. *Let $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{R}_i)_{i\in\Pi})$ be a quantitative reachability game, and let $(\sigma_i)_{i\in\Pi}$ be a secure equilibrium in $(\mathcal{G}, v_0)$ with cost profile $(x_i)_{i\in\Pi}$. Let h be a j-promising history w.r.t. $(\sigma_i)_{i\in\Pi}$ for some* *player $j \in \Pi$. We assume that $\Pi = \{1, \ldots, n\}$. If*

$$x_1 \leq \ldots \leq x_k \leq |h| < |h| + |V| \leq x_{k+1} \leq \ldots \leq x_l < +\infty$$
$$\text{and} \quad x_{l+1} = \ldots = x_n = +\infty,$$

where $0 \leq k \leq l \leq n$, then the coalition $\Pi \setminus \{j\}$ has a memoryless winning strategy $\mu^v_{\mathsf{C}j}$ from $v = \mathsf{Last}(h)$ in the zero-sum qualitative reachability under safety game $\mathcal{G}^{-j} = (\mathcal{A}^{-j}, \mathsf{R}^{-j}, \mathsf{S}^{-j})$, where R^{-j} and S^{-j} are defined as follows:
- if $j \leq k$, then $\mathsf{R}^{-j} = \cup_{i>k} \mathsf{R}_i$, $\mathsf{S}^{-j} = V$,
- if $k < j \leq l$, then $\mathsf{R}^{-j} = V$, $\mathsf{S}^{-j} = V \setminus \mathsf{R}_j$,
- if $l < j \wedge \mathsf{Cost}(\rho_{\leq |h|}) \preceq_j \mathsf{Cost}(h)$, then $\mathsf{R}^{-j} = \cup_{\substack{i>k \\ i \neq j}} \mathsf{R}_i$, $\mathsf{S}^{-j} = V \setminus \mathsf{R}_j$,
- if $l < j \wedge \mathsf{Cost}(\rho_{\leq |h|}) \not\preceq_j \mathsf{Cost}(h)$, then $\mathsf{R}^{-j} = V$, $\mathsf{S}^{-j} = V \setminus \mathsf{R}_j$.

In this lemma, either all goal sets are visited by ρ and $l = n$, or $l < n$ and the last visited goal set is R_l. Also notice that $\mathsf{R}^{-j} \neq \emptyset$ in all cases. Indeed, $k \neq n$ as h is j-promising, and then the set R^{-j} in the case $j \leq k$ of this lemma is not empty. In the third case, it is not empty either, otherwise we would have $k + 1 = l + 1 = n = j$ but such a situation is impossible because h is j-promising w.r.t. $(\sigma_i)_{i\in\Pi}$ (see the last case of Definition 5.2.16) and $(\sigma_i)_{i\in\Pi}$ is a secure equilibrium.

Proof of Lemma 5.2.17. Suppose that the hypotheses of the lemma are satisfied. By contradiction, let us assume that the coalition $\Pi \setminus \{j\}$ has no winning strategy from vertex v in the game $\mathcal{G}^{-j} = (\mathcal{A}^{-j}, \mathsf{R}^{-j}, \mathsf{S}^{-j})$, i.e. no winning strategy from v to reach R^{-j} while staying in S^{-j}. By Theorem 2.2.15, it implies that player j has a memoryless winning strategy μ^v_j from v to stay outside R^{-j} or to reach $V \setminus \mathsf{S}^{-j}$. Recall that h is consistent with σ_{-j} as it is j-promising w.r.t. $(\sigma_i)_{i\in\Pi}$. Let ρ' be the play with prefix h that is consistent with σ_{-j}, and with μ^v_j from v (see Figure 5.4). In the four cases of the lemma, we then prove that $(x_i)_{i\in\Pi} \prec_j (\mathsf{Cost}_i(\rho'))_{i\in\Pi}$, meaning that player j has a $\prec_j$-profitable deviation w.r.t. $(\sigma_i)_{i\in\Pi}$, which is impossible.

– Case $j \leq k$.

The strategy μ^v_j enables to avoid all goal sets R_i where $i > k$. As h is j-promising, we have that $\mathsf{Cost}_j(h) = x_j$ and $\forall i \in \Pi$, $\mathsf{Cost}_i(h) \geq x_i$.

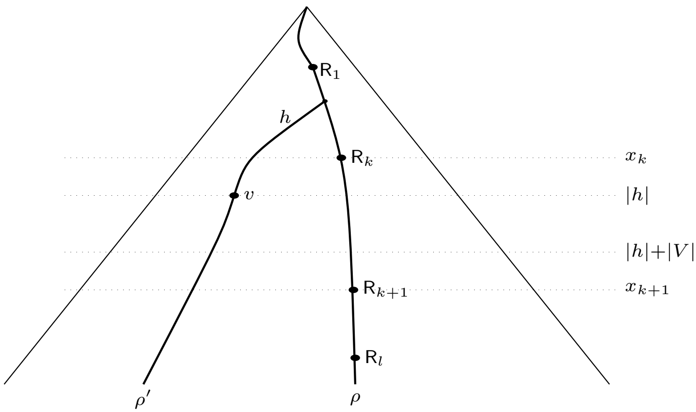


Figure 5.4: Play ρ and its deviation ρ' with prefix h.

By construction of ρ' and as $x_k \leq |h| < x_{k+1}$, we have that

$$\begin{aligned}&\mathsf{Cost}_j(\rho') = \mathsf{Cost}_j(h) = x_j,\\&\forall i \leq k,\ \mathsf{Cost}_i(\rho') \geq x_i,\\&\forall i > k,\ \mathsf{Cost}_i(\rho') = +\infty.\end{aligned}$$

Then for all $i \in \Pi$, we have that $\mathsf{Cost}_i(\rho') \geq x_i$. It remains to show that the cost of one player is strictly increased in ρ' compared with ρ. In the case where $x_{k+1} < +\infty$, i.e. $k < l$, we have in particular that $x_l < +\infty$ and $\mathsf{Cost}_l(\rho') = +\infty$. And in the case where $x_{k+1} = +\infty$ ($k = l$), we have that $(x_i)_{i\in\Pi} \prec_j \mathsf{Cost}(h)$ (by definition of j-promising), i.e. there exists $i \in \Pi$ such that $x_i < \mathsf{Cost}_i(h)$. Either $\mathsf{Cost}_i(h) = \mathsf{Cost}_i(\rho')$ and then $x_i < \mathsf{Cost}_i(\rho')$, or $\mathsf{Cost}_i(h) = +\infty > \mathsf{Cost}_i(\rho')$ and so $x_i \leq |h| < \mathsf{Cost}_i(\rho')$. In both cases, it implies that $(x_i)_{i\in\Pi} \prec_j (\mathsf{Cost}_i(\rho'))_{i\in\Pi}$.

– Case $k < j \leq l$.

As μ_j^v is memoryless, this strategy enables player j to reach his goal set R_j from v within $|V|$ steps. Thus, we have that

$$\mathsf{Cost}_j(\rho') < |h| + |V| \leq x_{k+1} \leq x_j$$

since $k < j \leq l$, and so, $(x_i)_{i\in\Pi} \prec_j (\mathsf{Cost}_i(\rho'))_{i\in\Pi}$.

– Case $l < j$ and $\mathsf{Cost}(\rho_{\leq |h|}) \preceq_j \mathsf{Cost}(h)$.

The strategy μ_j^v enables to avoid all goal sets R_i where $i > k$ and $i \neq j$, or to visit the goal set R_j. On one hand, if ρ' visits R_j, then

$$\mathsf{Cost}_j(\rho') < +\infty = x_j$$

as $j > l$, and so, $(x_i)_{i \in \Pi} \prec_j (\mathsf{Cost}_i(\rho'))_{i \in \Pi}$. On the other hand, if ρ' does not visit R_j, then ρ' does not visit either any R_i with $i > k$. Since $\mathsf{Cost}(\rho_{\leq |h|}) \preceq_j \mathsf{Cost}(h)$, the situation is quite similar to the first case, and we can deduce that

$$\begin{aligned}
&\mathsf{Cost}_j(\rho') = x_j = +\infty, \\
&\forall i \leq k,\ \mathsf{Cost}_i(\rho') \geq x_i, \\
&\forall i > k,\ \mathsf{Cost}_i(\rho') = +\infty.
\end{aligned}$$

Thus, for all $i \in \Pi$, we have that $\mathsf{Cost}_i(\rho') \geq x_i$. Moreover, exactly like in the case $j \leq k$, we can show that there exists $i \in \Pi$ such that $x_i < \mathsf{Cost}_i(\rho')$. Then it implies that $(x_i)_{i \in \Pi} \prec_j (\mathsf{Cost}_i(\rho'))_{i \in \Pi}$.

– Case $l < j$ and $\mathsf{Cost}(\rho_{\leq |h|}) \not\preceq_j \mathsf{Cost}(h)$.

Like in the second case, the strategy μ_j^v enables player j to reach his goal set R_j from v. Then we have that

$$\mathsf{Cost}_j(\rho') < +\infty = x_j$$

and so, $(x_i)_{i \in \Pi} \prec_j (\mathsf{Cost}_i(\rho'))_{i \in \Pi}$. □

Removing a cycle. The next lemma states that it is possible to modify the strategy profile of a secure equilibrium in a way to eliminate an unnecessary cycle[7] in its outcome. In the notations of this lemma, notice that β is the eliminated cycle (condition $\mathsf{Last}(\alpha) = \mathsf{Last}(\alpha\beta)$), notice also that a new goal set is visited after $\alpha\beta\gamma$ (condition $\mathsf{Visit}(\rho) \neq \mathsf{Visit}(\alpha)$). The elimination of the cycle is possible by modifying the strategies of the coalitions into strategies as described in Lemma 5.2.17.

7. See Definition 4.1.13.

Lemma 5.2.18. *Let $(\sigma_i)_{i\in\Pi}$ be a secure equilibrium in a quantitative reachability game $(\mathcal{G}, v_0)$, with outcome ρ. Suppose that $\rho = \alpha\beta\gamma\tilde{\rho}$, with $\alpha, \beta, \gamma \in V^+$, $|\gamma| \geq |V|$ and $\tilde{\rho} \in V^\omega$, such that*

$$\begin{aligned}\mathsf{Visit}(\alpha) &= \mathsf{Visit}(\alpha\beta\gamma)\\ \mathsf{Visit}(\rho) &\neq \mathsf{Visit}(\alpha)\\ \mathsf{Last}(\alpha) &= \mathsf{Last}(\alpha\beta).\end{aligned}$$

Then there exists a secure equilibrium in $(\mathcal{G}, v_0)$ with outcome $\alpha\gamma\tilde{\rho}$.

Proof. Suppose that the hypotheses of the lemma are fulfilled. We denote by $(x_i)_{i\in\Pi}$ the cost profile of ρ. Let us assume w.l.o.g. that $\Pi = \{1, \ldots, n\}$ and

$$x_1 \leq \ldots \leq x_k \leq |\alpha| < |\alpha\beta\gamma| \leq x_{k+1} \leq \ldots \leq x_l < +\infty \quad \text{and} \quad x_{l+1} = \ldots = x_n = +\infty,$$

where $0 \leq k < l \leq n$ (remark that $k < l$ as $\mathsf{Visit}(\rho) \neq \mathsf{Visit}(\alpha)$).

Let us define the required secure equilibrium $(\tau_i)_{i\in\Pi}$ with the aim to get the outcome $\alpha\gamma\tilde{\rho}$ by eliminating β in ρ. For every $i \in \Pi$ and every history $h \in \mathsf{Hist}_i$, we set

$$\tau_i(h) := \begin{cases} \sigma_i(\alpha\beta\delta) & \text{if } h = \alpha\delta \text{ for } \delta \in V^*,\\ \mu^v_{i,P(h)}(h) & \text{if } \alpha \not\leq h,\ P(h) \neq \bot, i \text{ and } \exists\, h'v \leq h \text{ s.t. } h'v \text{ is}\\ & \quad P(h)\text{-promising w.r.t. } (\sigma_i)_{i\in\Pi} \text{ and } |h'v| = |\alpha|,\\ \sigma_i(h) & \text{otherwise.}\end{cases}$$

In this definition, the punishment function P is defined as in the proof of Proposition 5.2.11, Part (i) (adapted to the play $\alpha\gamma\tilde{\rho}$). Moreover, when a player j deviates, each player $i \neq j$ plays according to σ_i, except in the case of a j-promising history h of length $|\alpha|$ from which he plays according to $\mu^v_{\mathtt{C}j}$, with $v = \mathsf{Last}(h)$ (see Lemma 5.2.17). Notation $\mu^v_{i,j}$ means the memoryless strategy of player i induced by $\mu^v_{\mathtt{C}j}$.

We observe that the outcome of $(\tau_i)_{i\in\Pi}$ is the play $\pi = \alpha\gamma\tilde{\rho}$ (see Figures 5.5 and 5.6). Let us write its cost profile as $(y_1, \ldots, y_n)$. It follows that for all $i \in \Pi$, $y_i \leq x_i$. More precisely,

$$\text{- if } i \leq k, \text{ then } y_i = x_i; \tag{5.5}$$

$$\text{- if } k < i \leq l, \text{ then } y_i = x_i - (|\beta| + 1); \tag{5.6}$$

$$\text{- if } i > l, \text{ then } y_i = x_i = +\infty. \tag{5.7}$$

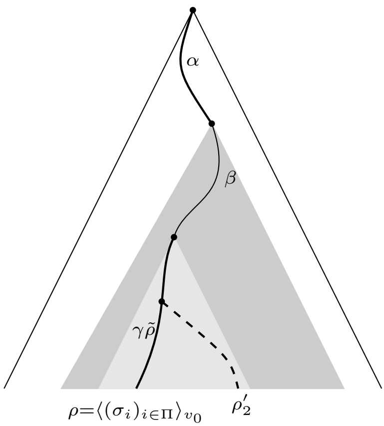


Figure 5.5: Play ρ.

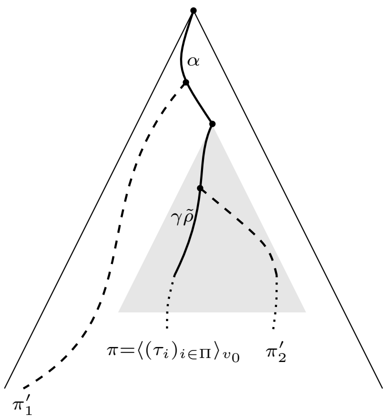


Figure 5.6: Play π and deviations.

Assume that there exists a $\prec_j$-profitable deviation τ'_j for player j w.r.t. $(\tau_i)_{i\in\Pi}$. Let π' be the outcome of the strategy profile (τ'_j, τ_{-j}) from v_0, and $(y'_1, \ldots, y'_n)$ its cost profile. Then we know that $(y_1, \ldots, y_n) \prec_j (y'_1, \ldots, y'_n)$. Two possible situations occur according to where player j deviates from π. We show that the first situation is impossible. In the second one, we construct a $\prec_j$-profitable deviation σ'_j for player j w.r.t. $(\sigma_i)_{i\in\Pi}$, and then get a contradiction with $(\sigma_i)_{i\in\Pi}$ being a secure equilibrium.

(i) The history α is not a prefix of π' (see the play π'_1 in Figure 5.6).

Let us consider the prefix h of π' of length $|\alpha|$. We first state that h cannot visit R_j in a way that $\mathsf{Cost}_j(h) < x_j$, because h is consistent with σ_{-j} (by definition of $(\tau_i)_{i\in\Pi}$), and $(\sigma_i)_{i\in\Pi}$ is a secure equilibrium. Therefore, h is a j-promising history w.r.t. $(\sigma_i)_{i\in\Pi}$, as τ'_j is a $\prec_j$-profitable deviation w.r.t. $(\tau_i)_{i\in\Pi}$. By definition of $(\tau_i)_{i\in\Pi}$, π' is consistent with $\mu^v_{\mathtt{C}j}$ from $v = \mathsf{Last}(h)$. We consider the four possible cases of Lemma 5.2.17:

– $j \leq k$.

We have that $y_j = y'_j$. The coalition $\Pi \setminus \{j\}$ forces the play π' to visit R_i, for a certain $i > k$ (let us remind that $k < n$), before depth $|\alpha| + |V|$ as $\mu^v_{\mathtt{C}j}$ is memoryless. And so, $y'_i < |\alpha| + |V| \leq |\alpha| + |\gamma| \leq y_{k+1} \leq y_i$ (as $|\alpha\beta\gamma| \leq x_{k+1}$ and by Equation (5.6)). This contradicts the fact that $(y_1, \ldots, y_n) \prec_j (y'_1, \ldots, y'_n)$.

– $k < j \leq l$.

The coalition $\Pi\setminus\{j\}$ prevents the play π' from visiting R_j, and so, $y'_j = +\infty$. As $y_j < +\infty$, it cannot be the case that $(y_1, \ldots, y_n) \prec_j (y'_1, \ldots, y'_n)$.

– $l < j$ and $\mathsf{Cost}(\rho_{\leq|h|}) \preceq_j \mathsf{Cost}(h)$.

The coalition $\Pi \setminus \{j\}$ forces the play π' to visit R_i, for a certain $i > k$, $i \neq j$, before depth $|\alpha| + |V|$, while avoiding the visit of R_j (then, $y_j = y'_j = +\infty$). As in the first case, this leads to a contradiction with the fact that $(y_1, \ldots, y_n) \prec_j (y'_1, \ldots, y'_n)$.

– $l < j$ and $\mathsf{Cost}(\rho_{\leq|h|}) \not\preceq_j \mathsf{Cost}(h)$.

Like in the second case, the coalition $\Pi \setminus \{j\}$ prevents the play π' from visiting R_j, and so, $y_j = y'_j = +\infty$. Moreover, the hypothesis $\mathsf{Cost}(\rho_{\leq|h|}) \not\preceq_j \mathsf{Cost}(h)$ implies that $(y_1, \ldots, y_n) \prec_j (y'_1, \ldots, y'_n)$ cannot be true.

(ii) The history α is a prefix of π' (see the play π'_2 in Figure 5.6).

We define for all histories $h \in \mathsf{Hist}_j$:

$$\sigma'_j(h) := \begin{cases} \sigma_j(h) & \text{if } \alpha\beta \not\leq h, \\ \tau'_j(\alpha\delta) & \text{if } h = \alpha\beta\delta. \end{cases}$$

Let us set $\rho' = \langle \sigma'_j, \sigma_{-j} \rangle_{v_0}$ of cost profile $(x'_1, \ldots, x'_n)$. As player j deviates after α with the strategy τ'_j, one can prove that

$$\pi' = \alpha\tilde{\pi}' \quad \text{and} \quad \rho' = \alpha\beta\tilde{\pi}'$$

by definition of $(\tau_i)_{i\in\Pi}$ (see the play ρ'_2 in Figure 5.5). Since $\mathsf{Visit}(\alpha) = \mathsf{Visit}(\alpha\beta)$, Equations (5.5), (5.6) and (5.7) also stand by replacing x_i with x'_i and y_i with y'_i (but the value of l might be different). Then

$$(x_1, \dots, x_n) \prec_j (x'_1, \dots, x'_n) \quad \text{iff} \quad (y_1, \dots, y_n) \prec_j (y'_1, \dots, y'_n),$$

which proves that σ'_j is a $\prec_j$-profitable deviation for player j w.r.t. $(\sigma_i)_{i\in\Pi}$, and this is a contradiction. □

Goal- and deviation-optimised secure equilibrium. The following lemma uses the ideas developed in the proof of Lemma 5.2.18 to show that any secure equilibrium can be transformed into one that is deviation-optimised. It is the last step before proving Proposition 5.2.14, and finally Part (*ii*) of Proposition 5.2.11.

Lemma 5.2.19. *If there exists a secure equilibrium in an initialised quantitative reachability game, then there exists a deviation-optimised secure equilibrium with the same outcome.*

Proof. Let $(\sigma_i)_{i\in\Pi}$ be a secure equilibrium in a quantitative reachability game $(\mathcal{G}, v_0)$ with outcome ρ, and let α be the prefix of ρ of length $\max\{\mathsf{Cost}_i(\rho) \mid \mathsf{Cost}_i(\rho) < +\infty\}$. It follows that $\mathsf{Visit}(\rho) = \mathsf{Visit}(\alpha)$. Then, we define the required secure equilibrium $(\tau_i)_{i\in\Pi}$ exactly like in the proof of Lemma 5.2.18. We only remove the first line of the definition: $\tau_i(h) = \sigma_i(\alpha\beta\delta)$ if $h = \alpha\delta$. One can be convinced that $(\tau_i)_{i\in\Pi}$ and $(\sigma_i)_{i\in\Pi}$ have the same outcome ρ. We prove in the exact same way that $(\tau_i)_{i\in\Pi}$ is a secure equilibrium in $(\mathcal{G}, v_0)$ (here, $k = l$).

Let us now show that $(\tau_i)_{i\in\Pi}$ is deviation-optimised by means of Lemma 5.2.12. Let τ'_j be a strategy of some player j such that the play $\rho' = \langle \tau'_j, \tau_{-j}\rangle_{v_0}$ verifies

(*i*) $\mathsf{Cost}_j(\rho) = \mathsf{Cost}_j(\rho')$,

(*ii*) $\forall\, i \in \Pi$ such that $\mathsf{Cost}_i(\rho) < +\infty$, it holds that $\mathsf{Cost}_i(\rho) \leq \mathsf{Cost}_i(\rho')$,

(*iii*) $\exists\, i \in \Pi$ $\mathsf{Cost}_i(\rho) < \mathsf{Cost}_i(\rho')$.

We must prove that there exists $l \in \Pi$ such that $\mathsf{Cost}_l(\rho) = +\infty$ and $\mathsf{Cost}_l(\rho') \leq d_{dev} = \max\{\mathsf{Cost}_i(\rho) \mid \mathsf{Cost}_i(\rho) < +\infty\} + |V|$. Notice that $\mathsf{Cost}(\rho) = \mathsf{Cost}(\alpha)$.

On one hand, suppose that $\mathsf{Cost}(\alpha) \not\prec_j \mathsf{Cost}(\rho'_{\leq|\alpha|})$. By (i), (ii) and (iii), the only possibility is to have some l such that $\mathsf{Cost}_l(\alpha) = +\infty$ and $\mathsf{Cost}_l(\rho'_{\leq|\alpha|}) < +\infty$, that is, $\mathsf{Cost}_l(\rho) = +\infty$ and $\mathsf{Cost}_l(\rho') \leq |\alpha| < d_{dev}$.

On the other hand, if $\mathsf{Cost}(\alpha) \prec_j \mathsf{Cost}(\rho'_{\leq|\alpha|})$, then according to the last case of Definition 5.2.16, $\rho'_{\leq|\alpha|}$ is j-promising w.r.t. $(\sigma_i)_{i\in\Pi}$. Indeed, $\rho'_{\leq|\alpha|}$ is consistent with σ_{-j}, and there exists $i \in \Pi$ such that $\mathsf{Cost}_i(\rho) = +\infty$ (otherwise it would contradict the fact that $(\sigma_i)_{i\in\Pi}$ is a secure equilibrium). By definition of $(\tau_i)_{i\in\Pi}$, ρ' is thus consistent with $\mu^v_{\mathtt{C}j}$ from vertex $v = \rho'_{|\alpha|}$. Thus, by Lemma 5.2.17 (first case or third case), there exists l such that $\mathsf{Cost}_l(\rho) = +\infty$ and $\mathsf{Cost}_l(\rho') < |\alpha| + |V| = d_{dev}$ (as $\mu^v_{\mathtt{C}j}$ is memoryless).

In both cases, by Lemma 5.2.12, we proved that $(\tau_i)_{i\in\Pi}$ is deviation-optimised. □

We are now able to prove Proposition 5.2.14, which states that if there exists a secure equilibrium in a quantitative reachability game, then there exists one which is goal-optimised and deviation-optimised.

Proof of Proposition 5.2.14. Let $(\sigma_i)_{i\in\Pi}$ be a secure equilibrium in a multiplayer quantitative reachability game $(\mathcal{G}, v_0)$ with outcome ρ and cost profile $(x_i)_{i\in\Pi}$. Let us assume w.l.o.g. that $\Pi = \{1, \ldots, n\}$ and

$$x_1 \leq \ldots \leq x_l < x_{l+1} = \ldots = x_n = +\infty$$

where $0 \leq l \leq n$. Let us set $x_0 = 0$. For all $k \in \{0, 1, \ldots, l-1\}$ such that $(x_{k+1} - x_k) \geq 2 \cdot |V|$ and while it is still the case, we apply the following procedure to get a goal-optimised secure equilibrium.

Consider such a $k \in \{0, 1, \ldots, l-1\}$. Then, we can write $\rho = \alpha\beta\gamma\tilde{\rho}$, with $\alpha, \beta, \gamma \in V^+$, $|\gamma| \geq |V|$ and $\tilde{\rho} \in V^\omega$, and such that

$$\begin{aligned} &x_k \leq |\alpha\beta\gamma| \leq x_{k+1} \\ &\mathsf{Visit}(\alpha) = \mathsf{Visit}(\alpha\beta\gamma) = \{1, \ldots, k\} \\ &\mathsf{Last}(\alpha) = \mathsf{Last}(\alpha\beta). \end{aligned}$$

Let us remark that $\mathsf{Visit}(\rho) \neq \mathsf{Visit}(\alpha)$ as $k < l$. By Lemma 5.2.18, there exists a secure equilibrium in $(\mathcal{G}, v_0)$ with outcome $\alpha\gamma\tilde{\rho}$. Its cost profile $(y_i)_{i\in\Pi}$ is such that

$$\begin{aligned} y_i &= x_i && \text{for } i \leq k\,, \\ y_i &< x_i && \text{for } k < i \leq l\,, \\ y_i &= x_i = +\infty && \text{for } i > l\,. \end{aligned}$$

By applying finitely many times this procedure, we can assume w.l.o.g. that $(\sigma_i)_{i\in\Pi}$ is a secure equilibrium with a cost profile $(x_1, \dots, x_n)$ such that

$$\begin{aligned} x_i &< i \cdot 2 \cdot |V| && \text{for } i \leq l\,, \\ x_i &= +\infty && \text{for } i > l\,, \end{aligned}$$

meaning that $(\sigma_i)_{i\in\Pi}$ is a goal-optimised secure equilibrium.

Moreover, by Lemma 5.2.19, there exists a deviation-optimised secure equilibrium with the same outcome, i.e. a goal-optimised and deviation-optimised secure equilibrium. And this concludes the proof. □

Remark 5.2.20. Regarding the costs, this proof shows that if there exists a secure equilibrium with cost profile $(a_i)_{i\in\Pi}$ in a game $(\mathcal{G}, v_0)$, then there exists a goal-optimised and deviation-optimised secure equilibrium with cost profile $(b_i)_{i\in\Pi}$ in $(\mathcal{G}, v_0)$, such that for all $i \in \Pi$, $b_i \leq a_i$. In particular, the cost profile is usually not preserved.

Finally, on the basis of Proposition 5.2.14, we are able to prove Part (*ii*) of Proposition 5.2.11: if there exists a secure equilibrium in a quantitative reachability $(\mathcal{G}, v_0)$, then there exists a goal-optimised and deviation-optimised secure equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$, for $d = d_{goal}(\mathcal{G}) + 3 \cdot |V|$.

Proof of Proposition 5.2.11, Part (ii). Let $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{R}_i)_{i\in\Pi})$ be a multiplayer quantitative game, and let $(\sigma_i)_{i\in\Pi}$ be a secure equilibrium in the game $(\mathcal{G}, v_0)$ with outcome ρ. By Proposition 5.2.14, we can suppose w.l.o.g. that $(\sigma_i)_{i\in\Pi}$ is goal-optimised and deviation-optimised. Let us define the strategy profile $(\tau_i)_{i\in\Pi}$ in $\mathsf{Trunc}_d(\mathcal{T})$ as the strategy profile $(\sigma_i)_{i\in\Pi}$ restricted to the finite tree $\mathsf{Trunc}_d(\mathcal{T})$. We prove that $(\tau_i)_{i\in\Pi}$

is a secure equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$, which is clearly goal-optimised ($d > d_{goal}(\mathcal{G})$).

For a contradiction, assume that player j has a $\prec_j$-profitable deviation τ'_j w.r.t. $(\tau_i)_{i\in\Pi}$. Let us denote $\pi = \langle(\tau_i)_{i\in\Pi}\rangle_{v_0}$ and $\pi' = \langle\tau'_j, \tau_{-j}\rangle_{v_0}$ in $\mathsf{Trunc}_d(\mathcal{T})$. We extend arbitrarily τ'_j in $\mathcal{T}$, into a strategy denoted σ'_j, and let $\rho' = \langle\sigma'_j, \sigma_{-j}\rangle_{v_0}$. Let us remark that π (resp. π') is a prefix of ρ (resp. ρ') of length $d > d_{goal}(\mathcal{G})$, and thus, in particular $\mathsf{Cost}(\rho) = \mathsf{Cost}(\pi)$. Moreover, it is impossible that $\mathsf{Cost}_j(\pi) > \mathsf{Cost}_j(\pi')$, otherwise we would have $\mathsf{Cost}_j(\rho) > \mathsf{Cost}_j(\rho')$ and so, get a contradiction with the fact that $(\sigma_i)_{i\in\Pi}$ is a secure equilibrium in $\mathcal{T}$. Then, player j gets the same cost $\mathsf{Cost}_j(\pi) = \mathsf{Cost}_j(\pi')$ and

$$\big(\forall i \in \Pi \;\; \mathsf{Cost}_i(\pi) \leq \mathsf{Cost}_i(\pi')\big) \quad \wedge \quad \big(\exists i \in \Pi \;\; \mathsf{Cost}_i(\pi) < \mathsf{Cost}_i(\pi')\big).$$

We now show that $\mathsf{Cost}_j(\rho) = \mathsf{Cost}_j(\rho')$. In the case where $\mathsf{Cost}_j(\pi) = \mathsf{Cost}_j(\pi') = +\infty$ ($= \mathsf{Cost}_j(\rho)$), we must have $\mathsf{Cost}_j(\rho') = +\infty$. Otherwise, it would contradict the fact that $(\sigma_i)_{i\in\Pi}$ is a secure equilibrium in $\mathcal{T}$. In the case where $\mathsf{Cost}_j(\pi) = \mathsf{Cost}_j(\pi') < +\infty$, then $\mathsf{Cost}_j(\rho) = \mathsf{Cost}_j(\rho')$ (as π and π' are prefixes of ρ and ρ' respectively). Moreover, since τ'_j is a $\prec_j$-profitable deviation w.r.t. $(\tau_i)_{i\in\Pi}$, it follows that for all $i \in \Pi$ such that $\mathsf{Cost}_i(\rho) < +\infty$, we have that $\mathsf{Cost}_i(\rho) \leq \mathsf{Cost}_i(\rho')$, and there exists $i \in \Pi$ such that $\mathsf{Cost}_i(\rho) < \mathsf{Cost}_i(\rho')$. As $(\sigma_i)_{i\in\Pi}$ is deviation-optimised, Lemma 5.2.12 implies that there exists some $l \in \Pi$ such that $\mathsf{Cost}_l(\rho) = +\infty$ and $\mathsf{Cost}_l(\rho') < d_{dev} = \max\{\mathsf{Cost}_i(\rho) \mid \mathsf{Cost}_i(\rho) < +\infty\} + |V|$. As $d_{dev} \leq d_{goal}(\mathcal{G}) + |V| < d$, we have that $\mathsf{Cost}_l(\pi) = \mathsf{Cost}_l(\rho) = +\infty$ and $\mathsf{Cost}_l(\pi') = \mathsf{Cost}_l(\rho') < d_{dev}$. This gives a contradiction with the fact that τ'_j is a $\prec_j$-profitable deviation w.r.t. $(\tau_i)_{i\in\Pi}$ in $\mathsf{Trunc}_d(\mathcal{T})$. Therefore, $(\tau_i)_{i\in\Pi}$ is a secure equilibrium in this game. On the other hand, the previous argument also shows that $(\tau_i)_{i\in\Pi}$ is deviation-optimised. □

Remark 5.2.21. This proof shows in particular that if there exists a goal-optimised and deviation-optimised secure equilibrium in $(\mathcal{G}, v_0)$, then there exists a goal-optimised and deviation-optimised secure equilibrium in $\mathsf{Trunc}_d(\mathcal{T})$ *with the same cost profile*. Together with Remark 5.2.20, we then proved the following result: if there exists a secure equilibrium with cost profile $(a_i)_{i\in\Pi}$ in $(\mathcal{G}, v_0)$, then there exists a goal-optimised and deviation-optimised secure equilibrium with cost profile $(b_i)_{i\in\Pi}$ in $\mathsf{Trunc}_d(\mathcal{T})$, such that for all $i \in \Pi$, $b_i \leq a_i$.

Remarks 5.2.13 and 5.2.21 imply the proposition below.

Proposition 5.2.22. *Given an initialised multiplayer quantitative reachability game and a tuple of thresholds $(t_i)_{i\in\Pi} \in (\mathbb{R} \cup \{+\infty\})^{|\Pi|}$, one can decide in* ExpSpace *whether there exists a secure equilibrium with cost profile $(c_i)_{i\in\Pi}$ such that for all $i \in \Pi$, $c_i \leq t_i$.*

The decision problem related to Proposition 5.2.22 is equivalent to decide whether there exists a goal-optimised and deviation-optimised secure equilibrium with cost profile $(a_i)_{i\in\Pi}$ in $\mathsf{Trunc}_d(\mathcal{T})$ where $d = d_{goal}(\mathcal{G}) + 3 \cdot |V|$, such that for all $i \in \Pi$, $a_i \leq t_i$. Notice that d does not depend on $(t_i)_{i\in\Pi}$.

Remark 5.2.23. Proposition 4.1.4 states that any Nash equilibrium in a multiplayer quantitative reachability game $(\mathcal{G}, v_0)$ is also a Nash equilibrium in the corresponding qualitative game $(\overline{\mathcal{G}}, v_0)$ (see Section 4.1.2 for more details). However, this proposition is false for secure equilibria. To see that, let us consider the two-player quantitative reachability game $\mathcal{G} = (\{1,2\}, \mathcal{A}, (\mathsf{R}_1, \mathsf{R}_2))$ whose arena $\mathcal{A}$ is depicted in Figure 5.7, and such that $\mathsf{R}_1 = \{B, E\}$ and $\mathsf{R}_2 = \{C\}$ (same game as in Example 4.1.3 on page 64).

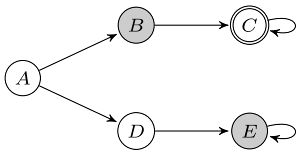


Figure 5.7: A two-player quantitative reachability game.

Let σ_1 be the positional strategy of player 1 defined by $\sigma_1(A) = B$. The strategy σ_1 with outcome ABC^ω in $(\mathcal{G}, A)$ is a secure equilibrium in the quantitative game $(\mathcal{G}, A)$ but not in the qualitative game $(\overline{\mathcal{G}}, A)$, since player 1 can reach his goal set and prevent player 2 from reaching his if he chooses the edge (A, D).

In particular, this means that existence of secure equilibria in multiplayer quantitative games would not directly imply the same result in the corresponding multiplayer qualitative games. Recall that both results are open problems to our knowledge.

Remark 5.2.24. The results we obtained about secure equilibria also hold in games with general quantitative reachability objectives where there is a unique non-zero natural price on every edge. Indeed, it suffices to replace any edge of price $c \in \mathbb{N}_0$ by a path of length c composed of c new edges (of price 1), and then apply our results on this new game.

However, the existence of a secure equilibrium in initialised (two-player) quantitative reachability games with tuples of prices on edges is an open problem.

Remark 5.2.25. It would be tempting to try to prove the existence of secure equilibria in *general cost games* with the same techniques as the ones used in Section 4.4. However, our definition of the Nash equilibrium in the proof of Proposition 4.4.6,[8] is (in general) not a secure equilibrium. To see this, let us consider the two-player quantitative reachability game $(\mathcal{G}, A)$, whose arena is depicted in Figure 5.8, and where $\mathsf{R}_1 = \{A, C\}$ and $\mathsf{R}_2 = \{C\}$. Note that player 2 does not really play in $\mathcal{G}$, only player 1 has a choice to make: he can choose the edge (B, C) or the edge (B, D).

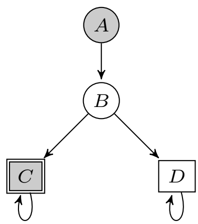


Figure 5.8: Quantitative reachability game with $\mathsf{R}_1 = \{A, C\}$, $\mathsf{R}_2 = \{C\}$.

We have seen in Section 4.4 that the cost functions of quantitative reachability games are always prefix-linear and positionally coalition-

8. As a reminder, Proposition 4.4.6 states that there exists a Nash equilibrium in every initialised multiplayer cost game where each cost function is prefix-linear and *positionally* coalition-determined.

determined, we can then follow the proof of Proposition 4.4.6. We study the two Min-Max cost games $\mathcal{G}^1$ and $\mathcal{G}^2$ (as described in Definition 4.4.4). Let $\sigma_1^\star$ be a positional strategy of player 1 in $\mathcal{G}^1$ such that $\sigma_1^\star(B) = C$, and $\sigma_{\mathtt{C}2}^\star$ be a positional strategy of player 1 in $\mathcal{G}^2$ such that $\sigma_{\mathtt{C}2}^\star(B) = D$. These strategies are optimal in the two respective games. Then, we define a Nash equilibrium in $(\mathcal{G}, A)$ in the same way as in the proof of Proposition 4.4.6. It means here that player 1 chooses the edge (B, C) (according to $\sigma_1^\star$).[9] Actually, this is not a secure equilibrium in $(\mathcal{G}, A)$ because player 1 can strictly increase player 2's cost while keeping his own cost, by choosing the edge (B, D) instead of (B, C).

9. In this example, the strategy $\sigma_{\mathtt{C}2}^\star$ is useless since player 2 has no choice to make in this game, and then can not deviate from a strategy.

Chapter 6

Subgame Perfect Equilibrium

In this chapter, based on [BBDG12, BBDG13], we show the existence of a *subgame perfect equilibrium* in quantitative reachability games (see Section 6.1), and then in quantitative reachability games with tuples of prices on edges (see Section 6.2).

The definition of subgame perfect equilibrium in multiplayer cost games, as well as some related notations, are given in Section 2.3.2 (see Definition 2.3.20, in particular). We also remind that the definition of quantitative reachability games (with tuples of prices on edges) and some related notations can be found in Section 4.1.1 (4.3).

6.1 Quantitative Reachability Objectives

In this section, we show the existence of subgame perfect equilibria in multiplayer quantitative reachability games. We then positively answer Problem 3 for subgame perfect equilibria.

Theorem 6.1.1. *In every initialised multiplayer quantitative reachability game, there exists a subgame perfect equilibrium.*

The proof uses techniques completely different from the ones given in Chapters 4 and 5 for the existence of Nash and secure equilibria. The way of defining a subgame perfect equilibrium in the game is similar to the one of the proof of Theorem 2.3.24 ([FL83, Har85]). But the way of proving that it is indeed a subgame perfect equilibrium is different, as any cost function $\mathrm{RP}_{\mathrm{Min}}$ encoding a quantitative reachability objective is not real-valued (which is one of the hypotheses of Theorem 2.3.24).

Let $(\mathcal{G}, v_0)$ be a game and $\mathcal{T}$ be the corresponding game played on the unravelling of G from v_0. Kuhn's theorem (and in particular Corollary 2.3.23) guarantees the existence of a subgame perfect equilibrium in each finite game $\mathsf{Trunc}_n(\mathcal{T})$ for every depth $n \in \mathbb{N}$. Given a sequence of such equilibria, the key point is to derive the existence of a subgame perfect equilibrium in the infinite game $\mathcal{T}$. This is possible by the following lemma.

Lemma 6.1.2. *Let $(\mathcal{G}, v_0)$ be a multiplayer cost game, and $\mathcal{T}$ be the corresponding game played on the unravelling of G from v_0. Let $(\sigma^n)_{n\in\mathbb{N}}$ be a sequence of strategy profiles such that for every $n \in \mathbb{N}$, σ^n is a strategy profile in the truncated game $\mathsf{Trunc}_n(\mathcal{T})$. Then there exists a strategy profile $\sigma^\star$ in the game $\mathcal{T}$ with the property:*

$$\forall\, d \in \mathbb{N},\ \exists\, n \geq d,\ \sigma^\star \text{ and } \sigma^n \text{ coincide on histories of length up to } d. \tag{6.1}$$

Proof. We give a tree structure, denoted by Γ, to the set of all strategy profiles in the games $\mathsf{Trunc}_n(\mathcal{T})$, $n \in \mathbb{N}$: the nodes of Γ are the strategy profiles, and we draw an edge from a strategy profile σ in $\mathsf{Trunc}_n(\mathcal{T})$ to a strategy profile σ' in $\mathsf{Trunc}_{n+1}(\mathcal{T})$ if and only if σ is the restriction of σ' to histories of length less than n. It means that the nodes at depth d correspond to strategy profiles of $\mathsf{Trunc}_d(\mathcal{T})$. We then consider the tree Γ' derived from Γ where we only keep the nodes σ^n, $n \in \mathbb{N}$, and their ancestors. Since Γ' has finite outdegree, it has an infinite path by König's lemma. This path goes through infinitely many nodes that are ancestors of nodes in the set $\{\sigma^n, n \in \mathbb{N}\}$. Therefore there exists a strategy profile $\sigma^\star$ in the infinite game $\mathcal{T}$ (given by the previous infinite path in Γ') with property (6.1). □

Before proving Theorem 6.1.1, let us introduce the following notation. Given a strategy profile $\sigma = (\sigma_i)_{i\in\Pi}$, we write $\sigma|_h$ for $(\sigma_i|_h)_{i\in\Pi}$, and $h\langle\sigma|_h\rangle_v$ for the play in $(\mathcal{G}, v_0)$ with prefix h that is consistent with $\sigma|_h$ from v.

Proof of Theorem 6.1.1. Let $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{R}_i)_{i\in\Pi})$ be a multiplayer quantitative reachability game, v_0 be an initial vertex, and $\mathcal{T}$ be the corresponding game played on the unravelling of G from v_0. For all $n \in \mathbb{N}$, we consider the finite game $\mathsf{Trunc}_n(\mathcal{T})$ and get a subgame perfect equilibrium $\sigma^n = (\sigma_i^n)_{i\in\Pi}$ in this game by Corollary 2.3.23. According to Lemma 6.1.2, there exists a strategy profile $\sigma^\star$ in the game $\mathcal{T}$ with property (6.1).

It remains to show that $\sigma^\star$ is a subgame perfect equilibrium in $\mathcal{T}$, and thus in $(\mathcal{G}, v_0)$. Let hv be a history of this game (with $v \in V$). We have to prove that $\sigma^\star|_h$ is a Nash equilibrium in the subgame $(\mathcal{T}|_h, v)$. As a contradiction, suppose that there exists a profitable deviation σ'_j for some player $j \in \Pi$ w.r.t. $\sigma^\star|_h$ in $(\mathcal{T}|_h, v)$. This means that $\mathsf{Cost}_j(\rho) > \mathsf{Cost}_j(\rho')$ for $\rho = h\langle\sigma^\star|_h\rangle_v$ and $\rho' = h\langle\sigma'_j|_h, \sigma^\star_{-j}|_h\rangle_v$, that is, ρ' visits R_j for the first time at a certain depth d, such that $|h| < d < +\infty$, and ρ visits R_j at a depth strictly greater than d (see Figure 6.1). Thus:

$$\mathsf{Cost}_j(\rho) > \mathsf{Cost}_j(\rho') = d.$$

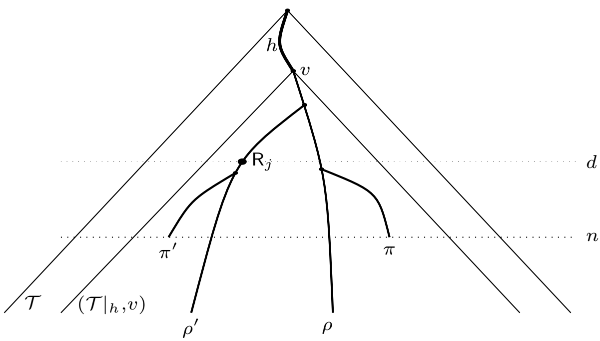


Figure 6.1: The game $\mathcal{T}$ with its subgame $(\mathcal{T}|_h, v)$.

According to property (6.1), there exists $n \geq d$ such that $\sigma^\star$ coincide with σ^n on histories of length up to d. It follows that for $\pi = h\langle \sigma^n|_h \rangle_v$ and $\pi' = h\langle \sigma'_j|_h, \sigma^n_{-j}|_h \rangle_v$, we have that (see Figure 6.1)

$$\mathsf{Cost}_j(\pi') = \mathsf{Cost}_j(\rho') = d \quad \text{and} \quad \mathsf{Cost}_j(\pi) > d,$$

as π' and ρ' coincide up to depth d. And so, σ'_j is a profitable deviation for player j w.r.t. $\sigma^n|_h$ in $(\mathsf{Trunc}_n(\mathcal{T})|_h, v)$, which leads to a contradiction with the fact that σ^n is a subgame perfect equilibrium in $\mathsf{Trunc}_n(\mathcal{T})$ by hypothesis. □

Alternative Proof of Lemma 6.1.2

The compactness of the set of strategy profiles in a finite graph gives an alternative proof to Lemma 6.1.2. This is the same kind of method used for the proof of Theorem 2.3.24 ([FL83, Har85]). Roughly, the ideas to show that this set is a compact space are the following ones. First, notice that the set of (finite) histories in a finite graph is countable, then, the set S_i of strategies of a player i is homeomorphic to a countable product of finite sets, which implies that S_i is a compact space for the product topology. By choosing an adequate metric on S_i, one can show that the set $S = \prod_{i \in \Pi} S_i$ of strategy profiles is a compact metric space.

Thanks to this, we can prove Lemma 6.1.2. Let $(\sigma^n)_{n \in \mathbb{N}}$ be a sequence of strategy profiles such that for every $n \in \mathbb{N}$, σ^n is a strategy profile in the truncated game $\mathsf{Trunc}_n(\mathcal{T})$. For all $n \in \mathbb{N}$ and $i \in \Pi$, we extend the strategy σ^n_i in an arbitrary way to get a strategy defined on every history of the game $\mathcal{G}$. We then obtain a sequence of strategy profiles of $\mathcal{G}$. By compactness, there exists a convergent subsequence. If the limit strategy profile is denoted by $\sigma^\star$, one can show that property (6.1) holds. This result is also a direct consequence of the compactness of the set of infinite trees with bounded outdegree [Kec95].

In the vein of such a proof, we give some words about the "convergence phenomenon" [1] of a sequence $(\sigma^n)_{n \in \mathbb{N}}$ such that for all $n \in \mathbb{N}$, σ^n is a

1. We studied this "convergence phenomenon" with the hope of adapting the proof of Theorem 6.1.1 to show the existence of secure equilibria in multiplayer quantitative reachability games (but finally it did not help).

subgame perfect equilibrium in the finite game $\mathsf{Trunc}_n(\mathcal{T})$, derived from a quantitative reachability game $\mathcal{G}$. For the sake of simplicity, we assume that the sequence $(\sigma^n)_{n\in\mathbb{N}}$ is converging, and we denote by $\sigma^\star$ the limit strategy profile in $\mathcal{T}$. Given a play ρ, we say that a player is *winning* if ρ visits his goal set. We denote by $\Omega(\sigma^n)$ (resp. $\Omega(\sigma^\star)$) the number of winning players of the outcome $\langle\sigma^n\rangle_{v_0}$ (resp. $\langle\sigma^\star\rangle_{v_0}$) in $\mathsf{Trunc}_n(\mathcal{T})$ (resp. in $\mathcal{T}$). Recall that v_0 is the initial vertex of $\mathcal{T}$. Then, it holds that $0 \leq \Omega(\sigma^n), \Omega(\sigma^\star) \leq |\Pi|$, where $|\Pi|$ is the number of players in the game. A question we can ask is: does the sequence $(\Omega(\sigma^n))_{n\in\mathbb{N}}$ converge to $\Omega(\sigma^\star)$? First, it is important to notice that this sequence does not always converge.

Example 6.1.3. Let us consider the two-player quantitative reachability game $\mathcal{G} = (\{1,2\}, \mathcal{A}, (\mathsf{R}_1, \mathsf{R}_2))$, whose arena $\mathcal{A}$ is depicted in Figure 6.2 (player 1 controls both vertices), and such that $\mathsf{R}_1 = \{A\}$ and $\mathsf{R}_2 = \{B\}$. Let us fix the initial vertex A.

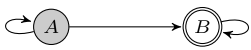


Figure 6.2: Quantitative reachability game with $\mathsf{R}_1 = \{A\}$ and $\mathsf{R}_2 = \{B\}$.

For every $n \in \mathbb{N}_0$ that is even, let σ_1^n be the strategy defined, for any history h ending in A, by

$$\sigma_1^n(h) = \begin{cases} A & \text{if } h = A^j, \text{ with } j < n, \\ B & \text{otherwise.} \end{cases} \tag{6.2}$$

For every odd natural n, σ_1^n is a positional strategy defined by $\sigma_1^n(A) = A$. Then, if n is even (resp. odd), the outcome of σ_1^n is A^nB (resp. A^{n+1}) in $\mathsf{Trunc}_n(\mathcal{T})$. Moreover, σ_1^n is a subgame perfect equilibrium in this game, for all $n \in \mathbb{N}_0$, and the sequence $(\sigma_1^n)_{n\in\mathbb{N}}$ converges to the positional strategy defined by $\sigma_1^\star(A) = A$. But the sequence $(\Omega(\sigma_1^n))_{n\in\mathbb{N}}$ does not converge, since $\Omega(\sigma_1^n) = 2$ if n is even, and $\Omega(\sigma_1^n) = 1$ otherwise.

In fact, the answer to the question "if the sequence $(\Omega(\sigma^n))_{n\in\mathbb{N}}$ is converging, does it converge to $\Omega(\sigma^\star)$?" is still no, as shown by the following example.

Example 6.1.4. Let us come back to the two-player quantitative reachability game $\mathcal{G} = (\{1,2\}, \mathcal{A}, (\mathsf{R}_1, \mathsf{R}_2))$ of Example 6.1.3. For all $n \in \mathbb{N}_0$, let $\tilde{\sigma}_1^n$ be the strategy defined as in Equation (6.2). This strategy is a subgame perfect equilibrium in $\mathsf{Trunc}_n(\mathcal{T})$ with outcome $A^n B$. One can show that the sequence $(\tilde{\sigma}_1^n)_{n\in\mathbb{N}}$ converges to $\sigma_1^\star$, where $\sigma_1^\star$ is the positional strategy given by $\sigma_1^\star(A) = A$. But we have that $\Omega(\tilde{\sigma}_1^n) = 2$ for all $n \in \mathbb{N}_0$, and $\Omega(\sigma_1^\star) = 1$, which implies that $(\Omega(\tilde{\sigma}_1^n))_{n\in\mathbb{N}}$ does not converge to $\Omega(\sigma_1^\star)$.

Nevertheless, the next property holds.

Proposition 6.1.5. *For all $n \in \mathbb{N}$, let σ^n be a subgame perfect equilibrium in $\mathsf{Trunc}_n(\mathcal{T})$. We assume that $(\sigma^n)_{n\in\mathbb{N}}$ converges to the strategy profile $\sigma^\star$ of $\mathcal{T}$. If $(\Omega(\sigma^n))_{n\in\mathbb{N}}$ converges to some value l, then $\Omega(\sigma^\star) \leq l$.*

Proof. By contradiction, assume that $\Omega(\sigma^\star) > l$. As $(\Omega(\sigma^n))_{n\in\mathbb{N}}$ converges to l and can only take a finite number of values, then $(\Omega(\sigma^n))_{n\in\mathbb{N}}$ is ultimately constant: there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $\Omega(\sigma^n) = l$. According to Property (6.1), with $d = n_0$, there exists $n_1 \geq n_0$ such that $\sigma^\star$ coincide with σ^{n_1} on histories of length up to n_0. Then $\Omega(\sigma^{n_1}) = l < \Omega(\sigma^\star)$. It implies that the outcome of $\sigma^\star$ visits the goal set of a player for the first time at some depth $n_0' > n_0$. We repeat the argument (with $d = n_0'$, and so on) until we get a contradiction with the fact that $\Omega(\sigma^\star)$ is finite. □

If we assume that the strategy profile σ^n maximises the number of winning players in $\mathsf{Trunc}_n(\mathcal{T})$ for all $n \in \mathbb{N}$, Example 6.1.4 shows that the answer to the question "if the sequence $(\Omega(\sigma^n))_{n\in\mathbb{N}}$ is converging, does it converge to $\Omega(\sigma^\star)$?" is still no. On the other hand, if we minimise the number of winning players, the answer is yes.

Proposition 6.1.6. *For all $n \in \mathbb{N}$, let σ^n be a subgame perfect equilibrium in $\mathsf{Trunc}_n(\mathcal{T})$ that minimises the number of winning players. We assume that $(\sigma^n)_{n\in\mathbb{N}}$ converges to the strategy profile $\sigma^\star$ of $\mathcal{T}$. If $(\Omega(\sigma^n))_{n\in\mathbb{N}}$ converges to some value l, then $\Omega(\sigma^\star) = l$.*

Proof. By Proposition 6.1.5, we know that $\Omega(\sigma^\star) \leq l$. It then remains to show the other inequality.

By the same arguments as before, $(\Omega(\sigma^n))_{n\in\mathbb{N}}$ is ultimately constant: there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $\Omega(\sigma^n) = l$. By contradiction, assume that $\Omega(\sigma^\star) < l$. Let us consider the strategy profile $\sigma^\star$ restricted to the game $\mathsf{Trunc}_{n_0}(\mathcal{T})$, and denote it by $\sigma^\star[n_0]$. One can show that this is a subgame perfect equilibrium in this game. Indeed, a profitable deviation in $\mathsf{Trunc}_{n_0}(\mathcal{T})$ would imply a profitable deviation in $\mathcal{T}$, and then contradicts the fact that $\sigma^\star$ is a subgame perfect equilibrium in $\mathcal{T}$ (see the proof of Theorem 6.1.1).

By hypothesis, we have that $\Omega(\sigma^\star) < l = \Omega(\sigma^{n_0})$. As $\Omega(\sigma^\star[n_0]) \leq \Omega(\sigma^\star)$, it contradicts the fact that σ^{n_0} is a subgame perfect equilibrium in $\mathsf{Trunc}_{n_0}(\mathcal{T})$ that minimises the number of winning players. □

We now state a property which somewhat binds the *sets* of winning players $\mathsf{Type}(\sigma^n)$ and $\mathsf{Type}(\sigma^\star)$. Recall that $\mathsf{Type}(\sigma) \subseteq \Pi$ and notice that $|\mathsf{Type}(\sigma)| = \Omega(\sigma)$ for any strategy profile σ, when the same initial vertex is fixed.

Proposition 6.1.7. *For all $n \in \mathbb{N}$, let σ^n be a subgame perfect equilibrium in $\mathsf{Trunc}_n(\mathcal{T})$ that minimises the number of winning players. We assume that for all $n \in \mathbb{N}$, $\mathsf{Type}(\sigma^n) = \mathsf{T}$ for a certain*[2] *set $\mathsf{T} \subseteq \Pi$ of players, and that $(\sigma^n)_{n\in\mathbb{N}}$ converges to the strategy profile $\sigma^\star$ of $\mathcal{T}$. Then, $\mathsf{Type}(\sigma^\star) = \mathsf{T}$.*

Proof. First assume by contradiction that $\mathsf{Type}(\sigma^\star) \not\subseteq \mathsf{T}$. Then, there exists $j \in \mathsf{Type}(\sigma^\star)$ such that $j \notin \mathsf{T}$. Let d be the depth at which the outcome of $\sigma^\star$ visits R_j for the first time. According to Property (6.1), there exists $n \geq d$ such that $\sigma^\star$ coincide with σ^n on histories of length up to d, which contradicts the facts that $\mathsf{Type}(\sigma^n) = \mathsf{T}$ and $j \notin \mathsf{T}$. Then, $\mathsf{Type}(\sigma^\star) \subseteq \mathsf{T}$.

By Proposition 6.1.6, it holds that $|\mathsf{Type}(\sigma^\star)| = |\mathsf{T}|$, since $(\Omega(\sigma^n))_{n\in\mathbb{N}}$ converges to $|\mathsf{T}|$ (recall that $\Omega(\sigma^n) = |\mathsf{Type}(\sigma^n)|$ and $\mathsf{Type}(\sigma^n) = \mathsf{T}$ for all $n \in \mathbb{N}$). Then, $\mathsf{Type}(\sigma^\star) = \mathsf{T}$. □

Remark 6.1.8. The problem of deciding, given an initialised multiplayer quantitative reachability game, and thresholds $(t_i)_{i\in\Pi} \in (\mathbb{R} \cup \{+\infty\})^{|\Pi|}$,

2. It is possible since $\mathsf{Type}(\sigma^n) \subseteq \Pi$ and $|\Pi| < +\infty$.

whether there exists a subgame perfect equilibrium with cost profile at most $(t_i)_{i\in\Pi}$, is NP-hard (this can be derived from the proof of [Umm05, Proposition 6.29]).

6.2 General Quantitative Reachability Objectives

As an extension to multiplayer quantitative reachability games, we consider *multiplayer quantitative reachability games with tuples of prices on edges* (as defined in Section 4.3). In this framework, we also prove the existence of a subgame perfect equilibrium. The proof is similar to the one of Theorem 6.1.1, the only difference lies in the choice of the considered depth d of the game $\mathsf{Trunc}_d(\mathcal{T})$.

Theorem 6.2.1. *In every initialised multiplayer quantitative reachability game with tuples of prices on edges, there exists a subgame perfect equilibrium.*

Let us introduce some notations that will be useful for the proof of this theorem. We define $\mathsf{c}_{\mathsf{min}} := \min_{i\in\Pi}\min_{e\in E}\mathsf{Cost}_i(e)$, $\mathsf{c}_{\mathsf{max}} := \max_{i\in\Pi}\max_{e\in E}\mathsf{Cost}_i(e)$ and $\mathsf{K} := \left\lceil \frac{\mathsf{c}_{\mathsf{max}}}{\mathsf{c}_{\mathsf{min}}} \right\rceil$. It is clear that $\mathsf{c}_{\mathsf{min}}, \mathsf{c}_{\mathsf{max}} > 0$ and $\mathsf{K} \geq 1$.

Proof of Theorem 6.2.1. Let $\mathcal{G} = (\Pi, \mathcal{A}, (\phi_i)_{i\in\Pi}, (\mathsf{R}_i)_{i\in\Pi})$ be some multiplayer quantitative reachability game with tuples of prices on edges, and $\mathcal{T}$ be the corresponding game played on the unravelling of G from an initial vertex v_0. For all $n \in \mathbb{N}$, we consider the finite game $\mathsf{Trunc}_n(\mathcal{T})$ and get a subgame perfect equilibrium $\sigma^n = (\sigma_i^n)_{i\in\Pi}$ in this game by Corollary 2.3.23. According to Lemma 6.1.2, there exists a strategy profile $\sigma^\star$ in the game $\mathcal{T}$ with property (6.1).

We then show that $\sigma^\star$ is a subgame perfect equilibrium in $\mathcal{T}$, and thus in $(\mathcal{G}, v_0)$. Let hv be a history of this game ($v \in V$). We have to prove that $\sigma^\star|_h$ is a Nash equilibrium in the subgame $(\mathcal{T}|_h, v)$. As a contradiction, suppose that there exists a profitable deviation σ'_j for some player $j \in \Pi$ w.r.t. $\sigma^\star|_h$ in $(\mathcal{T}|_h, v)$. This means that $\mathsf{Cost}_j(\rho) > \mathsf{Cost}_j(\rho')$

for $\rho = h\langle \sigma^\star|_h \rangle_v$ and $\rho' = h\langle \sigma'_j|_h, \sigma^\star_{-j}|_h \rangle_v$. Thus ρ' visits R_j for the first time at a certain depth d', such that $|h| < d' < +\infty$.

We define some depth d depending on the fact that ρ visits R_j or not.

$$d = \begin{cases} \max\{d', d''\} & \text{if } \rho \text{ visits } \mathsf{R}_j \text{ for the first time at depth } d'', \\ d' \cdot \mathsf{K} & \text{if } \rho \text{ does not visit } \mathsf{R}_j. \end{cases}$$

According to property (6.1), there exists $n \geq d$ such that $\sigma^\star$ coincide with σ^n on histories of length up to d. For $\pi = h\langle \sigma^n|_h \rangle_v$ and $\pi' = h\langle \sigma'_j|_h, \sigma^n_{-j}|_h \rangle_v$, since $d \geq d'$, it follows that:

$$\mathsf{Cost}_j(\pi') = \mathsf{Cost}_j(\rho').$$

If ρ visits R_j, then it holds that $\mathsf{Cost}_j(\pi) = \mathsf{Cost}_j(\rho)$ by definition of d, and so $\mathsf{Cost}_j(\pi) > \mathsf{Cost}_j(\pi')$. If ρ does not visit R_j, then the following inequalities hold:

$$\mathsf{Cost}_j(\pi') \leq d' \cdot \mathsf{c}_{\mathsf{max}} \leq d \cdot \mathsf{c}_{\mathsf{min}} < \mathsf{Cost}_j(\pi).$$

The first inequality comes from the fact that π' visits R_j at depth d', the second one from the definition of d, and the last one from the fact that if π visits R_j, it must happen after depth d (as ρ does not visit R_j).

In both cases $\mathsf{Cost}_j(\pi) > \mathsf{Cost}_j(\pi')$, and we conclude that σ'_j is a profitable deviation for player j w.r.t. $\sigma^n|_h$ in $(\mathsf{Trunc}_n(\mathcal{T})|_h, v)$, which leads to a contradiction with the fact that σ^n is a subgame perfect equilibrium in $\mathsf{Trunc}_n(\mathcal{T})$ by hypothesis. □

Remark 6.2.2. One can transform the cost functions $(\mathsf{Cost}_i)_{i\in\Pi}$ of quantitative reachability games in the following way: for any player i and any play ρ,

$$\mathsf{Cost}'_i(\rho) = \begin{cases} 1 - \frac{1}{c+1} & \text{if } \mathsf{Cost}_i(\rho) = c \in \mathbb{R}^+, \\ 1 & \text{if } \mathsf{Cost}_i(\rho) = +\infty. \end{cases}$$

These new cost functions $(\mathsf{Cost}'_i)_{i\in\Pi}$ are real-valued and continuous. Furthermore, a subgame perfect equilibrium in a game with the cost functions $(\mathsf{Cost}_i)_{i\in\Pi}$ is a subgame perfect equilibrium in this game with the new cost functions $(\mathsf{Cost}'_i)_{i\in\Pi}$, and conversely. Then, Theorems 6.1.1 and 6.2.1 are consequences of Theorem 2.3.24 ([FL83, Har85]).

Remark 6.2.3. One could be tempted to prove the existence of subgame perfect equilibria in *general cost games* with the same techniques as the ones used in Section 4.4. But like for secure equilibria (see Remark 5.2.25), the proof of Proposition 4.4.6 does not directly induce the existence of subgame perfect equilibria in multiplayer cost games: our definition of the Nash equilibrium in this proof is (in general) not a subgame perfect equilibrium. Let us show this with the following example.

We consider the two-player quantitative reachability game $(\mathcal{G}, A)$, whose arena is depicted in Figure 6.3, and where $\mathsf{R}_1 = \mathsf{R}_2 = \{D, E\}$.

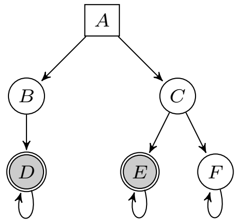


Figure 6.3: Quantitative reachability game with $\mathsf{R}_1 = \mathsf{R}_2 = \{D, E\}$.

Since the cost functions of quantitative reachability games are always prefix-linear and positionally coalition-determined (see Section 4.4), we can follow the proof of Proposition 4.4.6. So, we study the two Min-Max cost games $\mathcal{G}^1$ and $\mathcal{G}^2$ (as described in Definition 4.4.4). In the game $\mathcal{G}^1$, we define two positional strategies $\sigma_1^\star$ and $\sigma_{\mathtt{C}1}^\star$ for player Min (player 1 in $\mathcal{G}$) and player Max (player 2 in $\mathcal{G}$) respectively, as: $\sigma_1^\star(C) = E$ and $\sigma_{\mathtt{C}1}^\star(A) = C$. One can be convinced that $\sigma_1^\star$ and $\sigma_{\mathtt{C}1}^\star$ are optimal strategies in $\mathcal{G}^1$. In the game $\mathcal{G}^2$, we define two positional strategies $\sigma_2^\star$ and $\sigma_{\mathtt{C}2}^\star$ for player Min (player 2 in $\mathcal{G}$) and player Max (player 1 in $\mathcal{G}$) respectively, as: $\sigma_2^\star(A) = B$ and $\sigma_{\mathtt{C}2}^\star(C) = F$. We claim that these two strategies are optimal in $\mathcal{G}^2$.

If we define a Nash equilibrium (τ_1, τ_2) in $(\mathcal{G}, A)$ exactly as in the proof of Proposition 4.4.6, depending[3] on these strategies $\sigma_1^\star$, $\sigma_{\mathtt{C}1}^\star$, $\sigma_2^\star$ and $\sigma_{\mathtt{C}2}^\star$, then (τ_1, τ_2) is not a subgame perfect equilibrium in $(\mathcal{G}, A)$.

3. As a reminder, τ_1 prescribes to play according to $\sigma_1^\star$, and switch to $\sigma_{\mathtt{C}2}^\star$ if player 2 deviates. And symmetrically for the strategy τ_2 of player 2.

Indeed, $(\tau_1|_A, \tau_2|_A)$ is not a Nash equilibrium in the subgame $(\mathcal{G}|_A, C)$: player 1 punishes player 2 by choosing the edge (C, F) (according to $\sigma^{\star}_{\mathtt{C}2}$) whereas player 1 could pay a smaller cost by choosing the edge (C, E).

Furthermore, this Nash equilibrium also gives a counter-example of subgame perfect equilibrium for other classical punishments (see [OR94], e.g., punish the last player who has deviated and only for a finite number of steps).

Chapter 7

Subgame Perfect Secure Equilibrium

In this chapter, based on [BBDG13], we first introduce the new concept of *subgame perfect secure equilibrium* (in Section 7.1). Then we study this notion in cost games played on finite trees (Section 7.2), and in quantitative reachability games (in Section 7.3).

7.1 Definition

The concept of subgame perfect secure equilibrium is a new notion that combines both concepts of subgame perfect equilibrium and secure equilibrium in the following way. A strategy profile is a subgame perfect secure equilibrium in a game if it is a *secure* equilibrium in every subgame.

Let us remind that the definition of subgame in multiplayer cost games, as well as some related notations, are given in Section 2.3.2 (see just before Definition 2.3.20). We here give the definition of secure equilibrium in a subgame.

Given a history hv of a cost game $(\mathcal{G}, v_0)$ (with $v \in V$), we say that $(\sigma_i|_h)_{i\in\Pi}$ is a secure equilibrium in the subgame $(\mathcal{G}|_h, v)$ if, for every player $j \in \Pi$, there does not exist any strategy σ'_j of player j such that:

$\mathsf{Cost}|_h(\langle(\sigma_i|_h)_{i\in\Pi}\rangle_v) \prec_j \mathsf{Cost}|_h(\langle\sigma'_j|_h, \sigma_{-j}|_h\rangle_v)$,[1] or in an equivalent way, $\mathsf{Cost}(h\langle(\sigma_i|_h)_{i\in\Pi}\rangle_v) \prec_j \mathsf{Cost}(h\langle\sigma'_j|_h, \sigma_{-j}|_h\rangle_v)$.

Definition 7.1.1. Given a multiplayer cost game $(\mathcal{G}, v_0)$, a strategy profile $(\sigma_i)_{i\in\Pi}$ of $\mathcal{G}$ is a *subgame perfect secure equilibrium* of $(\mathcal{G}, v_0)$ if $(\sigma_i|_h)_{i\in\Pi}$ is a secure equilibrium in $(\mathcal{G}|_h, v)$, for every history hv of $(\mathcal{G}, v_0)$, with $v \in V$.

Notice that a subgame perfect secure equilibrium is a secure equilibrium, as well as a subgame perfect equilibrium.

In order to understand the differences between the various notions of equilibria, we provide three simple examples of games limited to two players and to finite trees.

Example 7.1.2. Let $\mathcal{G} = (\{1,2\}, \mathcal{A}, (\mathsf{R}_1, \mathsf{R}_2))$ be the two-player quantitative reachability game whose arena is depicted in Figure 7.1 ($V_1 = \{A, D, E, F\}$ and $V_2 = \{B, C\}$), and where $\mathsf{R}_1 = \{D, F\}$ and $\mathsf{R}_2 = \{F\}$. The number 2 labelling the edge (B, D) is a shortcut to indicate that there are in fact two consecutive edges from B to D (through one intermediate vertex). We also consider two other two-player quantitative reachability games $\mathcal{G}'$ and $\mathcal{G}''$, whose arenas are depicted in Figure 7.2 (notice that the number 2 has disappeared from the edge (B, D)) and in Figure 7.3 respectively. For the game $\mathcal{G}'$, the players' goal sets are $\mathsf{R}'_1 = \{D, F\}$ and $\mathsf{R}'_2 = \{F\}$, and for the game $\mathcal{G}''$, the players' goal sets are $\mathsf{R}''_1 = \{D, F\}$ and $\mathsf{R}''_2 = \{E, F\}$.

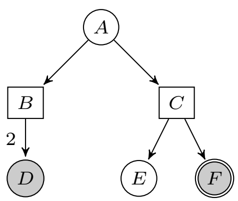


Figure 7.1: Game $\mathcal{G}$.

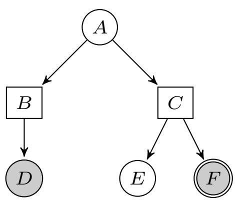


Figure 7.2: Game $\mathcal{G}'$.

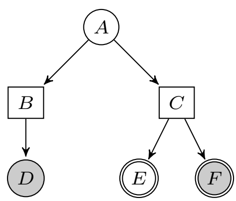


Figure 7.3: Game $\mathcal{G}''$.

In the games $\mathcal{G}$, $\mathcal{G}'$ and $\mathcal{G}''$, we define[2] two positional strategies σ_1, σ'_1

1. For a play ρ, we denote by $\mathsf{Cost}|_h(\rho)$ the cost profile $(\mathsf{Cost}_i|_h(\rho))_{i\in\Pi}$.
2. Notice that the players' possible choices to make in the arenas of the games $\mathcal{G}$,

of player 1 and two positional strategies σ_2, σ_2' of player 2 in the following way: $\sigma_1(A) = B$, $\sigma_1'(A) = C$, $\sigma_2(C) = E$ and $\sigma_2'(C) = F$.

In $(\mathcal{G}, A)$, one can easily check that the strategy profile (σ_1, σ_2) is a secure equilibrium (and thus a Nash equilibrium) with cost profile $(3, +\infty)$. Such a secure equilibrium exists because player 2 threatens player 1 to choose the edge (C, E) (thus preventing player 1 from having a cost of 2). But this threat is not credible for player 2 since by acting this way, player 2 gets an infinite cost instead of a cost of 2 (that he could obtain by reaching F). For this reason, (σ_1, σ_2) is not a subgame perfect equilibrium (and thus not a subgame perfect secure equilibrium). However, one can check that the strategy profile (σ_1', σ_2') is a subgame perfect secure equilibrium.

In $(\mathcal{G}', A)$, one can verify that the strategy profile (σ_1', σ_2') is a subgame perfect equilibrium which is not a secure equilibrium (and thus not a subgame perfect secure equilibrium). A subgame perfect secure equilibrium for $(\mathcal{G}', A)$ is given by the strategy profile (σ_1, σ_2').

In $(\mathcal{G}'', A)$, one can check that the strategy profile (σ_1, σ_2') is both a subgame perfect equilibrium and a secure equilibrium. However it is not a subgame perfect secure equilibrium. In particular, this shows that being a subgame perfect secure equilibrium is *not* equivalent to be a subgame perfect equilibrium and a secure equilibrium. On the other hand, (σ_1, σ_2) is a subgame perfect secure equilibrium in $(\mathcal{G}'', A)$.

A part of our work is to investigate *interesting* concepts of equilibria in multiplayer cost games. In particular, in quantitative reachability games, each player aims at reaching his goal set as soon as possible. Having that in mind, a play where a goal set is visited for the first time after *cycles* were no new goal set is visited does not seem to be a desirable behaviour (recall Definition 4.1.13 of *unnecessary cycle*). It appears thus reasonable to seek equilibrium concepts with outcomes that do not present this undesirable feature.

Example 7.1.3. Let us exhibit an example of this phenomenon on the two-player quantitative reachability game $(\mathcal{G}, A)$ whose arena is depicted in Figure 7.4, and where $\mathsf{R}_1 = \{A\}$ and $\mathsf{R}_2 = \{B\}$. For $n > 1$, let us

---

$\mathcal{G}'$ and $\mathcal{G}''$ are the same.

consider the play $A^n B^\omega$. Along this play, the cycles A^{n-1}, for $n > 1$, are unnecessary cycles. Indeed, once R_1 is visited (in A), looping n times in A just delays the apparition of R_2 (in B). However, for each $n > 1$, one can build a subgame perfect equilibrium (σ_1^n, σ_2) whose outcome is $A^n B^\omega$ and cost profile is $(0, n)$, as follows:

$$\sigma_1^n(h) = \begin{cases} A & \text{if } h = A^j, \text{ with } j < n, \\ B & \text{otherwise.} \end{cases}$$

This allows us to conclude that the notion of subgame perfect equilibrium does not prevent the existence of outcomes with unnecessary cycles. We can notice that, for any $n > 1$, (σ_1^n, σ_2) is not a secure equilibrium (and thus not a subgame perfect secure equilibrium) in $(\mathcal{G}, A)$. However, we will see in the next example that secure equilibria can also allow this kind of undesirable behaviours.

Figure 7.4: Subgame perfect equilibrium with outcome $A^n B^\omega$.

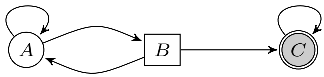


Figure 7.5: Secure equilibrium with outcome $A^n B C^\omega$.

Let us consider the two-player quantitative reachability game $(\mathcal{G}', A)$ whose arena is depicted in Figure 7.5, and where $\mathsf{R}'_1 = \mathsf{R}'_2 = \{C\}$. For $n > 1$, the cycles A^{n-1} are unnecessary along the play $A^n B C^\omega$. However, for each $n > 1$, we can build a secure equilibrium (σ_1^n, σ_2^n) whose outcome is $A^n B C^\omega$ and cost profile is $(n+1, n+1)$, as follows:

$$\sigma_1^n(h) = \begin{cases} A & \text{if } h = A^j, \text{ with } j < n, \\ B & \text{otherwise,} \end{cases} \quad ; \quad \sigma_2^n(h) = \begin{cases} C & \text{if } h = A^n B, \\ A & \text{otherwise.} \end{cases}$$

For each $n > 1$, the fact that (σ_1^n, σ_2^n) is a secure equilibrium in $(\mathcal{G}', A)$ is based on the following threat of player 2 against player 1: player 2 pretends that he will only decide to visit vertex C if player 1 has visited vertex A exactly n times. This behaviour is not credible since player 2's interest is to reach vertex C as soon as possible. In other words, we have that (σ_1^n, σ_2^n) is not a subgame perfect equilibrium (and thus not a subgame perfect secure equilibrium) in $(\mathcal{G}', A)$.

Note that in $(\mathcal{G}, A)$, the only subgame perfect secure equilibrium is the one where player 1 always chooses the edge (A, A), leading to the outcome A^ω and cost profile $(0, +\infty)$. In $(\mathcal{G}', A)$, the strategy profile where player 1 always chooses the edge (A, B) and player 2 always chooses the edge (B, C) is the only subgame perfect secure equilibrium. Its outcome is the play ABC^ω and its cost profile is $(2, 2)$.

These examples motivate the introduction of the notion of subgame perfect secure equilibrium. We believe that this notion can help in avoiding the undesirable behaviours of unnecessary cycles in quantitative reachability games (but it is an open problem). More generally, a deeper understanding of the studied equilibria whose outcomes have unnecessary cycles could be very useful.

7.2 Games with Various Objectives Played on Finite Trees

In this section, we state that there always exists a subgame perfect secure equilibrium in a multiplayer cost game whose graph is a *finite tree*. This in fact follows from a variant of Kuhn's theorem (Theorem 2.3.22).

Given two cost profiles x and y, let $\precsim_j$ be the relation defined by $x \precsim_j y$ iff $x \prec_j y$ or $x = y$, where $\prec_j$ is the relation of Equation (5.1) (used in the definition of secure equilibrium). One can show that in the two-player case (see Equation (5.2)), $\precsim_j$ is a preference relation (it is total, reflexive and transitive). However, when there are more than two players, $\precsim_j$ is no longer total[3], so Kuhn's theorem can not be applied. Nevertheless, it is proved in [LR09] that, when the binary relations on cost profiles are only transitive, Kuhn's theorem can be rewritten as follows.

Theorem 7.2.1 ([LR09]). *Given a multiplayer cost game $\mathcal{G} = (\Pi, \mathcal{A}, (\mathsf{Cost}_i)_{i\in\Pi})$ whose graph is a* finite tree*, and a transitive binary relation $\prec_i$ on cost profiles for each player $i \in \Pi$, there exists a strategy*

3. For example, $(0, 1, 0) \not\precsim_1 (0, 0, 1)$ and $(0, 0, 1) \not\precsim_1 (0, 1, 0)$.

profile $(\sigma_i)_{i\in\Pi}$ *such that for every history* hv *in* $\mathcal{G}$, *with* $v \in V$, *and for every player* $j \in \Pi$, *there does not exist a strategy* σ'_j *of player* j, *such that*

$$\mathsf{Cost}(\rho) \prec_j \mathsf{Cost}(\rho')$$

where $\rho = h\langle(\sigma_i|_h)_{i\in\Pi}\rangle_v$ *and* $\rho' = h\langle\sigma'_j|_h, \sigma_{-j}|_h\rangle_v$.

As one can show the transitivity of each relation $\prec_j$ of Equation (5.1) (used in the definition of secure equilibrium), the next corollary holds.

Corollary 7.2.2. *In every multiplayer cost game whose graph is a* finite tree, *there exists a subgame perfect secure equilibrium.*

7.3 Quantitative Reachability Objectives

In this section, we show the existence of a subgame perfect secure equilibrium in every initialised two-player quantitative reachability game. We then positively answer Problem 3 for subgame perfect secure equilibria, but in the two-player case only. The multiplayer case is still an open problem.

Theorem 7.3.1. *In every initialised two-player quantitative reachability game, there exists a subgame perfect secure equilibrium.*

The main ideas of the proof are similar to the ones for Theorem 6.1.1 (stating the existence of a subgame perfect equilibrium in quantitative reachability games).

Proof of Theorem 7.3.1. Let $\mathcal{G} = (\{1,2\}, \mathcal{A}, (\mathsf{R}_1, \mathsf{R}_2))$ be a two-player quantitative reachability game, v_0 be an initial vertex, and $\mathcal{T}$ be the corresponding game played on the unravelling of G from v_0. For every $n \in \mathbb{N}$, we consider the finite game $\mathsf{Trunc}_n(\mathcal{T})$ and get a subgame perfect secure equilibrium $\sigma^n = (\sigma_1^n, \sigma_2^n)$ in this game by Corollary 7.2.2. According to Lemma 6.1.2 (on page 158), there exists a strategy profile $\sigma^\star$ in the game $\mathcal{T}$ such that $\sigma^\star$ has property (6.1).

We show that $\sigma^\star = (\sigma_1^\star, \sigma_2^\star)$ is a subgame perfect secure equilibrium in $\mathcal{T}$, and thus in $(\mathcal{G}, v_0)$. Let hv be a history of this game ($v \in V$). We have to prove that $\sigma^\star|_h$ is a secure equilibrium in the subgame $(\mathcal{T}|_h, v)$. As a contradiction, suppose that there exists a $\prec_j$-profitable deviation σ'_j for some player $j \in \{1,2\}$ w.r.t. $\sigma^\star|_h$ in $(\mathcal{T}|_h, v)$. Let us assume w.l.o.g. that $j = 1$. As $\sigma^\star|_h$ is a Nash equilibrium in $(\mathcal{T}|_h, v)$ (see the proof of Theorem 6.1.1), we know that

$$\mathsf{Cost}_1(\rho) = \mathsf{Cost}_1(\rho') \text{ and } \mathsf{Cost}_2(\rho) < \mathsf{Cost}_2(\rho') \tag{7.1}$$

where $\rho = h\langle\sigma^\star_1|_h, \sigma^\star_2|_h\rangle_v$ and $\rho' = h\langle\sigma'_1|_h, \sigma^\star_2|_h\rangle_v$. Thus it implies that $\mathsf{Cost}_2(\rho)$ is finite. Let d be the maximum between $\mathsf{Cost}_1(\rho)$ and $\mathsf{Cost}_2(\rho)$ if $\mathsf{Cost}_1(\rho)$ is finite, or $\mathsf{Cost}_2(\rho)$ otherwise. Remark that $d > |h|$. According to property (6.1), there exists $n \geq d$ such that the strategy profiles $\sigma^\star$ and σ^n coincide on histories of length up to d.

Let us show that σ'_1 would then be a $\prec_1$-profitable deviation for player 1 w.r.t. $\sigma^n|_h$ in $(\mathsf{Trunc}_n(\mathcal{T})|_h, v)$. In this aim we first prove that

$$\mathsf{Cost}_2(\pi) < \mathsf{Cost}_2(\pi') \tag{7.2}$$

where $\pi = h\langle\sigma^n_1|_h, \sigma^n_2|_h\rangle_v$ and $\pi' = h\langle\sigma'_1|_h, \sigma^n_2|_h\rangle_v$ are finite plays in $\mathsf{Trunc}_n(\mathcal{T})$ (see Figure 7.6). By definition of d and according to property (6.1), we have that $\mathsf{Cost}_2(\pi) = \mathsf{Cost}_2(\rho) \leq d$. If $\mathsf{Cost}_2(\rho') = \mathsf{Cost}_2(\pi')$, Equation (7.1) implies that $\mathsf{Cost}_2(\pi) < \mathsf{Cost}_2(\pi')$. Otherwise, we have that $\mathsf{Cost}_2(\pi') > d$ as ρ' and π' coincide until depth d (by property (6.1)), and then $\mathsf{Cost}_2(\pi) \leq d < \mathsf{Cost}_2(\pi')$.

We now consider $\mathsf{Cost}_1(\pi)$ and $\mathsf{Cost}_1(\pi')$. Let us study the next two cases.

- If $\mathsf{Cost}_1(\rho) < +\infty$, then we have that

$$\mathsf{Cost}_1(\pi) = \mathsf{Cost}_1(\pi') \tag{7.3}$$

  because $\mathsf{Cost}_1(\rho') = \mathsf{Cost}_1(\rho) = \mathsf{Cost}_1(\pi) = \mathsf{Cost}_1(\pi') \leq d$ by Equation (7.1), property (6.1) and definition of d.
- If $\mathsf{Cost}_1(\rho) = +\infty$, then we show that $\mathsf{Cost}_1(\pi) = +\infty$, and as a consequence we get that

$$\mathsf{Cost}_1(\pi) \geq \mathsf{Cost}_1(\pi'). \tag{7.4}$$

  As a contradiction suppose that $\mathsf{Cost}_1(\pi) < +\infty$. Let ρ_d be the first vertex of ρ that belongs to R_2 (we remind that $\mathsf{Cost}_2(\rho) = d$).

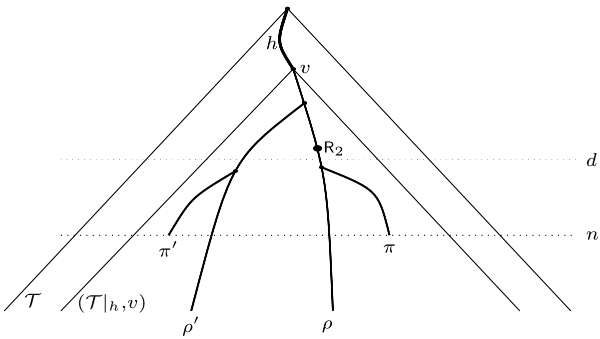


Figure 7.6: The game $\mathcal{T}$ with its subgame $(\mathcal{T}|_h, v)$.

We consider the (zero-sum qualitative) reachability game $\mathcal{G}^1 = (\mathcal{A}^1, \mathsf{R}_1)$, where $\mathcal{A}^1 = (V, (V_1, V_2), E)$, and player 1 aims at reaching R_1 while player 2 wants to prevent this. Suppose that player 1 has a winning strategy to reach his goal from vertex ρ_d in the game $\mathcal{G}^1$. Then this contradicts the fact that $\sigma^\star$ is a subgame perfect equilibrium in $\mathcal{T}$ (see the proof of Theorem 6.1.1). Therefore, by determinacy of $\mathcal{G}^1$ (Theorem 2.2.15), player 2 has a winning strategy from vertex ρ_d to prevent player 1 from reaching R_1. But in this case, this strategy is a $\prec_2$-profitable deviation w.r.t. $\sigma^n|_h$ in $(\mathsf{Trunc}_n(\mathcal{T})|_h, v)$, because player 2 can keep his cost while strictly increasing player 1's cost. This is impossible as σ^n is a subgame perfect secure equilibrium in $\mathsf{Trunc}_n(\mathcal{T})$. Thus, we must have that $\mathsf{Cost}_1(\pi) = +\infty$.

In all possible situations, we proved that σ'_1 is a $\prec_1$-profitable deviation for player 1 w.r.t. $\sigma^n|_h$ in $(\mathsf{Trunc}_n(\mathcal{T})|_h, v)$ because either $\mathsf{Cost}_1(\pi) = \mathsf{Cost}_1(\pi')$ and $\mathsf{Cost}_2(\pi) < \mathsf{Cost}_2(\pi')$, or $\mathsf{Cost}_1(\pi) > \mathsf{Cost}_1(\pi')$ (see (7.2–7.4)). So we get a contradiction with the fact that σ^n is a subgame perfect secure equilibrium in $\mathsf{Trunc}_n(\mathcal{T})$ by hypothesis. □

Unfortunately the proof does not seem to extend to the multiplayer case. Indeed we face the same kind of problems encountered in Section 5.2.1, where the existence of secure equilibria is proved for two-player games and left open for multiplayer games.

Remark 7.3.2. The existence of a subgame perfect secure equilibrium also holds in initialised two-player games with general quantitative reachability objectives where there is a unique non-zero natural price on every edge. Indeed, it suffices to replace any edge of price $c \in \mathbb{N}_0$ by a path of length c composed of c new edges (of price 1), and then apply Theorem 7.3.1 on this new game.

However, the existence of a subgame perfect secure equilibrium in initialised (two-player) quantitative reachability games with tuples of prices on edges is an open problem.

Remark 7.3.3. We know (see Remark 5.2.23 on page 154) that a secure equilibrium in a quantitative reachability game $(\mathcal{G}, v_0)$ is (in general) not a secure equilibrium in the corresponding qualitative game $(\overline{\mathcal{G}}, v_0)$. Then, Theorem 7.3.1 does *not* directly imply the existence of a subgame perfect secure equilibrium in two-player qualitative non-zero-sum games with reachability objectives. To our knowledge, the existence of a subgame perfect secure equilibrium in qualitative non-zero-sum games is an open question.

Chapter 8

Conclusion and Future Work

Let us conclude this thesis with a summary of our results and several perspectives for future work.

In Chapter 4, we stated the existence of a Nash equilibrium in large classes of multiplayer cost games. In Chapter 5, we considered the notion of secure equilibria in the quantitative framework, and proved the existence of secure equilibria in two-player quantitative reachability games. In Chapter 6, we gave an alternative proof to [FL83, Har85] for the existence of subgame perfect equilibria in quantitative reachability games. Finally, in Chapter 7, we introduced the new concept of subgame perfect secure equilibrium, and showed its existence in two-player quantitative reachability games.

Table 8.1 gives an overview of existence results for these four kinds of equilibria in several classes of cost games. In this table, 'NE' (resp. 'SE', 'SPE', 'SPSE') means 'Nash (resp. secure, subgame perfect, subgame perfect secure) equilibrium'. Moreover, an entry with a "Yes" and a reference implies that the existence result of the corresponding kind of equilibrium in the corresponding class of cost games has been proved, in the given reference(s). If "Yes" is in bold, then it corresponds to one of the results of this thesis. An entry “???” means that it is an open problem to our knowledge. Concerning the existence of secure (resp. subgame perfect secure) equilibria in cost games with real-valued, continuous cost functions (and then, in multiplayer quantitative reachability games), the entry is “Yes?” (resp. “Yes??”) because after some discussions with János Flesch, Jeroen Kuipers, Gijs Schoenmakers, and Koos Vrieze (from Maastricht university), we think that they have a sketch of proof (resp. some roads to explore) for this result.

Table 8.1: Results and open questions.

	Existence of			
	NE	SE	SPE	SPSE
Two-player quantitative reachability games	**Yes** [BBD10, BBD12]	**Yes** [BBD10, BBD12]	**Yes** [BBDG12, BBDG13]	**Yes** [BBDG12, BBDG13]
Multiplayer quantitative reachability games	**Yes** [BBD10, BBD12]	Yes?	**Yes** [BBDG12, BBDG13]	Yes??
Cost games as in Theorem 4.4.14	**Yes** [BDS13]	???	???	???
Cost games with real-valued, continuous cost functions	Yes [FL83, Har85]	Yes?	Yes [FL83, Har85]	Yes??

A first step to extend the results of Section 5.2 (and 7.3) could be to study the existence of (subgame perfect) secure equilibria in *multiplayer* quantitative reachability games, and in quantitative reachability games *with tuples of prices on edges*. A possible extension of the results of Section 4.4 could be to know if *secure (subgame perfect, subgame perfect* *secure) equilibria* always exist in (some subclasses of) the classes of cost games we defined in this section.

On the other hand, we do not know much about the new concept of *subgame perfect secure equilibrium.* It could be investigated further to find some properties that it satisfies.[1] Moreover, the definition of subgame perfect secure equilibrium naturally applies in the qualitative framework, and so, the existence of this equilibrium in *qualitative games* could be explored.

More generally, the results of this thesis can be extended into two main directions. The first one is about the *game model*: changes can be made in relation to the arena, the cost functions, the type of strategies,... The second direction concerns the *solution concept*: study other equilibria or other kinds of notions.

Regarding the game model, we could assume some hypotheses about the *game graph* and see if some proofs could not be simpler or generalised.

In order to extend the results of Section 4.4, a potential future work could be to identify other *classes* of multiplayer cost games where existence of (simple) Nash equilibria holds. Another possibility consists in fixing some particular cost functions. For example, the objective of a player could be a boolean combination of reachability and safety objectives, and a *multi-dimensional cost function* could be considered.

In this work, we have only considered pure strategies, but *mixed strategies* could be an interesting road to explore. In the same way, the game model could benefit from *randomised aspects*, while allowing for efficient verification and synthesis algorithms. It could also be appealing to see how our techniques apply to different kinds of models, like *timed games* [BBM10a], or *concurrent games* [UW11, KLŠT12]. Very recently, new quantitative objectives have been introduced in [CDRR13], but only zero-sum games have been regarded. The study of equilibria in the non-zero-sum variant of the latter paper is a challenging question.

In the direction of *solution concepts*, there are also many problems to study. Solution concepts represent relevant formalisms to express the

1. For example, in Section 7.1, we wonder if this notion could avoid unnecessary cycles in outcomes in quantitative reachability games.

properties of game models. Compared to zero-sum games, where giving one winning strategy is enough, non-zero-sum games have much richer solution concepts, including (but not exclusively) equilibria. New solution concepts could be introduced, possibly inspired from *verification* (in order to, for instance, express the long-term stability of a system in an open environment). Since computational tractability is a crucial factor in verification, the new solution concepts should be *decidable*.

Whereas some notions of equilibria may fail to capture interesting properties to be verified on particular computer systems, *logical formalisms* combine rich and adequate expressiveness. Several flavours of *temporal logics* for non-zero-sum games have been defined over the last few years (like the strategy logic [CHP10]). These formalisms have the required expressiveness, but unfortunately, they do not enjoy efficient algorithms for verification, and do not consider randomised strategies, nor any quantitative property. It could be promising to explore *variations* on these logics in order to arrive to a suitable combination of expressiveness and tractability.

Once a solution concept has been chosen, a characterisation of the *memory* usage could be done, as well as a study of the needed memory structure (is a finite automaton sufficient to compute the memory of the strategies, or are more complicated structures necessary?).

Other interesting questions involve *decision problems* related to games, and their associated *complexity*. For example, finding an efficient algorithm to decide, given a game and a tuple of thresholds, whether there exists a simple [2] equilibrium such that the costs of the players are below the thresholds.

Another line of research could be the description of the set of all equilibria outcomes in a game as, for example, the language accepted by an *automaton* (like in [KLŠT12]) or maybe a *regular expression*. It might also be possible to use *tree automata* in order to identify equilibria in a game (in the same spirit as [GU08] for qualitative games). Among the set of equilibria, we then could find the "best" equilibria according to a

2. In terms of the memory needed by the strategies, the complexity to describe them,. . .

certain criterion.

Most of the results in this thesis concern the existence of equilibria. Some of their proofs are constructive, but others are not. It could be interesting to get an *efficient construction* of equilibria, and of constrained equilibria (regarding the costs of the players, the memory needed by the strategies,. . .).

Bibliography

[A+81] Robert J Aumann et al. Survey of repeated games. *Essays in game theory and mathematical economics in honor of Oskar Morgenstern*, 4:11–42, 1981.

[AKW08] Rajeev Alur, Aditya Kanade, and Gera Weiss. Ranking automata and games for prioritized requirements. In *CAV*, volume 5123 of *Lecture Notes in Computer Science*, pages 240–253. Springer, 2008.

[BBD10] Thomas Brihaye, Véronique Bruyère, and Julie De Pril. Equilibria in quantitative reachability games. In *CSR*, volume 6072 of *Lecture Notes in Computer Science*, pages 72–83. Springer, 2010.

[BBD12] Thomas Brihaye, Véronique Bruyère, and Julie De Pril. On equilibria in quantitative games with reachability/safety objectives. *CoRR*, abs/1205.4889, 2012. Accepted for publication in Theory of Computing Systems.

[BBDG12] Thomas Brihaye, Véronique Bruyère, Julie De Pril, and Hugo Gimbert. Subgame perfection for equilibria in quantitative reachability games. In *FoSSaCS*, volume 7213 of *Lecture Notes in Computer Science*, pages 286–300. Springer, 2012.

[BBDG13] Thomas Brihaye, Véronique Bruyère, Julie De Pril, and Hugo Gimbert. On (subgame perfect) secure equilibrium in quantitative reachability games. *Logical Methods in Computer Science*, 9(1), 2013.

[BBM10a] Patricia Bouyer, Romain Brenguier, and Nicolas Markey. Computing equilibria in two-player timed games *via* turn-based finite games. In *FORMATS*, volume 6246 of *Lecture Notes in Computer Science*, pages 62–76. Springer, 2010.

[BBM10b] Patricia Bouyer, Romain Brenguier, and Nicolas Markey. Nash equilibria for reachability objectives in multi-player timed games. In *CONCUR*, volume 6269 of *Lecture Notes in Computer Science*, pages 192–206. Springer, 2010.

[BBMU11] Patricia Bouyer, Romain Brenguier, Nicolas Markey, and Michael Ummels. Nash equilibria in concurrent games with Büchi objectives. In *FSTTCS*, volume 13 of *LIPIcs*, pages 375–386. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2011.

[BCHJ09] Roderick Bloem, Krishnendu Chatterjee, Thomas A. Henzinger, and Barbara Jobstmann. Better quality in synthesis through quantitative objectives. In *CAV*, volume 5643 of *Lecture Notes in Computer Science*, pages 140–156. Springer, 2009.

[BDS13] Thomas Brihaye, Julie De Pril, and Sven Schewe. Multiplayer cost games with simple Nash equilibria. In *LFCS*, volume 7734 of *Lecture Notes in Computer Science*, pages 59–73. Springer, 2013.

[BFL+08] Patricia Bouyer, Ulrich Fahrenberg, Kim Guldstrand Larsen, Nicolas Markey, and Jirí Srba. Infinite runs in weighted timed automata with energy constraints. In *FORMATS*, volume 5215 of *Lecture Notes in Computer Science*, pages 33–47. Springer, 2008.

[BG09] Endre Boros and Vladimir Gurvich. Why chess and back gammon can be solved in pure positional uniformly optimal strategies. Rutcor Research Report 21-2009, Rutgers University, 2009.

[BPS07] Nicole Berline, Alain Plagne, and Claude Sabbah. *Théorie des jeux: introduction à la théorie des jeux répétés*. Ecole Polytechnique, 2007.

[CDRR13] Krishnendu Chatterjee, Laurent Doyen, Mickael Randour, and Jean-François Raskin. Looking at mean-payoff and total-payoff through windows. *CoRR*, abs/1302.4248, 2013.

[CGP01] Edmund M. Clarke, Orna Grumberg, and Doron Peled. *Model checking.* MIT Press, 2001.

[CH06] Krishnendu Chatterjee and Thomas A. Henzinger. Finitary winning in omega-regular games. In *TACAS*, volume 3920 of *Lecture Notes in Computer Science*, pages 257–271. Springer, 2006.

[CH07] Krishnendu Chatterjee and Thomas A. Henzinger. Assume-guarantee synthesis. In *TACAS*, volume 4424 of *Lecture Notes in Computer Science*, pages 261–275. Springer, 2007.

[CHJ04] Krishnendu Chatterjee, Thomas A. Henzinger, and Marcin Jurdzinski. Games with secure equilibria. In *FMCO*, volume 3657 of *Lecture Notes in Computer Science*, pages 141–161. Springer, 2004.

[CHJ05] Krishnendu Chatterjee, Thomas A. Henzinger, and Marcin Jurdzinski. Mean-payoff parity games. In *LICS*, pages 178–187. IEEE Computer Society, 2005.

[CHP10] Krishnendu Chatterjee, Thomas A. Henzinger, and Nir Piterman. Strategy logic. *Inf. Comput.*, 208(6):677–693, 2010.

[CMJ04] Krishnendu Chatterjee, Rupak Majumdar, and Marcin Jurdzinski. On Nash equilibria in stochastic games. In *CSL*, volume 3210 of *Lecture Notes in Computer Science*, pages 26–40. Springer, 2004.

[CR10] Krishnendu Chatterjee and Vishwanath Raman. Assume-guarantee synthesis for digital contract signing. *CoRR*, abs/1004.2697, 2010.

[DGP09] Constantinos Daskalakis, Paul W Goldberg, and Christos H Papadimitriou. The complexity of computing a Nash equilibrium. *SIAM Journal on Computing*, 39(1):195–259, 2009.

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

[FL83] Drew Fudenberg and David Levine. Subgame-perfect equilibria of finite-and infinite-horizon games. *Journal of Economic Theory*, 31(2):251–268, 1983.

[FL08] Drew Fudenberg and David Levine. A long-run collaboration on long-run games. *World Scientific Publishing Co*, 2008.

[FT91] Drew Fudenberg and Jean Tirole. *Game theory*. MIT Press, 1991.

[FV97] Jerzy Filar and Koos Vrieze. *Competitive Markov decision processes*. Springer Verlag, 1997.

[FZ12] Nathanaël Fijalkow and Martin Zimmermann. Cost-parity and cost-Streett games. *CoRR*, abs/1207.0663, 2012.

[GTW02] Erich Grädel, Wolfgang Thomas, and Thomas Wilke, editors. *Automata, logics, and infinite games*, volume 2500 of *Lecture Notes in Computer Science*. Springer, 2002.

[GU08] Erich Grädel and Michael Ummels. Solution concepts and algorithms for infinite multiplayer games. In *New Perspectives on Games and Interaction*, volume 4 of *Texts in Logic and Games*, pages 151–178. Amsterdam University Press, 2008.

[Har68] John C. Harsanyi. Games with incomplete information played by "Bayesian" players, I–III. *Management science*, 14(3):pages 159–182, 320–334, 486–502, 1967–1968.

[Har85] Christopher Harris. Existence and characterization of perfect equilibrium in games of perfect information. *Econometrica: Journal of the Econometric Society*, pages 613–628, 1985.

[HTW08] Florian Horn, Wolfgang Thomas, and Nico Wallmeier. Optimal strategy synthesis in request-response games. In *ATVA*, volume 5311 of *Lecture Notes in Computer Science*, pages 361–373. Springer, 2008.

[Kec95] Alexander S. Kechris. *Classical descriptive set theory.* Springer-Verlag, 1995.

[KLŠT12] Miroslav Klimoš, Kim Larsen, Filip Štefaňák, and Jeppe Thaarup. Nash equilibria in concurrent priced games. In *LATA*, volume 7183 of *Lecture Notes in Computer Science*, pages 363–376. Springer, 2012.

[Kuh53] Harold W. Kuhn. Extensive games and the problem of information. *Classics in Game Theory*, pages 46–68, 1953.

[LR09] Stéphane Le Roux. Acyclic preferences and existence of sequential Nash equilibria: A formal and constructive equivalence. In *TPHOLs*, volume 5674 of *Lecture Notes in Computer Science*, pages 293–309. Springer, 2009.

[Mar75] Donald A. Martin. Borel determinacy. *Ann. of Math. (2)*, 102(2):363–371, 1975.

[Mer86] Jean-François Mertens. Repeated games. CORE reprint No. 8624, Université catholique de Louvain, Center for Operations Research and Econometrics (CORE), 1986.

[Nas50] John F. 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.

[NRTV07] Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V Vazirani. *Algorithmic game theory.* Cambridge University Press, 2007.

[OR94] Martin J. Osborne and Ariel Rubinstein. *A course in game theory.* MIT Press, Cambridge, MA, 1994.

[PP04] Dominique Perrin and Jean-Eric Pin. *Infinite words: automata, semigroups, logic and games.* Academic Press, 2004.

[PS09] Soumya Paul and Sunil Simon. Nash equilibrium in generalised Muller games. In *FSTTCS*, volume 4 of *LIPIcs*, pages 335–346. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2009.

[PS11] Roger A. Purves and William D. Sudderth. Perfect information games with upper semicontinuous payoffs. *Mathematics of Operations Research*, 36(3):468–473, 2011.

[Sel65] Reinhard Selten. Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit. *Zeitschrift für die gesamte Staatswissenschaft*, 121:301–324 and 667–689, 1965.

[Sha53] Lloyd S. Shapley. Stochastic games. *Proceedings of the National Academy of Sciences of the United States of America*, 39(10):1095, 1953.

[Tho95] Wolfgang Thomas. On the synthesis of strategies in infinite games. In *STACS*, pages 1–13, 1995.

[Tho08] Wolfgang Thomas. Church's problem and a tour through automata theory. In *Pillars of Computer Science*, volume 4800 of *Lecture Notes in Computer Science*, pages 635–655. Springer, 2008.

[TR98] Frank Thuijsman and Thirukkannamangai E. S. Raghavan. Perfect information stochastic games and related classes. *International Journal of Game Theory*, 26(3):403–408, 1998.

[Tri09] Ashutosh Trivedi. *Competative optimisation on timed automata.* PhD thesis, University of Warwick, 2009.

[Umm05] Michael Ummels. Rational behaviour and strategy construction in infinite multiplayer games. Diploma thesis, RWTH Aachen, 2005.

[Umm06] Michael Ummels. Rational behaviour and strategy construction in infinite multiplayer games. In *FSTTCS*, 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 *FoSSaCS*, volume 4962 of *Lecture Notes in Computer Science*, pages 20–34. Springer, 2008.

[UW11] Michael Ummels and Dominik Wojtczak. The complexity of Nash equilibria in limit-average games. In *CONCUR*, volume 6901 of *Lecture Notes in Computer Science*, pages 482–496. Springer, 2011.

[vNM44] John von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, New Jersey, 1944.

[Zim09] Martin Zimmermann. Time-optimal winning strategies for poset games. In *CIAA*, volume 5642 of *Lecture Notes in Computer Science*, pages 217–226. Springer, 2009.