Stochastic Multiplayer Games
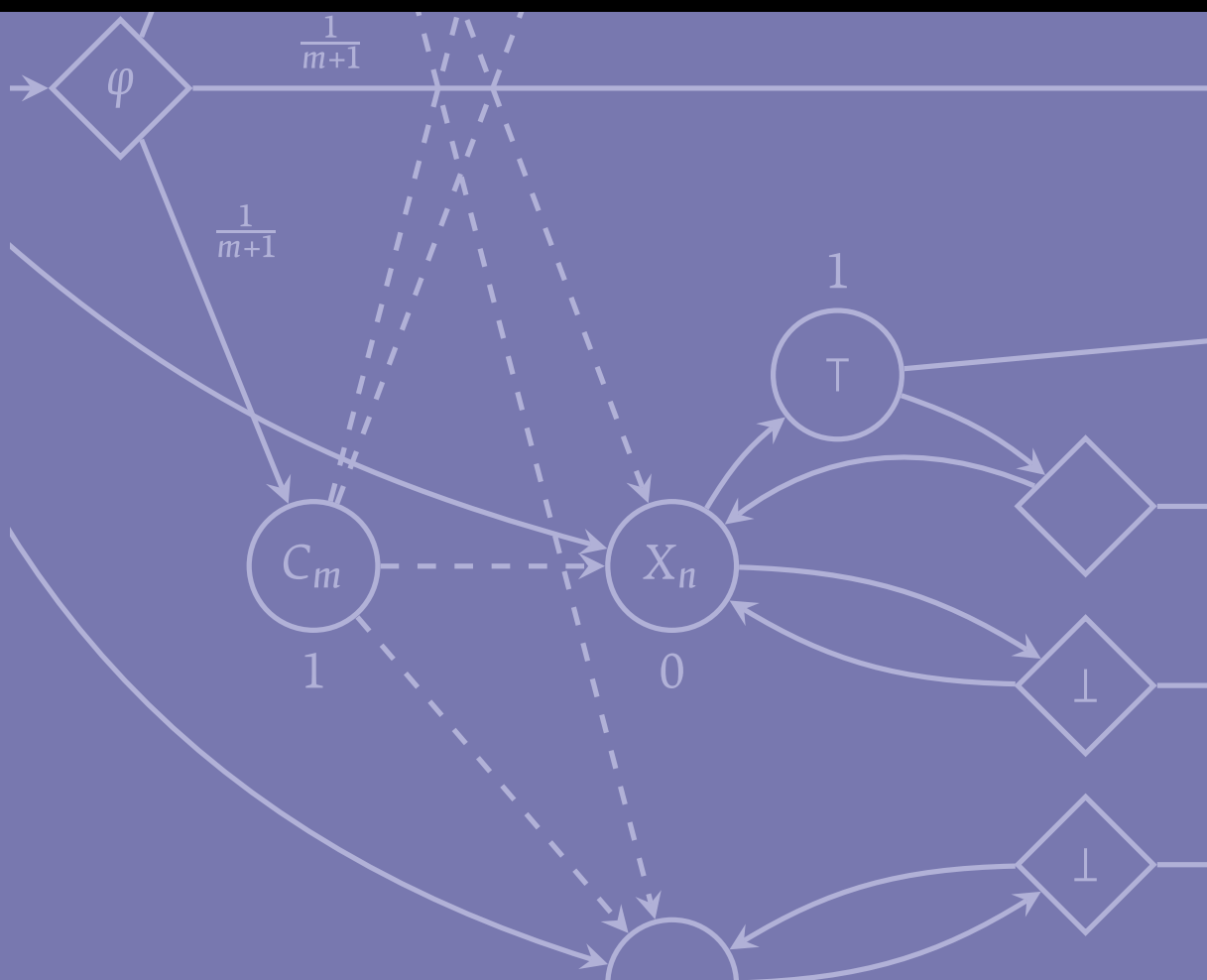# Theory and Algorithms

Michael Ummels

Stochastic Multiplayer Games
Theory and Algorithms

Stochastic Multiplayer Games: Theory and Algorithms

vorgelegt von

Diplom-Informatiker

Michael Ummels

aus Köln

Preface

The last decades have seen an immense amount of research on the algorithmic content of game theory. On the one hand, a new subject called *algorithmic game theory* has emerged that is concerned with the study of the algorithmic theory of *finite* games with multiple players. On the other hand, *infinite* and, in particular, *stochastic* two-player zero-sum games have become an important tool for the verification of open systems, which interact with their environment.

The aim of this work is to bring together algorithmic game theory with the games that are used in verification by extending the algorithmic theory of stochastic two-player zero-sum games to incorporate multiple players, whose objectives are not necessarily conflicting. In particular, this work contains a comprehensive study of the complexity of the most prominent solution concepts that are applicable in this setting, namely Nash and subgame-perfect equilibria.

This book is the result of my doctoral studies at RWTH Aachen University. I am indebted to my primary supervisor Erich Grädel for giving me the opportunity to pursue these studies, for introducing me to the scientific community and for giving me advice just when I needed it. I am equally grateful to my secondary supervisor Wolfgang Thomas for his constant support and encouragement.

Marcin Jurdziński did not hesitate to act as an external reviewer for this thesis. I thank him not only for his careful reading and numerous remarks, but also for giving an inspiring talk on branching vector addition systems, which indirectly led to the resolution of a problem that was left open in the original version of this thesis.

A substantial part of this book is based on joint work with Dominik Wojtczak. I am indebted to him for our numerous illuminating discussions, for his insights and ideas, and—last but not least—for hosting me in Edinburgh, Amsterdam and Oxford.

Among the various other people who contributed to this work, I would like to thank in particular Łukasz Kaiser for many enlightening discussions and for discovering Proposition 3.18. Special thanks also go to Florian Horn for many interesting discussions, to János Flesch for pointing out Proposition 3.13, and to Peter Bro Miltersen for drawing my attention to Corollary 4.4. Moreover, I am grateful to Hugo Gimbert and Eilon Solan for answering my questions and to Rohit Chadha, Tobias Ganzow, Jörg Olschewski and Edeline Wong for their comments on preliminary drafts of this work.

Finally, I would like to thank Sam Ross-Gower for designing the cover of this book, and Donald Knuth and Leslie Lamport for creating (LA)TEX.

Paris, November 2010

Contents

List of Figures

List of Tables

List of Algorithms

1

Introduction

In this first chapter, we introduce games and equilibria, present the main contributions of this work and discuss related work. Finally, the end of this chapter contains an outline of the rest of this book.

1.1 Games and equilibria

Generally speaking, game theory is occupied with understanding the phenomena that occur when rational entities interact. As a distinct field of study, game theory came into being in 1944, when von Neumann & Morgenstern published their seminal monograph, although it can be traced back to 1838 when Cournot published his work on *duopolies*; other early contributors were Zermelo (1913) and Borel (1921). Since then, game theory has found applications in fields as diverse as economics, sociology, biology, logic and—last but not least—computer science.

Matrix games

According to von Neumann & Morgenstern, a game is described by a k-dimensional matrix that consists of the payoffs (one for each player) for each possible combination of *strategies*. Consider, for example, the following situation (Osborne & Rubinstein 1994; Luce & Raiffa 1957). A couple wishes to attend a concert of classical music. Their main goal is to go out together, but

Table 1.1. Bach or Stravinsky.

	Bach	Stravinsky
Bach	2,1	0,0
Stravinsky	0,0	1,2

one of them prefers Bach, whereas the other person prefers Stravinsky. The payoff matrix for this game would then look like the one shown in Table 1.1.

Von Neumann & Morgenstern dealt primarily with two-player games that are completely antagonistic; what one player gains is the other player's loss. Formally, they require that for every pair of strategies, the payoffs of the two players sum up to 0 (or to any other constant value); this is why such games are called *two-player zero-sum games*.

The game of Bach or Stravinsky is obviously not zero-sum. What is a solution of such a game? Intuitively, there are two possible rational outcomes, in which both persons attend together a concert with music composed by either Bach or Stravinsky. If they go to different concerts, then each of them has an incentive to go to the respective other concert since their main concern is to enjoy a concert together. In general, a profile of strategies, one for each player, is a *Nash equilibrium* (Nash 1950) if no player can increase her payoff by unilaterally switching to a different strategy. Hence, the game of Bach or Stravinsky has two Nash equilibria: (Bach,Bach) and (Stravinsky,Stravinsky).

Note that a Nash equilibrium makes no statement on how the players arrive at the equilibrium. Moreover, a serious problem with Nash equilibria is that they are not *orthogonal*; if, for instance, one player arrives at the conclusion that (Bach,Bach) is the preferred equilibrium and therefore picks the strategy Bach, while the other player picks Stravinsky because she thinks that (Stravinsky,Stravinsky) is the preferred equilibrium, then the resulting pair of strategies is *not* an equilibrium. Hence, in general, the players have to *coordinate* their strategies in order to arrive at a Nash equilibrium.

Now consider a different situation, where two players have to choose either Head or Tail; if the choices are the same, the first player has to pay 1 € to the second player; if the choices differ, the second player has to pay 1 € to the first player. The payoff matrix of this game is depicted in Table 1.2: the rows of the matrix correspond to strategies of the first player. At first glance, it seems that this game does not have a Nash equilibrium; if the

Table 1.2. Matching pennies.

	Head	Tail
Head	1,−1	−1, 1
Tail	−1, 1	1,−1

choices are the same, then the second player will change her strategy, and if the choices differ, then the first player will change her strategy. However, there is an equilibrium in a different kind of strategy; If both players *randomise* their strategies and play Head or Tail with probability $\frac{1}{2}$ each, then each player receives an *expected payoff* of $\frac{1}{2}$ against every strategy of the other player, and we have a Nash equilibrium.

Formally, a *mixed strategy* is a probability distribution over the basic, so-called *pure*, strategies, and a *Nash equilibrium in mixed strategies* is a profile of mixed strategies such that no player can increase her expected payoff by unilaterally switching to another (mixed) strategy; in fact, it is easy to see that, in order to have a Nash equilibrium, it suffices that no player can gain from switching to a different *pure* strategy. Nash's theorem (1950) states that every game with finitely many players and finitely many pure strategies for each player has a Nash equilibrium in mixed strategies; for two-player zero-sum games, the existence of an equilibrium in mixed strategies already follows from the *minimax theorem* (von Neumann 1928).

Games in extensive form

Games in matrix form model one-shot events; both players choose their strategies at once and independently of each other, and the game is over. In practice, interaction occurs usually over time in a sequential fashion. This aspect is taken into account by games *in extensive form*. Consider, for example, the sequential version of matching pennies where the second player makes her decision only after the first player has made hers and announced it to the second player. Such a game is naturally represented by a tree such as the one in Figure 1.1.

A pure strategy in a game in extensive form selects, for each node in the tree that is labelled by the respective player, a possible action. If the tree is finite, then there is only a finite number of such strategies, and Nash's theorem guarantees the existence of a Nash equilibrium in mixed strategies.
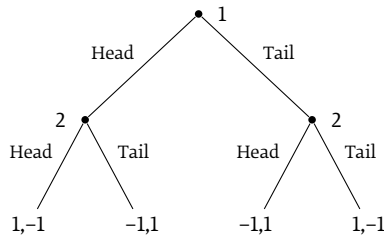
Figure 1.1. Matching pennies as a game in extensive form.

In fact, Kuhn (1950) showed that every finite game (of *perfect information*) in extensive form has an equilibrium in pure strategies, which can be found by a simple *backward induction*. Intuitively, this result relies on the fact that games in extensive form are *turn-based*: at every node of the tree there is a unique player who makes a decision, whereas in a matrix game the players make their decisions simultaneously. We will present several variants of Kuhn's theorem for stochastic games in Chapter 3.

In our example of matching pennies in extensive form, the second player can always make her choice dependent on the first player's choice; if the first player selects Head, she will select Tail, and if the first player selects Tail, she will select Head. If paired with any of the two pure strategies of the first player, we have a Nash equilibrium in pure strategies.

For games in extensive form, it turned out that Nash equilibria may lack credibility because players are able to change their strategy during the course of the game.[1] Hence, researchers have come up with more restricted solution concepts for games in extensive form. In particular, the notion of a *subgame-perfect equilibrium*, introduced by Selten (1965), addresses this deficiency and plays a central role in this work.

Stochastic games

Arguably, most—if not all—real-world systems are influenced by events of a probabilistic nature. Shapley (1953) was the first to define a game model that incorporates probabilistic choices: *Shapley games* are played by a finite number of players on a finite state space, and in each state, each player chooses one of finitely many actions; the resulting profile of actions determines a *reward* for each player and a probability distribution on successor states.

---

[1] We will see an example of such a game in Chapter 3.

In principle, a stochastic game proceeds ad infinitum. The payoff that each player receives is given by a function of the infinite stream of rewards for this player: Shapley considered games where payoffs are *discounted* sums of rewards; other popular payoff functions are the *limit average* of the rewards or the *total* sum of the rewards (see Filar & Vrieze 1997).

A pure strategy in a stochastic game assigns an action to each possible sequence of states visited so far, whereas a *randomised strategy* (the analogue of a mixed strategy for stochastic games) assigns a probability distribution on actions to each such sequence. Hence, every player has at her command an, in general, infinite number of pure strategies, and Nash's theorem is not applicable. Nevertheless, in the case of discounted payoffs, there always exists a Nash equilibrium in randomised strategies (Fink 1964). There is even a Nash equilibrium where the strategies only depend on the current state and not on the full history of visited states; we call such strategies *stationary*. For games with limit-average payoffs, Nash equilibria do, in general, not exist. However, Vielle (2000a,b) proved that every two-player stochastic game with limit-average payoffs has an *ε-equilibrium*, i.e. a profile of strategies where each player can gain at most ε from deviating, for all ε > 0. Whether ε-equilibria exist in stochastic games with more than two players and limit-average payoffs is an open question (Neyman & Sorin 2003).

Games in computer science

The first time that games were used as a tool to solve a (theoretical) problem in computer science was in 1969, when Büchi & Landweber solved *Church's problem* (Church 1957, 1963). Church asked whether it is possible, given a circuit or a logical formula that describes a binary relation on infinite sequences, to synthesise a circuit that computes for every input sequence an output sequence such that the output is in relation to the input (or to determine that such a circuit does not exist). Here, the circuit should compute the output *on the fly*, i.e. the ith letter of the output may only depend on the first i letters of the input. This scenario can naturally be modelled by a two-player game of infinite duration, where the players alternate in choosing letters from the two sequences. Büchi and Landweber showed that these games are *determined*, i.e. that either one of the two players has a winning strategy, and that one can compute a winning strategy that can be realised by a finite-state transducer.

The games that arise from Church's problem are games in extensive form, played on the (infinite) unravelling of a finite graph; each vertex carries the information which player has to output a letter and which letter has been output last. These *graph games* can also be used to solve Church's problem in the more general setting of a *reactive system* (*plant*) interacting with its environment (Abadi et al. 1989; Pnueli & Rosner 1989; Ramadge & Wonham 1989). The task is to synthesise a controller for the system such that the system behaves correctly for all possible behaviours of the environment (or to detect that this is impossible). If one combines the possible behaviours of the system and the environment into one game, such a controller corresponds to a winning strategy for one player in this game.

Graph games also play an important role in the automated verification of systems with respect to logical specifications, known as *model checking* (Clarke et al. 1999; Baier & Katoen 2008). For instance, the question whether a formula in Hennessy-Milner logic (Hennessy & Milner 1985) holds in a finite transition system can be reduced to the question whether one player has a winning strategy in a *reachability game* on a certain graph, which can easily be constructed from the formula and the system.

For more complex logics such as the modal μ-calculus L_μ (Kozen 1983), the games that arise from the model-checking problem have more complicated winning conditions which refer to the set of vertices occurring infinitely often in a play. These games are called *parity games*, and we will discuss them in the next chapter. Let us only mention at this point that the computational complexity of model checking L_μ hinges on the complexity of deciding which player has a winning strategy in a parity game (see Grädel 2007), a problem which is not known to be solvable in polynomial time.

There are many more areas of computer science where games have entered the picture. For instance, the semantics of a computational model can often be naturally defined as a game. This can not only be done for models that are close in spirit to games such as *alternating Turing machines* (Chandra et al. 1981), but also for certain functional programming languages such as PCF, for which game-theoretic semantics provided the first fully abstract model (Abramsky et al. 2000; Hyland & Ong 2000).

An emerging area of computer science whose subjects are games themselves is *algorithmic game theory* (Nisan et al. 2007), which is concerned with the computational content of game theory. In particular, algorithmic game theory has dealt with the computational complexity of finding equilibria.

Daskalakis et al. (2006, 2009) established that finding an ε-equilibrium of a finite matrix game is complete for the class PPAD (a class of function problems situated between FP and FNP). Later, Chen & Deng (2006) strengthened their result by establishing that computing a Nash equilibrium of a finite two-player matrix game is also PPAD-complete.

The stochastic games we study in this work are closer to games in extensive form than Shapley games. In particular, they differ from the classical setting in two aspects: First, they are turn-based. Second, there are no immediate rewards on transitions; instead, every player has a certain *objective*, which is a (Borel) set of desired plays. Intuitively, every player aims at generating a play that meets her objective. Hence, the payoff a player receives from a single play is just 0 or 1 (depending on whether the play fulfils the player's objective or not). In the two-player zero-sum variant, these games occur naturally when one wants to build a controller for a system interacting with a probabilistic environment (Baier et al. 2004).

1.2 The stochastic dining philosophers problem

For the kind of stochastic games we study here, most research has concentrated on the two-player zero-sum case; see Chapter 2 for a survey of results. To see why it is worthwhile to study games with multiple players in computer science, let us look at an example, which is a variant of the *dining philosophers problem*, originally introduced by Dijkstra (1971) to illustrate the difficulties of synchronisation in concurrent systems.

In the dining philosophers problem, there are $n + 1$ philosophers ($n \geq 1$) sitting at a round table with a bowl of rice in the middle. Between any two philosophers who sit next to each other lies a chopstick, which can be accessed by both of them. Since the table is round, there are as many chopsticks as there are philosophers; see Figure 1.2. To eat from the bowl, a philosopher needs to acquire both of the chopsticks he has access to. Hence, if one philosopher eats, then his two neighbours cannot eat at the same time. The life of a philosopher is rather simple and consists of thinking and eating; to survive, a philosopher needs to think and eat again and again. The task is to design a protocol that allows all of the philosophers to survive.

There are many solutions to the dining philosophers problem. For example, the philosophers could proceed in rounds: in each round, only one philosopher eats and all others think (see below).
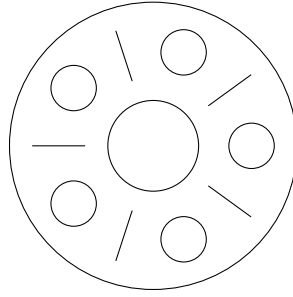
Figure 1.2. Dining philosophers.

Now, let us make the problem a little harder by removing one of the $n + 1$ chopsticks uniformly at random at the beginning of the game. Obviously, this makes it impossible for two philosophers to survive (since they only have access to one chopstick). More precisely, with probability $2/(n + 1)$ a philosopher will have access to only one chopstick and die.

What is a good protocol in such a situation? Clearly, we want each philosopher to survive with high probability (i.e. with probability $1 - 2/(n + 1)$). Moreover, it is natural to require that a philosopher who does not survive if he follows the protocol should not be able to survive by sabotaging the protocol (possibly inflicting harm on the other philosophers). This property is ensured if the proposed protocol forms a Nash equilibrium; being perfectly rational, no philosopher has an incentive to deviate from such a protocol.

Let us model the stochastic dining philosophers problem by a stochastic game. A state of the game comprises the state of each philosopher and the state of each chopstick; a philosopher may either think, eat, wait for the chopstick on his left (right) side, or wait to return the chopstick on his left (right) side, and a chopstick may either be missing, available, or occupied by the philosopher on its left (right) side. Since our model is turn-based, we also assume that there is a variable *turn*, whose value determines which philosopher may execute an action; after the action has been performed, the variable is reset randomly.

The complete game can be represented as the synchronous composition of processes $P_0, \ldots, P_n, C_0, \ldots, C_n, S$: process P_i models the ith philosopher, process C_i models the ith chopstick, and process S models the *scheduler*, which controls the turn variable and removes one chopstick uniformly at random at the beginning of the game. These processes and the actions they can execute
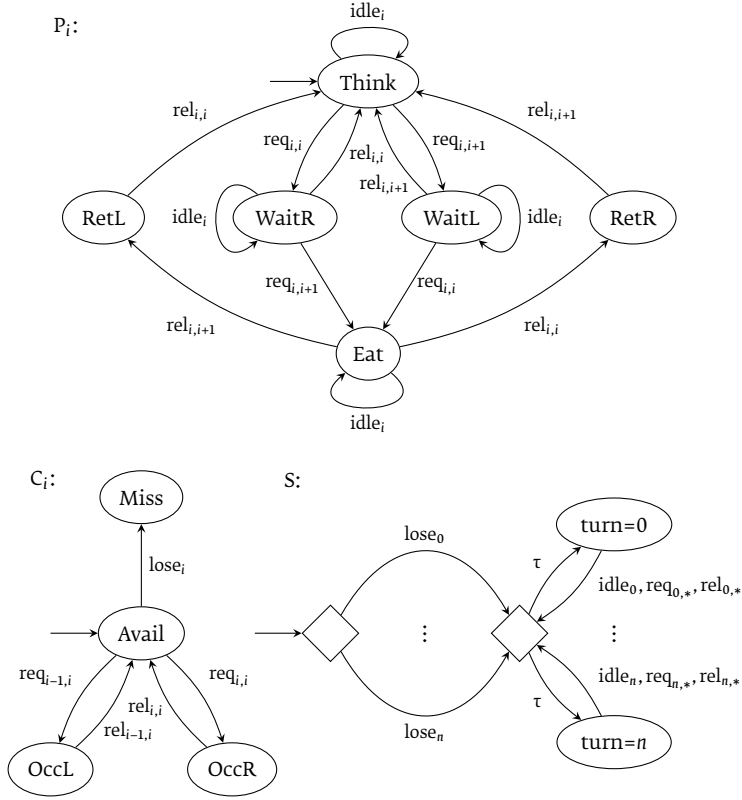
Figure 1.3. Processes for the stochastic dining philosophers problem.

are depicted in Figure 1.3 (arithmetical operations ought to be understood as modulo $n+1$). Diamond shaped vertices stand for states where a probabilistic choice is taken; with probability $1/(n+1)$ each, one of the outgoing transitions is selected, and the corresponding action is taken. Note that the actions $req_{i,j}$ and $rel_{i,j}$ are shared by the processes P_i, C_j and S, whereas the action $idle_i$ is only shared by P_i and S, and the action $lose_i$ is only shared by C_i and S; the symbol τ denotes an internal action.

For $n = 1$, a part of the complete game is depicted in Figure 1.4; the part of the game that is entered when the action $lose_0$ is taken is symmetric and not shown. In the figure, the vertex labelled $(wr, or, t, m, 0)$ represents, for instance, the state where the first philosopher waits for the chopstick on his right (i.e. he has acquired the chopstick on his left), the first chopstick

$lose_0$ …

$lose_1$

$rel_{0,0}$ $rel_{1,0}$

$idle_0$ $idle_1$

$t, a, t, m, 0$ $t, a, t, m, 1$

$req_{0,0}$ $req_{1,0}$

$idle_0$ $idle_1$ $idle_0$ $idle_1$

$wr, or, t, m, 0$ $wr, or, t, m, 1$ $t, ol, wl, m, 0$ $t, ol, wl, m, 1$
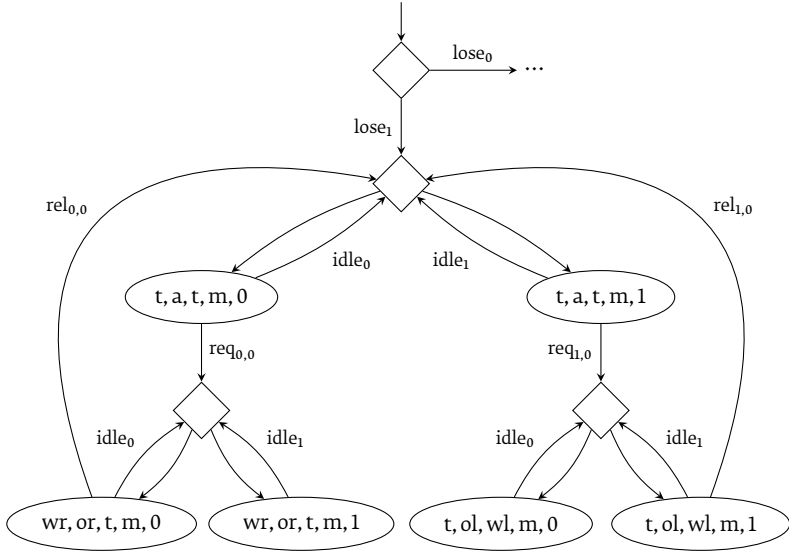
Figure 1.4. The stochastic dining philosophers game with two philosophers.

is occupied by the philosopher on its right (i.e. by the first philosopher), the second philosopher thinks, the second chopstick is missing (since the action $lose_1$ has been executed), and the first philosopher may execute an action. Note that no state where a philosopher eats is reachable from the initial state (which is not surprising given that a philosopher needs two chopsticks to eat). Hence, there does not exist a protocol where a philosopher survives with non-zero probability.

For $n > 1$, the stochastic dining philosophers game has several equilibria: in some of them, each philosopher survives with probability 0; in others, the probability of survival is non-zero. For instance, consider the greedy (albeit foolish) strategy where a philosopher first tries to acquire the left chopstick and subsequently the right chopstick; once a philosopher has acquired both chopsticks, he continues eating forever. In particular, a chopstick that has been acquired once is never released. Clearly, every philosopher survives with probability 0 if all philosophers adhere to this strategy. Yet, this profile of strategies constitutes a Nash equilibrium; if one philosopher changes his strategy and returns his two chopsticks to resume thinking, with positive probability one of his neighbours (adhering to the greedy strategy) picks up one of these chopsticks and never hands it back. Hence, the probability

that the philosopher who has changed his strategy can go from thinking to eating and back at least k times tends to 0.

Now consider the strategy where a philosopher only acquires a chopstick if both chopsticks adjacent to him are present and no other philosopher is holding a chopstick; once a philosopher has eaten, he puts both chopsticks back on the table, so that he can resume thinking (the order in which chopsticks are put up and down is arbitrary). With probability 1, a philosopher who is not missing a chopstick will survive if all philosophers adhere to this strategy. On the other hand, if a philosopher is missing one of his chopsticks, he will starve and die. Hence, since the probability of missing a chopstick is $2/(n+1)$, each philosopher survives with probability $1 - 2/(n+1)$ with this profile of strategies. Moreover, we have a Nash equilibrium since there is no way for a philosopher who is missing a chopstick to survive.

Clearly, the latter equilibrium is preferable to the former because the probability of survival is greater. Moreover, the equilibrium strategies have the attractive property that the chosen action only depends on the current state of the game; we call such strategies *positional*. In Chapter 4, we will see that deciding the existence of an equilibrium in positional strategies is NP-complete.

1.3 Contributions

The first step in analysing a mathematical concept is to prove its existence. For Nash equilibria in stochastic games, existence was proven by Chatterjee et al. (2004b). However, their proof contains an inaccuracy, which we address in this work. By contrast, subgame-perfect equilibria do, in general, not exist in stochastic games. Nevertheless, we show that they do exist in the special case of *deterministic* games with Borel objectives.

From a computer science point of view, the mere existence of an object is not sufficient; we also want to compute it. We observe that, for games with parity objectives, we can verify in polynomial time whether a given strategy profile is a Nash or subgame-perfect equilibrium. This puts the problem of computing a Nash equilibrium of a stochastic game with parity objectives into the class FNP of function problems for which a possible solution can be verified in polynomial time. In particular, there exists a polynomial-space algorithm for computing an arbitrary Nash equilibrium of a stochastic game with parity objectives.

With the stochastic dining philosophers example in mind, we argue that it makes sense to measure the computational complexity of equilibria not only in terms of how hard it is to compute an arbitrary equilibrium (as in algorithmic game theory), but also of how hard it is to compute an equilibrium with a certain payoff. More precisely, we permit the placing of a lower and an upper threshold on the payoff of each player. The corresponding decision problem is:

> Given a game with k players and payoff thresholds $\bar{x}, \bar{y} \in [0,1]^k$, decide whether the game has an equilibrium whose payoff lies in between $\bar{x}$ and $\bar{y}$.

Depending on whether we ask for a Nash or a subgame-perfect equilibrium, we already obtain two different decision problems. It turns out that it also makes a difference in what types of strategies the equilibrium is realised. In this work, we consider six types of strategies: positional strategies, stationary strategies (which can be randomised), pure finite-state strategies, randomised finite-state-strategies (both of which may depend on some finite information about the sequence of states seen so far), arbitrary pure strategies and arbitrary randomised strategies (both of which may depend on the full sequence of states seen so far).

We show that the complexity of the decision problem is highly dependent on the type of strategies that one allows for the equilibrium: The problem is typically decidable if we look for equilibria in positional or stationary strategies, but it becomes undecidable if we allow arbitrary (randomised or pure) strategies or finite-state strategies. In fact, we prove that it is not possible to decide the existence of an equilibrium where a designated player wins with probability 1 for these types of strategies (for all other players there is no constraint on the payoff).

In order to perform a more refined complexity analysis, we need to restrict the type of objectives; we show that for the typical objectives used as acceptance conditions for automata on infinite words, deciding whether there exists a positional equilibrium whose payoff lies in between $\bar{x}$ and $\bar{y}$ is NP-complete, whereas we can only give a Pspace upper bound for the stationary case. However, we prove that the latter problem is at least as hard as the *square root sum problem*, a problem about exact numerical computations which is not known to lie inside the polynomial hierarchy. Hence, our Pspace upper bound seems hard to improve.

In fact, all the lower bounds we have mentioned so far hold for stochastic games with a very restricted type of objectives, namely *simple reachability objectives*. In particular, this type of objectives is subsumed by all the objectives that play a role in verification. Moreover, the payoff function defined by simple reachability objectives is a special case of the limit-average payoff function with binary rewards on transitions; hence, our lower bounds also hold for stochastic games with limit-average payoffs.

Although it is, in general, not possible to decide the existence of an equilibrium with a certain payoff, we prove decidability for several fragments of the original decision problem: First, we show that the problem becomes decidable when one looks for an equilibrium where each player either wins or loses with probability 1. Second, we prove decidability for the restriction where one requires all but one player to win with probability 1; additionally, for the payoff of the remaining player, we can specify a lower threshold. Finally, we show that the problem is decidable for deterministic games if we restrict ourselves to binary thresholds.

For all of the fragments we study, we classify the complexity of the problem with respect to the type of objectives. In many cases, it turns out that their complexity is comparable to the complexity of solving two-player zero-sum stochastic games with the same type of objectives. In other cases, the problems become harder; for instance, deciding whether in a deterministic game with co-Büchi objectives there exists a Nash equilibrium that is winning for the first player is NP-complete, whereas the corresponding decision problem for two-player zero-sum games is solvable in polynomial time. In addition, we show that for all of the fragments we consider it does not make a difference whether one considers randomised or pure strategies; in fact, in most cases, pure finite-state strategies are sufficient.

Most of the results presented in Chapter 4 and some of the results presented in Chapter 5 were obtained in collaboration with Dominik Wojtczak. Preliminary expositions of most of the results presented in this work were published in the proceedings of various conferences and workshops (Ummels 2008; Grädel & Ummels 2008; Ummels & Wojtczak 2009a, 2009b).

1.4 Related work

In algorithmic game theory, the predominant question has been the complexity of computing equilibria as a function problem. The decision version,

where one asks whether there exists an equilibrium with certain proper-
ties, has attracted considerably less interest. Surely, one reason for this lack
of interest is that it was realised early on that such problems are usually
NP-hard for finite matrix games (Gilboa & Zemel 1989). In particular, decid-
ing whether in a two-player matrix game there exists a Nash equilibrium
where the first player's payoff is greater than a given threshold is NP-hard
(Conitzer & Sandholm 2003), even if the payoff matrix is binary (Codenotti &
Štefankovič 2005). Neither of these results implies one of our results since
our games are turn-based.

A more restricted model of stochastic games, where questions like ours
have been studied, are *Markov decision processes* (MDPs) with multiple objec-
tives. These games can be considered as stochastic games where only one
player can influence the outcome of the game. For MDPs with multiple
ω-regular objectives, Etessami et al. (2008) showed that questions like the
one we ask are decidable. Their result relies on the fact that, for MDPs
with multiple simple reachability objectives, stationary strategies suffice to
achieve a payoff that is higher than a given threshold. Unfortunately, this
property does not extend to our model: we give an example of a stochastic
game with simple reachability objectives where every Nash equilibrium
in which the first player wins with probability 1 requires infinite memory
(see Proposition 4.12).

1.5 Outline

In Chapter 2, we define the game model that underlies this work and survey
results on two-player zero-sum stochastic games.

Chapter 3 contains our results on the existence of Nash and subgame-
perfect equilibria in stochastic games. In that chapter, we also analyse
the complexity of computing an equilibrium with an arbitrary payoff and
introduce the various decision problems associated with Nash and subgame-
perfect equilibria in different types of strategies.

In Chapter 4, we present our results on the complexity of Nash and
subgame-perfect equilibria: Sections 4.1 and 4.2 deal with equilibria in posi-
tional and stationary strategies respectively, for which we prove decidability
results. Finally, Sections 4.3 and 4.4 are concerned with equilibria in arbi-
trary (pure or randomised) strategies and finite-state strategies respectively,
for which we prove undecidability.

In Chapter 5, we look at several fragments of the original decision problem for Nash equilibria and prove their decidability. Section 5.1 covers the fragment where one restricts to equilibria in which each player either wins or loses almost surely; Section 5.2 deals with the special case where all but one player are required to win with probability 1, and Section 5.3 contains our results on deterministic games.

Finally, in Chapter 6, we list some open problems and point out possible extensions to this work.

For readers who do not have the necessary background on probability or complexity theory, Appendix A provides a brief introduction to the relevant concepts. Additionally, Appendix B surveys results on Markov chains and Markov decision processes that are essential for this work.

2

Stochastic Games

In this chapter, we introduce stochastic games formally, and we summarise the main results on two-player zero-sum stochastic games.

Notation

We denote by $\mathbb{N} = \{0, 1, \ldots\}$ the set of all natural numbers (including 0), by $\mathbb{R}$ the set of all real numbers, and by $[0, 1]$ the set of all $x \in \mathbb{R}$ such that $0 \leq x \leq 1$. Given a set A, we denote by $\mathcal{P}(A)$ its power set, and by $\mathcal{D}(A)$ the set of all (discrete) *probability distributions* over A, i.e. functions $p \colon A \to [0, 1]$ such that $p(a) = 0$ for all but countably many $a \in A$ and $\sum_{a \in A} p(a) = 1$. Moreover, we denote by A^* and A^ω the set of all finite, respectively infinite, sequences over A; the empty sequence is denoted by ε, and we set $A^+ := A^* \setminus \{\varepsilon\}$. The length of a finite sequence x is denoted by $|x|$, and we write $x \prec y$ ($x \preceq y$) if x is a proper (non-proper) prefix of y. Given an infinite sequence $\alpha = \alpha(0)\alpha(1)\ldots$, we denote by $\alpha|_k = \alpha(0)\ldots\alpha(k-1)$ its prefix of length $k \in \mathbb{N}$ and by $\mathrm{Inf}(\alpha)$ the set of elements occurring infinitely often in α. Finally, for $X \subseteq A^\omega$ and $x \in X^*$, we denote by $x^{-1}X$ the set $\{\alpha \in A^\omega : x \cdot \alpha \in X\}$.

2.1 Arenas and objectives

Let us start by giving a formal definition of the game model that underlies this work. The games we are interested in are played by multiple players

taken from a finite set Π of players; we usually refer to them as player 0, player 1, player 2, and so on.

The *arena* of the game is basically a directed, coloured graph. Intuitively, the players take turns to form an infinite path through the arena, a *play*. Additionally, there is an element of chance involved: at some vertices, it is not a player who decides how to proceed but *nature*, who chooses a successor vertex according to a probability distribution. To model this scenario, we partition the set V of vertices into sets V_i of vertices *controlled* by player $i \in \Pi$ and a set of *stochastic vertices*, and we extend the edge relation to a transition relation that takes probabilities into account. Formally, an arena for a game with players in Π consists of:

- a nonempty, countable set V of *vertices* or *states*,
- for each player i a set $V_i \subseteq V$ of vertices *controlled* by player i,
- a *transition relation* $\Delta \subseteq V \times ([0,1] \cup \{\bot\}) \times V$, and
- a *colouring function* $\chi \colon V \to C$ into a set C of *colours*.

We make the assumption that every vertex is controlled by at most one player: $V_i \cap V_j = \varnothing$ if $i \neq j$; vertices that are not controlled by any player are called *stochastic*. Moreover, we require that $\bot$ appears in a transition $(v, p, w) \in \Delta$ if and only if v is a controlled vertex, and that transition probabilities are unique: if v is a stochastic vertex and w is an arbitrary vertex, then there exists precisely one $p \in [0,1]$ such that $(v, p, w) \in \Delta$; we denote this probability by $\Delta(w \mid v)$. For computational purposes, we assume that these probabilities are rational numbers. Naturally, for each stochastic vertex v the probabilities on outgoing transitions must sum up to 1: $\sum_{w \in V} \Delta(w \mid v) = 1$. Finally, for $v \in V$, let

$$v\Delta := \{w \in V : \text{there exists } 0 \neq p \in [0,1] \cup \{\bot\} \text{ such that } (v, p, w) \in \Delta\}$$

be the set of possible successor vertices; for technical reasons, we assume that for each controlled vertex $v \in \bigcup_{i \in \Pi} V_i$ the set $v\Delta$ is finite and nonempty.

The description of a game is completed by specifying an *objective* for each player. On an abstract level, these are just arbitrary sets of infinite sequences of colours, i.e. subsets of C^ω. Since we want to assign a probability to them, we assume that objectives are Borel sets (see Appendix A), if not stated otherwise. Since objectives specify which plays are winning for a player, they are also called *winning conditions*.

In general, we identify an objective Win $\subseteq C^\omega$ over colours with the corresponding objective $\chi^{-1}(\text{Win}) := \{\pi \in V^\omega : \chi(\pi) \in \text{Win}\} \subseteq V^\omega$ over vertices (which is also Borel since χ, as a mapping $V^\omega \to C^\omega$, is continuous). In fact, for the mathematical treatment of stochastic games, it is perfectly safe to assume that $C = V$ and that χ is the identity function. The reason that we allow objectives to refer to a colouring of the vertices is that the number of colours can be much smaller than the number of vertices, and it is possible that an objective can be represented more succinctly as an objective over colours rather than as an objective over vertices.

If Π is a finite set of players, $(V, (V_i)_{i \in \Pi}, \Delta, \chi)$ is an arena and $(\text{Win}_i)_{i \in \Pi}$ is a collection of objectives, we call the tuple $\mathcal{G} = (\Pi, V, (V_i)_{i \in \Pi}, \Delta, \chi, (\text{Win}_i)_{i \in \Pi})$ a *stochastic multiplayer game* (SMG). An SMG is *finite* if its arena is finite.

A *play* of $\mathcal{G}$ is an infinite path through the arena of $\mathcal{G}$, i.e. an infinite sequence $\pi = \pi(0)\pi(1)\dots$ of vertices such that $\pi(k+1) \in \pi(k)\Delta$ for each $k \in \mathbb{N}$. Finite prefixes of plays are called *histories*. We say that a play π of $\mathcal{G}$ is *won* by player i if the corresponding sequence of colours fulfils player i's objective, i.e. if $\chi(\pi) \in \text{Win}_i$; the *payoff* of a play π is the vector $\bar{x} \in \{0,1\}^\Pi$ defined by $x_i = 1$ if and only if $\chi(\pi) \in \text{Win}_i$.

Often, it is convenient to designate an initial vertex $v_0 \in V$; we call the pair $(\mathcal{G}, v_0)$ an *initialised* SMG. A play or a history of an initialised SMG $(\mathcal{G}, v_0)$ is just a play, respectively a history, of $\mathcal{G}$ that starts in v_0. In the following, we will refer to both SMGs and initialised SMGs as SMGs; it should always be clear from the context whether the game is initialised or not.

The SMG model may be generalised to allow for concurrent behaviour. In this case, each player has at her command a number of *actions*, one of which she has to pick whenever the play arrives at a vertex. The joint profile of actions, chosen by the players simultaneously, determines a probability distribution on successor vertices. The resulting model, named *concurrent games* by de Alfaro et al. (2007), is closer to the original model by Shapley (1953), but lacks many of the attractive properties of our model.

Although they are devoid of concurrency, SMGs provide a versatile model and generalise various other stochastic models, each of them the subject of intensive research. First, there are *Markov chains*, the basic model for stochastic processes, in which no control is possible. These are SMGs where the set Π of players is empty and (consequently) there are only stochastic vertices.

If we extend Markov chains by a single controller, we arrive at the model of a *Markov decision process* (MDP), a model introduced by Bellman (1957) and

heavily used in operations research. Formally, an MDP is an SMG with only one player (and only one objective).

Finally, in a (perfect-information) *stochastic two-player zero-sum game* (S2G), there are only two players, player 0 and player 1, who have opposing objectives: one player wants to fulfil her objective, while the other one wants to prevent her from doing so. Hence, one player's objective is the complement of the other player's objective. Due to their competitive nature, these games are also known as *competitive Markov decision processes* (see Filar & Vrieze 1997).

The SMG model also incorporates several non-stochastic models. In particular, we call an SMG *deterministic* if it contains no stochastic vertices. In the two-player zero-sum setting, the resulting model has found applications in logic and controller synthesis (see Section 1.1).

Types of objectives

We have introduced objectives as abstract sets of infinite sequences. In order to be amenable to algorithmic manipulation, we need to restrict to a class of objectives representable by finite objects. The objectives we consider for this purpose are standard in logic and verification (see Grädel et al. 2002); for all of them, we require that the set C of colours the objective refers to is *finite*.

- A *reachability objective* is given by a set $F \subseteq C$ of *good* colours, and the objective requires that a good colour is seen at least once. The corresponding subset of C^ω is $\mathrm{Reach}(F) := \{\alpha \in C^\omega : \alpha(k) \in F \text{ for some } k \in \mathbb{N}\}$.
- A *safety objective* is also given by a set $F \subseteq C$ of good colours, but this time the objective requires that *only* good colours are seen. The corresponding subset of C^ω is $\mathrm{Safe}(F) := \{\alpha \in C^\omega : \alpha(k) \in F \text{ for all } k \in \mathbb{N}\}$.
- A *Büchi objective* is again given by a set $F \subseteq C$ of good colours, but it requires that a good colour is seen infinitely often. The corresponding subset of C^ω is $\mathrm{Büchi}(F) := \{\alpha \in C^\omega : \mathrm{Inf}(\alpha) \cap F \neq \varnothing\}$.
- A *co-Büchi objective* is also given by a set $F \subseteq C$ of good colours; this time, the objective requires that *from some point onwards* only good colours are seen. The corresponding subset of C^ω is $\mathrm{coBüchi}(F) = \{\alpha \in C^\omega : \mathrm{Inf}(\alpha) \subseteq F\}$.
- A *parity objective* is given by a *priority function* $\Omega: C \to \{0, \ldots, d\}$, where $d \in \mathbb{N}$, which assigns to each colour a certain *priority*. The objective requires that the least priority that occurs infinitely often is even. The corresponding subset of C^ω is $\mathrm{Parity}(\Omega) = \{\alpha \in C^\omega : \min(\mathrm{Inf}(\Omega(\alpha))) \text{ is even}\}$.

- A *Streett objective* is given by a set Ω of *Streett pairs* (F, G), where $F, G \subseteq C$. The objective requires that, for each of the pairs, if a colour on the left-hand side is seen infinitely often, then so is a colour on the right-hand side of this pair. The corresponding subset of C^ω is $\mathrm{Streett}(\Omega) = \{\alpha \in C^\omega : \mathrm{Inf}(\alpha) \cap F = \varnothing \text{ or } \mathrm{Inf}(\alpha) \cap G \neq \varnothing \text{ for all } (F, G) \in \Omega\}$.

- A *Rabin objective* is given by a set Ω of *Rabin pairs* (F, G), where $F, G \subseteq C$. The objective requires that for some pair a colour on the left-hand side is seen infinitely often while all colours on the right-hand side of this pair are seen only finitely often. The corresponding subset of C^ω is $\mathrm{Rabin}(\Omega) = \{\alpha \in C^\omega : \mathrm{Inf}(\alpha) \cap F \neq \varnothing \text{ and } \mathrm{Inf}(\alpha) \cap G = \varnothing \text{ for some } (F, G) \in \Omega\}$.

- A *Muller objective* is given by a family $\mathcal{F}$ of *accepting sets* $F \subseteq C$, and it requires that the set of colours seen infinitely often equals one of these accepting sets. The corresponding subset of C^ω is $\mathrm{Muller}(\mathcal{F}) = \{\alpha \in C^\omega : \mathrm{Inf}(\alpha) \in \mathcal{F}\}$.

Parity, Streett, Rabin and Muller objectives are of particular relevance because they provide a standard form for arbitrary ω-regular objectives; any game with arbitrary ω-regular objectives can be reduced to one with parity, Streett, Rabin or Muller objectives (over a larger arena) by taking the product of its original arena with a suitable deterministic word automaton for each player's objective (see Thomas 1990).

In this work, for reasons that will become clear later, we are particularly attracted to objectives that are invariant under adding and removing finite prefixes; we call such objectives *prefix-independent*. More formally, an objective is prefix-independent if for each $\alpha \in C^\omega$ and $x \in C^*$ the sequence α satisfies the objective if and only if the sequence $x \cdot \alpha$ does. Note that, if $\mathrm{Win} \subseteq C^\omega$ is a prefix-independent objective over colours, then the corresponding objective $\chi^{-1}(\mathrm{Win})$ over vertices is also prefix-independent.

Of the objectives listed above, only reachability and safety objectives are, in general, not prefix-independent. However, many of our results (in particular, many of the lower bounds we prove) apply to games with a prefix-independent form of reachability, which we call *simple reachability*. For such an objective, we assume that each vertex is coloured by itself, i.e. $C = V$, and χ is the identity mapping. The simple reachability objective for a set $F \subseteq V$ coincides with the reachability objective for F, but we require that each $v \in F$ is a *terminal vertex*: $v\Delta = \{v\}$. For any such set F, we have $\pi(k) \in F$ for some $k \in \mathbb{N}$ if and only if $\mathrm{Inf}(\pi) \cap F \neq \varnothing$ (or equivalently, $\mathrm{Inf}(\pi) \subseteq F$). Hence, simple reachability objectives are prefix-independent.
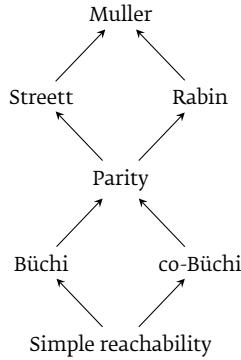
Figure 2.1. A hierarchy of prefix-independent objectives.

For S2Gs, the distinction between reachability and simple reachability is not important: every S2G with a reachability objective can easily be transformed into an equivalent S2G with a simple reachability objective. For SMGs, we believe that any such transformation requires exponential time: Deciding whether in a deterministic game with simple reachability objectives there exists a play that fulfils each of the objectives can be done in polynomial time, whereas the same problem is NP-complete for deterministic games with arbitrary reachability objectives (see Ummels 2005).

The resulting hierarchy of objectives is depicted in Figure 2.1. As explained above, a simple reachability objective can be viewed as a (co-)Büchi objective. Any (co-)Büchi objective is equivalent to a parity objective with only two priorities, and any parity objective is equivalent to both a Streett and a Rabin objective; in fact, the intersection (union) of two parity objectives is equivalent to a Streett (Rabin) objective. Moreover, any Streett or Rabin objective is equivalent to a Muller objective, although the translation from a set of Streett/Rabin pairs to an equivalent family of accepting sets is, in general, exponential. Finally, the complement of a Büchi (Streett) objective is equivalent to a co-Büchi (Rabin) objective, and vice versa, whereas the complement of a parity (Muller) objective is also a parity (Muller) objective. In fact, any objective that is equivalent to both a Streett and a Rabin objective is equivalent to a parity objective (Zielonka 1998).

To denote the class of SMGs (S2Gs) with a certain type of objectives, we prefix the name SMG (S2G) with the names of the objectives; for instance, we use the term *Streett-Rabin SMG* to denote SMGs where each player has a Streett or

a Rabin objective. For S2Gs, we adopt the convention to name the objective of player 0 first; hence, in a Streett-Rabin S2G player 0 has a Streett objective, while player 1 has a Rabin objective. Inspired by Condon (1992), we will refer to SMGs with simple reachability objectives and S2Gs with a (simple) reachability objective for player 0 as *simple stochastic multiplayer games* (SSMGs) and *simple stochastic two-player zero-sum games* (SS2Gs), respectively.

2.2 Strategies and strategy profiles

Randomised and pure strategies

The notion of a strategy lies at the heart of game theory. Formally, a (*randomised*) *strategy* of player i in an SMG $\mathcal{G}$ is a mapping $\sigma\colon V^*V_i \to \mathcal{D}(V)$ assigning to each possible sequence $xv \in V^*V_i$ of vertices ending in a vertex controlled by player i a (discrete) probability distribution over V such that $\sigma(xv)(w) > 0$ only if $(v, \bot, w) \in \Delta$. Instead of $\sigma(xv)(w)$, we usually write $\sigma(w \mid xv)$. We say that a play π of $\mathcal{G}$ is *compatible* with a strategy σ of player i if $\sigma(\pi(k+1) \mid \pi(0)\dots\pi(k)) > 0$ for all $k \in \mathbb{N}$ with $\pi(k) \in V_i$. Similarly, a history $x = v_0\dots v_n$ is compatible with σ if $\sigma(v_{k+1} \mid v_0\dots v_k) > 0$ for all $0 \le k < n$.

A (*randomised*) *strategy profile* of $\mathcal{G}$ is a tuple $\bar{\sigma} = (\sigma_i)_{i\in\Pi}$ where σ_i is a strategy of player i in $\mathcal{G}$. We say that a play or a history of $\mathcal{G}$ is compatible with a strategy profile $\bar{\sigma}$ if it is compatible with each σ_i. Given a strategy profile $\bar{\sigma} = (\sigma_j)_{j\in\Pi}$ and a strategy τ of player i, we denote by $(\bar{\sigma}_{-i}, \tau)$ the strategy profile obtained from $\bar{\sigma}$ by replacing σ_i with τ.

A strategy σ of player i is called *pure* or *deterministic* if for each $xv \in V^*V_i$ there exists $w \in v\Delta$ with $\sigma(w \mid xv) = 1$; note that a pure strategy of player i can be identified with a function $\sigma\colon V^*V_i \to V$. A strategy profile $\bar{\sigma} = (\sigma_i)_{i\in\Pi}$ is called *pure* (or *deterministic*) if each σ_i is pure.

The probability measure induced by a strategy profile

Given a game $\mathcal{G}$ and a strategy profile $\bar{\sigma} = (\sigma_i)_{i\in\Pi}$ of $\mathcal{G}$, the *conditional probability* of $w \in V$ given $xv \in V^*V$ is the number $\sigma_i(w \mid xv)$ if $v \in V_i$ and the unique $p \in [0,1]$ such that $(v, p, w) \in \Delta$ if v is a stochastic vertex; let us denote this probability by $\bar{\sigma}(w \mid xv)$. Given an initial vertex $v_0 \in V$, the probabilities $\bar{\sigma}(w \mid xv)$ give rise to a probability measure: the probability of a basic cylinder set $v_0\dots v_k \cdot V^\omega$ equals the product $\prod_{j=1}^{k} \bar{\sigma}(v_j \mid v_0\dots v_{j-1})$; basic cylinder sets that start in a vertex different from v_0 have probability 0. This definition
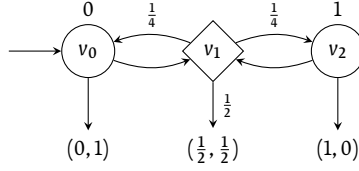
Figure 2.2. An example of a two-player SSMG.

induces a probability measure on the algebra of cylinder sets, which—by Carathéodory's extension theorem (Theorem A.5)—can be extended to a probability measure on the Borel σ-algebra over V^ω; we denote the extended measure by $\mathrm{Pr}_{v_0}^{\bar{\sigma}}$. Finally, by viewing the colouring function $\chi\colon V \to C$ as a continuous function $V^\omega \to C^\omega$, we obtain a probability measure on the Borel σ-algebra over C^ω; we abuse notation and denote this measure also by $\mathrm{Pr}_{v_0}^{\bar{\sigma}}$.

For a strategy profile $\bar{\sigma}$, we are mainly interested in the probabilities $p_i \coloneqq \mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Win}_i)$ of winning. We call p_i the (*expected*) *payoff* of $\bar{\sigma}$ for player i (from v_0) and the vector $(p_i)_{i\in\Pi}$ the (expected) payoff of $\bar{\sigma}$ (from v_0). Note that, if $\bar{\sigma}$ is a pure strategy profile of a deterministic game, then its payoff is just the payoff of the unique play π of $(\mathcal{G}, v_0)$ that is compatible with each σ_i. Finally, we say that a history xv of $(\mathcal{G}, v_0)$ is *consistent* with $\bar{\sigma}$ if $\mathrm{Pr}_{v_0}^{\bar{\sigma}}(xv \cdot V^\omega) > 0$, i.e. if the basic cylinder induced by this history has positive probability. Note that each history that is consistent with $\bar{\sigma}$ is also compatible with $\bar{\sigma}$.

Example 2.1. Let $\mathcal{G}$ be the SSMG depicted in Figure 2.2 according to the following conventions, to which we adhere throughout this work: Vertices controlled by players are drawn as circles, where the player who controls a vertex is given by the label next to it. Stochastic vertices are drawn as diamonds, and transition probabilities are given by labels on edges (the default being $\frac{1}{2}$). If there is a designated initial vertex, it is marked by a dangling incoming edge. Finally, terminal vertices are generally depicted by their associated payoff vector. As syntactic sugar, we allow arbitrary vectors of rational probabilities as payoffs; this does not increase the power of the model since such a payoff vector can easily be realised by an SSMG consisting exclusively of stochastic and terminal vertices.

Now consider the strategy profile $\bar{\sigma}$ defined by $\bar{\sigma}(v_1 \mid xv_0) = \bar{\sigma}(v_1 \mid xv_2) = 1$ for each $x \in V^*$. Starting from the initial vertex v_0 of $\mathcal{G}$, the payoff of this strategy profile is $(\frac{1}{2}, \frac{1}{2})$ because the probability of reaching the terminal vertex that has this payoff equals 1.

In order to apply known results about Markov chains, we can also view the stochastic process induced by a strategy profile $\bar{\sigma}$ as a countable Markov chain $\mathcal{G}^{\bar{\sigma}}$, defined as follows: The set of states of $\mathcal{G}^{\bar{\sigma}}$ is the set V^+ of all nonempty sequences of vertices in $\mathcal{G}$. The only transitions from a state xv, where $x \in V^*$ and $v \in V$, are to states of the form xvw, where $w \in V$, and such a transition occurs with probability $p > 0$ if and only if either v is stochastic and $(v, p, w) \in \Delta$, or $v \in V_i$ and $\sigma_i(w \mid xv) = p$. Finally, the colouring χ of vertices is extended to a colouring of states by setting $\chi(xv) = \chi(v)$ for all $x \in V^*$ and $v \in V$. With this definition, we can recover the payoff of $\bar{\sigma}$ for player i as the probability of the event $\chi^{-1}(\mathrm{Win}_i)$ in $(\mathcal{G}^{\bar{\sigma}}, v_0)$.

For each player i, the Markov decision process $\mathcal{G}^{\bar{\sigma}_{-i}}$ is defined just as $\mathcal{G}^{\bar{\sigma}}$, but states $xv \in V^*V_i$ are controlled by player i (the sole player in $\mathcal{G}^{\bar{\sigma}_{-i}}$), and there is a transition from such a state to any state of the form xvw, where $w \in V$, such that $(v, \bot, w) \in \Delta$; player i's objective is the same as in $\mathcal{G}$.

Strategies with memory

A *memory structure* for a game $\mathcal{G}$ with vertices in V is a triple $\mathfrak{M} = (M, \delta, m_0)$, where M is a set of *memory states*, $\delta \colon M \times V \to M$ is the *update function*, and $m_0 \in M$ is the *initial memory*. A *(randomised) strategy with memory* $\mathfrak{M}$ of player i is a function $\sigma \colon M \times V_i \to \mathcal{D}(V)$ such that $\sigma(m, v)(w) > 0$ only if $w \in vE$. The strategy σ is a *pure strategy with memory* $\mathfrak{M}$ if additionally the following property holds: for all $m \in M$ and $v \in V$ there exists $w \in V$ such that $\sigma(m, v)(w) = 1$. Hence, a pure strategy with memory $\mathfrak{M}$ can be described by a function $\sigma \colon M \times V_i \to V$. Finally, a *(pure) strategy profile with memory* $\mathfrak{M}$ is a tuple $\bar{\sigma} = (\sigma_i)_{i \in \Pi}$ such that each σ_i is a (pure) strategy with memory $\mathfrak{M}$ of player i.

A (pure) strategy σ with memory $\mathfrak{M}$ of player i defines a (pure) strategy of player i in the usual sense as follows: Let $\delta^*(x)$ be the memory state after $x \in V^*$, defined inductively by $\delta^*(\varepsilon) = m_0$ and $\delta^*(xv) = \delta(\delta^*(x), v)$ for $x \in V^*$ and $v \in V$. If $v \in V_i$, then the distribution (successor vertex) chosen by the strategy σ for the sequence xv is $\sigma(\delta^*(x), v)$. Vice versa, every strategy (profile) of $\mathcal{G}$ can be viewed as a strategy (profile) with memory $\mathfrak{M} := (V^*, \cdot, \varepsilon)$.

A *finite-state strategy (profile)* is a strategy (profile) with memory $\mathfrak{M}$ for a finite memory structure $\mathfrak{M}$. Note that a strategy profile is finite-state if and only if each of its strategies is finite-state. If $|M| = 1$, we call a strategy (profile) with memory $\mathfrak{M}$ *stationary*. Moreover, we call a strategy (profile) that is both pure and stationary a *positional* strategy (profile). A stationary strategy of

Table 2.1. Types of strategies in stochastic games.

	Pure	Randomised
Stationary	$V_i \to V$	$V_i \to \mathcal{D}(V)$
With memory $\mathfrak{M}$	$M \times V_i \to V$	$M \times V_i \to \mathcal{D}(V)$
General	$V^* V_i \to V$	$V^* V_i \to \mathcal{D}(V)$

player i can be described by a function $\sigma \colon V_i \to \mathcal{D}(V)$, and a positional strategy by a function $\sigma \colon V_i \to V$.

If $\bar{\sigma} = (\sigma_i)_{i \in \Pi}$ is a strategy profile with memory $\mathfrak{M}$, we modify the Markov chain $\mathcal{G}^{\bar{\sigma}}$ by taking $M \times V$ as its domain. The transition relation is defined as follows: there is a transition from (m, v) to (n, w) with probability $p > 0$ if and only if $\delta(m, v) = n$ and either v is a stochastic vertex of $\mathcal{G}$ and $(v, p, w) \in \Delta$ or $v \in V_i$ and $\sigma_i(m, v)(w) = p$. Finally, a state (m, v) has the same colour as the vertex v in $\mathcal{G}$. Analogously, we modify the Markov decision process $\mathcal{G}^{\bar{\sigma}_{-i}}$ by using $M \times V$ as its domain: vertices $(m, v) \in M \times V_i$ are controlled by player i, and there is a transition from such a vertex (m, v) to $(n, w) \in M \times V$ if and only if $n = \delta(m, v)$ and $(v, \bot, w) \in \Delta$. Note that the arenas of both $\mathcal{G}^{\bar{\sigma}}$ and $\mathcal{G}^{\bar{\sigma}_{-i}}$ are finite if the memory $\mathfrak{M}$ and the original arena of $\mathcal{G}$ are finite.

All the types of strategies we consider in this work and their representations are summarised in Table 2.1.

Residual games and strategies

Given an SMG $\mathcal{G}$ and a sequence $x \in V^*$ (which is usually a history), the *residual game* $\mathcal{G}[x]$ has the same arena as $\mathcal{G}$ but different objectives: if the objective of player i in $\mathcal{G}$ is $\mathrm{Win}_i \subseteq C^\omega$, then her objective in $\mathcal{G}[x]$ is given by the set $\chi(x)^{-1}\mathrm{Win}_i = \{\alpha \in C^\omega : \chi(x) \cdot \alpha \in \mathrm{Win}_i\}$. In particular, if all objectives in $\mathcal{G}$ are prefix-independent, then $\mathcal{G}[x] = \mathcal{G}$.

If player i plays according to a strategy σ in $\mathcal{G}$, then the natural choice for her strategy in $\mathcal{G}[x]$ is the *residual strategy* $\sigma[x]$, defined by $\sigma[x](yv) = \sigma(xyv)$. If $\bar{\sigma} = (\sigma_i)_{i \in \Pi}$ is a strategy profile, then the *residual strategy profile* $\bar{\sigma}[x]$ is just the profile of the residual strategies $\sigma_i[x]$. The following lemma, taken from (Zielonka 2004), shows how to compute probabilities with respect to a residual strategy profile.

Lemma 2.2. Let $\bar{\sigma}$ be a strategy profile of an SMG $(\mathcal{G}, v_0)$, and let $xv \in V^* V$. If $X \subseteq V^\omega$ is a Borel set, then $\mathrm{Pr}^{\bar{\sigma}}_{v_0}(X \cap xv \cdot V^\omega) = \mathrm{Pr}^{\bar{\sigma}}_{v_0}(xv \cdot V^\omega) \cdot \mathrm{Pr}^{\bar{\sigma}[x]}_v(x^{-1}X)$.

2.3 Subarenas and end components

Algorithms for stochastic games often employ a *divide-and-conquer* approach and compute a solution for a complex game from the solutions of several smaller games. These smaller games are usually obtained from the original game by restricting to a *subarena*. Formally, given an SMG $\mathcal{G}$, a set $U \subseteq V$ is a subarena if:

- $U \neq \emptyset$,
- $v\Delta \cap U \neq \emptyset$ for each $v \in U$, and
- $v\Delta \subseteq U$ for each stochastic vertex $v \in U$.

Clearly, if U is a subarena, then the restriction of $\mathcal{G}$ to vertices in U is again an SMG, which we denote by $\mathcal{G} \upharpoonright U$. Formally,

$$\mathcal{G} \upharpoonright U := (\Pi, U, (V_i \cap U)_{i \in \Pi}, \Delta \cap (U \times ([0,1] \cup \{\bot\}) \times U), \chi_U, (\mathrm{Win}_i)_{i \in \Pi}),$$

where $\chi_U \colon U \to C \colon u \mapsto \chi(u)$ is the restriction of the colouring function to U.

Of particular interest are the strongly connected subarenas of a game because they can arise as the sets $\mathrm{Inf}(\pi)$ of vertices visited infinitely often in a play; we call these sets *end components*. Formally, $\emptyset \neq U \subseteq V$ is an end component if U is a subarena and every vertex $w \in U$ is reachable from every other vertex $v \in U$ (i.e. there exists a sequence $v = v_1, v_2, \ldots, v_n = w$ such that $v_{i+1} \in v_i\Delta$ for each $0 < i < n$). An end component U is *maximal* in a set $S \subseteq V$ if there is no end component U' such that $U \subsetneq U' \subseteq S$. For any finite subset $S \subseteq V$, the set of all end components maximal in S can be computed in quadratic time (see Appendix B for the algorithm).

The theory of end components has been developed by de Alfaro (1997, 1998) and Courcoubetis & Yannakakis (1995, 1998). The central fact about end components in finite SMGs is that, under any strategy profile, the set of vertices visited infinitely often is almost surely an end component (cf. Lemma B.11).

Lemma 2.3. Let $\mathcal{G}$ be a finite SMG, and let $\bar{\sigma}$ be a strategy profile of $\mathcal{G}$. Then $\mathrm{Pr}_v^{\bar{\sigma}}(\{\pi \in V^\omega : \mathrm{Inf}(\pi) \text{ is an end component}\}) = 1$ for each vertex $v \in V$.

Moreover, for any end component U, we can construct a stationary strategy profile, or alternatively a pure finite-state strategy profile, that when started in U guarantees almost surely to visit all (and only) vertices in U infinitely often (cf. Lemma B.12).

Lemma 2.4. Let $\mathcal{G}$ be a finite SMG, and let U be an end component of $\mathcal{G}$. There exists both a stationary and a pure finite-state strategy profile $\bar{\sigma}$ of $\mathcal{G}$ such that $\Pr_v^{\bar{\sigma}}(\{\pi \in V^\omega : \mathrm{Inf}(\pi) = U\}) = 1$ for every vertex $v \in U$.

Given an SMG $\mathcal{G}$ with objectives representable as Muller objectives given by the family $\mathcal{F}_i$ of accepting sets, we say that an end component U is *winning* for player i if $\chi(U) \in \mathcal{F}_i$; the *payoff* of U is the vector $\bar{x} \in \{0,1\}^\Pi$, defined by $x_i = 1$ if and only if U is winning for player i.

2.4 Values, determinacy and optimal strategies

The notions of the value and an *optimal strategy* are central for the theory of two-player zero-sum games. However, they can also be applied to SMGs.

Given a strategy τ of player i in $\mathcal{G}$ and a vertex $v \in V$, the *value* of τ from v is the number $\mathrm{val}^\tau(v) := \inf_{\bar{\sigma}} \Pr_v^{\bar{\sigma}_{-i},\tau}(\mathrm{Win}_i)$, where $\bar{\sigma}$ ranges over all strategy profiles of $\mathcal{G}$. Moreover, the *value* of $\mathcal{G}$ for player i from v is the supremum of these values: $\mathrm{val}_i^{\mathcal{G}}(v) := \sup_\tau \mathrm{val}^\tau(v)$, where τ ranges over all strategies of player i in $\mathcal{G}$. Intuitively, $\mathrm{val}_i^{\mathcal{G}}(v)$ is the maximal payoff that player i can ensure when the game starts from v.

Given an initial vertex $v_0 \in V$, a strategy τ of player i in $\mathcal{G}$ is called (*almost-surely*) *winning* if $\mathrm{val}^\tau(v_0) = 1$. More generally, τ is called *optimal* if $\mathrm{val}^\tau(v_0) = \mathrm{val}_i^{\mathcal{G}}(v_0)$. For $\varepsilon > 0$, it is called ε-*optimal* if $\mathrm{val}^\tau(v_0) \geq \mathrm{val}_i^{\mathcal{G}}(v_0) - \varepsilon$. A *globally* ($\varepsilon$-)*optimal* strategy is a strategy that is (ε-)optimal for every possible initial vertex $v_0 \in V$. Note that optimal strategies do not need to exist since the supremum in the definition of $\mathrm{val}_i^{\mathcal{G}}$ is not necessarily attained; in this case, only ε-optimal strategies do exist. However, if for every possible initial vertex there exists an (ε-)optimal strategy, then there also exists a globally (ε-) optimal strategy.

Before we state the most important result on stochastic two-player zero-sum games, we define two other notions of optimality, which will be useful for proving the existence of certain equilibria in the next chapter: We say that a strategy τ of player i in $(\mathcal{G}, v_0)$ is *residually optimal* if the residual strategy $\tau[x]$ is optimal in the residual game $(\mathcal{G}[x], v)$ for every history xv of $(\mathcal{G}, v_0)$. More generally, τ is *strongly optimal* if $\tau[x]$ is optimal in $(\mathcal{G}[x], v)$ for every history xv of $(\mathcal{G}, v_0)$ that is compatible with τ. Note that a positional strategy profile that is globally optimal is also residually optimal. Apart from being relevant for the existence of equilibria, strongly and residually optimal strategies

have been considered as *best-effort* strategies in two-player zero-sum games (Faella 2009).

Determining values and finding optimal strategies in SMGs actually reduces to performing the same tasks in S2Gs. Formally, given an SMG $\mathcal{G}$, define for each player i the *coalition game* $\mathcal{G}_i$ to be the same game as $\mathcal{G}$ but with only two players: player i acting as player 0 and the coalition player $\Pi \setminus \{i\}$ acting as player 1. The coalition controls all vertices that in $\mathcal{G}$ are controlled by some player $j \neq i$, and its objective is the complement of player i's objective in $\mathcal{G}$. Clearly, $\mathcal{G}_i$ is an S2G, and $\mathrm{val}^{\mathcal{G}_i}(v) = \mathrm{val}_i^{\mathcal{G}}(v)$ for every vertex v. Moreover, any (residually, strongly, ε-) optimal strategy for player i in $(\mathcal{G}, v_0)$ is (residually, strongly, ε-) optimal in $(\mathcal{G}_i, v_0)$, and vice versa. Hence, when we study values and optimal strategies, we can restrict our investigation to S2Gs.

A celebrated theorem due to Martin (1998) and Maitra & Sudderth (1998) (see also Maitra & Sudderth 2003) states that S2Gs with Borel objectives are *determined*: $\mathrm{val}_0^{\mathcal{G}} = 1 - \mathrm{val}_1^{\mathcal{G}}$ (where the equality holds pointwise).[1] The number $\mathrm{val}^{\mathcal{G}}(v) := \mathrm{val}_0^{\mathcal{G}}(v)$ is consequently called the *value* of $\mathcal{G}$ from v. In fact, an inspection of the proof shows that—for the kind of games we study in this work—both players not only have randomised ε-optimal strategies but pure ε-optimal strategies.

Theorem 2.5 (Martin; Maitra & Sudderth). Every S2G with Borel objectives is determined; for all $\varepsilon > 0$, both players have ε-optimal pure strategies.

For finite S2Gs with prefix-independent objectives, we can show a stronger result than Theorem 2.5: in these games, both players not only have ε-optimal pure strategies but optimal ones (Gimbert & Horn 2010). In fact, their proof reveals not only the existence of optimal strategies but the existence of residually optimal strategies; for an alternative proof of the following theorem, see Section 2.6.

Theorem 2.6 (Gimbert & Horn). There exist residually optimal pure strategies in every finite S2G with prefix-independent objectives.

As witnessed by the following two examples, Theorem 2.6 fails if either the objective is not prefix-independent or the arena is not finite, even if there is only one player.

---

[1] Martin proved the theorem originally for *Blackwell games*; Maitra & Sudderth adapted his proof to stochastic games.
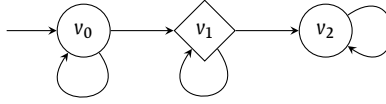
Figure 2.3. An MDP with no optimal strategy.
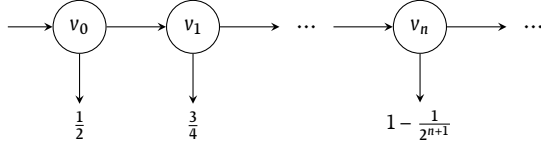
Figure 2.4. Another MDP with no optimal strategy.

Example 2.7. Consider the MDP $\mathcal{G}$ depicted in Figure 2.3 where player 0 wins if the number of visits to vertex v_0 is finite but strictly greater than the number of visits to vertex v_1. We claim that $(\mathcal{G}, v_0)$ does not admit an optimal strategy. First, for each $n \in \mathbb{N}$, consider the pure strategy σ_n of moving from v_0 to v_1 after completing precisely n loops around v_0. Clearly, $\Pr_{v_0}^{\sigma_n}(\mathrm{Win}) = 1 - 1/2^n$, and therefore $\mathrm{val}^{\mathcal{G}}(v_0) = 1$. However, no strategy τ achieves this value: if $\tau(v_0 \,|\, (v_0)^{n+1}) = 1$ for all $n \in \mathbb{N}$, then obviously $\Pr_{v_0}^{\tau}(\mathrm{Win}) = 0$; otherwise, consider the least $n \in \mathbb{N}$ such that $p := \tau(v_1 \,|\, (v_0)^{n+1}) > 0$; we have $\Pr_{v_0}^{\tau}(\mathrm{Win}) \leq 1 - p/2^n < 1$.

Example 2.8. Consider the MDP $\mathcal{G}$ depicted in Figure 2.4; every play that visits each vertex v_i is losing. Again, we claim that $(\mathcal{G}, v_0)$ does not admit an optimal strategy. First, for each $n \in \mathbb{N}$, consider the positional strategy σ_n of "leaving the game" at vertex v_n. Clearly, $\Pr_{v_0}^{\sigma_n}(\mathrm{Win}) = 1 - 1/2^{n+1}$, and therefore $\mathrm{val}^{\mathcal{G}}(v_0) = 1$. But again, no strategy τ achieves this value: if $\tau(v_{n+1} \,|\, v_0 \ldots v_n) = 1$ for all $n \in \mathbb{N}$, then $\Pr_{v_0}^{\tau}(\mathrm{Win}) = 0$; otherwise, consider the least $n \in \mathbb{N}$ such that $p := 1 - \tau(v_{n+1} \,|\, v_0 \ldots v_n) > 0$; then $\Pr_{v_0}^{\tau}(\mathrm{Win}) \leq 1 - p/2^{n+1} < 1$.

For deterministic two-player zero-sum games with Borel objectives, every value is either 0 or 1, and every ε-optimal strategy is already optimal. In particular, from every vertex either one of the two players has a winning strategy. This follows easily from Theorem 2.5 because any pure strategy profile of a deterministic game gives payoff 0 or 1 to each player. The determinacy of deterministic two-player zero-sum games was proven earlier than the corresponding result for stochastic games, also by Martin (1975, 1985).

In fact, the proof of Theorem 2.5 relies on the determinacy of deterministic two-player zero-sum games.

Theorem 2.9 (Martin). Every deterministic two-player zero-sum game with Borel objectives is determined. From each vertex, either player 0 or player 1 has a pure winning strategy.

In fact, in every deterministic two-player zero-sum game with Borel objectives there exists a pair of residually optimal pure strategies, i.e. a pair (σ, τ) of pure strategies such that, for each history xv of the game, either $\sigma[x]$ or $\tau[x]$ is winning in the residual game $(\mathcal{G}[x], v)$.

Corollary 2.10. There exist residually optimal pure strategies in any deterministic two-player zero-sum game with Borel objectives.

Proof. Let $(\mathcal{G}, v_0)$ be a deterministic two-player zero-sum game with a Borel objective Win $\subseteq V^\omega$ for player 0. Since the class of Borel sets is closed under complementation, it suffices to show that player 0 has a residually optimal pure strategy. With Win, the set x^{-1}Win is Borel for each $x \in V^*$. Hence, by Theorem 2.9, for each history xv of $(\mathcal{G}, v_0)$, we can fix a pure strategy σ^x of player 0 that is optimal in the residual game $(\mathcal{G}[x], v)$; note that we can assume that σ^x is independent of v. We have to combine these strategies in an appropriate way to a residually optimal strategy σ. (Let us point out that the trivial combination, namely $\sigma(xv) := \sigma^x(v)$, does not work, in general.)

We say that a decomposition $x = x_1 \cdot x_2$ is *good* with respect to vertex v if $\sigma^{x_1}[x_2]$ is winning in $(\mathcal{G}[x], v)$. If the strategy σ^x is winning in $(\mathcal{G}[x], v)$, then the decomposition $x = x \cdot \varepsilon$ is good with respect to v; so, a good decomposition exists in this case. For each history xv, if σ^x is winning in $(\mathcal{G}[x], v)$, we choose the good (with respect to vertex v) decomposition $x = x_1 \cdot x_2$ with minimal x_1, and set $\sigma(xv) := \sigma^{x_1}(x_2 v)$; otherwise, we set $\sigma(xv) := \sigma^x(v)$.

To show that σ is residually optimal, it suffices to show that, for each history xv of $(\mathcal{G}, v_0)$, the strategy $\sigma[x]$ is winning in $(\mathcal{G}[x], v)$ whenever the strategy σ^x is. Hence, assume that σ^x is winning in $(\mathcal{G}[x], v)$, and let π be a play starting in $\pi(0) = v$ that is compatible with $\sigma[x]$. We need to show that $\pi \in x^{-1}$Win.

We claim that for each $k \in \mathbb{N}$ there exists a decomposition of the form $x \cdot \pi|_k = x_1 \cdot (x_2 \cdot \pi|_k)$ that is good with respect to $\pi(k)$. For $k = 0$, this is obviously true. For $k > 0$, assume that there exists a decomposition $x \cdot \pi|_{k-1} = x_1 \cdot (x_2 \cdot \pi|_{k-1})$ that is good with respect to $\pi(k-1)$, and consider the one where

x_1 is minimal. Then $\pi(k) = \sigma(x \cdot \pi|_k) = \sigma^{x_1}(x_2 \cdot \pi|_k)$, and $x \cdot \pi|_k = x_1 \cdot (x_2 \cdot \pi|_k)$ is a good decomposition with respect to $\pi(k)$.

Now, consider the sequence $x_1^0, x_1^1, \ldots$ of prefixes of the decompositions $x \cdot \pi|_k = x_1^k \cdot (x_2^k \cdot \pi|_k)$ that are good with respect to $\pi(k)$ and where x_1^k is minimal. We have $x_1^0 \succeq x_1^1 \succeq \cdots$ because for each $k > 0$ the decomposition $x \cdot \pi_k = x_1^{k-1} \cdot (x_2^{k-1} \cdot \pi_k)$ is also good with respect to $\pi(k)$. Since $\prec$ is well-founded, there must exist $k \in \mathbb{N}$ such that $x_1^k = x_1^j$ and $x_2^k = x_2^j$ for each $j \geq k$. But then the play $\pi(k)\pi(k+1)\ldots$ is compatible with $\sigma^{x_1^k}[x_2^k \cdot \pi|_k]$, which is winning in $(\mathcal{G}[x \cdot \pi|_k], \pi(k))$. Hence, $\pi(k)\pi(k+1)\ldots \in (x \cdot \pi|_k)^{-1}\mathrm{Win}$ and $\pi \in x^{-1}\mathrm{Win}$. $\qquad\square$

For deterministic games, the payoff of a strategy profile is well-defined even if the game has non-Borel objectives. Does Theorem 2.9 hold for such games as well? Unfortunately, the answer is negative: Gale & Stewart (1953) gave an example of a deterministic two-player zero-sum game with a non-Borel objective where none of the two players has a pure winning strategy.

For finite S2Gs with ω-regular objectives, more attractive strategies than arbitrary pure strategies suffice for optimality. In particular, in any finite Rabin-Streett S2G there exists a globally optimal positional strategy for player 0 (Klarlund 1994; Chatterjee et al. 2005).

Theorem 2.11 (Klarlund; Chatterjee et al.). In any finite Rabin-Streett S2G, player 0 has a globally optimal positional strategy.

It follows from Theorem 2.11 that the values of a finite Rabin-Streett S2G are rational of polynomial bit complexity in the size of the arena: Given a positional strategy profile $\bar{\sigma}$ of $\mathcal{G}$, the finite MDP $\mathcal{G}^{\bar{\sigma}_{-1}}$ is not larger than the game $\mathcal{G}$. Moreover, if σ_0 is globally optimal, then for every vertex v the value of $\mathcal{G}$ from v and the value of $\mathcal{G}^{\bar{\sigma}_{-1}}$ from v sum up to 1. But the values of a Streett MDP form the optimal solution of a linear programme of polynomial size (see Appendix B) and are therefore rational of polynomial bit complexity.

Of course, it also follows from Theorem 2.11 that finite parity S2Gs are *positionally determined*: both players have globally optimal positional strategies. This result was first proven for deterministic two-player zero-sum parity games (even over infinite arenas) independently by Emerson & Jutla (1991) and Mostowski (1991). For SS2Gs, the existence of optimal positional strategies follows from a more general result of Liggett & Lippman (1969). Independently, McIver & Morgan (2002), Chatterjee et al. (2004a) and Zielonka (2004) extended these results to parity S2Gs.

Corollary 2.12. In any finite parity S2G, both players have globally optimal positional strategies.

Since every finite S2G with ω-regular objectives can be reduced to one with parity objectives, we can conclude from Corollary 2.12 that both players have residually optimal pure finite-state strategies in finite S2Gs with arbitrary ω-regular objectives.

Corollary 2.13. In any finite S2G with ω-regular objectives, both players have residually optimal pure finite-state strategies.

Corollary 2.13 generalises the well-known theorem by Büchi & Landweber (1969) that both players have optimal pure finite-state strategies in every deterministic two-player zero-sum game with ω-regular objectives.

2.5 Algorithmic problems

Throughout this section, we only consider finite two-player zero-sum games. The main computational problems for these games are computing the value and optimal strategies for one or both players. Instead of computing the value exactly, we can ask whether the value is greater than some given rational probability p, a problem which we call the *quantitative decision problem*:

Given an S2G $\mathcal{G}$, a vertex v and a rational number $p \in [0, 1]$, decide whether $\mathrm{val}^{\mathcal{G}}(v) \geq p$.

In many cases, it suffices to know whether the value is 1, i.e. whether player 0 has a strategy to win the game almost surely (asymptotically, at least). We call the resulting decision problem the *qualitative decision problem*.

Clearly, if we can solve the quantitative decision problem, we can approximate the values $\mathrm{val}^{\mathcal{G}}(v)$ up to any desired precision by using binary search. In fact, for parity S2Gs it turns out that it suffices to solve the decision problems, since the other problems (computing the values and optimal strategies) are polynomial-time equivalent to the quantitative decision problem.

Proposition 2.14. Either none or all of the following problems are solvable in polynomial time:

1. the quantitative decision problem for parity S2Gs,
2. computing the values $\mathrm{val}^{\mathcal{G}}(v)$ of a parity S2G,
3. computing globally optimal positional strategies in a parity S2G.

Proof. (1. ⇒ 2.) Assume that we have a polynomial-time algorithm for the quantitative decision problem. Since the values of a finite parity S2G are always rational of bit complexity polynomial in the size of the game, binary search for the value $\mathrm{val}^{\mathcal{G}}(v)$ terminates after polynomially many steps with the exact value of $\mathrm{val}^{\mathcal{G}}(v)$.

(2. ⇒ 3.) Assume that there exists a polynomial-time algorithm for computing the values. Then, given a parity S2G $\mathcal{G}$, we can find a globally optimal positional strategy for player 0 by the following procedure: In the case that every vertex controlled by player 0 has only one outgoing transition, we are done. Otherwise, let $v \in V_0$ be a vertex with $|v\Delta| > 1$. Since there exists an optimal positional strategy, there must exist a transition $(v, \bot, w) \in \Delta$ such that the values of the game do not change when all other transitions $(v, \bot, w')$ are removed. Using the polynomial-time algorithm for computing the values, we can find such a transition. Now, we can iterate the procedure on the (smaller) game that is obtained from $\mathcal{G}$ by removing all other transitions that originate in v.

(3. ⇒ 1.) Assume that there exists a polynomial-time algorithm for computing globally optimal positional strategies. To determine $\mathrm{val}^{\mathcal{G}}(v)$, we can then compute a pair (σ, τ) of such strategies, one for each player, and construct the Markov chain $\mathcal{G}^{(\sigma,\tau)}$. The value $\mathrm{val}^{\mathcal{G}}(v)$ equals the probability of reaching from v a bottom SCC of $\mathcal{G}^{(\sigma,\tau)}$ in which the least priority is even. By solving a system of linear equations, we can easily compute this probability and check whether it is greater than p (see Appendix B). □

For a Markov decision process whose objective can be represented as a Muller objective, we can compute the values by an analysis of its end components: for a given initial vertex v, the value of the MDP from v is the maximal probability of reaching a winning end component from v. Once all vertices that reside in winning end components have been identified, these probabilities can be computed in polynomial time via linear programming.

For MDPs with Rabin or Muller objectives, it is easy to see that the union of all winning end components can be computed in polynomial time (see Appendix B); for MDPs with Streett objectives, Chatterjee et al. (2005) gave a polynomial-time algorithm for computing this set. Hence, for MDPs with any of these objectives, the quantitative decision problem is solvable in polynomial time.

Theorem 2.15 (de Alfaro; Chatterjee et al.). The quantitative decision problem is in P for Streett, Rabin or Muller MDPs.

It follows from Theorems 2.11 and 2.15 that the quantitative decision problem for Rabin-Streett S2Gs is in NP: to decide whether $\mathrm{val}^{\mathcal{G}}(v) \geq p$, it suffices to guess a positional strategy for player 0 and to check whether in the resulting Streett MDP the value from v is $\geq p$. By determinacy, this result implies that the quantitative decision problem is in coNP for Streett-Rabin S2Gs and in NP ∩ coNP for parity S2Gs.

Corollary 2.16. The quantitative decision problem is

- in NP for Rabin-Streett S2Gs,
- in coNP for Streett-Rabin S2Gs, and
- in NP ∩ coNP for parity S2Gs.

A corresponding NP-hardness result for deterministic Rabin-Streett S2Gs was established by Emerson & Jutla (1999). In particular, this hardness result also holds for the qualitative decision problem. Moreover, by determinacy, this result can be turned into a coNP-hardness result for (deterministic) Streett-Rabin S2Gs.

For S2Gs with Muller objectives, Chatterjee (2007) showed that the quantitative decision problem falls into PSPACE; for deterministic games, a polynomial-space algorithm had been given earlier by McNaughton (1993). A matching lower bound for deterministic games with Muller objectives was provided by Hunter & Dawar (2005).

Theorem 2.17 (Chatterjee). The quantitative decision problem is in PSPACE for Muller S2Gs.

Theorem 2.18 (Hunter & Dawar). The qualitative decision problem for deterministic Muller S2Gs is PSPACE-hard.

Theorem 2.18 does not hold if the Muller objectives are given by a family of subsets of *vertices*: Horn (2008a,b) showed that the qualitative decision problem for *explicit* Muller S2Gs is in P, and that the quantitative problem is in NP ∩ coNP.

Another class of S2Gs for which the qualitative decision problem is in P is, for each $d \in \mathbb{N}$, the class Parity[d] of all parity S2Gs whose priority function refers to at most d priorities (de Alfaro & Henzinger 2000). In particular, the qualitative decision problem for SS2Gs as well as (co-)Büchi S2Gs is in P.

Table 2.2. The complexity of deciding the value in S2Gs.

	Qualitative	Quantitative
SS2Gs	P-complete	NP ∩ coNP
Parity[d]	P-complete	NP ∩ coNP
Parity	UP ∩ coUP	NP ∩ coNP
Rabin-Streett	NP-complete	NP-complete
Streett-Rabin	coNP-complete	coNP-complete
Muller	Pspace-complete	Pspace-complete

For general parity S2Gs, however, the qualitative decision problem is only known to lie in UP ∩ coUP (Jurdziński 1998; Chatterjee et al. 2003).

Theorem 2.19 (Jurdziński; Chatterjee et al.). The qualitative decision problem is in UP ∩ coUP for parity S2Gs.

Theorem 2.20 (de Alfaro & Henzinger). For each $d \in \mathbb{N}$, the qualitative decision problem is in P for parity S2Gs with at most d priorities.

Table 2.2 summarises the results about the complexity of the quantitative and the qualitative decision problem for S2Gs. P-hardness (via Logspace-reductions) for all these problems follows from the fact that *and-or graph reachability* is P-complete (Immerman 1981).

The results summarised in Table 2.2 leave open the possibility that at least one of the following problems is decidable in polynomial time:

1. the qualitative decision problem for parity S2Gs,
2. the quantitative decision problem for SS2Gs,
3. the quantitative decision problem for parity S2Gs.

Note that, given that all of them are contained in both NP and coNP, it is unlikely that one of them is NP-hard or coNP-hard; such a result would imply that NP = coNP, and the polynomial hierarchy would collapse.

For the first problem, Chatterjee et al. (2003) gave a polynomial-time reduction to the qualitative decision problem for *deterministic* two-player zero-sum parity games. Hence, solving the qualitative decision problem for parity S2Gs is not harder than deciding which of the two players has a winning strategy in a deterministic two-player zero-sum parity game. Whether the latter problem is decidable in polynomial time is a long-standing open

problem. Several years after Emerson & Jutla (1991) put the problem into NP ∩ coNP, Jurdziński (1998) improved this bound slightly to UP ∩ coUP. Together with Paterson and Zwick (2008), he also gave an algorithm that decides the winner in subexponential time; a *randomised* subexponential algorithm had been given earlier by Björklund et al. (2003).

Another line of research has identified structural subclasses of graphs on which deterministic parity games can be solved efficiently. In particular, deterministic two-player zero-sum parity games can be solved in polynomial time on graphs of bounded tree width (Obdržálek 2003), bounded entanglement (Berwanger & Grädel 2005), bounded DAG width (Berwanger et al. 2006; Obdržálek 2006), bounded Kelly width (Hunter & Kreutzer 2007) and bounded clique width (Obdržálek 2007). However, Friedmann (2009) recently showed that the most promising candidate for a polynomial-time algorithm for the general case so far, the *discrete strategy improvement* algorithm due to Vöge & Jurdziński (2000), requires exponential time in the worst case.

Regarding the second problem, only some progress towards a polynomial-time algorithm has been made since Condon (1992) proved membership in NP ∩ coNP; for instance, Björklund & Vorobyov (2005) gave a randomised subexponential algorithm for solving SS2Gs, and Gimbert & Horn (2009) showed that the quantitative decision problem for SS2Gs is fixed-parameter tractable with respect to the number of stochastic vertices as the parameter.

For the third problem, Andersson & Miltersen (2009) recently established a polynomial-time Turing reduction to the second. Hence, there exists a polynomial-time algorithm for 2. if and only if there exists one for 3. In particular, a polynomial-time algorithm for 2. would also give a polynomial-time algorithm for 1. However, to the best of our knowledge, it is plausible that the qualitative decision problem for parity S2Gs is in P while the quantitative decision problem for SS2Gs is not.

2.6 Existence of residually optimal strategies

The goal of this section is to prove the existence of residually optimal pure strategies in finite S2Gs with prefix-independent objectives (Theorem 2.6). Although they did not state this explicitly, Gimbert & Horn (2010) actually proved this (stronger) result in their proof for the existence of optimal strategies in these games. We present an alternative proof of Theorem 2.6, which uses a concept introduced by Chatterjee et al. (2005). Our starting points are

the following two propositions, proved in (Gimbert & Horn 2010) using the notion of *reset strategies*.

Proposition 2.21. Let $\mathcal{G}$ be a finite S2G with prefix-independent objectives. If $\mathrm{val}^{\mathcal{G}}(v) = 1$, then player 0 has a winning strategy in $(\mathcal{G}, v)$.

Proposition 2.22. Let $\mathcal{G}$ be a finite S2G with prefix-independent objectives. If $\mathrm{val}^{\mathcal{G}}(v) > 0$ for all vertices v, then $\mathrm{val}^{\mathcal{G}}(v) = 1$ for all vertices v.

In order to apply Proposition 2.21, we partition the state space into regions of states with equal value and show that a residually optimal strategy can be obtained from a winning strategy in each of these regions. Formally, a *value class* of $\mathcal{G}$ is a maximal subset U of V such that $\mathrm{val}^{\mathcal{G}}$ is constant on U, i.e. $U = \{v \in V : \mathrm{val}^{\mathcal{G}}(v) = r\} \neq \varnothing$ for some $r \in [0, 1]$. We call a value class U *positive* if $\mathrm{val}^{\mathcal{G}}(U) > 0$. If both players play optimally, a value class can only be left through a stochastic vertex; we denote by $\mathrm{Bnd}(U)$ the set of all stochastic vertices $v \in U$ with $v\Delta \not\subseteq U$. Note that, since the value of a stochastic vertex is a weighted average of the values at its successors, every vertex in $\mathrm{Bnd}(U)$ must have both a successor with a higher value and one with a lower value.

Due to the possibility that $\mathrm{Bnd}(U) \neq \varnothing$, a value class U is, in general, not a subarena. However, the value classes of $\mathcal{G}$ are subarenas of the game $\overline{\mathcal{G}}$ that is derived from $\mathcal{G}$ by turning every vertex $v \in V$ such that $v \in \mathrm{Bnd}(U)$ for a value class U of $\mathcal{G}$ into a terminal vertex that is winning for player 0. Moreover, all vertices in the subgame $\overline{\mathcal{G}} \upharpoonright U$ have value 1 if U is a positive value class.

Lemma 2.23. Let $\mathcal{G}$ be a finite S2G with prefix-independent objectives, and let U be a positive value class of $\mathcal{G}$. Then $\mathrm{val}^{\overline{\mathcal{G}} \upharpoonright U}(v) = 1$ for all $v \in U$.

Proof. Let $\mathrm{Win} \subseteq V^{\omega}$ be the objective of player 0, let $U = \{v \in V : \mathrm{val}^{\mathcal{G}}(v) = r\}$ for $r > 0$, and denote by s the highest value of a vertex $v \notin U$ that is a successor of a vertex $u \in U \cap V_0$, i.e. $v \in u\Delta$; if no such vertex exists, we set $s := 0$. Since U is a positive value class, we have $s < r$. By Proposition 2.22, we only need to show that there is no vertex $u \in U$ with $\mathrm{val}^{\overline{\mathcal{G}} \upharpoonright U}(u) = 0$. Towards a contradiction, assume there is such a vertex u. Then, by Proposition 2.21, player 1 would have a strategy τ such that $\mathrm{Pr}_u^{\sigma, \tau}(\mathrm{Win} \cup \mathrm{Reach}(\mathrm{Bnd}(U))) = 0$ for all strategies σ of player 0 in $\overline{\mathcal{G}} \upharpoonright U$. Now, let $0 < \varepsilon < r - s$ and fix a globally ε-optimal strategy τ_{ε} of player 1 in $\mathcal{G}$. We devise a new strategy τ^* of player 1 in $\mathcal{G}$ as follows: as long as the play stays in U, τ^* behaves like τ; as soon as

the play leaves U, τ^* starts to behave like τ_ε. Formally, we set $\tau^*(xv) = \tau(xv)$ for histories xv that stay in U and $\tau^*(xvy) = \tau_\varepsilon(vy)$ for histories of the form $xvy \in V^* \cdot V \cdot V^*$ with $x \in U^*$ and $v \in V \setminus U$.

We claim that $\sup_\sigma \Pr_u^{\sigma,\tau^*}(\text{Win}) \le s + \varepsilon$ and therefore $\text{val}^{\mathcal{G}}(u) < r$, a contradiction to $\text{val}^{\mathcal{G}}(u) = r$. Let σ be a strategy of player 0 in $\mathcal{G}$. By the definition of τ^*, we have $\Pr_u^{\sigma,\tau^*}(\text{Win} \cap U^\omega) = 0$ and $\Pr_u^{\sigma,\tau^*}(U^* \cdot v \cdot V^\omega) > 0$ only if $v \in U$ or $v \in u\Delta$ for some $u \in U \cap V_0$. Hence,

$$
\begin{aligned}
&\Pr_u^{\sigma,\tau^*}(\text{Win}) \\
&= \Pr_u^{\sigma,\tau^*}(\text{Win} \cap U^\omega) + \sum_{xv \in U^*(V \setminus U)} \Pr_u^{\sigma,\tau^*}(\text{Win} \cap xv \cdot V^\omega) \\
&= \sum_{xv \in U^*(V \setminus U)} \Pr_u^{\sigma,\tau^*}(xv \cdot V^\omega) \cdot \Pr_v^{\sigma[x],\tau^*[x]}(\text{Win}) \\
&= \sum_{xv \in U^*(V \setminus U)} \Pr_u^{\sigma,\tau^*}(xv \cdot V^\omega) \cdot \Pr_v^{\sigma[x],\tau_\varepsilon}(\text{Win}) \\
&\le \sum_{xv \in U^*(V \setminus U)} \Pr_u^{\sigma,\tau^*}(xv \cdot V^\omega) \cdot (\text{val}^{\mathcal{G}}(v) + \varepsilon) \\
&\le \sum_{xv \in U^*(V \setminus U)} \Pr_u^{\sigma,\tau^*}(xv \cdot V^\omega) \cdot (s + \varepsilon) \\
&\le s + \varepsilon.
\end{aligned}
$$

Since σ was chosen arbitrarily, we get that $\sup_\sigma \Pr_u^{\sigma,\tau^*}(\text{Win}) \le s + \varepsilon$. $\qquad\square$

By Lemma 2.23 and Proposition 2.21, player 0 has a winning strategy in $\overline{\mathcal{G}} \restriction U$ if U is a positive value class of $\mathcal{G}$. To prove Theorem 2.6, we show that we can compose these strategies to a residually optimal strategy in $\mathcal{G}$.

Proof (of Theorem 2.6). Let $\text{Win} \subseteq V^\omega$ be the objective of player 0. It suffices to prove that player 0 has a residually optimal strategy; the claim for player 1 follows from exchanging the players' roles. Let $U_1, \ldots, U_k$ be an enumeration of the positive value classes of $\mathcal{G}$ such that $\text{val}^{\mathcal{G}}(U_i) < \text{val}^{\mathcal{G}}(U_j)$ for $i < j$, and let U_0 be the set of all vertices with value 0. For each $i = 1, \ldots, k$, let σ_i be a winning strategy in the game $\overline{\mathcal{G}} \restriction U_i$, and let σ_0 be an arbitrary strategy of player 0 in $\mathcal{G} \restriction U_0$. Define a strategy σ of player 0 in $\mathcal{G}$ by setting $\sigma(xv) = \sigma_i(yv)$ if $v \in U_i$ and y is the longest suffix of x that is contained in U_i. In order to prove that σ is residually optimal, let xv be a history of $(\mathcal{G}, v_0)$, and let τ be a strategy of player 1 in $\mathcal{G}$. We claim that

$$
\Pr_v^{\sigma[x],\tau}(\text{Win} \cup \text{coBüchi}(U_0)) = 1. \tag{2.1}
$$

It follows from Lemma 2.23 and the definition of $\text{Bnd}(U_i)$ that

1. $\text{Pr}_v^{\sigma[x],\tau}(V^\omega \setminus \text{Win} \cap \text{Büchi}(U_i \setminus \text{Bnd}(U_i)) \setminus \text{Büchi}(\text{Bnd}(U_i))) = 0$, and
2. $\text{Pr}_v^{\sigma[x],\tau}(\text{Büchi}(\text{Bnd}(U_i)) \setminus \text{Büchi}(\bigcup_{j>i} U_j)) = 0$

for all $i = 1, \dots, k$. Using these two facts, we can establish (2.1) by proving that the complementary event $V^\omega \setminus \text{Win} \cap \text{Büchi}(V \setminus U_0)$ occurs with probability 0:

$$
\text{Pr}_v^{\sigma[x],\tau}(V^\omega \setminus \text{Win} \cap \text{Büchi}(V \setminus U_0))
$$

$$
= \sum_{i=1}^k \text{Pr}_v^{\sigma[x],\tau}(V^\omega \setminus \text{Win} \cap \text{Büchi}(U_i) \setminus \text{Büchi}(\bigcup_{j>i} U_j))
$$

$$
= \sum_{i=1}^k \text{Pr}_v^{\sigma[x],\tau}(V^\omega \setminus \text{Win} \cap \text{Büchi}(U_i \setminus \text{Bnd}(U_i)) \setminus \text{Büchi}(\text{Bnd}(U_i) \cup \bigcup_{j>i} U_j))
$$

$$
+ \sum_{i=1}^k \text{Pr}_v^{\sigma[x],\tau}(V^\omega \setminus \text{Win} \cap \text{Büchi}(\text{Bnd}(U_i)) \setminus \text{Büchi}(\bigcup_{j>i} U_j))
$$

$$
\leq \sum_{i=1}^k \text{Pr}_v^{\sigma[x],\tau}(V^\omega \setminus \text{Win} \cap \text{Büchi}(U_i \setminus \text{Bnd}(U_i)) \setminus \text{Büchi}(\text{Bnd}(U_i)))
$$

$$
+ \sum_{i=1}^k \text{Pr}_v^{\sigma[x],\tau}(\text{Büchi}(\text{Bnd}(U_i)) \setminus \text{Büchi}(\bigcup_{j>i} U_j))
$$

$$
= 0 \, .
$$

It remains to be shown that $\sigma[x]$ is optimal in $(\mathcal{G}, v)$. Consider the random variables $\Theta_n : V^\omega \to V$, where $n \in \mathbb{N}$, defined by $\Theta_n(\pi) = \pi(n)$. The *expectation* of $\text{val}^\mathcal{G}(\Theta_n)$ under the probability measure $\text{Pr}_v^{\sigma[x],\tau}$ equals

$$
f(n) := \sum_{w \in V} \text{Pr}_v^{\sigma[x],\tau}(\Theta_n = w) \cdot \text{val}^\mathcal{G}(w) \, .
$$

It is easy to see that, by the definition of σ, we have $f(n) \leq f(n+1)$ for all $n \in \mathbb{N}$. Hence, $f^* := \lim_n f(n)$ exists, and we have $f(n) \leq f^*$ for all $n \in \mathbb{N}$. Moreover, since $f(n) \leq 1 - \text{Pr}_v^{\sigma[x],\tau}(\Theta_n \in U_0)$ for all $n \in \mathbb{N}$, we have $f^* = \limsup_n f(n) \leq 1 - \liminf_n \text{Pr}_v^{\sigma[x],\tau}(\Theta_n \in U_0) \leq 1 - \text{Pr}_v^{\sigma[x],\tau}(\text{coBüchi}(U_0))$. By (2.1), we have $\text{Pr}_v^{\sigma[x],\tau}(\text{Win}) \geq 1 - \text{Pr}_v^{\sigma[x],\tau}(\text{coBüchi}(U_0))$ and therefore $\text{val}^\mathcal{G}(v) = f(0) \leq f^* \leq \text{Pr}_v^{\sigma[x],\tau}(\text{Win})$. Since τ was chosen arbitrarily, we get that $\text{val}^{\sigma[x]}(v) \geq \text{val}^\mathcal{G}(v)$. $\qquad\square$

3

Equilibria

In this chapter, we introduce the equilibrium concepts that we consider in this work, i.e. Nash and subgame-perfect equilibria, and prove their existence for (subclasses of) stochastic games. Towards the end of this chapter, we turn to computational questions and introduce the decision problems that will occupy us for the rest of this work.

3.1 Definitions and basic properties

To capture rational behaviour of (selfish) players, John Nash (1950) introduced the notion of, what is now called, a Nash equilibrium. Formally, given a strategy profile $\bar{\sigma}$ of a game $(\mathcal{G}, v_0)$, we call a strategy τ of player i in $\mathcal{G}$ a *best response* to $\bar{\sigma}$ if τ maximises the expected payoff of player i: $\Pr_{v_0}^{\bar{\sigma}_{-i},\tau'}(\mathrm{Win}_i) \leq \Pr_{v_0}^{\bar{\sigma}_{-i},\tau}(\mathrm{Win}_i)$ for all strategies τ' of player i. A strategy profile $\bar{\sigma} = (\sigma_i)_{i \in \Pi}$ is a *Nash equilibrium* if each σ_i is a best response to $\bar{\sigma}$.

In a Nash equilibrium, no player can improve her payoff by unilaterally switching to a different strategy. In fact, to have a Nash equilibrium, it suffices that no player can gain from switching to a pure strategy.

Proposition 3.1. A strategy profile $\bar{\sigma}$ of a game $(\mathcal{G}, v_0)$ is a Nash equilibrium if and only if, for each player i and for each *pure* strategy τ of player i in $\mathcal{G}$, $\Pr_{v_0}^{\bar{\sigma}_{-i},\tau}(\mathrm{Win}_i) \leq \Pr_{v_0}^{\bar{\sigma}}(\mathrm{Win}_i)$.

Proof. Clearly, if $\bar{\sigma}$ is a Nash equilibrium, then $\mathrm{Pr}_{v_0}^{\bar{\sigma}_{-i},\tau}(\mathrm{Win}_i) \le \mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Win}_i)$ for each pure strategy τ of player i in $\mathcal{G}$. Now, assume that $\bar{\sigma}$ is not a Nash equilibrium. Hence, $p := \sup_\tau \mathrm{Pr}_{v_0}^{\bar{\sigma}_{-i},\tau}(\mathrm{Win}_i) = \mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Win}_i) + \varepsilon$ for some player i and some $\varepsilon > 0$. Consider the Markov decision process $\mathcal{G}^{\bar{\sigma}_{-i}}$. Clearly, the value of $\mathcal{G}^{\bar{\sigma}_{-i}}$ from v_0 equals p. By Theorem 2.5, there exists an $\varepsilon/2$-optimal pure strategy τ in $(\mathcal{G}^{\bar{\sigma}_{-i}}, v_0)$. Since the arena of $\mathcal{G}^{\bar{\sigma}_{-i}}$ is a forest, we can assume that τ is a positional strategy, which can be viewed as a pure strategy in $\mathcal{G}$. We have $\mathrm{Pr}_{v_0}^{\bar{\sigma}_{-i},\tau}(\mathrm{Win}_i) \ge p - \varepsilon/2 > p - \varepsilon = \mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Win}_i)$. $\qquad\square$

For two-player zero-sum games, a Nash equilibrium is nothing other than a pair of optimal strategies.

Proposition 3.2. Let $(\mathcal{G}, v_0)$ be an S2G. A strategy profile (σ, τ) of $(\mathcal{G}, v_0)$ is a Nash equilibrium if and only if both σ and τ are optimal. In particular, every Nash equilibrium of $(\mathcal{G}, v_0)$ has payoff $(\mathrm{val}^{\mathcal{G}}(v_0), 1 - \mathrm{val}^{\mathcal{G}}(v_0))$.

Proof. ($\Rightarrow$) Assume that both σ and τ are optimal, but that (σ, τ) is not a Nash equilibrium. Hence, one of the players, say player 1, can improve her payoff by playing another strategy τ'. Hence, $\mathrm{val}^{\mathcal{G}}(v_0) = \mathrm{Pr}_{v_0}^{\sigma,\tau}(\mathrm{Win}_0) > \mathrm{Pr}_{v_0}^{\sigma,\tau'}(\mathrm{Win}_0)$. However, since σ is optimal, $\mathrm{val}^{\mathcal{G}}(v_0) \le \mathrm{Pr}_{v_0}^{\sigma,\tau'}(\mathrm{Win}_0)$, a contradiction. The reasoning in the case that player 0 can improve is analogous.

($\Leftarrow$) Let (σ, τ) be a Nash equilibrium of $(\mathcal{G}, v_0)$, and let us first assume that σ is not optimal, i.e. $\mathrm{val}^{\sigma}(v_0) < \mathrm{val}^{\mathcal{G}}(v_0)$. By the definition of $\mathrm{val}^{\mathcal{G}}$, there exists another strategy σ' of player 0 such that $\mathrm{val}^{\sigma}(v_0) < \mathrm{val}^{\sigma'}(v_0) \le \mathrm{val}^{\mathcal{G}}(v_0)$. Moreover, since (σ, τ) is a Nash equilibrium,

$$\mathrm{Pr}_{v_0}^{\sigma,\tau}(\mathrm{Win}_0) \le \mathrm{val}^{\sigma}(v_0) < \mathrm{val}^{\sigma'}(v_0) = \inf_{\tau'} \mathrm{Pr}_{v_0}^{\sigma',\tau'}(\mathrm{Win}_0) \le \mathrm{Pr}_{v_0}^{\sigma',\tau}(\mathrm{Win}_0).$$

Thus, player 0 can improve her payoff by playing σ' instead of σ, a contradiction to the fact that (σ, τ) is a Nash equilibrium. The argumentation in the case that τ is not optimal is analogous. $\qquad\square$

In general, a Nash equilibrium can give a player a higher payoff than her value. However, the payoff a player receives in a Nash equilibrium can never be lower than her value, and this is true for every history that is consistent with the equilibrium.

Lemma 3.3. Let $(\mathcal{G}, v_0)$ be an SMG with objectives $\mathrm{Win}_i \subseteq V^\omega$. If $\bar{\sigma}$ is a Nash equilibrium of $(\mathcal{G}, v_0)$, then $\mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Win}_i \mid xv \cdot V^\omega) \ge \mathrm{val}_i^{\mathcal{G}[x]}(v)$ for each player i and every history xv that is consistent with $\bar{\sigma}$.

Proof. Assume there exists a history xv of $(\mathcal{G}, v_0)$ such that $\Pr_{v_0}^{\bar{\sigma}}(xv \cdot V^{\omega}) > 0$, but $p := \Pr_{v_0}^{\bar{\sigma}}(\text{Win}_i \mid xv \cdot V^{\omega}) < \text{val}_i^{\mathcal{G}[x]}(v)$. By the definition of $\text{val}_i^{\mathcal{G}[x]}$, there exists a strategy τ of player i in $\mathcal{G}[x]$ such that $\text{val}^{\tau}(v) > p$. We define a new strategy σ' for player i in $\mathcal{G}$ as follows: σ' is defined as σ_i for histories that do not begin with xv. For histories of the form xvy, however, we set $\sigma'(xvy) = \tau(vy)$. Clearly, $\Pr_{v_0}^{\bar{\sigma}_{-i}, \sigma'}(xv \cdot V^{\omega}) = \Pr_{v_0}^{\bar{\sigma}}(xv \cdot V^{\omega})$. Moreover, we claim that

$$\Pr_{v_0}^{\bar{\sigma}_{-i}, \sigma'}(X \setminus xv \cdot V^{\omega}) = \Pr_{v_0}^{\bar{\sigma}}(X \setminus xv \cdot V^{\omega}) \tag{3.1}$$

for every Borel set $X \subseteq V^{\omega}$. Eq. (3.1) is obviously true if X is a cylinder set. To prove (3.1) for all Borel sets, by the monotone class theorem (Theorem A.1), it suffices to prove that whenever we are given Borel sets $X_0, X_1, \ldots \subseteq V^{\omega}$ with $X_0 \subseteq X_1 \subseteq \cdots$ or $X_0 \supseteq X_1 \supseteq \cdots$ such that each X_n fulfils (3.1), then the set $\lim_n X_n := \bigcup_{n \in \mathbb{N}} X_n$ or $\lim_n X_n := \bigcap_{n \in \mathbb{N}} X_n$, respectively, also fulfils (3.1). Hence, assume that $X_0 \subseteq X_1 \subseteq \cdots$ or $X_0 \supseteq X_1 \supseteq \cdots$ and that each X_n fulfils (3.1). Clearly, $(\lim_n X_n) \setminus xv \cdot V^{\omega} = \lim_n (X_n \setminus xv \cdot V^{\omega})$. Moreover, since probability measures are continuous from below and above,

$$\Pr_{v_0}^{\bar{\sigma}_{-i}, \sigma'}(\lim_n (X_n \setminus xv \cdot V^{\omega}))$$
$$= \lim_n \Pr_{v_0}^{\bar{\sigma}_{-i}, \sigma'}(X_n \setminus xv \cdot V^{\omega})$$
$$= \lim_n \Pr_{v_0}^{\bar{\sigma}}(X_n \setminus xv \cdot V^{\omega})$$
$$= \Pr_{v_0}^{\bar{\sigma}}(\lim_n (X_n \setminus xv \cdot V^{\omega})),$$

which proves that (3.1) also holds for $\lim_n X_n$. Using Lemma 2.2, we can conclude that

$$\Pr_{v_0}^{\bar{\sigma}_{-i}, \sigma'}(\text{Win}_i)$$
$$= \Pr_{v_0}^{\bar{\sigma}_{-i}, \sigma'}(\text{Win}_i \setminus xv \cdot V^{\omega}) + \Pr_{v_0}^{\bar{\sigma}_{-i}, \sigma'}(\text{Win}_i \cap xv \cdot V^{\omega})$$
$$= \Pr_{v_0}^{\bar{\sigma}}(\text{Win}_i \setminus xv \cdot V^{\omega}) + \Pr_{v}^{\bar{\sigma}[x]_{-i}, \sigma'[x]}(x^{-1}\text{Win}_i) \cdot \Pr_{v_0}^{\bar{\sigma}_{-i}, \sigma'}(xv \cdot V^{\omega})$$
$$= \Pr_{v_0}^{\bar{\sigma}}(\text{Win}_i \setminus xv \cdot V^{\omega}) + \Pr_{v}^{\bar{\sigma}[x]_{-i}, \tau}(x^{-1}\text{Win}_i) \cdot \Pr_{v_0}^{\bar{\sigma}}(xv \cdot V^{\omega})$$
$$\geq \Pr_{v_0}^{\bar{\sigma}}(\text{Win}_i \setminus xv \cdot V^{\omega}) + \text{val}^{\tau}(v) \cdot \Pr_{v_0}^{\bar{\sigma}}(xv \cdot V^{\omega})$$
$$> \Pr_{v_0}^{\bar{\sigma}}(\text{Win}_i \setminus xv \cdot V^{\omega}) + p \cdot \Pr_{v_0}^{\bar{\sigma}}(xv \cdot V^{\omega})$$
$$= \Pr_{v_0}^{\bar{\sigma}}(\text{Win}_i \setminus xv \cdot V^{\omega}) + \Pr_{v_0}^{\bar{\sigma}}(\text{Win}_i \mid xv \cdot V^{\omega}) \cdot \Pr_{v_0}^{\bar{\sigma}}(xv \cdot V^{\omega})$$
$$= \Pr_{v_0}^{\bar{\sigma}}(\text{Win}_i \setminus xv \cdot V^{\omega}) + \Pr_{v_0}^{\bar{\sigma}}(\text{Win}_i \cap xv \cdot V^{\omega})$$
$$= \Pr_{v_0}^{\bar{\sigma}}(\text{Win}_i).$$
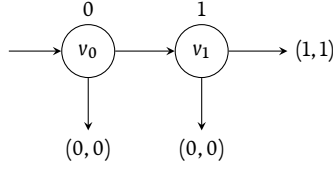
Figure 3.1. A two-player reachability game with an irrational Nash equilibrium.

Hence, player i can improver her payoff by switching to σ_i', a contradiction to $\bar{\sigma}$ being a Nash equilibrium. $\qquad\square$

As demonstrated by the following example, some Nash equilibria lack rationality in games that progress over time—such as the games we study in this work.

Example 3.4. Consider the deterministic two-player reachability game $(\mathcal{G}, v_0)$ depicted in Figure 3.1. Intuitively, the only rational outcome of this game should be the play leading to the terminal vertex with payoff $(1, 1)$. However, there are two Nash equilibria in this game:

- both players move "right" and win;
- both players move "down" and lose.

Clearly, the first strategy profile is a Nash equilibrium. For the second profile, note that player 1 cannot get a better payoff by changing her strategy since v_1 is never reached from v_0 if player 0 moves down.

The justification for the second Nash equilibrium in Example 3.5 is that player 1 threatens to move down if the game reaches v_1. However, this threat is not credible: if the game reaches v_1, then the only rational choice for player 1 is to move right because this is the only way for her to win. An equilibrium concept that eliminates such threats was introduced by Selten (1965). Formally, a strategy profile $\bar{\sigma}$ of a game $(\mathcal{G}, v_0)$ is a *subgame-perfect equilibrium* if $\bar{\sigma}[x]$ is a Nash equilibrium of the residual game $(\mathcal{G}[x], v)$ for every history xv of $(\mathcal{G}, v_0)$.

In a subgame-perfect equilibrium, every strategy is not only a best response after the initial history but after every possible history of the game (including histories that are not consistent with the equilibrium profile).

Example 3.5. Consider the same game as in Example 3.4. The Nash equilibrium where both players move down is not a subgame-perfect equilibrium because moving down is not a best response after the history $v_0 v_1$.

Recall that for two-player zero-sum games, Nash equilibria correspond to pairs of optimal strategies (Proposition 3.2). Similarly, subgame-perfect equilibria correspond to pairs of residually optimal strategies.

Proposition 3.6. Let $(\mathcal{G}, v_0)$ be a two-player zero-sum game. A strategy profile (σ, τ) of $(\mathcal{G}, v_0)$ is a subgame-perfect equilibrium if and only if both σ and τ are residually optimal.

Proof. Similar to the proof of Proposition 3.2. □

3.2 Existence of Nash equilibria

It follows from Theorem 2.6 and Proposition 3.2 that every finite two-player zero-sum stochastic game with prefix-independent objectives has a Nash equilibrium in pure strategies. The question arises whether this is still true if the two-player zero-sum assumption is relaxed.

By Lemma 3.3, a strategy profile $\bar{\sigma}$ can only be a Nash equilibrium if $\Pr_{v_0}^{\bar{\sigma}}(\text{Win}_i \mid xv \cdot V^{\omega}) \geq \text{val}_i^{\mathcal{G}}(v)$ for each player i and for each history xv consistent with $\bar{\sigma}$. The next lemma shows that, conversely, we can turn every strategy profile that fulfils this property into a Nash equilibrium. The proof uses so-called *threat* strategies (or *trigger* strategies), which are added on top of the given strategy profile: each player threatens to change her behaviour when one of the other players deviates from the prescribed strategy profile. Before being applied to stochastic games, this concept proved fruitful in the related area of *repeated games* (see Osborne & Rubinstein 1994, Chapter 8, and Aumann 1981).

Lemma 3.7. Let $(\mathcal{G}, v_0)$ be a finite SMG with prefix-independent objectives $\text{Win}_i \subseteq V^{\omega}$. If $\bar{\sigma}$ is a pure strategy profile such that $\Pr_{v_0}^{\bar{\sigma}}(\text{Win}_i \mid xv \cdot V^{\omega}) \geq \text{val}_i^{\mathcal{G}}(v)$ for each player i and for each history xv of $(\mathcal{G}, v_0)$ that is consistent with $\bar{\sigma}$, then $(\mathcal{G}, v_0)$ has a pure Nash equilibrium $\bar{\sigma}^*$ with $\Pr_{v_0}^{\bar{\sigma}} = \Pr_{v_0}^{\bar{\sigma}^*}$.

Proof. By Theorem 2.6, for each player i we can fix a globally optimal pure strategy τ_i of the coalition $\Pi \setminus \{i\}$ in the coalition game $\mathcal{G}_i$; denote by $\tau_{j,i}$ the corresponding pure strategy of player $j \neq i$ in $\mathcal{G}$. To simplify notation, we also define $\tau_{i,i}$ to be an arbitrary pure strategy of player i in $\mathcal{G}$. Player i's equilibrium strategy σ_i^* is defined as follows: For histories xv that are compatible with $\bar{\sigma}$, we set $\sigma_i^*(xv) = \sigma_i(xv)$. If xv is not compatible with $\bar{\sigma}$, then decompose x into $x = x_1 \cdot x_2$, where x_1 is the longest prefix of x that is compatible

with $\bar{\sigma}$, and let j be the player who has deviated, i.e. x_1 ends in V_j; we set $\sigma_i^*(xv) = \tau_{i,j}(x_2 v)$. Intuitively, σ_i^* behaves like σ_i as long as no other player j deviates from playing σ_j, in which case σ_i^* starts to behave like $\tau_{i,j}$.

Note that $\mathrm{Pr}_{v_0}^{\bar{\sigma}^*} = \mathrm{Pr}_{v_0}^{\bar{\sigma}}$. We claim that $\bar{\sigma}^*$ is additionally a Nash equilibrium of $(\mathcal{G}, v_0)$. Let $i \in \Pi$, and let ρ be a pure strategy of player i in $\mathcal{G}$; by Proposition 3.1, it suffices to show that $\mathrm{Pr}_{v_0}^{\bar{\sigma}_{-i}^*,\rho}(\mathrm{Win}_i) \leq \mathrm{Pr}_{v_0}^{\bar{\sigma}^*}(\mathrm{Win}_i)$.

Let us call a history $xv \in V^* V_i$ a *deviation history* if xv is compatible with both $\bar{\sigma}$ and $(\bar{\sigma}_{-i}, \rho)$, but $\sigma_i(xv) \neq \rho(xv)$; we denote the set of all deviation histories consistent with $\bar{\sigma}$ by D. Clearly, $\mathrm{Pr}_{v_0}^{\bar{\sigma}}(xv \cdot V^\omega) = \mathrm{Pr}_{v_0}^{\bar{\sigma}^*}(xv \cdot V^\omega) = \mathrm{Pr}_{v_0}^{\bar{\sigma}_{-i}^*,\rho}(xv \cdot V^\omega)$ for all $xv \in D$.

Claim. $\mathrm{Pr}_{v_0}^{\bar{\sigma}_{-i}^*,\rho}(X \setminus D \cdot V^\omega) = \mathrm{Pr}_{v_0}^{\bar{\sigma}}(X \setminus D \cdot V^\omega)$ for every Borel set $X \subseteq V^\omega$.

Proof. The proof of this claim uses the monotone class theorem and resembles the proof of the corresponding claim in the proof of Lemma 3.3.

Claim. $\mathrm{Pr}_{v_0}^{\bar{\sigma}_{-i}^*,\rho}(\mathrm{Win}_i \mid xv \cdot V^\omega) \leq \mathrm{val}_i^{\mathcal{G}}(v)$ for every $xv \in D$.

Proof. By the definition of the strategies $\tau_{j,i}$, we have that $\mathrm{Pr}_v^{(\tau_{j,i})_{j \neq i},\rho}(\mathrm{Win}_i) \leq \mathrm{val}_i^{\mathcal{G}}(v)$ for *every* vertex $v \in V$ and *every* strategy ρ of player i. Moreover, if xv is a deviation history, then for each player $j \neq i$ the residual strategy $\sigma_j^*[xv]$ is equal to $\tau_{j,i}$ on histories that start in $w := \rho(xv)$. Hence, by Lemma 2.2 and since Win_i is prefix-independent,

$$
\begin{aligned}
& \mathrm{Pr}_{v_0}^{\bar{\sigma}_{-i}^*,\rho}(\mathrm{Win}_i \mid xv \cdot V^\omega) \\
={} & \mathrm{Pr}_{v_0}^{\bar{\sigma}_{-i}^*,\rho}(\mathrm{Win}_i \mid xvw \cdot V^\omega) \\
={} & \mathrm{Pr}_{v_0}^{\bar{\sigma}_{-i}^*,\rho}(\mathrm{Win}_i \cap xvw \cdot V^\omega) / \mathrm{Pr}_{v_0}^{\bar{\sigma}_{-i}^*,\rho}(xvw \cdot V^\omega) \\
={} & \mathrm{Pr}_w^{\bar{\sigma}_{-i}^*[xv],\rho[xv]}(\mathrm{Win}_i) \\
={} & \mathrm{Pr}_w^{(\tau_{j,i})_{j \neq i},\rho[xv]}(\mathrm{Win}_i) \\
\leq{} & \mathrm{val}_i^{\mathcal{G}}(w) \\
\leq{} & \mathrm{val}_i^{\mathcal{G}}(v).
\end{aligned}
$$

Using the previous two claims, we prove that $\mathrm{Pr}_{v_0}^{\bar{\sigma}_{-i}^*,\rho}(\mathrm{Win}_i) \leq \mathrm{Pr}_{v_0}^{\bar{\sigma}^*}(\mathrm{Win}_i)$ as follows:

$$
\begin{aligned}
& \mathrm{Pr}_{v_0}^{\bar{\sigma}_{-i}^*,\rho}(\mathrm{Win}_i) \\
={} & \mathrm{Pr}_{v_0}^{\bar{\sigma}_{-i}^*,\rho}(\mathrm{Win}_i \setminus D \cdot V^\omega) + \sum_{xv \in D} \mathrm{Pr}_{v_0}^{\bar{\sigma}_{-i}^*,\rho}(\mathrm{Win}_i \cap xv \cdot V^\omega)
\end{aligned}
$$

$$= \mathrm{Pr}^{\bar{\sigma}}_{v_0}(\mathrm{Win}_i \setminus D \cdot V^\omega) + \sum_{xv \in D} \mathrm{Pr}^{\bar{\sigma}^*_{-i},\rho}_{v_0}(\mathrm{Win}_i \cap xv \cdot V^\omega)$$

$$= \mathrm{Pr}^{\bar{\sigma}}_{v_0}(\mathrm{Win}_i \setminus D \cdot V^\omega) + \sum_{xv \in D} \mathrm{Pr}^{\bar{\sigma}^*_{-i},\rho}_{v_0}(\mathrm{Win}_i \mid xv \cdot V^\omega) \cdot \mathrm{Pr}^{\bar{\sigma}^*_{-i},\rho}_{v_0}(xv \cdot V^\omega)$$

$$= \mathrm{Pr}^{\bar{\sigma}}_{v_0}(\mathrm{Win}_i \setminus D \cdot V^\omega) + \sum_{xv \in D} \mathrm{Pr}^{\bar{\sigma}^*_{-i},\rho}_{v_0}(\mathrm{Win}_i \mid xv \cdot V^\omega) \cdot \mathrm{Pr}^{\bar{\sigma}}_{v_0}(xv \cdot V^\omega)$$

$$\leq \mathrm{Pr}^{\bar{\sigma}}_{v_0}(\mathrm{Win}_i \setminus D \cdot V^\omega) + \sum_{xv \in D} \mathrm{val}^{\mathcal{G}}_i(v) \cdot \mathrm{Pr}^{\bar{\sigma}}_{v_0}(xv \cdot V^\omega)$$

$$\leq \mathrm{Pr}^{\bar{\sigma}}_{v_0}(\mathrm{Win}_i \setminus D \cdot V^\omega) + \sum_{xv \in D} \mathrm{Pr}^{\bar{\sigma}}_{v_0}(\mathrm{Win}_i \mid xv \cdot V^\omega) \cdot \mathrm{Pr}^{\bar{\sigma}}_{v_0}(xv \cdot V^\omega)$$

$$= \mathrm{Pr}^{\bar{\sigma}}_{v_0}(\mathrm{Win}_i \setminus D \cdot V^\omega) + \sum_{xv \in D} \mathrm{Pr}^{\bar{\sigma}}_{v_0}(\mathrm{Win}_i \cap xv \cdot V^\omega)$$

$$= \mathrm{Pr}^{\bar{\sigma}}_{v_0}(\mathrm{Win}_i)$$

$$= \mathrm{Pr}^{\bar{\sigma}^*}_{v_0}(\mathrm{Win}_i). \qquad \square$$

A variant of Lemma 3.7 handles games with prefix-independent ω-regular objectives and finite-state strategies.

Lemma 3.8. Let $(\mathcal{G}, v_0)$ be a finite SMG with prefix-independent ω-regular objectives $\mathrm{Win}_i \subseteq V^\omega$. If $\bar{\sigma}$ is a pure finite-state strategy profile such that $\mathrm{Pr}^{\bar{\sigma}}_{v_0}(\mathrm{Win}_i \mid xv \cdot V^\omega) \geq \mathrm{val}^{\mathcal{G}}_i(v)$ for each player i and for each history xv consistent with $\bar{\sigma}$, then $(\mathcal{G}, v_0)$ has a pure finite-state Nash equilibrium $\bar{\sigma}^*$ with $\mathrm{Pr}^{\bar{\sigma}}_{v_0} = \mathrm{Pr}^{\bar{\sigma}^*}_{v_0}$.

Proof. The proof is analogous to the proof of Lemma 3.7. Since, by Corollary 2.13, there exist optimal pure finite-state strategies in every finite SMG with ω-regular objectives, the strategies $\tau_{j,i}$ defined there can be assumed to be pure finite-state strategies. Consequently, the equilibrium profile $\bar{\sigma}^*$ can be implemented using finite-state strategies as well. $\qquad \square$

Using Lemma 3.7 and Theorem 2.6, we can easily prove the existence of pure Nash equilibria in finite SMGs with prefix-independent objectives.

Theorem 3.9. There exists a pure Nash equilibrium in any finite SMG with prefix-independent objectives.

Proof. Let $\mathcal{G}$ be a finite SMG with prefix-independent objectives $\mathrm{Win}_i \subseteq V^\omega$ and initial vertex v_0. By Theorem 2.6, each player i has a strongly optimal strategy σ_i in $\mathcal{G}$. Let $\bar{\sigma} = (\sigma_i)_{i \in \Pi}$. For every history xv that is consistent with $\bar{\sigma}$ and each player i, we have $\mathrm{Pr}^{\bar{\sigma}}_{v_0}(\mathrm{Win}_i \mid xv \cdot V^\omega) = \mathrm{Pr}^{\bar{\sigma}[x]}_v(\mathrm{Win}_i) \geq \mathrm{val}^{\mathcal{G}}_i(v)$. By Lemma 3.7, this implies that $(\mathcal{G}, v_0)$ has a pure Nash equilibrium. $\qquad \square$
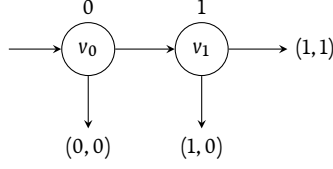
Figure 3.2. A two-player game with a pair of optimal strategies that cannot be extended to a Nash equilibrium.

For finite SMGs with ω-regular objectives, we can even show the existence of a pure finite-state equilibrium

Theorem 3.10. There exists a pure finite-state Nash equilibrium in any finite SMG with ω-regular objectives.

Proof. Since any SMG with ω-regular objectives can be reduced to one with parity objectives, it suffices to consider parity SMGs. For these games, the claim follows from Corollary 2.12 and Lemma 3.8 using the same argumentation as in the proof of Theorem 3.9. □

For deterministic games, one can prove the existence of a Nash equilibrium even if the game has an infinite arena and arbitrary Borel objectives. We will prove an even stronger theorem, namely the existence of subgame-perfect equilibria in these games, in the next section (Theorem 3.15).

Theorem 3.11. There exists a pure Nash equilibrium in any deterministic game with Borel objectives.

Theorem 3.10, Theorem 3.11 and a variant of Theorem 3.9 appeared originally in (Chatterjee et al. 2004b). However, their proof contains an inaccuracy: Essentially, they claim that any profile of optimal strategies can be extended to a Nash equilibrium with the same payoff (by adding threat strategies on top). This is, in general, not true, as the following example demonstrates.

Example 3.12. Consider the deterministic two-player game $(\mathcal{G}, v_0)$ depicted in Figure 3.2. Clearly, the value $\mathrm{val}_0^{\mathcal{G}}(v_0)$ for player 0 from v_0 is 1, and player 0's optimal strategy σ is to play from v_0 to v_1. For player 1, the value from v_0 is 0, and both of her positional strategies are optimal (albeit not necessarily globally optimal). In particular, her strategy τ of playing from v_1 to the terminal vertex with payoff $(1, 0)$ is optimal. The payoff of the strategy profile (σ, τ) is $(1, 0)$. However, there is no Nash equilibrium of $(\mathcal{G}, v_0)$ with payoff $(1, 0)$:
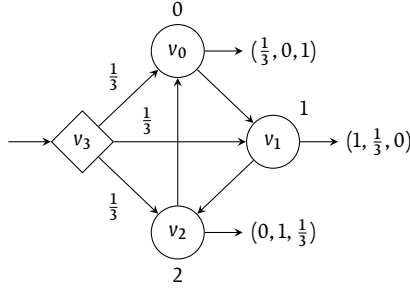
Figure 3.3. An SSMG with no stationary Nash equilibrium.

In any Nash equilibrium of $(\mathcal{G}, v_0)$, player 0 will move from v_0 to v_1 with probability 1. To have a Nash equilibrium, player 1 must play from v_1 to the terminal vertex with payoff $(1, 1)$ with probability 1; hence, every Nash equilibrium of this game has payoff $(1, 1)$.

Note that Theorem 3.10 only guarantees the existence of a pure finite-state Nash equilibrium, even for games with objectives where each player is guaranteed to have a positional optimal strategy. The question arises whether we can also guarantee the existence of a positional Nash equilibrium in such games. Kuipers et al. (2009) proved that this is not the case. In fact, they gave an example of a finite three-player SSMG that has no *stationary* Nash equilibrium.[1]

Proposition 3.13 (Kuipers et al.). There exists a finite SSMG that has no stationary Nash equilibrium.

Proof. Consider the three-player SSMG $\mathcal{G}$ depicted in Figure 3.3. We claim that $(\mathcal{G}, v_3)$ does not admit a stationary Nash equilibrium. Towards a contradiction, assume that $\bar{\sigma} = (\sigma_0, \sigma_1, \sigma_2)$ is a stationary Nash equilibrium, and denote by $p_i := \sigma_i(v_{i+1 \bmod 3} \mid v_i)$ the probability that player i stays inside the cycle. Since the game is symmetric, we can assume without loss of generality that $p_1 = \min\{p_0, p_1, p_2\}$. Clearly, $p_1 < 1$ since otherwise each player would receive payoff 0 but could improve her payoff by leaving the cycle. Now, since $p_1 \leq p_2$, player 0's only best response to $\bar{\sigma}$ is the strategy that plays to v_1 with probability 1; this gives player 1 payoff $\geq \frac{1}{2}$ because the probability of reaching the terminal vertex with payoff 1 is higher than the probability of reaching

---

[1] A similar game has been described by Boros & Gurvich (2003).
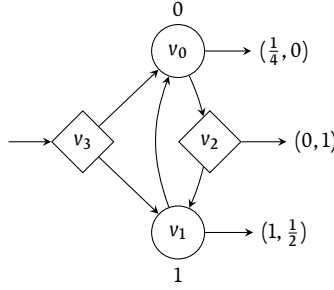
Figure 3.4. A two-player SSMG with no positional Nash equilibrium.

the terminal vertex with payoff 0. Hence, $p_0 = 1$. Since $p_0 = 1$, player 2's only best response to $\bar{\sigma}$ is to leave the cycle with probability 1. Hence, $p_2 = 0$ and, due to the minimality of p_1, also $p_1 = 0$. But then $\bar{\sigma}$ is not a Nash equilibrium because player 1 is better off by playing to v_2 with probability 1. $\qquad\square$

The existence of stationary Nash equilibria in finite *two-player* SSMGs is open. As the following proposition shows, the stronger statement that every such game has a positional Nash equilibrium remains false.

Proposition 3.14. There exists a finite two-player SSMG that has no positional Nash equilibrium.

Proof. Consider the SSMG $(\mathcal{G}, v_3)$ depicted in Figure 3.4. It is easy to see that none of the four positional strategy profiles in this game constitutes a Nash equilibrium. Note however that the stationary strategy profile (σ, τ) defined by $\sigma(v_2 \mid v_0) = \tau(v_0 \mid v_1) = \frac{2}{3}$ is a Nash equilibrium of $(\mathcal{G}, v_3)$.

For *deterministic* SSMGs, the existence of positional Nash equilibria is open; Boros & Gurvich (2003) only proved their existence for certain special cases such as games with only two players. For deterministic two-player parity games, we will prove the existence of positional equilibria (even subgame-perfect ones) in the next section (Theorem 3.17).

3.3 Existence of subgame-perfect equilibria

The main result presented in this section is the existence of subgame-perfect equilibria in deterministic games with Borel objectives (Ummels 2005);

in our presentation of the result, we follow Grädel & Ummels (2008). The construction is conceptually similar to the one used for proving the existence of Nash equilibria, but more involved: in particular, we will employ a fixed-point construction.

Theorem 3.15. There exists a pure subgame-perfect equilibrium in any deterministic game with Borel objectives.

Proof. Let $(\mathcal{G}, v_0)$ be a deterministic game with Borel objectives $\text{Win}_i \subseteq V^\omega$. Without loss of generality, assume that the arena of $\mathcal{G}$ is a tree with v_0 as its root: this can be achieved by unravelling the arena from v_0; the resulting arena is bisimilar to the original one.

For each ordinal α, we define a set $\Delta^\alpha \subseteq \Delta$, beginning with $\Delta^0 = \Delta$ and $\Delta^\gamma = \bigcap_{\alpha < \gamma} \Delta^\alpha$ for limit ordinals γ. To define $\Delta^{\alpha+1}$ from Δ^α, consider for each player i the two-player zero-sum game $\mathcal{G}_i^\alpha$ obtained from the coalition game $\mathcal{G}_i$ by restricting to transitions in Δ^α. Denote by $r_0, r_1, \ldots$ the *roots* of $\mathcal{G}_i^\alpha$, i.e. the vertices that have no predecessor with respect to the transition relation Δ^α, and let $x_0 r_0, x_1 r_0, \ldots$ be the unique histories of $(\mathcal{G}, v_0)$ ending in $r_0, r_1, \ldots$ (where $r_0 = v_0$ and $x_0 = \varepsilon$). By Corollary 2.10, and since the arena of $\mathcal{G}_i^\alpha$ is a forest, for every $k = 0, 1, \ldots$ there exist residually optimal positional strategies $\sigma_i^{\alpha,k}$ and $\tau_i^{\alpha,k}$ for player i and the coalition $\Pi \setminus \{i\}$, respectively, in the game $(\mathcal{G}_i^\alpha[x_k], r_k)$. Let σ_i^α and τ_i^α be the respective unions of these strategies, i.e. $\sigma_i^\alpha(v) = \sigma_i^{\alpha,k}(v)$ and $\tau_i^\alpha(v) = \tau_i^{\alpha,k}(v)$ for the unique $k \in \mathbb{N}$ such that v lies in the tree with root r_k; the strategies σ_i^α and τ_i^α are residually optimal in $(\mathcal{G}_i^\alpha[x_k], r_k)$ for each $k \in \mathbb{N}$. The set $\Delta^{\alpha+1}$ is obtained from Δ^α by removing all edges that are not taken by a winning strategy. Formally, if X_i^α is the set of all $v \in V_i$ such that σ_i^α is winning in $(\mathcal{G}_i^\alpha[x], v)$, where xv is the unique history of $(\mathcal{G}, v_0)$ ending in v, then

$$\Delta^{\alpha+1} = \Delta^\alpha \cap \bigcap_{i \in \Pi} \{(v, w) \in \Delta : v \notin X_i^\alpha \text{ or } w = \sigma_i^\alpha(v)\}.$$

Obviously, the sequence $(\Delta^\alpha)_{\alpha \in \text{On}}$ is non-increasing. Thus we can fix the least ordinal ξ with $\Delta^\xi = \Delta^{\xi+1}$ and define $\sigma_i := \sigma_i^\xi$ and $\tau_i := \tau_i^\xi$. Moreover, for each player $j \neq i$, let $\tau_{j,i}$ be the positional strategy of player j in $\mathcal{G}$ induced by τ_i. Intuitively, player i's equilibrium strategy σ_i^* works as follows: player i plays σ_i as long as no other player deviates; whenever some player $j \neq i$ deviates from her equilibrium strategy, player i switches to $\tau_{i,j}$. Formally, define $\delta(v) \in \Pi \cup \{\bot\}$ for each $v \in V$ by setting $\delta(v_0) = \bot$ and

$$\delta(v) = \begin{cases} \bot & \text{if } \delta(u) = \bot \text{ and } v = \sigma_i(u), \\ \delta(u) & \text{if } i \neq \delta(u) \neq \bot \text{ and } v = \tau_{i,\delta(u)}(u), \\ i & \text{otherwise,} \end{cases}$$

for $u \in V_i$ and $v \in u\Delta$. Then, for $v \in V_i$, we set $\sigma_i^*(v) = \sigma_i(v)$ if $\delta(v) \in \{i, \bot\}$ and $\sigma_i^*(v) = \tau_{i,\delta(v)}(v)$ otherwise.

It remains to be shown that $\bar{\sigma}^* := (\sigma_i^*)_{i \in \Pi}$ is a subgame-perfect equilibrium of $(\mathcal{G}, v_0)$. First note that σ_i is winning in $(\mathcal{G}_i^\xi[x], v)$ if σ_i^α is winning in $(\mathcal{G}_i^\alpha[x], v)$ for some ordinal α because, if σ_i^α is winning in $(\mathcal{G}_i^\alpha[x], v)$, then every play of $(\mathcal{G}_i^{\alpha+1}[x], v)$ is compatible with σ_i^α and therefore won by player i. Since $\Delta^\xi \subseteq \Delta^{\alpha+1}$, this also holds for every play of $(\mathcal{G}_i^\xi[x], v)$. Now, let v be any vertex of $\mathcal{G}$, and let xv be the unique history of $(\mathcal{G}, v_0)$ ending in v. We claim that $\bar{\sigma}^*$ is a Nash equilibrium of $(\mathcal{G}[x], v)$: Let σ' be a strategy of any player i in $\mathcal{G}$, and denote by π and π' the unique plays of $(\mathcal{G}[x], v)$ compatible with $\bar{\sigma}^*$ and $(\bar{\sigma}_{-i}^*, \sigma')$, respectively; we need to show that $\pi \in x^{-1}\text{Win}_i$ or $\pi' \notin x^{-1}\text{Win}_i$.

The claim is trivial if $\pi = \pi'$. Hence, assume that $\pi \neq \pi'$ and fix the least $n \in \mathbb{N}$ such that $\pi(n+1) \neq \pi'(n+1)$; clearly, $\pi(n) \in V_i$ and $\sigma'(\pi(n)) \neq \sigma_i^*(\pi(n))$. Without loss of generality, we can assume that $n = 0$ and thus $\pi(n) = v$. We distinguish whether σ_i is winning in $(\mathcal{G}_i^\xi[x], v)$ or not.

First, assume that σ_i is winning in $(\mathcal{G}_i^\xi[x], v)$. By the definition of the strategies σ_j^*, the play π is a play of $(\mathcal{G}_i^\xi[x], v)$. We claim that π is compatible with σ_i, which implies that $\pi \in x^{-1}\text{Win}_i$. Otherwise, fix the least $k \in \mathbb{N}$ such that $\pi(k) \in V_i$ and $\sigma_i(\pi(k)) \neq \pi(k+1)$. Since σ_i is winning in $(\mathcal{G}_i^\xi[x], v)$, this strategy is also winning in $(\mathcal{G}_i^\xi[x \cdot \pi|_k], \pi(k))$. But then $(\pi(k), \pi(k+1)) \in \Delta^\xi \setminus \Delta^{\xi+1}$, a contradiction to $\Delta^\xi = \Delta^{\xi+1}$.

Now, assume that σ_i is not winning in $(\mathcal{G}_i^\xi[x], v)$. By determinacy and since σ_i and τ_i are residually optimal, τ_i is winning in $(\mathcal{G}_i^\xi[x], v)$. Since $\sigma'(v) \neq \sigma_i^*(v)$, player i has deviated; hence, π' is compatible with τ_i. We claim that π' is a play of $(\mathcal{G}_i^\xi[x], v)$. Since τ_i is winning in $(\mathcal{G}_i^\xi[x], v)$, this implies that $\pi' \notin x^{-1}\text{Win}_i$. Otherwise, fix the least $k \in \mathbb{N}$ such that $(\pi'(k), \pi'(k+1)) \notin \Delta^\xi$ and the ordinal α such that $(\pi'(k), \pi'(k+1)) \in \Delta^\alpha \setminus \Delta^{\alpha+1}$. Hence, σ_i^α is winning in $(\mathcal{G}_i^\alpha[x \cdot \pi'|_k], \pi'(k))$, which implies that σ_i is winning in $(\mathcal{G}_i^\xi[x \cdot \pi'|_k], \pi'(k))$. Since π' is compatible with τ_i, this implies that τ_i is not winning in $(\mathcal{G}_i^\xi[x], v)$, a contradiction. $\qquad\qquad\square$

Similar to the situation for Nash equilibria in stochastic games, Theorem 3.15 can be strengthened for finite games with ω-regular objectives.

Theorem 3.16. There exists a pure finite-state subgame-perfect equilibrium in any finite deterministic game with ω-regular objectives.

Proof. Again, it suffices to consider games with parity objectives. For such games, the existence of globally optimal positional strategies allows us to perform the construction used in the proof of Theorem 3.15 directly on the arena of the game (regardless of whether it is a tree or not). It is easy to see that the resulting subgame-perfect equilibrium $\bar{\sigma}^*$ can be implemented as a strategy profile with memory of size $(|\Pi| + 1) \cdot |V|$. □

Finally, for games with only two players and parity objectives, we can prove the existence of a positional subgame-perfect equilibrium, even for games with an infinite arena.

Theorem 3.17. There exists a positional subgame-perfect equilibrium in any deterministic two-player parity game.

Proof. Let $(\mathcal{G}, v_0)$ be a deterministic two-player (not necessarily zero-sum) parity game. As pointed out in the proof of Theorem 3.16, the construction used in the proof of Theorem 3.15 can be performed directly on the arena of $\mathcal{G}$. Moreover, since the games $\mathcal{G}_i^{\alpha}[x]$ are all deterministic two-player zero-sum parity games, both strategies $\tau_{1,0}$ and $\tau_{0,1}$, as defined in the proof of Theorem 3.15, can be assumed to be positional. It is easy to see that the strategy profile $(\tau_{1,0}, \tau_{0,1})$ is a subgame-perfect equilibrium of $(\mathcal{G}, v_0)$. □

In contrast to the situation for Nash equilibria, Theorems 3.15 and 3.16 fail for stochastic games, as was demonstrated by Solan & Vieille (2003).

Proposition 3.18 (Solan & Vieille). There exists a finite two-player Büchi SMG that has no subgame-perfect equilibrium.

Proof. Consider the SSMG $(\mathcal{G}, v_0)$ depicted in Figure 3.5, where player 1 wins additionally all plays that visit v_0 infinitely often (a Büchi objective) or, equivalently, all plays that do not end in a terminal vertex (a safety objective). We claim that $(\mathcal{G}, v_0)$ has no subgame-perfect equilibrium.

Towards a contradiction, assume that (σ, τ) is a subgame-perfect equilibrium of $(\mathcal{G}, v_0)$, and let $\alpha_k := \sigma(v_1 \mid (v_0 v_1)^k v_0)$ and $\beta_k := \tau(v_0 \mid (v_0 v_1)^k)$ for each $k \in \mathbb{N}$. Additionally, we define $x_k := \prod_{i=k}^{\infty} \alpha_i$ and $y_k = \prod_{i=k}^{\infty} \beta_i$. We distinguish whether $x_k > \frac{1}{2}$ for some k or not.
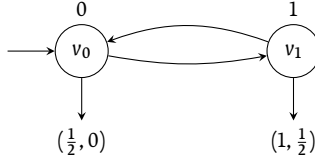
Figure 3.5. A Büchi SMG with no subgame-perfect equilibrium.

First, assume that $x_k > \frac{1}{2}$ for some k (and consequently $x_i > \frac{1}{2}$ for each $i \geq k$). We claim that $\beta_i = 1$ for each $i > k$. Otherwise, $y_{k+1} < 1$, and the expected payoff for player 1 after history $(v_0 v_1)^k v_0$ would be

$$\leq x_k y_{k+1} + (1 - y_{k+1}) \cdot \tfrac{1}{2}$$
$$< x_k y_{k+1} + x_k (1 - y_{k+1})$$
$$= x_k .$$

But with the strategy of playing to v_0 with probability 1 all the time, player 1 would receive expected payoff x_k.

Hence, $\beta_i = 1$ for each $i > k$, and the expected payoff for player 0 after history $(v_0 v_1)^k v_0$ equals $(1 - x_k) \cdot \frac{1}{2} < \frac{1}{4}$. But then, she could improve after this history by leaving the game, which would give her payoff $\frac{1}{2}$, a contradiction.

Now assume that $x_k \leq \frac{1}{2}$ for all k. Then there must exist infinitely many k such that $\alpha_k < 1$; we claim that $\beta_k = 0$ for any such k. Otherwise, the expected payoff for player 1 after history $(v_0 v_1)^k$ would be

$$\leq (1 - \beta_k) \cdot \tfrac{1}{2} + \alpha_k \beta_k \big(x_{k+1} y_{k+1} + (1 - y_{k+1}) \cdot \tfrac{1}{2} \big)$$
$$\leq (1 - \beta_k) \cdot \tfrac{1}{2} + \alpha_k \beta_k \big(y_{k+1} \cdot \tfrac{1}{2} + (1 - y_{k+1}) \cdot \tfrac{1}{2} \big)$$
$$= (1 - \beta_k) \cdot \tfrac{1}{2} + \alpha_k \beta_k \cdot \tfrac{1}{2}$$
$$< \tfrac{1}{2} .$$

However, by leaving the game, player 1 could get payoff $\frac{1}{2}$ immediately.

Hence, we can fix $k_1 < k_2$ such that $\alpha_{k_1} < 1$, $\beta_{k_2} = 0$ and $\alpha_k = 1$ for all $k_1 < k < k_2$. The expected payoff for player 0 after history $(v_0 v_1)^{k_1} v_0$ equals $(1 - \alpha_{k_1}) \cdot \frac{1}{2} + \alpha_{k_1} < 1$. But then player 0 could improve by moving to v_1 with probability 1, in which case she receives payoff 1, again a contradiction. □

The existence of subgame-perfect equilibria in finite SSMGs remains open. However, Flesch et al. (2010) proved the existence of subgame-perfect

ε-equilibria in these games for all $\varepsilon > 0$. Moreover, they showed that subgame-perfect equilibria do exist in deterministic SSMGs with arbitrary nonnegative payoffs on terminal vertices.

3.4 Computing equilibria

The first computational problem coming to mind when one considers equilibria is computing an equilibrium for a given game. For this problem to be meaningful, we need to make sure that both the possible inputs and the possible outputs are representable by finite means. In order to ensure this, we will restrict the inputs to finite SMGs with ω-regular objectives, and the outputs to equilibria in pure finite-state strategies. Moreover, for the sake of simplicity, we concentrate on parity SMGs.[2]

Computing Nash equilibria

For Nash equilibria, it is easy to see that the problem of computing an equilibrium lies in the class FNP of function problems where a potential solution can be verified in polynomial time.

Theorem 3.19. The problem of computing a pure finite-state Nash equilibrium (of polynomial size) in a finite parity SMG is in FNP.

Proof. To prove membership in FNP, we need to show that, given a finite parity SMG $(\mathcal{G}, v_0)$ and a pure strategy profile $\bar{\sigma}$ with finite memory $\mathfrak{M} = (M, \delta, m_0)$, we can decide in polynomial time whether $\bar{\sigma}$ is a Nash equilibrium of the game. This can be achieved as follows: First, for each player i, we calculate the payoff z_i of $\bar{\sigma}$ by computing the probability of the event $\chi^{-1}(\mathrm{Win}_i)$ in the Markov chain $(\mathcal{G}^{\bar{\sigma}}, (v_0, m_0))$. To check whether $\bar{\sigma}$ is a Nash equilibrium, we additionally need to compute for each player i the value r_i of the MDP $\mathcal{G}^{\bar{\sigma}_{-i}}$ from (v_0, m_0). Clearly, $\bar{\sigma}$ is a Nash equilibrium if and only if $r_i \leq z_i$ for each player i. Since we can compute the values of an MDP (or a Markov chain) with a parity objective in polynomial time, all this can be done in polynomial time. □

Arguably more interesting is the following theorem which essentially states that we can reduce the problem of computing a Nash equilibrium to

---

[2] One problem with computing equilibria for games with more complex objectives is that optimal strategies might be of exponential size (Dziembowski et al. 1997; Horn 2005).

the problem of computing optimal strategies. For any class $\mathcal{C}$ of parity S2Gs, let $\mathcal{C}^*$ be the class of all parity SMGs $\mathcal{G}$ such that for each player i the coalition game $\mathcal{G}_i$ is in $\mathcal{C}$.

Theorem 3.20. Let $\mathcal{C}$ be any class of finite parity S2Gs. There exists a polynomial-time Turing reduction from the problem of computing a Nash equilibrium for games in $\mathcal{C}^*$ to the problem of computing globally optimal positional strategies for games in $\mathcal{C}$.

Proof. We describe a deterministic polynomial-time algorithm for computing Nash equilibria for games in $\mathcal{C}^*$ with access to an oracle for computing globally optimal positional strategies for games in $\mathcal{C}$. On input $(\mathcal{G}, v_0)$, where $\mathcal{G} \in \mathcal{C}^*$, the algorithm starts by requesting from the oracle, for each player i, globally optimal positional strategies σ_i and τ_i for both players in the coalition game $\mathcal{G}_i \in C$. Then, the algorithm constructs a finite-state Nash equilibrium of $(\mathcal{G}, v_0)$ by combining the strategies σ_i and τ_i in the way it is done in the proof of Lemma 3.7, which can be done in polynomial time. □

Since optimal strategies can be computed in polynomial time for deterministic two-player zero-sum parity games with a bounded number of priorities, Theorem 3.20 implies that a Nash equilibrium of a deterministic multiplayer parity game with a bounded number of priorities can be computed in polynomial time. We will prove a stronger result below, namely that we can even compute a subgame-perfect equilibrium of such a game in polynomial time. Finally, it follows from Theorem 3.20 that computing a finite-state Nash equilibrium in a parity SMG can be done in polynomial time if and only if the quantitative decision problem for parity S2Gs and related problems are solvable in polynomial time.

Corollary 3.21. Either none or all of the following problems are solvable in polynomial time:

1. the quantitative decision problem for parity S2Gs,
2. computing the values of a parity S2G,
3. computing globally optimal positional strategies in a parity S2G,
4. computing a pure finite-state Nash equilibrium of a parity SMG,
5. computing a finite-state Nash equilibrium of a parity SMG.

Proof. The polynomial-time equivalence of 1., 2. and 3. is the subject of Proposition 2.14. That 4. can be done in polynomial time if 3. can follows from

---

Algorithm 3.1. Computing the set of consistent memory-vertex pairs.

Input: SMG $\mathcal{G} = (\Pi, V, (V_i)_{i \in \Pi}, \Delta, \chi, (\mathrm{Win}_i)_{i \in \Pi})$, $v_0 \in V$, memory $\mathfrak{M} = (M, \delta, m_0)$
Output: $\{(m, v) \in M \times V : \text{exists history } xv \text{ of } (\mathcal{G}, v_0) \text{ with } \delta^*(m_0, x) = m\}$

$X := \{(m_0, v_0)\}$
repeat
    $X' := X$
    $X := X \cup \{(n, w) \in M \times V : \text{exists } (m, v) \in X \text{ with } \delta(m, v) = n \text{ and } w \in v\Delta\}$
until $X = X'$
output X

Theorem 3.20, and that 5. can be done in polynomial time if 4. can is trivial. Finally, to compute $\mathrm{val}^{\mathcal{G}}(v)$ for a parity S2G $\mathcal{G}$, we can compute a finite-state Nash equilibrium (σ, τ) of $(\mathcal{G}, v)$. It follows from Proposition 3.2 that the payoff of (σ, τ) for player 0 equals $\mathrm{val}^{\mathcal{G}}(v)$. This payoff can be computed in polynomial time from (σ, τ) by analysing the generated Markov chain. Hence, 2. can be done in polynomial time if 5. can. □

Computing subgame-perfect equilibria

For subgame-perfect equilibria, the problem of computing a pure finite-state equilibrium of polynomial size in a parity SMG can again easily be put into FNP. The restriction to polynomial size is important: we do not know whether the existence of a pure finite-state subgame-perfect equilibrium in a parity SMG implies the existence of one with polynomial size.

Theorem 3.22. The problem of computing a pure finite-state subgame-perfect equilibrium of polynomial size in a finite parity SMG is in FNP.[3]

Proof. We need to show that, given a finite parity SMG $(\mathcal{G}, v_0)$ and a pure strategy profile $\bar{\sigma}$ with finite memory $\mathfrak{M} = (M, \delta, m_0)$, we can decide in polynomial time whether $\bar{\sigma}$ is a subgame-perfect equilibrium of the game. Our algorithm starts by computing the set C of *consistent* pairs of a memory state and a vertex, i.e. the set of all pairs $(m, v) \in M \times V$ such that there exists a history xv of $(\mathcal{G}, v_0)$ with $\delta^*(m_0, x) = m$. This can be achieved in polynomial time (for any kind of SMG) by Algorithm 3.1.

---

[3] More precisely, the problem of computing a pure finite-state subgame-perfect equilibrium of size at most $p(n)$ in a finite parity SMG of size n is in FNP for any polynomial p.

After having computed the set C, the algorithm proceeds by computing (in polynomial time) for each $i \in \Pi$ and $(m, v) \in C$ the probability $z_i(m, v)$ of the event $\chi^{-1}(\text{Win}_i)$ in the Markov chain $(\mathcal{G}^{\bar{\sigma}}, (m, v))$ and the value $r_i(m, v)$ of the MDP $\mathcal{G}^{\bar{\sigma}_{-i}}$ from (m, v). Clearly, $\bar{\sigma}$ is a subgame-perfect equilibrium of $(\mathcal{G}, v_0)$ if and only if $r_i(m, v) \leq z_i(m, v)$ for each $i \in \Pi$ and each $(m, v) \in C$. $\quad\square$

For deterministic games, we know how to construct a finite-state subgame-perfect equilibrium (Theorem 3.16). It is easy to see that the equilibrium can be computed in polynomial time if globally optimal positional strategies can be computed in polynomial time. For a class $\mathcal{C}$ of parity S2Gs, the class $\mathcal{C}^*$ is defined as above.

Theorem 3.23. Let $\mathcal{C}$ be any class of finite deterministic two-player zero-sum parity games. There exists a polynomial-time Turing reduction from the problem of computing a subgame-perfect equilibrium for games in $\mathcal{C}^*$ to the problem of computing globally optimal positional strategies for games in $\mathcal{C}$.

Theorem 3.23 makes the entire machinery that has been developed for solving (subclasses of) deterministic two-player zero-sum parity games available for the computation of subgame-perfect equilibria in deterministic multiplayer parity games. For example, the deterministic subexponential algorithm due to Jurdziński et al. (2008) can be adapted to compute subgame-perfect equilibria. Moreover, we can compute subgame-perfect equilibria in polynomial time for games on arenas that admit a polynomial-time algorithm for solving deterministic two-player zero-sum parity games, such as the ones mentioned in Section 2.5. In particular, we can compute a subgame-perfect equilibrium of a deterministic multiplayer parity game with a bounded number of priorities in polynomial time.

Corollary 3.24. For each $d \in \mathbb{N}$, there exists a polynomial-time algorithm for computing a subgame-perfect equilibrium of a finite deterministic multiplayer parity game with at most d priorities.

Finally, it follows from Theorem 3.23 that computing a Nash or subgame-perfect equilibrium of a deterministic multiplayer parity game is polynomial-time equivalent to deciding the winner of a deterministic two-player zero-sum parity game.

Corollary 3.25. Either none or all of the following problems are solvable in polynomial time:

1. deciding whether player 0 has a winning strategy in a deterministic two-player zero-sum parity game,
2. computing globally optimal positional strategies in a deterministic two-player zero-sum parity game,
3. computing a pure finite-state subgame-perfect equilibrium of a deterministic multiplayer parity game,
4. computing a finite-state subgame-perfect equilibrium of a deterministic multiplayer parity game,
5. computing a pure finite-state Nash equilibrium of a deterministic multiplayer parity game,
6. computing a finite-state Nash equilibrium of a deterministic multiplayer parity game.

Proof. The polynomial-time equivalence of 1. and 2. is standard (see Proposition 2.14). That 3. can be done in polynomial time if 2. can follows from Theorem 3.23. That 4. and 5. can be done in polynomial time if 3. can and that 6. can be done in polynomial time if 4. or 5. can is obvious. Finally, it follows from Proposition 3.2 that 1. can be done in polynomial time if 6. can. □

3.5 Decision problems

In applications, computing an arbitrary equilibrium is often not enough. For instance, in the stochastic dining philosophers problem, introduced in Section 1.2, we are after an equilibrium where each philosopher survives with high probability. In order to compute a "good" equilibrium, we permit the placing of a constraint on the payoff of the equilibrium. More precisely, for each player one may put both a lower and an upper threshold on her payoff. For any solution concept, the corresponding decision problem can be phrased as follows (the ordering ≤ is applied componentwise):

> Given a finite SMG $(\mathcal{G}, v_0)$ and $\bar{x}, \bar{y} \in [0,1]^{\Pi}$, decide whether there exists a solution with payoff $\geq \bar{x}$ and $\leq \bar{y}$.

To obtain meaningful results, we assume that all transition probabilities in $\mathcal{G}$ as well as the thresholds $\bar{x}$ and $\bar{y}$ are rational numbers (with numerator and denominator given in binary) and that all objectives are ω-regular. For the two solution concepts we study in this work, namely Nash and subgame-perfect equilibria, we obtain the decision problems NE and SPE.

Note that we have not put any restriction on the type of strategies that realise the equilibrium. It is natural to restrict the search space to equilibria in pure, finite-state, pure finite-state, stationary, or even positional strategies. For Nash equilibria, let us call the resulting decision problems PureNE, FinNE, PureFinNE, StatNE and PosNE, respectively; for subgame-perfect equilibria, the corresponding problems are PureSPE, FinSPE, PureFinSPE, StatSPE and PosSPE, respectively.

Often, we are not interested in the exact payoff of a solution, but only in which players win or lose almost surely. For any of the aforementioned decision problems, we obtain its *qualitative fragment* by requiring the thresholds $\bar{x}$ and $\bar{y}$ to be binary:

Given a finite SMG $(\mathcal{G}, v_0)$ and $\bar{x}, \bar{y} \in \{0, 1\}^\Pi$, decide whether there exists a solution with payoff $\geq \bar{x}$ and $\leq \bar{y}$.

It will turn out that the difficulty of the general problems manifests itself in this fragment: almost all of the lower bounds on the complexity of NE, SPE and their relatives we are going to prove in Chapter 4 can be obtained by a reduction to the qualitative fragment. In fact, in most cast cases, we show hardness for the following problem:

Given a finite SMG $(\mathcal{G}, v_0)$, decide whether there exists a solution where player 0 wins almost surely.

We already know that there exist SSMGs that have a stationary Nash equilibrium but no positional one. Hence, the problems StatNE and PosNE are distinct. Another extension of PosNE is PureFinNE. In fact, these extensions are incomparable, even if we consider only SSMGs. This has to be compared with the situation for SS2Gs, where all these problems coincide because SS2Gs admit globally optimal positional strategies.

Proposition 3.26. There exists a finite SSMG with a stationary subgame-perfect equilibrium where player 0 wins almost surely, but with no pure Nash equilibrium where player 0 wins with positive probability.

Proof. Consider the three-player SSMG depicted in Figure 3.6. Clearly, the stationary strategy profile where from vertex v_2 player 0 selects both outgoing transitions with probability $\frac{1}{2}$ each, player 1 plays from v_0 to v_1 and player 2 plays from v_1 to v_2 is a subgame-perfect equilibrium where player 0 wins almost surely. However, for any pure strategy profile where player 0 wins
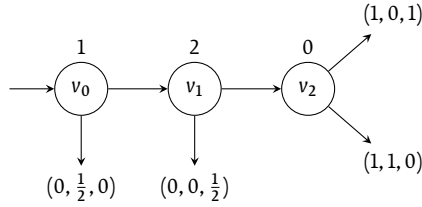
Figure 3.6. An SSMG that has a stationary subgame-perfect equilibrium where player 0 wins almost surely but no pure Nash equilibrium where player 0 wins with positive probability.

with positive probability (i.e. with probability 1), either player 1 or player 2 receives payoff 0 and could improve her payoff by switching her strategy at v_0 or v_1, respectively. □

Proposition 3.27. There exists a finite SSMG with a pure finite-state subgame-perfect equilibrium, but with no stationary Nash equilibrium.

Proof. Consider the game $(\mathcal{G}, v_3)$ given in the proof of Proposition 3.13 and depicted in Figure 3.3. We have already shown that this game does not admit a stationary Nash equilibrium. Now consider the pure strategy profile $\bar{\sigma}$ where player i leaves the cycle after history xv_i if and only if x is of even length. We claim that $\bar{\sigma}$ is a subgame-perfect equilibrium of $(\mathcal{G}, v_3)$. By symmetry, we only need to show that player 1 cannot improve her payoff after any history. Let xv be a history of $(\mathcal{G}, v_3)$; without loss of generality, $v = v_1$. If x is even, then player 1 receives payoff $\frac{1}{3}$ after history xv, but would receive payoff 0 by staying in the cycle. On the other hand, if x is odd, then player 1 receives payoff 1 after history xv, which is the best she can get. □

The complete taxonomy of the decision problems related to Nash and subgame-perfect equilibria is depicted in Figure 3.7. An edge from problem A to problem B means that A is a subset of B, i.e. all positive instances of A are positive instances of B. All inclusions are strict; this follows from Example 3.4, Propositions 3.26 and 3.27 as well as a result on finite-state Nash equilibria (Proposition 4.12), which we will present in the next chapter. Note however that an edge from A to B does, in general, not imply that one problem is computationally harder than the other (in the sense that there is a computable reduction between these problems). Hence, decidability has to be studied separately for each of these problems.
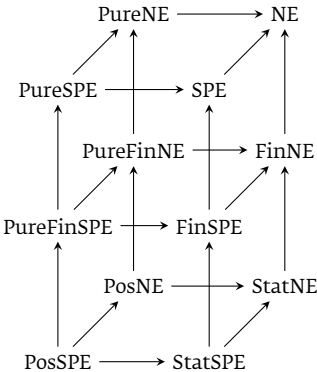
Figure 3.7. The different decision problems related to Nash and subgame-perfect equilibria.

4

Complexity of Equilibria

The aim of this chapter is to establish tight bounds on the complexity of the decision problems we have introduced in Chapter 3. All upper bounds apply to SMGs with Streett-Rabin or Muller objectives, while all lower bounds apply to SSMGs. Throughout this chapter, all games are finite.

4.1 Positional equilibria

In this section, we analyse the complexity of the presumably simplest of the decision problems introduced so far: PosNE and PosSPE. Not surprisingly, both these problems are decidable; in fact, they are NP-complete for all types of objectives we consider in this work. Let us start by proving membership in NP; it suffices to consider Streett-Rabin and Muller SMGs.

Theorem 4.1. PosNE is in NP for Streett-Rabin SMGs and Muller SMGs.

Proof. The proof is similar to the proof of Theorem 3.19. To decide PosNE, on input $\mathcal{G}, v_0, \bar{x}, \bar{y}$, we can guess a positional strategy profile $\bar{\sigma}$, i.e. a mapping $\bigcup_{i \in \Pi} V_i \to V$; then, we verify whether $\bar{\sigma}$ is a Nash equilibrium with the desired payoff. To do this, we first compute the payoff z_i of $\bar{\sigma}$ for each player i by computing the probability of the event Win_i in the (finite) Markov chain $(\mathcal{G}^{\bar{\sigma}}, v_0)$. Once each z_i is computed, we can easily check whether $x_i \le z_i \le y_i$. To verify that $\bar{\sigma}$ is a Nash equilibrium, we additionally compute, for each player i, the value r_i of the (finite) MDP $\mathcal{G}^{\bar{\sigma}_{-i}}$ from v_0. Clearly, $\bar{\sigma}$ is a Nash

equilibrium if and only if $r_i \leq z_i$ for each player i. Since we can compute the value of an MDP (or a Markov chain) with a Streett, Rabin or Muller objective in polynomial time (Theorem 2.15), all these checks can be carried out in polynomial time.   □

Theorem 4.2. PosSPE is in NP for Streett-Rabin SMGs and Muller SMGs.

Proof. The proof is virtually identical to the proof of Theorem 4.1. Since a stationary strategy profile $\bar{\sigma}$ is a subgame-perfect equilibrium of a Muller SMG $(\mathcal{G}, v_0)$ if and only if $\bar{\sigma}$ is a Nash equilibrium of $(\mathcal{G}, v)$ for every vertex $v \in V$ reachable from v_0, we only have to adapt the algorithm as follows: For each player i, instead of computing only the payoff z_i of $\bar{\sigma}$ in $(\mathcal{G}, v_0)$ and the value r_i of the MDP $\mathcal{G}^{\bar{\sigma}_{-i}}$ from v_0, we compute for each $v \in V$ the payoff z_v^i of $\bar{\sigma}$ for player i in $(\mathcal{G}, v)$ and the value r_v^i of $\mathcal{G}^{\bar{\sigma}_{-i}}$ from v. Finally, we compute (in polynomial time) the set R of vertices that are reachable from v_0 and check whether $r_v^i \leq z_v^i$ for each $v \in R$. Clearly, the resulting algorithm still runs in polynomial time.   □

   To establish NP-completeness, we still need to show NP-hardness. In fact, the reduction we are going to present does not only work for PosNE and PosSPE, but also for the problems StatNE and StatSPE, where we allow arbitrary stationary equilibria.

Theorem 4.3. PosNE, StatNE, PosSPE and StatSPE are NP-hard, even for SSMGs with only two players (three players for the qualitative fragments).

Proof. The proof is by reduction from SAT. Let $\varphi = C_1 \wedge \cdots \wedge C_m$, where $m \geq 1$, be a formula in conjunctive normal form over propositional variables $X_1, \ldots, X_n$; without loss of generality, we assume that each clause is nonempty. Our aim is to construct a two-player SSMG $(\mathcal{G}, v_0)$ such that the following statements are equivalent:

1. φ is satisfiable;
2. $(\mathcal{G}, v_0)$ has a positional subgame-perfect equilibrium with payoff $(1, \frac{1}{2})$;
3. $(\mathcal{G}, v_0)$ has a stationary Nash equilibrium with payoff $(1, \frac{1}{2})$.

Provided that the game can be constructed in polynomial time, these equivalences establish all desired reductions. The game $\mathcal{G}$ is depicted in Figure 4.1. The game proceeds from the initial vertex v_0 to X_i or $\neg X_i$ with probability $1/2^{i+1}$ each, and to vertex φ with probability $1/2^{n+1}$; with the remaining probability of $1/2^{n+1}$ the game proceeds to a terminal vertex with payoff $(1, 0)$.
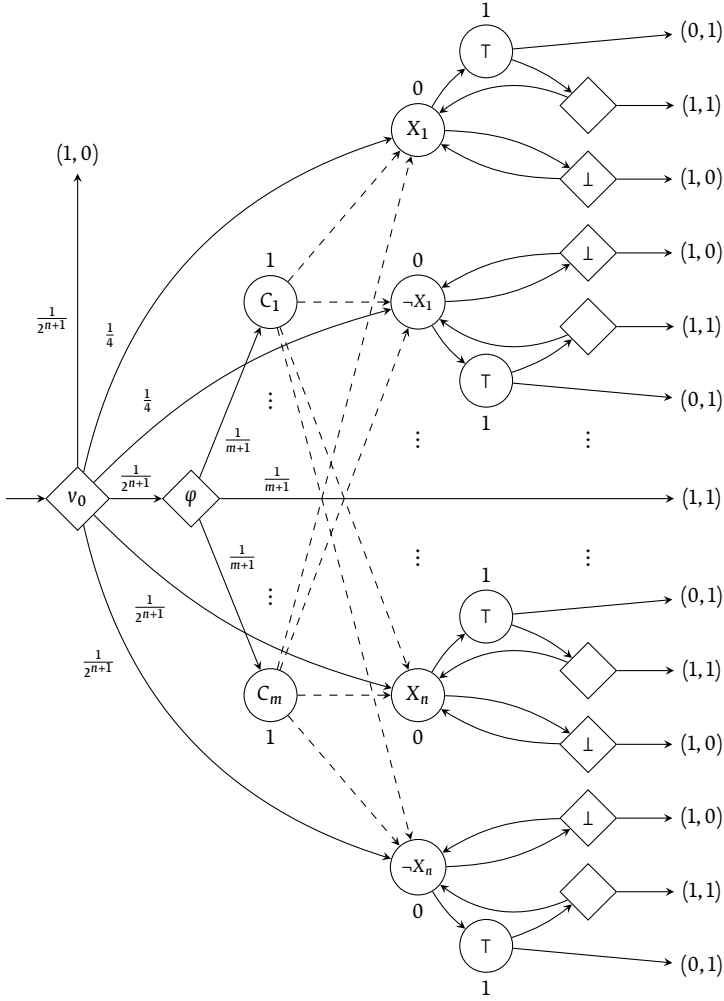
Figure 4.1. Reducing SAT to PosNE, StatNE, PosSPE and StatSPE.

From φ, the game proceeds to each vertex C_j with probability $1/(m+1)$; with the remaining probability of $1/(m+1)$, the game proceeds to a terminal vertex with payoff $(1,1)$. From vertex C_j (controlled by player 1), there is a transition to a literal L, i.e. $L = X_i$ or $L = \neg X_i$, if and only if L occurs inside the clause C_j. Obviously, the game $\mathcal{G}$ can be constructed from φ in polynomial time. We conclude the proof by showing that 1.–3. are equivalent.

$(1. \Rightarrow 2.)$ Assume that $\alpha \colon \{X_1, \ldots, X_n\} \to \{\text{true}, \text{false}\}$ is a satisfying assignment of φ. In the positional subgame-perfect equilibrium of $(\mathcal{G}, v_0)$, player 0 moves from a literal L to the neighbouring $\top$-labelled vertex if and only if L is mapped to true by α. Player 1 moves from vertex C_j to a (fixed) literal L that is contained in C_j and mapped to true by α (which is possible since α is a satisfying assignment); at $\top$-labelled vertices, player 1 never leaves the game. Obviously, player 0 wins almost surely in this strategy profile. For player 1, the payoff equals

$$\frac{1}{2^{n+1}} + \sum_{i=1}^{n} \frac{1}{2^{i+1}} = \frac{1}{2^{n+1}} + \frac{1}{2}\left(\sum_{i=1}^{n}\frac{1}{2^i}\right) = \frac{1}{2^{n+1}} + \frac{1}{2}\left(1 - \frac{1}{2^n}\right) = \frac{1}{2},$$

where the first summand is the probability of going from the initial vertex to φ, from where player 1 wins almost surely since from every clause vertex she plays to a "true" literal. Obviously, changing her strategy at any vertex cannot give her a better payoff. Therefore, we have a subgame-perfect equilibrium.

$(2. \Rightarrow 3.)$ Trivial.

$(3. \Rightarrow 1.)$ Let $\bar{\sigma} = (\sigma_0, \sigma_1)$ be a stationary Nash equilibrium of $(\mathcal{G}, v_0)$ with payoff $(1, \frac{1}{2})$. Our first aim is to show that σ_0 is actually a positional strategy. Towards a contradiction, assume that there exists a literal L such that $\sigma_0(L)$ assigns probability $0 < q < 1$ to the neighbouring $\top$-labelled vertex. Since player 0 wins almost surely, player 1 never leaves the game. Hence, the expected payoff for player 1 from vertex L equals q. However, if she left the game at the $\top$-labelled vertex, she would receive payoff $2q/(1+q) > q$. Therefore, $\bar{\sigma}$ is not a Nash equilibrium, a contradiction.

Since σ_0 is a positional strategy, we can define a *pseudo assignment* $\alpha \colon \{X_1, \neg X_1, \ldots, X_n, \neg X_n\} \to \{\text{true}, \text{false}\}$ by setting $\alpha(L) = \text{true}$ if σ_0 prescribes to go from vertex L to the neighbouring $\top$-labelled vertex. Our next aim is to show that α is actually an assignment: $\alpha(X_i) = \text{true}$ if and only if $\alpha(\neg X_i) = \text{false}$. To see this, note that we can compute player 1's expected payoff as follows:

$$\frac{1}{2} = \frac{p}{2^{n+1}} + \sum_{i=1}^{n} \frac{a_i}{2^{i+1}}, \quad a_i = \begin{cases} 0 & \text{if } \alpha(X_i) = \alpha(\neg X_i) = \text{false}, \\ 1 & \text{if } \alpha(X_i) \neq \alpha(\neg X_i), \\ 2 & \text{if } \alpha(X_i) = \alpha(\neg X_i) = \text{true}, \end{cases}$$

where p is the expected payoff for player 1 from vertex φ. By the construction of $\mathcal{G}$, we have $p > 0$, and the equality only holds if $p = 1$ and $a_i = 1$ for all $i = 1, \ldots, n$, which proves that α is an assignment.

Finally, we claim that α satisfies φ. If this were not the case, there would exist a clause C such that player 1's expected payoff from vertex C equals 0 and therefore $p < 1$. This is a contradiction to $p = 1$, as shown above.

To show that the qualitative fragments of PosNE and StatNE are also NP-hard, it suffices to modify the game $\mathcal{G}$ as follows: First, we add one new player, player 2, who wins at precisely those terminal vertices where player 1 loses. Second, we add two new vertices v_1 and v_2. At v_1, player 1 has the choice to leave the game; if she decides to stay inside the game, the play proceeds to v_2, where player 2 has the choice to leave the game; if she also decides to stay inside the game, the play proceeds to vertex v_0 from where the game continues normally; if player 1 or player 2 decide to leave the game, then each of them receives payoff $\frac{1}{2}$, but player 0 receives payoff 0. Let us denote the modified game by $\mathcal{G}'$. It is straightforward to see that the following statements are equivalent:

1. $(\mathcal{G}', v_1)$ has a stationary Nash equilibrium where player 0 wins almost surely;
2. $(\mathcal{G}, v_0)$ has a stationary Nash equilibrium with payoff $(1, \frac{1}{2})$;
3. φ is satisfiable;
4. $(\mathcal{G}, v_0)$ has a positional subgame-perfect equilibrium with payoff $(1, \frac{1}{2})$;
5. $(\mathcal{G}', v_1)$ has a positional subgame-perfect equilibrium where player 0 wins almost surely. □

Recall from Chapter 3 that not every SMG has a positional Nash equilibrium (Proposition 3.14). Hence, it is also a nontrivial problem to decide whether an SMG has a positional Nash equilibrium at all. It follows from Theorem 4.1 that, e.g. for SMGs with Muller objectives, there exists a nondeterministic polynomial-time algorithm for deciding this problem. On the other hand, this problem is NP-hard, even for three-player SSMGs.

Corollary 4.4. Deciding whether in a given three-player SSMG there exists a positional Nash equilibrium is NP-hard.

Proof. The proof is by reduction from the following problem, which we just proved to be NP-hard: Given a three-player SSMG $(\mathcal{G}, v_0)$, decide whether $(\mathcal{G}, v_0)$ has a positional Nash equilibrium where player 0 wins almost surely. In the following, let $(\mathcal{G}_1, v_1)$ be a fixed three-player SSMG that does not have a positional Nash equilibrium and where player 0 wins almost surely in every strategy profile (such a game can be derived from Proposition 3.14 by adding one more player). We need to show how to construct (in polynomial time) from a given three-player SSMG $(\mathcal{G}, v_0)$ a new three-player SSMG $(\widetilde{\mathcal{G}}, \widetilde{v}_0)$ such that $(\mathcal{G}, v_0)$ has a positional Nash equilibrium where player 0 wins almost surely if and only if $(\widetilde{\mathcal{G}}, \widetilde{v}_0)$ has a positional Nash equilibrium at all. The game $\widetilde{\mathcal{G}}$ is the disjoint union of $\mathcal{G}$ and $\mathcal{G}_1$ combined with one new vertex $\widetilde{v}_0$, controlled by player 0. At $\widetilde{v}_0$, player 0 can choose to move to v_0 or to v_1; in either case, the remaining play stays inside $\mathcal{G}$ or $\mathcal{G}_1$, respectively. Obviously, $\widetilde{\mathcal{G}}$ can be constructed from $\mathcal{G}$ in polynomial time. It remains to be shown that $(\mathcal{G}, v_0)$ has positional Nash equilibrium where player 0 wins almost surely if and only if $(\widetilde{\mathcal{G}}, \widetilde{v}_0)$ has a positional Nash equilibrium.

($\Rightarrow$). Assume that $(\mathcal{G}, v_0)$ has a positional Nash equilibrium where player 0 wins almost surely. This equilibrium can easily be extended to a positional Nash equilibrium of $(\widetilde{\mathcal{G}}, \widetilde{v}_0)$ by letting player 0 move from $\widetilde{v}_0$ to v_0 (and choosing an arbitrary positional strategy profile for $\mathcal{G}_1$).

($\Leftarrow$) Assume that $(\widetilde{\mathcal{G}}, \widetilde{v}_0)$ has a positional Nash equilibrium $\bar{\sigma}$. We claim that $\sigma_0(\widetilde{v}_0) = v_0$. Otherwise, $\bar{\sigma}$ would induce a positional Nash equilibrium of $(\mathcal{G}_1, v_1)$, a contradiction. Hence, $\sigma_0(\widetilde{v}_0) = v_0$, and $\bar{\sigma}$ induces a positional Nash equilibrium of $(\mathcal{G}, v_0)$. We claim that player 0 wins almost surely in this equilibrium. Otherwise, she could improve her payoff by playing from $\widetilde{v}_0$ to v_1 from where she wins with probability 1, a contradiction. □

4.2 Stationary equilibria

To prove the decidability of StatNE and StatSPE, we appeal to results established for the *existential theory of the reals*, the set of all existential first-order sentences (over the appropriate signature) that hold in the ordered field $\mathfrak{R} := (\mathbb{R}, +, \cdot, 0, 1, \leq)$. The best known upper bound for the complexity of the associated decision problem is PSPACE (Canny 1988), which leads to the following theorem.

Theorem 4.5. StatNE is in PSPACE for SMGs with Streett-Rabin or Muller objectives.

Proof. Since PSPACE = NPSPACE, it suffices to provide a nondeterministic algorithm with polynomial space requirements for deciding StatNE. On input $\mathcal{G}, v_0, \bar{x}, \bar{y}$, where without loss of generality $\mathcal{G}$ is an SMG with Muller objectives given by $\mathcal{F}_i \subseteq \mathcal{P}(C)$, the algorithm starts by guessing the *support* $S \subseteq V \times V$ of a stationary strategy profile $\bar{\sigma}$ of $\mathcal{G}$, i.e. $S = \{(v, w) \in V \times V : \bar{\sigma}(w \mid v) > 0\}$. From the set S alone, by standard graph algorithms, one can compute for each player i the following sets in polynomial time (see Appendix B):

1. the union F_i of all end components (i.e. bottom SCCs) U of the Markov chain $\mathcal{G}^{\bar{\sigma}}$ with $\chi(U) \in \mathcal{F}_i$,

2. the set R_i of vertices v such that $\Pr_v^{\bar{\sigma}}(\text{Reach}(F_i)) > 0$,

3. the union T_i of all end components of the MDP $\mathcal{G}^{\bar{\sigma}_{-i}}$ that are winning for player i.

After computing all these sets, the algorithm evaluates an existential first-order sentence ψ, which can be computed in polynomial time from $\mathcal{G}, v_0$, $\bar{x}, \bar{y}, (R_i)_{i \in \Pi}, (F_i)_{i \in \Pi}$ and $(T_i)_{i \in \Pi}$, over $\mathfrak{R}$ and returns the answer to this query.

How does ψ look like? Let $\bar{\alpha} = (\alpha_{vw})_{v,w \in V}$, $\bar{r} = (r_v^i)_{i \in \Pi, v \in V}$ and $\bar{z} = (z_v^i)_{i \in \Pi, v \in V}$ be three sets of variables, and let $V_* = \bigcup_{i \in \Pi} V_i$. The formula

$$\varphi(\bar{\alpha}) := \bigwedge_{v \in V_*} \left(\bigwedge_{w \in v\Delta} \alpha_{vw} \geq 0 \wedge \bigwedge_{w \in V \setminus v\Delta} \alpha_{vw} = 0 \wedge \sum_{w \in v\Delta} \alpha_{vw} = 1 \right) \wedge$$

$$\bigwedge_{\substack{v \in V \setminus V_* \\ w \in V}} \alpha_{vw} = \Delta(w \mid v) \wedge \bigwedge_{(v,w) \in S} \alpha_{vw} > 0 \wedge \bigwedge_{(v,w) \notin S} \alpha_{vw} = 0$$

states that the mapping $\bar{\sigma} \colon V \to \mathcal{D}(V)$, defined by $\bar{\sigma}(w \mid v) = \alpha_{vw}$, constitutes a valid stationary strategy profile of $\mathcal{G}$ whose support is S. Provided that $\varphi(\bar{\alpha})$ holds in $\mathfrak{R}$, the formula

$$\eta_i(\bar{\alpha}, \bar{z}) := \bigwedge_{v \in F_i} z_v^i = 1 \wedge \bigwedge_{v \in V \setminus R_i} z_v^i = 0 \wedge \bigwedge_{v \in V \setminus F_i} z_v^i = \sum_{w \in v\Delta} \alpha_{vw} \cdot z_w^i$$

states that $z_v^i = \Pr_v^{\bar{\sigma}}(\text{Win}_i)$ for each $v \in V$, where $\bar{\sigma}$ is defined as above. This follows from a well-known result about Markov chains, namely that the vector of the aforementioned probabilities is the unique solution of the given system of equations (see Appendix B). Finally, the formula

$$\vartheta_i(\bar\alpha, \bar r) := \bigwedge_{v \in V} r_v^i \geq 0 \,\wedge\, \bigwedge_{v \in T_i} r_v^i = 1 \,\wedge\, \bigwedge_{\substack{v \in V_i \\ w \in v\Delta}} r_v^i \geq r_w^i \,\wedge\, \bigwedge_{v \in V \setminus V_i} r_v^i = \sum_{w \in v\Delta} \alpha_{vw} \cdot r_w^i$$

states that $\bar r$ is a solution of the linear programme for computing the values of the MDP $\mathcal{G}^{\bar\sigma_{-i}}$ (see Appendix B). In particular, the formula is fulfilled if $r_v^i = \sup_\tau \Pr_v^{\bar\sigma_{-i}, \tau}(\mathrm{Reach}(T_i)) = \sup_\tau \Pr_v^{\bar\sigma_{-i}, \tau}(\mathrm{Win}_i)$ (the latter equality follows from Lemmas 2.3 and 2.4), and every other solution is greater than this one (in each component).

The desired sentence ψ is the existential closure of the conjunction of φ and, for each player i, the formulae η_i and ϑ_i combined with formulae stating that player i cannot improve her payoff and that the expected payoff for player i lies in between the given thresholds:

$$\psi := \exists\bar\alpha \, \exists\bar r \, \exists\bar z \left(\varphi(\bar\alpha) \wedge \bigwedge_{i \in \Pi} (\eta_i(\bar\alpha, \bar z) \wedge \vartheta_i(\bar\alpha, \bar r) \wedge r_{v_0}^i \leq z_{v_0}^i \wedge x_i \leq z_{v_0}^i \leq y_i) \right).$$

Clearly, ψ holds in $\mathfrak{R}$ if and only if $(\mathcal{G}, v_0)$ has a stationary Nash equilibrium with payoff at least $\bar x$ and at most $\bar y$ whose support is S. Consequently, the algorithm is correct. □

Theorem 4.6. StatSPE is in Pspace for SMGs with Streett-Rabin or Muller objectives.

Proof. Again, the proof is virtually identical to the proof of Theorem 4.5. As part of the preprocessing, we compute (in polynomial time) the set R of vertices reachable from v_0. Finally, instead of evaluating the sentence ψ, we evaluate the following sentence:

$$\psi' := \exists\bar\alpha \, \exists\bar r \, \exists\bar z \left(\varphi(\bar\alpha) \wedge \bigwedge_{i \in \Pi} (\eta_i(\bar\alpha, \bar z) \wedge \vartheta_i(\bar\alpha, \bar r) \wedge \bigwedge_{v \in R} r_v^i \leq z_v^i \wedge x_i \leq z_{v_0}^i \leq y_i) \right).$$

Clearly, ψ' holds in $\mathfrak{R}$ if and only if there exists a stationary subgame-perfect equilibrium of $\mathcal{G}$ with payoff at least $\bar x$ and at most $\bar y$ whose support is S. □

In the previous section, we showed that StatNE and StatSPE are NP-hard, leaving a considerable gap to our upper bound of Pspace. Towards gaining a better understanding of these problems, we relate StatNE and StatSPE to the *square root sum problem* (SqrtSum) of deciding, given numbers $d_1, \ldots, d_n, k \in \mathbb{N}$, whether $\sum_{i=1}^n \sqrt{d_i} \geq k$.

Recently, Allender et al. (2009) showed that SqrtSum belongs to the fourth level of the *counting hierarchy*, a slight improvement over the previously known
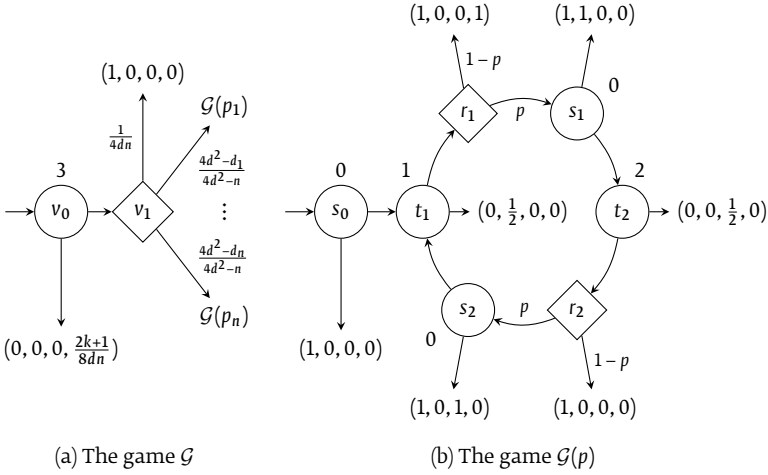
(a) The game $\mathcal{G}$ (b) The game $\mathcal{G}(p)$

Figure 4.2. Reducing SqrtSum to StatNE and StatSPE.

PSPACE upper bound. However, it has been an open question since the 1970s as to whether SqrtSum falls into the polynomial hierarchy (Garey et al. 1976; Etessami & Yannakakis 2010). We identify a polynomial-time reduction from SqrtSum to StatNE and StatSPE, even for four-player SSMGs.[1] Hence, StatNE and StatSPE are at least as hard as SqrtSum, and showing that StatNE or StatSPE resides inside the polynomial hierarchy would imply a major break-through in understanding the complexity of numerical computation.

Theorem 4.7. SqrtSum is polynomial-time reducible to both StatNE and StatSPE, even for four-player SSMGs.

Before we state the reduction, let us first examine the game $\mathcal{G}(p)$, where $\frac{1}{2} \leq p < 1$, which is depicted in Figure 4.2 (b).

Lemma 4.8. The maximal payoff player 3 receives in a stationary Nash or subgame-perfect equilibrium of $(\mathcal{G}(p), s_0)$ equals $(\sqrt{2 - 2p} - p + 1)/(2p + 2)$.

Proof. Note that a stationary strategy profile $\bar{\sigma}$ can only be a Nash equilibrium where player 3 receives payoff > 0 if player 1 plays from t_1 to r_1 with probabil-ity 1 and player 2 plays from t_2 to r_2 with probability 1 (or if t_2 is not reachable with $\bar{\sigma}$, but then player 3 receives payoff $\leq 1 - p$) because otherwise player 0

---

[1] Some authors define SqrtSum using $\leq$ instead of $\geq$. With this definition, we would reduce from the complement of SqrtSum instead.

would prefer to leave the game at v_0. Moreover, the maximum payoff for player 3 can only be attained when player 0 plays with probability 1 from s_0 to t_1 because, if player 0 plays from s_0 to t_1 with probability $0 < x < 1$, then setting x to 1 yields a Nash equilibrium with a better payoff for player 3. Hence, the only variable quantities are the probabilities x_1 and x_2 that player 0 plays from s_1 to t_2 respectively from s_2 to t_1. Given x_1 and x_2, we can compute the probabilities $p_1(x_1, x_2) := \Pr_{t_1}^{\bar{\sigma}}(\text{Win}_1)$ and $p_2(x_1, x_2) := \Pr_{t_2}^{\bar{\sigma}}(\text{Win}_2)$ as follows: $p_1(x_1, x_2) = p(1 - x_1)/(1 - x_1 x_2 p^2)$, and $p_2(x_1, x_2) = p(1 - x_2)/(1 - x_1 x_2 p^2)$. To have a Nash equilibrium, it must be the case that $p_1(x_1, x_2), p_2(x_1, x_2) \geq \frac{1}{2}$ since otherwise player 1 or player 2 would prefer to leave the game at t_1 or t_2, respectively, where they could obtain payoff $\frac{1}{2}$ immediately. Vice versa, if $p_1(x_1, x_2), p_2(x_1, x_2) \geq \frac{1}{2}$ then $\bar{\sigma}$ is a subgame-perfect equilibrium with expected payoff $(1 - p)/(1 - x_1 x_2 p^2) \geq 1 - p$ for player 3.

Hence, to determine the maximum payoff for player 3 in a stationary Nash or subgame-perfect equilibrium, we have to maximise $(1-p)/(1-x_1 x_2 p^2)$, the expected payoff for player 3, under the constraints $p_1(x_1, x_2), p_2(x_1, x_2) \geq \frac{1}{2}$ and $0 \leq x_1, x_2 \leq 1$. We claim that the maximum is reached only if $x_1 = x_2$. If e.g. $x_1 > x_2$, then we can achieve a higher payoff for player 3 by setting $x_2' := x_1$, and the constraints are still satisfied:

$$\frac{p(1 - x_2')}{1 - x_1 x_2' p^2} = \frac{p(1 - x_1)}{1 - x_1^2 p^2} \geq \frac{p(1 - x_1)}{1 - x_1 x_2 p^2} \geq \frac{1}{2} \,.$$

Thus, it suffices to maximise $(1 - p)/(1 - x^2 p^2)$ subject to $p(1 - x)/(1 - x^2 p^2) \geq \frac{1}{2}$ and $0 \leq x \leq 1$. Since $\frac{1}{2} \leq p < 1$, this is equivalent to maximising $(1-p)/(1-x^2 p^2)$ subject to $p^2 x^2 - 2px + 2p - 1 \geq 0$ and $0 \leq x \leq 1$. The roots of the former polynomial are $(1 \pm \sqrt{2 - 2p})/p$, but $(1 + \sqrt{2 - 2p})/p > 1$ for $\frac{1}{2} \leq p < 1$. Therefore, any solution must be less than (or equal to) $x_0 := (1 - \sqrt{2 - 2p})/p$. In fact, we always have $0 \leq x_0 < 1$ for $p \in (\frac{1}{2}, 1)$. Therefore, x_0 is the optimal solution, and the maximal payoff for player 3 does indeed equal

$$\frac{1 - p}{1 - x_0^2 p^2} = \frac{1 - p}{1 - (1 - \sqrt{2 - 2p})^2} = \frac{\sqrt{2 - 2p} - p + 1}{2p + 2} \,. \qquad \square$$

Proof (of Theorem 4.7). Given an instance $(d_1, \ldots, d_n, k)$ of SqrtSum, where without loss of generality $n > 0$, $d_i > 0$ for each $i = 1, \ldots, n$ and $k \leq d := \sum_{i=1}^{n} d_i$, we construct a four-player SSMG $(\mathcal{G}, v_0)$ such that the following statements are equivalent:

1. $\sum_{i=1}^{n} \sqrt{d_i} \geq k$;

2. $(\mathcal{G}, v_0)$ has a stationary subgame-perfect equilibrium where player 0 wins almost surely;

3. $(\mathcal{G}, v_0)$ has a stationary Nash equilibrium where player 0 wins almost surely.

Define $p_i := 1 - d_i/2d^2$ for $i = 1, \ldots, n$. Note that $\frac{1}{2} \leq p_i < 1$ since $0 < d_i \leq d \leq d^2$. For the reduction, we use n copies of the game $\mathcal{G}(p)$, where in the ith copy we set p to p_i. The complete game $\mathcal{G}$ is depicted in Figure 4.2 (a); it can obviously be constructed in polynomial time.

By Lemma 4.8, the maximal payoff player 3 receives in a stationary Nash or subgame-perfect equilibrium of $(\mathcal{G}(p_i), s_0)$ equals

$$\frac{\sqrt{2 - 2p_i} - p_i + 1}{2p_i + 2} = \frac{\sqrt{d_i}/d - (1 - d_i/2d^2) + 1}{4 - d_i/d^2} = \frac{d\sqrt{d_i} + d_i/2}{4d^2 - d_i} .$$

Consequently, the maximal payoff player 3 receives in a stationary Nash or subgame-perfect equilibrium of $(\mathcal{G}, v_1)$ equals

$$q := \sum_{i=1}^{n} \frac{4d^2 - d_i}{4d^2 n} \cdot \frac{d\sqrt{d_i} + d_i/2}{4d^2 - d_i} = \sum_{i=1}^{n} \frac{\sqrt{d_i}}{4dn} + \sum_{i=1}^{n} \frac{d_i}{8d^2 n} = \sum_{i=1}^{n} \frac{\sqrt{d_i}}{4dn} + \frac{1}{8dn} .$$

To complete the proof, we need to establish the equivalence of 1.–3.

(1. $\Rightarrow$ 2.) Assume that $\sum_{i=1}^{n} \sqrt{d_i} \geq k$. Then $q \geq (2k+1)/8dn$, and any stationary subgame-perfect equilibrium $\bar{\sigma}$ of $(\mathcal{G}, v_1)$ with this payoff for player 3 can be extended to a stationary subgame-perfect equilibrium of $(\mathcal{G}, v_0)$ where player 0 wins almost surely by setting $\bar{\sigma}(v_1 \mid v_0) = 1$.

(2. $\Rightarrow$ 3.) Trivial.

(3. $\Rightarrow$ 1.) Assume that $(\mathcal{G}, v_0)$ has a stationary Nash equilibrium where player 0 wins almost surely, but $\sum_{i=1}^{n} \sqrt{d_i} < k$. Then $q < (2k+1)/8dn$, and in every stationary Nash equilibrium of $(\mathcal{G}, v_0)$ player 3 leaves the game at v_0, which gives payoff 0 to player 0, a contradiction. $\qquad\square$

In Chapter 3, we have seen that not every SSMG admits a stationary Nash equilibrium (Proposition 3.13). As for positional equilibria, we can thus ask whether a given game has a stationary Nash equilibrium at all. With a construction similar to the one used in the proof of Corollary 4.4, we can infer from Theorems 4.3 and 4.7 that this problem is both NP-hard and SqrtSum-hard, even for four-player SSMGs.

Corollary 4.9. Deciding whether in a given four-player SSMG there exists a stationary Nash equilibrium is both NP-hard and SqrtSum-hard.

Remark. The positive results of Sections 4.1 and 4.2 can easily be extended to equilibria in pure or randomised strategies with a memory of a fixed size $k \in \mathbb{N}$: a nondeterministic algorithm can guess a memory $\mathfrak{M}$ of size k and then look for a positional, respectively stationary, equilibrium in the product of the original game $\mathcal{G}$ with the memory $\mathfrak{M}$. Hence, for any fixed $k \in \mathbb{N}$, we can decide in PSPACE (NP) the existence of a randomised (pure) equilibrium of size k with payoff $\geq \bar{x}$ and $\leq \bar{y}$.

4.3 Pure and randomised equilibria

In this section, we show that all of the following problems are undecidable: PureNE, PureSPE, NE and SPE. In fact, we prove that the qualitative fragments of these problems are not recursively enumerable. The proof proceeds by a single reduction from an undecidable problem about *deterministic two-counter machines*.

Let $\Gamma := \{\mathrm{inc}(j), \mathrm{dec}(j), \mathrm{zero}(j) : j = 1, 2\}$. A two-counter machine $\mathcal{M}$ is of the form $\mathcal{M} = (Q, q_0, \delta)$, where

- Q is a finite set of *states*,
- $q_0 \in Q$ is the *initial state*, and
- $\delta \subseteq Q \times \Gamma \times Q$ is a *transition relation*.

For $q \in Q$ let $\delta(q) := \{(\gamma, q') \in \Gamma \times Q : (q, \gamma, q') \in \delta\}$. We call $\mathcal{M}$ *deterministic* if for each $q \in Q$ either $\delta(q) = \varnothing$, or $\delta(q) = \{(\mathrm{inc}(j), q')\}$ for some $j \in \{1, 2\}$ and $q' \in Q$, or $\delta(q) = \{(\mathrm{zero}(j), q_1), (\mathrm{dec}(j), q_2)\}$ for some $j \in \{1, 2\}$ and $q_1, q_2 \in Q$.

A configuration of $\mathcal{M}$ is a triple $C = (q, i_1, i_2) \in Q \times \mathbb{N} \times \mathbb{N}$, where q denotes the current state and i_j denotes the current value of counter j. A configuration $C' = (q', i_1', i_2')$ is a *successor* of configuration $C = (q, i_1, i_2)$, denoted by $C \vdash C'$, if there exists a "matching" transition $(q, \gamma, q') \in \delta$. For example, $(q, i_1, i_2) \vdash (q', i_1 + 1, i_2)$ if and only if $(q, \mathrm{inc}(1), q') \in \delta$. The instruction $\mathrm{zero}(j)$ performs a *zero test*: $(q, i_1, i_2) \vdash (q', i_1, i_2)$ if and only if $i_1 = 0$ and $(q, \mathrm{zero}(1), q') \in \delta$, or $i_2 = 0$ and $(q, \mathrm{zero}(2), q') \in \delta$.

A *partial computation* of $\mathcal{M}$ is a finite or infinite sequence $\rho = \rho(0)\rho(1)\dots$ of configurations such that $\rho(0) \vdash \rho(1) \vdash \cdots$ and $\rho(0) = (q_0, 0, 0)$ (the *initial configuration*). A partial computation of $\mathcal{M}$ is a *computation* of $\mathcal{M}$ if it is infinite

or ends in a configuration C for which there is no C' with $C \vdash C'$. Note that each deterministic two-counter machine has a unique computation.

The *halting problem* is to decide, given a machine $\mathcal{M}$, whether the computation of $\mathcal{M}$ is finite. It is well-known that deterministic two-counter machines are Turing powerful, which makes the halting problem and its dual, the *non-halting problem*, undecidable, even when restricted to deterministic two-counter machines. In fact, the non-halting problem for deterministic two-counter machines is not recursively enumerable.

Theorem 4.10. PureNE, PureSPE, NE and SPE are not recursively enumerable, even for 10-player SSMGs.

To prove Theorem 4.10, we give a reduction from the non-halting problem for deterministic two-counter machines. Our aim is thus to compute from a machine $\mathcal{M}$ a 10-player SSMG $(\mathcal{G}, v_0)$ such that the following statements are equivalent:

1. the computation of $\mathcal{M}$ is infinite;
2. $(\mathcal{G}, v_0)$ has a pure subgame-perfect equilibrium in which player 0 wins almost surely;
3. $(\mathcal{G}, v_0)$ has a Nash equilibrium in which player 0 wins almost surely.

Without loss of generality, we assume that in $\mathcal{M}$ there is no zero test that is followed by another zero test: if $(\text{zero}(j), q') \in \delta(q)$, then $|\delta(q')| \leq 1$.

The game $\mathcal{G}$ is played by players 0, 1 and eight other players A_j^t and B_j^t, indexed by $j \in \{1, 2\}$ and $t \in \{0, 1\}$. Intuitively, player 0 and player 1 build up the computation of $\mathcal{M}$: player 0 updates the counters, and player 1 chooses transitions. The other players make sure that player 0 updates the counters correctly: if player 0 cheats or the computation halts, one of them will prefer to play a strategy that gives a bad payoff to player 0. More precisely, in every step of the computation, the players A_j^0 and A_j^1 make sure that the value of counter j is not too high, and the players B_j^0 and B_j^1 make sure that the value of counter j is not too low. Hereby, they alternate: the first step of the computation is monitored by the players A_j^0 and B_j^0, the second step by the players A_j^1 and B_j^1, the third step again by the players A_j^0 and B_j^0, and so on.

Let $\Gamma' := \Gamma \cup \{\text{init}\}$. For each $q \in Q$, each $\gamma \in \Gamma'$, each $j \in \{1, 2\}$ and each $t \in \{0, 1\}$, the game $\mathcal{G}$ contains the gadgets $S_{\gamma, q}^t$, I_q^t and $C_{\gamma, j}^t$, which are depicted in Figure 4.3. For better readability, terminal vertices are depicted as squares; the label indicates which players win. The initial vertex of $\mathcal{G}$ is $v_0 := v_{\text{init}, q_0}^0$.
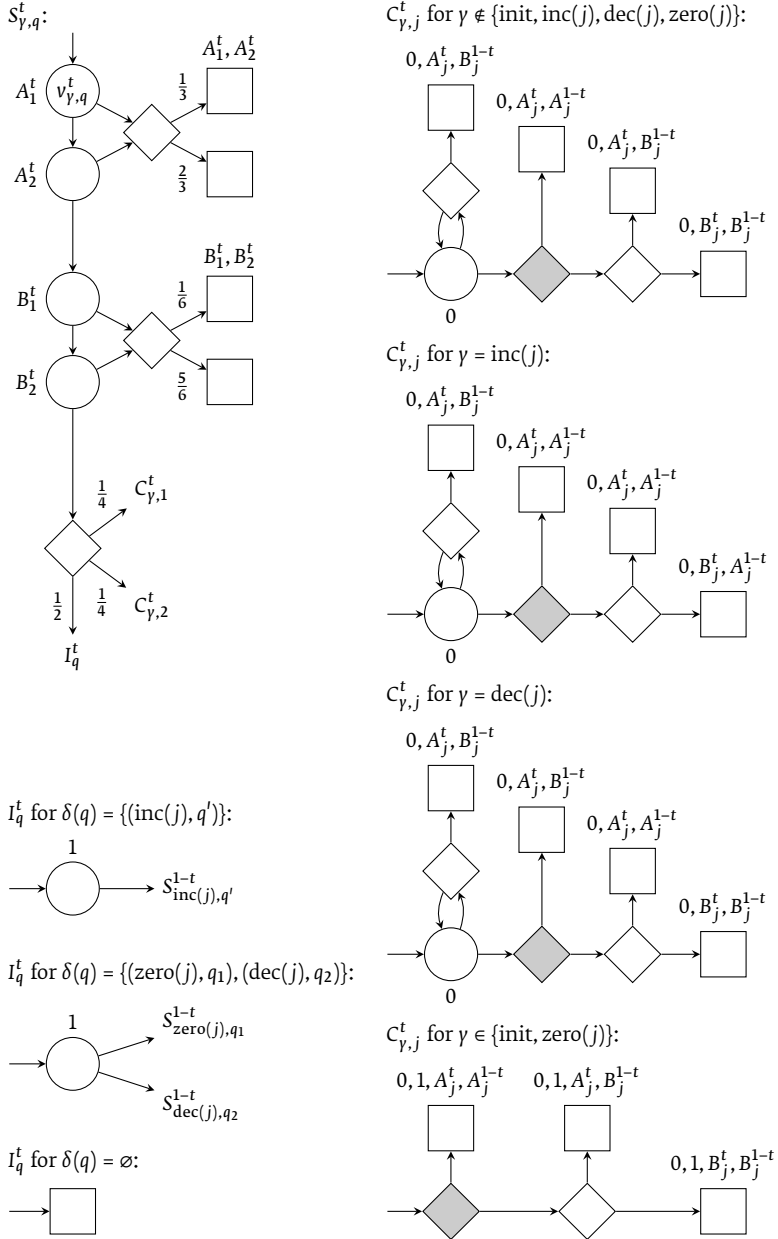
Figure 4.3. Simulating a two-counter machine.

Note that in the gadget $S^t_{\gamma,q}$, each of the players A^t_j and B^t_j may *quit* the game, which gives her a payoff of $\frac{1}{3}$ or $\frac{1}{6}$, respectively, but payoff 0 to players 0 and 1.

It will turn out that player 1 will play a pure strategy in any Nash equilibrium of $(\mathcal{G}, v_0)$ where player 0 wins almost surely, except possibly for histories that are not consistent with the equilibrium. Formally, we say that a strategy profile $\bar{\sigma}$ of $(\mathcal{G}, v_0)$ is *safe* if for all histories xv consistent with $\bar{\sigma}$ and ending in a vertex $v \in I^t_q$ there exists $w \in V$ with $\sigma_1(w \mid xv) = 1$.

For a safe strategy profile $\bar{\sigma}$ where player 0 wins almost surely, let $x_0 v_0 \prec x_1 v_1 \prec x_2 v_2 \prec \cdots$ (where $x_i \in V^*$, $v_i \in V$ and $x_0 = \varepsilon$) be the unique sequence containing all histories xv of $(\mathcal{G}, v_0)$ that are consistent with $\bar{\sigma}$ and end in a vertex v of the form $v = v^t_{\gamma,q}$. This sequence is infinite because player 0 wins almost surely. Additionally, let $q_0, q_1, \ldots$ be the corresponding sequence of states and $\gamma_0, \gamma_1, \ldots$ be the corresponding sequence of instructions, i.e. $v_n = v^0_{\gamma_n, q_n}$ or $v_n = v^1_{\gamma_n, q_n}$ for all $n \in \mathbb{N}$. For each $j \in \{1, 2\}$ and $n \in \mathbb{N}$, we set:

$$a^n_j := \Pr^{\bar{\sigma}}_{v_0}(\text{player } A^{n \bmod 2}_j \text{ wins} \mid x_n v_n \cdot V^\omega);$$
$$b^n_j := \Pr^{\bar{\sigma}}_{v_0}(\text{player } B^{n \bmod 2}_j \text{ wins} \mid x_n v_n \cdot V^\omega).$$

Note that at every terminal vertex of the counter gadgets $C^t_{\gamma,j}$ and $C^{1-t}_{\gamma,j}$ either player A^t_j or player B^t_j wins. For each j, the conditional probability that, given the history $x_n v_n$, we reach such a vertex is $\sum_{k \in \mathbb{N}} 1/2^k \cdot \frac{1}{4} = \frac{1}{2}$. Hence, $a^n_j = \frac{1}{2} - b^n_j$ for all $n \in \mathbb{N}$. We say that $\bar{\sigma}$ is *stable* if $a^n_j = \frac{1}{3}$ or, equivalently, $b^n_j = \frac{1}{6}$ for each $j \in \{1, 2\}$ and for all $n \in \mathbb{N}$.

Finally, for each $j \in \{1, 2\}$ and $n \in \mathbb{N}$, we define a number $c^n_j \in [0,1]$ as follows: After the history $x_n v_n$, with probability $\frac{1}{4}$ the play enters the counter gadget $C^{n \bmod 2}_{\gamma_n, j}$. The number c^n_j is defined as the probability of subsequently reaching a grey-coloured vertex. Note that, by the construction of $\mathcal{G}$, it holds that $c^n_j = 1$ if $\gamma_n = \text{zero}(j)$ or $\gamma_n = \text{init}$. In particular, $c^0_1 = c^0_2 = 1$.

Lemma 4.11. Let $\bar{\sigma}$ be a safe strategy profile of $(\mathcal{G}, v_0)$ in which player 0 wins almost surely. Then $\bar{\sigma}$ is stable if and only if

$$c^{n+1}_j = \begin{cases} \frac{1}{2} \cdot c^n_j & \text{if } \gamma_{n+1} = \text{inc}(j), \\ 2 \cdot c^n_j & \text{if } \gamma_{n+1} = \text{dec}(j), \\ c^n_j = 1 & \text{if } \gamma_{n+1} = \text{zero}(j), \\ c^n_j & \text{otherwise}, \end{cases} \tag{4.1}$$

for each $j \in \{1, 2\}$ and $n \in \mathbb{N}$.

To prove the lemma, consider a safe strategy profile $\bar{\sigma}$ of $(\mathcal{G}, v_0)$ in which player 0 wins almost surely. For each $j \in \{1, 2\}$ and $n \in \mathbb{N}$, set

$$p_j^n := \mathrm{Pr}_{v_0}^{\bar{\sigma}}(\text{player } A_j^{n \bmod 2} \text{ wins} \mid x_n v_n \cdot V^\omega \setminus x_{n+2} v_{n+2} \cdot V^\omega).$$

The following claim relates the numbers a_j^n and p_j^n.

Claim. Let $j \in \{1, 2\}$. Then $a_j^n = \frac{1}{3}$ for all $n \in \mathbb{N}$ if and only if $p_j^n = \frac{1}{4}$ for all $n \in \mathbb{N}$.

Proof. ($\Rightarrow$) Assume that $a_j^n = \frac{1}{3}$ for all $n \in \mathbb{N}$. We have $a_j^n = p_j^n + \frac{1}{4} \cdot a_j^{n+2}$ and therefore $\frac{1}{3} = p_j^n + \frac{1}{12}$ for all $n \in \mathbb{N}$. Hence, $p_j^n = \frac{1}{4}$ for all $n \in \mathbb{N}$.

($\Leftarrow$) Assume that $p_j^n = \frac{1}{4}$ for all $n \in \mathbb{N}$. Since $a_j^n = p_j^n + \frac{1}{4} \cdot a_j^{n+2}$ for all $n \in \mathbb{N}$, the numbers a_j^n must satisfy the following recurrence: $a_j^{n+2} = 4a_j^n - 1$. Since all the numbers a_j^n are probabilities, $0 \le a_j^n \le 1$ for all $n \in \mathbb{N}$. It is easy to see that the only values for a_j^0 and a_j^1 such that $0 \le a_j^n \le 1$ for all $n \in \mathbb{N}$ are $a_j^0 = a_j^1 = \frac{1}{3}$. But this implies that $a_j^n = \frac{1}{3}$ for all $n \in \mathbb{N}$. $\square$

Proof (of Lemma 4.11). By the previous claim, we only need to show that $p_j^n = \frac{1}{4}$ if and only if (4.1) holds. Let $j \in \{1, 2\}$, $n \in \mathbb{N}$ and $t = n \bmod 2$. The probability p_j^n can be expressed as the sum of the probability that the play reaches a terminal vertex that is winning for player A_j^t inside $C_{\gamma_n, j}^t$ and the probability that the play reaches such a vertex inside $C_{\gamma_{n+1}, j}^{1-t}$. The first probability does not depend on γ_n, but the second depends on γ_{n+1}. Let us consider the case $\gamma_{n+1} = \mathrm{inc}(j)$. In this case, the aforementioned sum is equal to

$$\frac{1}{4} \cdot \left(1 - \frac{1}{4} \cdot c_j^n\right) + \frac{1}{8} \cdot c_j^{n+1} = \frac{1}{4} - \frac{1}{16} \cdot c_j^n + \frac{1}{8} \cdot c_j^{n+1}.$$

Obviously, this sum is equal to $\frac{1}{4}$ if and only if $c_j^{n+1} = \frac{1}{2} \cdot c_j^n$. For any other value of γ_{n+1}, the argumentation is similar. $\square$

To establish the reduction, we need to show that the following statements are equivalent:

1. the computation of $\mathcal{M}$ is infinite;
2. $(\mathcal{G}, v_0)$ has a pure subgame-perfect equilibrium in which player 0 wins almost surely;
3. $(\mathcal{G}, v_0)$ has a Nash equilibrium in which player 0 wins almost surely.

(1. $\Rightarrow$ 2.) Assume that the computation $\rho = \rho(0)\rho(1) \ldots$ of $\mathcal{M}$ is infinite. Player 0's pure equilibrium strategy σ_0 can be described as follows: for a history that ends at the unique vertex $v \in C_{\gamma, j}^t$ controlled by player 0 after

visiting a vertex of the form $v_{\gamma',q}^t$ or $v_{\gamma',q}^{1-t}$ exactly $n > 0$ times and v exactly $k \geq 0$ times, player 0 plays to the grey-coloured successor vertex if k is greater than or equal to the value of counter j in configuration $\rho(n-1)$; otherwise, player 0 plays to the other successor vertex.

The only place where player 1 has a choice is the sole vertex inside the gadget I_q^t for $\delta(q) = \{(\text{zero}(j), q_1), (\text{dec}(j), q_2)\}$. If the play arrives at such a vertex after visiting a vertex of the form $v_{\gamma,q'}^t$ or $v_{\gamma,q'}^{1-t}$ exactly $n > 0$ times, then player 1's pure strategy σ_1 prescribes to play to $S_{\text{zero}(j),q_1}^{1-t}$ if the value of counter j in configuration $\rho(n-1)$ is zero and to $S_{\text{dec}(j),q_2}^{1-t}$ if the value of counter j in configuration $\rho(n-1)$ is non-zero.

Any other player's pure strategy is defined as follows: after a history ending in $S_{\gamma,q}^t$, the strategy prescribes to quit the game if and only if the history is not compatible with ρ (i.e. if the corresponding sequence of instructions does not match ρ).

Note that the resulting strategy profile $\bar{\sigma}$ is safe. Moreover, since player 0 and player 1 follow the computation of $\mathcal{M}$, a terminal vertex inside one of the counter gadgets $C_{\gamma,j}^t$ is reached with probability 1. Hence, player 0 wins almost surely in $\bar{\sigma}$. Finally, by the definition of $\bar{\sigma}$, (4.1) holds, and we can conclude from Lemma 4.11 that $\bar{\sigma}$ is stable.

We claim that $\bar{\sigma}$ is, in fact, a subgame-perfect equilibrium of $(\mathcal{G}, v_0)$: It is obvious that player 0 cannot improve her payoff, even for histories where she receives payoff < 1. If player 1 deviates, we reach a history that is not compatible with ρ. Hence, player A_1^0 or A_2^0 will quit the game, which ensures that player 1 will receive payoff 0 after this history. Finally, since $\bar{\sigma}$ is stable, none of the players A_j^t or B_j^t can improve her payoff after a history that is consistent with $\bar{\sigma}$. For all other histories, this follows immediately from the definition of $\bar{\sigma}$. For instance, if player A_1^0 changes her strategy after such a history and decides not to quit the game, then she will still receive payoff $\frac{1}{3}$, because player A_2^0 will still quit the game.

(2. $\Rightarrow$ 3.) Trivial.

(3. $\Rightarrow$ 1.) Assume that $\bar{\sigma}$ is a Nash equilibrium of $(\mathcal{G}, v_0)$ in which player 0 wins almost surely. In order to apply Lemma 4.11, we first prove that $\bar{\sigma}$ is safe. Towards a contradiction, assume that there exists a history xv ending in a vertex $v \in I_q^t$ such that $\Pr_{v_0}^{\bar{\sigma}}(xv \cdot V^\omega) > 0$ and $\sigma_1(xv)$ assigns probability > 0 to two distinct successor vertices. Hence, $\delta(q) = \{(\text{zero}(j), q_1), (\text{dec}(j), q_2)\}$ for some $j \in \{1, 2\}$ and $q_1, q_2 \in Q$. By our assumption that there are no consecutive zero tests and since player 0 wins almost surely,

$$\Pr_{v_0}^{\bar{\sigma}}(\text{player 1 wins} \mid xv \cdot v_{\text{zero}(j),q_1}^{1-t} \cdot V^{\omega}) \geq \tfrac{1}{4},$$

but

$$\Pr_{v_0}^{\bar{\sigma}}(\text{player 1 wins} \mid xv \cdot v_{\text{dec}(j),q_2}^{1-t} \cdot V^{\omega}) \leq \tfrac{1}{6}.$$

Hence, player 1 could improve her payoff by playing to $v_{\text{zero}(j),q_1}^{1-t}$ with probability 1, a contradiction to $\bar{\sigma}$ being a Nash equilibrium.

To apply Lemma 4.11 and obtain (4.1), it remains to be shown that $\bar{\sigma}$ is stable. In order to derive a contradiction, assume that there exists $j \in \{1, 2\}$ and $n \in \mathbb{N}$ such that either $a_j^n < \tfrac{1}{3}$ or $a_j^n > \tfrac{1}{3}$ (i.e. $b_j^n < \tfrac{1}{6}$). In the former case, player $A_j^{n \bmod 2}$ could improve her payoff by quitting the game after history $x_n v_n$, while in the latter case, player $B_j^{n \bmod 2}$ could improve her payoff by quitting the game, again a contradiction to $\bar{\sigma}$ being a Nash equilibrium.

From $c_j^0 = 1$ and (4.1), it follows that each c_j^n is of the form $c_j^n = 1/2^i$ where $i \in \mathbb{N}$. We denote by i_j^n the unique number i such that $c_j^n = 1/2^i$ and set $\rho(n) = (q_n, i_1^n, i_2^n)$ for each $n \in \mathbb{N}$. We claim that $\rho := \rho(0)\rho(1)\ldots$ is in fact the computation of $\mathcal{M}$. In particular, this computation is infinite. It suffices to verify the following two properties:

- $\rho(0) = (q_0, 0, 0)$;
- $\rho(n) \vdash \rho(n + 1)$ for all $n \in \mathbb{N}$.

The first property is immediate. To prove the second property, let $\rho(n) = (q, i_1, i_2)$ and $\rho(n + 1) = (q', i_1', i_2')$. Hence, v_n lies inside $S_{\gamma,q}^t$, and v_{n+1} lies inside $S_{\gamma',q'}^{1-t}$ for suitable γ, γ' and $t = n \bmod 2$. We only prove the claim for $\delta(q) = \{(\text{zero}(1), q_1), (\text{dec}(1), q_2)\}$; the other cases are similar. Note that, by the construction of the gadget I_q^t, it must be the case that either $q' = q_1$ and $\gamma' = \text{zero}(1)$, or $q' = q_2$ and $\gamma' = \text{dec}(1)$. By (4.1), if $\gamma' = \text{zero}(1)$, then $i_1' = i_1 = 0$ and $i_2' = i_2$, and if $\gamma' = \text{dec}(1)$, then $i_1' = i_1 - 1$ and $i_2' = i_2$. This implies $\rho(n) \vdash \rho(n + 1)$: on the one hand, if $i_1 = 0$, then $i_1' \neq i_1 - 1$, which implies $\gamma' \neq \text{dec}(1)$ and thus $\gamma' = \text{zero}(1)$, $q' = q_1$ and $i_1' = i_1 = 0$; on the other hand, if $i_1 > 0$, then $\gamma' \neq \text{zero}(1)$ and thus $\gamma' = \text{dec}(1)$, $q' = q_2$ and $i_1' = i_1 - 1$. $\qquad\square$

Remark. For the problems PureNE and PureSPE, we can strengthen Theorem 4.10 slightly by showing undecidability already for 9-player SSMGs. This can be achieved by merging player 0 and player 1 in the game described in the proof of Theorem 4.10: if player 0 is a priori restricted to play a pure strategy, she cannot cheat by playing to both $S_{\text{zero}(j),q_1}^t$ and $S_{\text{dec}(j),q_2}^t$ with positive probability.

The proof of Theorem 4.10 can also be viewed as a proof for the undecidability of a problem about the logic PCTL (*probabilistic computation tree logic*), introduced by Hansson & Jonsson (1994). PCTL is evaluated over labelled Markov chains and replaces the universal and existential path quantifiers of CTL by a family of probabilistic quantifiers $P^{\sim x}$, where $\sim$ is a comparison operator and $x \in [0,1]$ is a rational probability. For example, the formula $P^{=1/2}F\,Q$ holds in state v if (and only if) the probability of reaching a state labelled with Q from v equals $\frac{1}{2}$.

By employing a similar reduction to ours, Brázdil et al. (2006) proved the undecidability of the following problem: given a labelled Markov decision process $(\mathcal{G}, v_0)$ and a PCTL formula φ, decide whether the controller has a strategy σ such that the Markov chain $(\mathcal{G}^\sigma, v_0)$ is a model of φ. We can prove a stronger result, namely that there exists a *fixed* PCTL formula φ, which only contains the quantifiers $P^{=x}F$ and $P^{=x}G$, for which the problem is undecidable. It suffices to add propositions $A_1^0, A_1^1, A_2^0, A_2^1, Q, Q_1, Q_2, T, Z_0$ and Z_1 according to the following rules:

1. if v is a terminal vertex that is winning for player $A \in \{A_1^0, A_1^1, A_2^0, A_2^1\}$, then label v with A;

2. if $v \in I_q^t$, then label v with Q;

3. if $v = v_{\gamma,q}^t$ for $\gamma \neq \text{zero}(j)$, then label v with Q_1; if $v = v_{\text{zero}(j),q}^t$, then label v with Q_2;

4. if v is a terminal vertex that is winning for player 0, then label v with T;

5. if $v = v_{\gamma,q}^0$, then label v with Z_0; if $v = v_{\gamma,q}^1$, then label v with Z_1.

To obtain an MDP, we make all non-stochastic vertices controlled by player 0. Finally, the PCTL formula for which we prove undecidability is

$$P^{=1}F\,T \wedge \bigwedge_{t=0,1} P^{=1}G\,(Z_t \to P^{=1/3}F\,A_1^t \wedge P^{=1/3}F\,A_2^t) \wedge P^{=1}G\,(Q \to P^{=1}F\,Q_1 \vee P^{=1}F\,Q_2).$$

The first part of the formula states that player 0 wins almost surely, the second part requires the strategy to be stable, and the last part of the formula requires the strategy to be safe.

An immediate corollary to this result is that there exists a fixed formula of *stochastic game logic* (Baier et al. 2007) for which the model-checking problem (with respect to pure or randomised strategies) is undecidable.

4.4 Finite-state equilibria

We can use the construction in the proof of Theorem 4.10 to show that Nash and subgame-perfect equilibria may require infinite memory, even if we are only interested in whether a player wins with probability 0 or 1.

Proposition 4.12. There exists an SSMG that has a pure subgame-perfect equilibrium where player 0 wins almost surely but that has no finite-state Nash equilibrium where player 0 wins with positive probability.

Proof. Consider the game $(\mathcal{G}, v_0)$ constructed in the proof of Theorem 4.10 for the machine $\mathcal{M}$ with the single transition $(q_0, \mathrm{inc}(1), q_0)$. We modify this game by adding a new initial vertex v_1, which is controlled by a new player, player 2, and from where she can either move to v_0 or to a new terminal vertex where she receives payoff 1 and every other player receives payoff 0. Additionally, player 2 wins at every terminal vertex of the game $\mathcal{G}$ that is winning for player 0. Let us denote the modified game by $\mathcal{G}'$.

Since the computation of $\mathcal{M}$ is infinite, the game $(\mathcal{G}, v_0)$ has a pure subgame-perfect equilibrium where player 0 wins almost surely. This equilibrium induces a pure subgame-perfect equilibrium of $(\mathcal{G}', v_1)$ where both player 0 and player 2 win almost surely.

Now assume that there exists a finite-state Nash equilibrium of $(\mathcal{G}', v_1)$ where player 0 wins with positive probability. Such an equilibrium induces a finite-state Nash equilibrium $\bar{\sigma}$ of $(\mathcal{G}, v_0)$ where player 2, and thus also player 0, wins almost surely: otherwise, player 2 would prefer to quit the game. Using the same notation as in the proof of Theorem 4.10, it follows from Lemma 4.11 that $c_1^n = 1/2^n$ for each $n \in \mathbb{N}$. But this is impossible if $\bar{\sigma}$ is a finite-state strategy profile. $\square$

Propositions 3.26 and 4.12 (together with Example 3.4) imply that the decision problems NE, SPE, FinNE, FinSPE, PureNE, PureSPE, PureFinNE and PureFinSPE are pairwise distinct. Another way to see that PureNE and PureFinNE are distinct is to observe that PureFinNE is recursively enumerable: to decide whether an SMG $(\mathcal{G}, v_0)$ has a pure finite-state Nash equilibrium with payoff $\geq \bar{x}$ and $\leq \bar{y}$, one can just enumerate all possible pure finite-state profiles $\bar{\sigma}$ and check for each of them whether it constitutes a Nash equilibrium with the desired properties by analysing the finite Markov chain $\mathcal{G}^{\bar{\sigma}}$ and the finite MDPs $\mathcal{G}^{\bar{\sigma}-i}$. Hence, to prove that PureFinNE is undecidable, we cannot reduce from the non-halting problem. Instead, we reduce
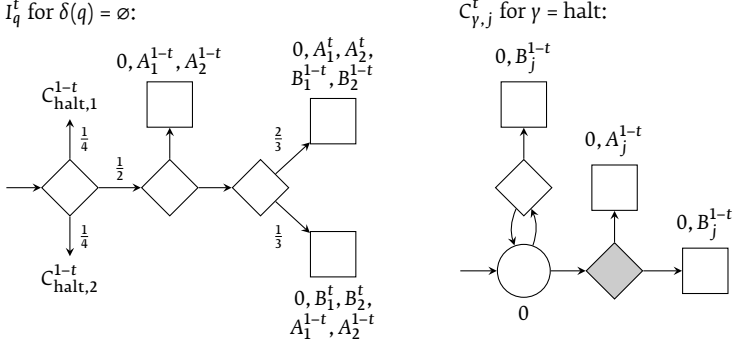
Figure 4.4. Reducing from the halting problem.

from the halting problem (which is recursively enumerable itself). The same reduction proves that PureFinSPE, FinNE and FinSPE are undecidable.

Theorem 4.13. PureFinNE, PureFinSPE, FinNE and FinSPE are undecidable, even for 14-player SSMGs.

Proof (Sketch). The construction is similar to the one for proving the undecidability of NE. Given a two-counter machine $\mathcal{M}$, we modify the SSMG $\mathcal{G}$ constructed in the proof of Theorem 4.10 by adding another counter (together with four more players for checking whether the counter is updated correctly) that has to be incremented in each step. Moreover, the gadget $I_{\gamma,q}$ for $\delta(q) = \emptyset$ is replaced by the gadget shown in Figure 4.4, and a new instruction halt is added, together with a suitable gadget $C_{\text{halt},j}^t$, also depicted in Figure 4.4. Let us denote the new game by $\mathcal{G}'$. Now, if $\mathcal{M}$ does not halt, any Nash equilibrium of $(\mathcal{G}', v_0)$ where player 0 wins with probability 1 needs infinite memory: to win almost surely, player 0 must follow the computation of $\mathcal{M}$ and increment the new counter at each step, which requires infinite memory. On the other hand, if $\mathcal{M}$ halts, there exists a pure finite-state subgame-perfect equilibrium of $(\mathcal{G}', v_0)$ in which player 0 wins almost surely. (The arguments for the existence of such an equilibrium are the same as in the proof of Theorem 4.10; since $\mathcal{M}$ halts, the equilibrium can be implemented with finite memory). □

Remark. With the same reasoning as for PureNE and PureSPE, we can eliminate one player in the reductions for PureFinNE and PureFinSPE. Hence, these problems are already undecidable for 13-player SSMGs.

Table 4.1. The complexity of NE, SPE and their relatives.

	Pure	Randomised
Stationary	NP-complete	PSPACE NP-hard + SqrtSum-hard
Finite-state	undecidable (r.e.)	undecidable
General	undecidable (not r.e.)	undecidable (not r.e.)

4.5 Summary of results

Table 4.1 summarises our findings on the complexity of NE, SPE and their relatives. The rows of the table correspond to the restrictions of strategies with respect to memory, whereas the columns of the table indicate whether randomisation is allowed or not. The complexity bounds shown hold for both Nash and subgame-perfect equilibria in SSMGs as well as SMGs with Streett, Rabin or Muller objectives. Moreover, each lower bound holds already for the qualitative fragment of the respective decision problem.

5

Decidable Fragments

This chapter is devoted to proving decidability results for fragments of NE. The first fragment, which we call the *strictly qualitative fragment*, arises from NE by restricting the thresholds to be the same binary payoff (i.e. each entry is either 0 or 1). For the second fragment, which we call the *positive-one fragment*, we require that the upper threshold is trivial and the that lower threshold is of the form $(p, 1, \ldots, 1)$, where $p \in [0, 1]$ is an arbitrary rational number. Hence, the strictly qualitative fragment asks for an equilibrium with a binary payoff $\bar{x}$, whereas the positive-one fragment asks for an equilibrium that is almost surely winning for all but one player and that is winning for the remaining player with probability $\geq p$. Finally, we show that the qualitative fragment of NE (and thereby PureNE) is decidable for deterministic games. As in the previous chapter, all games in this chapter are finite.

5.1 The strictly qualitative fragment

In this section, we prove that the problem StrQualNE is decidable. Formally, StrQualNE is the following decision problem:

> Given a finite SMG $(\mathcal{G}, v_0)$ and $\bar{x} \in \{0, 1\}^{\Pi}$, decide whether there exists a Nash equilibrium of $(\mathcal{G}, v_0)$ with payoff $\bar{x}$.

To prove decidability, we first characterise the existence of a Nash equilibrium with a binary payoff in games with prefix-independent objectives.

Characterisation of existence

Given an SMG $\mathcal{G}$ and a player i, we denote by W_i the set of all vertices $v \in V$ with $\mathrm{val}_i^{\mathcal{G}}(v) > 0$.

Proposition 5.1. Let $(\mathcal{G}, v_0)$ be a finite SMG with prefix-independent objectives $\mathrm{Win}_i \subseteq V^\omega$, and let $\bar{x} = (x_i)_{i \in \Pi} \in \{0,1\}^\Pi$. Then the following statements are equivalent:

1. $(\mathcal{G}, v_0)$ has a Nash equilibrium with payoff $\bar{x}$;
2. there exists a strategy profile $\bar{\sigma}$ of $(\mathcal{G}, v_0)$ with payoff $\bar{x}$ such that $\mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(W_i)) = 0$ for each player i with $x_i = 0$;
3. there exists a pure strategy profile $\bar{\sigma}$ of $(\mathcal{G}, v_0)$ with payoff $\bar{x}$ such that $\mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(W_i)) = 0$ for each player i with $x_i = 0$;
4. $(\mathcal{G}, v_0)$ has a pure Nash equilibrium with payoff $\bar{x}$.

If additionally all objectives are ω-regular, then each of the above statements is equivalent to each of the following statements:

5. there exists a pure finite-state strategy profile $\bar{\sigma}$ of $(\mathcal{G}, v_0)$ with payoff $\bar{x}$ such that $\mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(W_i)) = 0$ for each player i with $x_i = 0$;
6. $(\mathcal{G}, v_0)$ has a pure finite-state Nash equilibrium with payoff $\bar{x}$.

Proof. $(1. \Rightarrow 2.)$ Let $\bar{\sigma}$ be a Nash equilibrium of $(\mathcal{G}, v_0)$ with payoff $\bar{x}$. We claim that $\bar{\sigma}$ is already the strategy profile we are looking for: $\mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(W_i)) = 0$ for each player i with $x_i = 0$. Let $i \in \Pi$ be a player with $x_i = 0$. By Lemma 3.3 and since Win_i is prefix-independent, we have $0 = \mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Win}_i \mid xv \cdot V^\omega) \geq \mathrm{val}_i^{\mathcal{G}}(v)$ for all histories xv that are consistent with $\bar{\sigma}$. Hence, $v \in V \setminus W_i$ for all such histories xv, and $\mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(W_i)) = 0$.

$(2. \Rightarrow 3.)$ Let $\bar{\sigma}$ be a strategy profile of $(\mathcal{G}, v_0)$ with payoff $\bar{x}$ such that $\mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(W_i)) = 0$ for each player i with $x_i = 0$. Consider the MDP $\mathcal{M}$ that is obtained from $\mathcal{G}$ by removing all vertices $v \in V$ such that $v \in W_i$ for some player i with $x_i = 0$, merging all players into one, and imposing the objective

$$\mathrm{Win} = \bigcap_{\substack{i \in \Pi \\ x_i = 1}} \mathrm{Win}_i \cap \bigcap_{\substack{i \in \Pi \\ x_i = 0}} V^\omega \setminus \mathrm{Win}_i \,.$$

The MDP $\mathcal{M}$ is well-defined since its domain is a subarena of $\mathcal{G}$. Moreover, the value $\mathrm{val}^{\mathcal{M}}(v_0)$ of $\mathcal{M}$ from v_0 equals 1 because the strategy profile $\bar{\sigma}$ induces a strategy σ in $\mathcal{M}$ satisfying $\mathrm{Pr}_{v_0}^{\sigma}(\mathrm{Win}) = 1$. Since each of the

objectives Win_i is prefix-independent, so is the objective Win. Hence, by Theorem 2.6, $(\mathcal{M}, v_0)$ admits an optimal pure strategy τ. Since $\text{val}^{\mathcal{M}}(v_0) = 1$, we have $\text{Pr}_{v_0}^{\tau}(\text{Win}) = 1$, and τ induces a pure strategy profile of $(\mathcal{G}, v_0)$ with the desired properties.

$(3. \Rightarrow 4.)$ Let $\bar{\sigma}$ be a pure strategy profile of $(\mathcal{G}, v_0)$ with payoff $\bar{x}$ such that $\text{Pr}_{v_0}^{\bar{\sigma}}(\text{Reach}(W_i)) = 0$ for each player i with $x_i = 0$. We show that the requirements of Lemma 3.7 are fulfilled: $\text{Pr}_{v_0}^{\bar{\sigma}}(\text{Win}_i \mid xv \cdot V^{\omega}) \geq \text{val}_i^{\mathcal{G}}(v)$ for each player i and each history xv of $(\mathcal{G}, v_0)$ that is consistent with $\bar{\sigma}$. There are two cases: If $x_1 = 1$, then $\text{Pr}_{v_0}^{\bar{\sigma}}(\text{Win}_i \mid xv \cdot V^{\omega}) = 1$ for all histories xv consistent with $\bar{\sigma}$, and the inequality holds. Otherwise, $x_i = 0$ and $\text{Pr}_{v_0}^{\bar{\sigma}}(\text{Reach}(W_i)) = 0$. Hence, $\text{val}_i^{\mathcal{G}}(v) = 0$ for all histories xv consistent with $\bar{\sigma}$, and the inequality holds as well. Now, by Lemma 3.7, we can extend $\bar{\sigma}$ to a pure Nash equilibrium with payoff $\bar{x}$.

$(4. \Rightarrow 1.)$ Trivial.

Under the additional assumption that all objectives are ω-regular, the implications $(2. \Rightarrow 5.)$ and $(5. \Rightarrow 6.)$ are proven analogously (using Lemma 3.8 instead of Lemma 3.7); the implication $(6. \Rightarrow 1.)$ is trivial. $\qquad\square$

As an immediate consequence of Proposition 5.1, we can conclude that pure finite-state strategies are as powerful as arbitrary randomised strategies as far as the existence of a Nash equilibria with binary payoffs in finite SMGs with ω-regular objectives is concerned.

Corollary 5.2. Let $(\mathcal{G}, v_0)$ be a finite SMG with ω-regular objectives, and let $x \in \{0, 1\}^{\Pi}$. There exists a Nash equilibrium of $(\mathcal{G}, v_0)$ with payoff $\bar{x}$ if and only if there exists a pure finite-state Nash equilibrium of $(\mathcal{G}, v_0)$ with payoff $\bar{x}$.

Proof. The claim follows from Proposition 5.1 and the fact that every SMG with ω-regular objectives can be reduced to one with parity objectives. $\quad\square$

Computational complexity

We can now describe an algorithm that decides StrQualNE for SMGs with Muller objectives. The algorithm relies on Proposition 5.1, which allows us to reduce StrQualNE to an MDP problem.

Formally, given a Muller SMG $\mathcal{G} = (\Pi, V, (V_i)_{i \in \Pi}, \Delta, \chi, (\mathcal{F}_i)_{i \in \Pi})$ and a binary payoff $\bar{x} = (x_i)_{i \in \Pi}$, we define the Markov decision process $\mathcal{G}(\bar{x})$ as follows: Let $Z \subseteq V$ be the set of all vertices v such that $\text{val}_i^{\mathcal{G}}(v) = 0$ for each player i with $x_i = 0$; the set of vertices of $\mathcal{G}(\bar{x})$ is precisely the set Z, with the set of vertices

controlled by player 0 being $Z_0 := \bigcup_{i \in \Pi}(V_i \cap Z)$; if $Z = \varnothing$, we define $\mathcal{G}(\bar{x})$ to be a trivial MDP with the empty set as its objective. The transition relation of $\mathcal{G}(\bar{x})$ is the restriction of Δ to transitions between Z-states. Note that the transition relation of $\mathcal{G}(\bar{x})$ is well-defined since Z is a subarena of $\mathcal{G}$. Finally, the single objective in $\mathcal{G}(\bar{x})$ is $\mathrm{Reach}(T)$ where $T \subseteq Z$ is the union of all end components $U \subseteq Z$ with payoff $\bar{x}$.

Lemma 5.3. Let $(\mathcal{G}, v_0)$ be a finite Muller SMG, and let $\bar{x} \in \{0, 1\}^{\Pi}$. Then $(\mathcal{G}, v_0)$ has a Nash equilibrium with payoff $\bar{x}$ if and only if $\mathrm{val}^{\mathcal{G}(\bar{x})}(v_0) = 1$.

Proof. ($\Rightarrow$) Assume that $(\mathcal{G}, v_0)$ has a Nash equilibrium with payoff $\bar{x}$. By Proposition 5.1, there exists a strategy profile $\bar{\sigma}$ of $(\mathcal{G}, v_0)$ with payoff $\bar{x}$ such that $\mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(V \setminus Z)) = 0$. We claim that $\mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(T)) = 1$. Otherwise, by Lemma 2.3, there would exist an end component $U \subseteq Z$ such that $\mathrm{Pr}_{v_0}^{\bar{\sigma}}(\{\pi \in V^\omega : \mathrm{Inf}(\pi) = U\}) > 0$, and U is either not winning for some player i with $x_i = 1$ or it is winning for some player i with $x_i = 0$. But then $\bar{\sigma}$ cannot have payoff $\bar{x}$, a contradiction. Now, since $\mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(V \setminus Z)) = 0$, the strategy profile $\bar{\sigma}$ induces a strategy σ in $\mathcal{G}(\bar{x})$ such that $\mathrm{Pr}_{v_0}^{\sigma}(X) = \mathrm{Pr}_{v_0}^{\bar{\sigma}}(X)$ for every Borel set $X \subseteq Z^\omega$. In particular, $\mathrm{Pr}_{v_0}^{\sigma}(\mathrm{Reach}(T)) = 1$ and hence $\mathrm{val}^{\mathcal{G}(\bar{x})}(v_0) = 1$.

($\Leftarrow$) Assume that $\mathrm{val}^{\mathcal{G}(\bar{x})}(v_0) = 1$ (in particular, $v_0 \in Z$), and let σ be an optimal strategy in $(\mathcal{G}(\bar{x}), v_0)$. From σ, using Lemma 2.4, we can devise a strategy σ' such that $\mathrm{Pr}_{v_0}^{\sigma'}(\{\pi \in V^\omega : \mathrm{Inf}(\pi) \text{ has payoff } \bar{x}\}) = 1$. Finally, σ' can be extended to a strategy profile $\bar{\sigma}$ of $(\mathcal{G}, v_0)$ with payoff $\bar{x}$ such that $\mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(V \setminus Z)) = 0$. By Proposition 5.1, this implies that $(\mathcal{G}, v_0)$ has a Nash equilibrium with payoff $\bar{x}$. $\qquad\square$

Since the values of an MDP with a reachability objective can be computed in polynomial time, the difficult part lies in computing the MDP $\mathcal{G}(\bar{x})$ from $\mathcal{G}$ and $\bar{x}$ (i.e. its domain Z and the target set T). For Muller SMGs, polynomial space suffices to achieve this. In fact, StrQualNE is Pspace-complete for these games.

Theorem 5.4. StrQualNE is Pspace-complete for Muller SMGs.

Proof. Hardness follows from Theorem 2.18. To prove membership in Pspace, we describe a polynomial-space algorithm for deciding StrQualNE on Muller SMGs: On input $\mathcal{G}, v_0, \bar{x}$, the algorithm starts by computing for each player i with $x_i = 0$ the set of vertices v such that $\mathrm{val}_i^{\mathcal{G}}(v) = 0$, which can be done in polynomial space by Theorem 2.17. The intersection of these sets is the

domain Z of the Markov decision process $\mathcal{G}(\bar{x})$. If v_0 is not contained in this intersection, the algorithm immediately rejects. Otherwise, the algorithm determines the union T of all end components with payoff $\bar{x}$ contained in Z by enumerating all subsets of Z, one at a time, and checking which ones are end components with payoff $\bar{x}$. Finally, the algorithm computes (in polynomial time) the value $\mathrm{val}^{\mathcal{G}(\bar{x})}(v_0)$ of the MDP $\mathcal{G}(\bar{x})$ from v_0 and accepts if this value is 1. In all other cases, the algorithm rejects. The correctness of the algorithm follows immediately from Lemma 5.3. $\qquad\square$

For games with Streett objectives, StrQualNE becomes NP-complete; we start by proving the upper bound.

Theorem 5.5. StrQualNE is in NP for Streett SMGs.

Proof. We describe a nondeterministic polynomial-time algorithm for solving StrQualNE: On input $\mathcal{G}, v_0, \bar{x}$, the algorithm starts by guessing a subarena $Z' \subseteq V$ and for each player i with $x_i = 0$ a positional strategy τ_i of the coalition $\Pi \setminus \{i\}$ in the coalition game $\mathcal{G}_i$. In the next step, the algorithm checks (in polynomial time) whether $\mathrm{val}^{\tau_i}(v) = 1$ for each vertex $v \in Z'$ and each player i with $x_i = 0$. If not, the algorithm rejects immediately. Otherwise, the algorithm proceeds by guessing (at most) $n := |V|$ subsets $U_1, \ldots, U_n \subseteq Z'$ and checks whether they are end components with payoff $\bar{x}$ (which can be done in polynomial time). If yes, the algorithm sets $T' := \bigcup_{j=1}^n U_j$ and computes (in polynomial time) the value $\mathrm{val}^{\mathcal{G}(\bar{x})}(v_0)$ of the MDP $\mathcal{G}(\bar{x})$ from v_0 with Z' substituted for Z and T' substituted for T. If this value equals 1, the algorithm accepts; otherwise, it rejects.

It remains to be shown that the algorithm is correct: On the one hand, if $(\mathcal{G}, v_0)$ has a Nash equilibrium with payoff $\bar{x}$, then the run of the algorithm where it guesses $Z' = Z$, globally optimal positional strategies τ_i (which exist by Theorem 2.11) and end components U_i such that $T' = T$ will be accepting since then, by Lemma 5.3, $\mathrm{val}^{\mathcal{G}(\bar{x})}(v_0) = 1$. On the other hand, in any accepting run of the algorithm we have $Z' \subseteq Z$ and $T' \subseteq T$, and the computed value cannot be higher than $\mathrm{val}^{\mathcal{G}(\bar{x})}(v_0)$; hence, $\mathrm{val}^{\mathcal{G}(\bar{x})}(v_0) = 1$, and Lemma 5.3 guarantees the existence of a Nash equilibrium with payoff $\bar{x}$. $\qquad\square$

The matching lower bound does not only hold for StrQualNE, but also for the analogous problem for subgame-perfect equilibria, which we denote by StrQualSPE. Moreover, both these problems are NP-hard even for deterministic two-player Streett games.
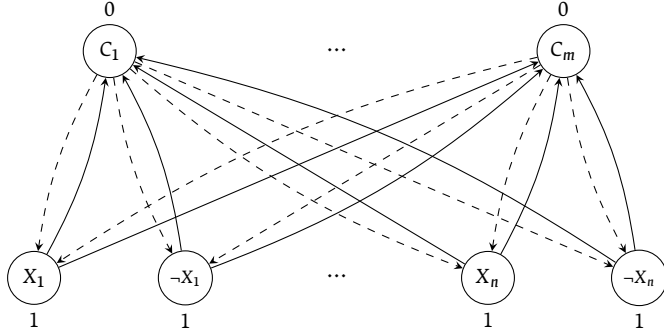
Figure 5.1. Reducing SAT to StrQualNE for games with Streett objectives.

Theorem 5.6. StrQualNE and StrQualSPE are NP-hard for deterministic two-player Streett games.

Proof. The proof is accomplished by a variant of the proof for NP-hardness of the qualitative decision problem for deterministic two-player zero-sum Rabin-Streett games (Emerson & Jutla 1999) and by a reduction from SAT. Given a Boolean formula $\varphi = C_1 \wedge \cdots \wedge C_m$ in conjunctive normal form, where without loss of generality each clause is nonempty, we construct a deterministic two-player Streett game $\mathcal{G}$ as follows: For each clause C, the game $\mathcal{G}$ has a vertex C, which is controlled by player 0, and for each literal L occurring in φ, there is a vertex L, which is controlled by player 1. There are edges from a clause to each literal that occurs in this clause, and from a literal to each clause occurring in φ. The structure of the game is depicted in Figure 5.1. Player 0's objective is given by the single Streett pair $(\varnothing, V)$, i.e. she wins every play of the game, whereas player 1's objective consists of all Streett pairs of the form $(\{X\}, \{\neg X\})$ or $(\{\neg X\}, \{X\})$, i.e. she wins if, for each variable X, either X and $\neg X$ are both visited infinitely often or neither of them is.

Clearly, $\mathcal{G}$ can be constructed from φ in polynomial time. We claim that the following statements are equivalent:

1. φ is satisfiable;
2. $(\mathcal{G}, C_1)$ has a subgame-perfect equilibrium with payoff $(1, 0)$;
3. $(\mathcal{G}, C_1)$ has a Nash equilibrium with payoff $(1, 0)$.

(1. $\Rightarrow$ 2.) Assume that φ is satisfiable, and consider the following positional strategy σ_0 of player 0: whenever the play reaches a clause, then σ_0 plays to

a literal that is mapped to true by the satisfying assignment. This strategy ensures that, for each variable X and after any initial history, at most one of the literals X or $\neg X$ is visited infinitely often. Hence, (σ_0, σ_1) is a subgame-perfect equilibrium of $(\mathcal{G}, C_1)$ with payoff $(1, 0)$ for every strategy σ_1 of player 1.

(2. $\Rightarrow$ 3.) Trivial.

(3. $\Rightarrow$ 1.) Let (σ_0, σ_1) be a Nash equilibrium of $(\mathcal{G}, C_1)$ with payoff $(1, 0)$, and assume that φ is not satisfiable. Consider the two-player zero-sum Rabin-Streett game $\widetilde{\mathcal{G}}$, which is derived from $\mathcal{G}$ by setting player 0's objective to the complement of player 1's objective. We claim that player 1 has a winning strategy in $(\widetilde{\mathcal{G}}, C_1)$, which she could use to improve her payoff in $(\mathcal{G}, C_1)$, a contradiction to (σ_0, σ_1) being a Nash equilibrium. By determinacy, we only need to show that player 0 does not have a winning strategy. Let τ be an optimal positional strategy of player 0 in $(\widetilde{\mathcal{G}}, C_1)$ (which exists by Theorem 2.11). Since φ is unsatisfiable, there must exist a variable X and clauses C and C' such that $\tau(C) = X$ and $\tau(C') = \neg X$. But player 1 can counter this strategy by playing from X to C' and from any other literal to C. Hence, τ is not winning. $\qquad\square$

For games with Rabin objectives, the situation is more delicate. One might think that, because of the duality of Rabin and Streett objectives, StrQualNE is in coNP for SMGs with Rabin objectives.[1] However, as we will see later, this is rather unlikely, and we can only show that the problem lies in the class $\mathrm{P}^{\mathrm{NP}[\log]}$ of problems solvable by a deterministic polynomial-time algorithm that may perform a logarithmic number of queries to an NP oracle (see Appendix A). In fact, the same upper bound holds for games with a Streett *or* a Rabin objective for each player.

Theorem 5.7. StrQualNE is in $\mathrm{P}^{\mathrm{NP}[\log]}$ for Streett-Rabin SMGs.

Proof. Let us describe a polynomial-time algorithm performing a logarithmic number of queries to an NP oracle for the problem. On input $\mathcal{G}, v_0, \bar{x}$, the algorithm starts by determining for each vertex v and each *Rabin player* i with $x_i = 0$ whether $\mathrm{val}_i^{\mathcal{G}}(v) = 0$. Naively implemented, this requires a super-logarithmic number of queries to the oracle. To reduce the number of queries, we use a neat trick, due to Hemachandra (1989). Let us denote by R and S the set of players $i \in \Pi$ with $x_i = 0$ who have a Rabin, respectively a Streett objective. Instead of looping through all pairs of a vertex and a player, we start by determining the number r of all pairs (v, i) such that $i \in R$ and $\mathrm{val}_i^{\mathcal{G}}(v) = 0$.

---

[1] In fact, Ummels & Wojtczak (2009b) claimed that the problem is in coNP.

It is not difficult to see that this number can be computed using binary search by performing only a logarithmic number of queries to an NP oracle, which we can use for deciding whether $\mathrm{val}_i^{\mathcal{G}}(v) > 0$ (Corollary 2.16). Then we perform one more query; we ask whether for each player $i \in R \cup S$ there exists a set $Z_i \subseteq V$ as well as sets $U_1, \ldots, U_{|V|} \subseteq V$ and positional strategies $(\sigma_i)_{i \in R}$ and $(\tau_i)_{i \in S}$, where σ_i is a strategy of player i and τ_i is a strategy of the coalition $\Pi \setminus \{i\}$ in the coalition game $\mathcal{G}_i$, with the following properties:

1. $Z := \bigcap_{i \in R \cup S} Z_i$ is a subarena of $\mathcal{G}$ with $v_0 \in Z$, and $\sum_{i \in R} |Z_i| = r$;
2. $\mathrm{val}^{\sigma_i}(v) > 0$ for each player $i \in R$ and each $v \in V \setminus Z_i$;
3. $\mathrm{val}^{\tau_i}(v) = 1$ for each player $i \in S$ and each $v \in Z_i$;
4. each U_j is an end component of $\mathcal{G} \restriction Z$ with payoff $\bar{x}$;
5. the value from v_0 of the MDP that is obtained from $\mathcal{G}$ by restricting to vertices inside Z and imposing the objective $\mathrm{Reach}(\bigcup\{U_1, \ldots, U_{|V|}\})$ equals 1.

This query can be decided by an NP oracle by guessing suitable sets and strategies and verifying 1.–5. in polynomial time. If the answer to the query is *yes*, the algorithm accepts; otherwise it rejects.

Obviously, the algorithm runs in polynomial time. To see that the algorithm is correct, first note that for each player $i \in R$ the set Z_i does not only include all $v \in V$ such that $\mathrm{val}_i^{\mathcal{G}}(v) = 0$, but also excludes all other vertices. Otherwise, there would exist a vertex $v \in Z_i$ with $\mathrm{val}_i^{\mathcal{G}}(v) > 0$. But then the number of pairs (v, i) with $i \in R$ and $\mathrm{val}_i^{\mathcal{G}}(v) = 0$ would be strictly less than r, a contradiction. Now, the correctness of the algorithm follows with the same reasoning as in the proof of Theorem 5.5. □

Remark. For a bounded number of players, StrQualNE is in coNP for SMGs with Rabin objectives.

Regarding lower bounds for StrQualNE in SMGs with Rabin objectives, we start by proving that the problem is coNP-hard, even for deterministic two-player games. Moreover, the same lower bound holds for StrQualSPE, the corresponding problem for subgame-perfect equilibria. In particular, unless NP = coNP, both StrQualNE and StrQualSPE are not in NP for SMGs with Rabin objectives.

Theorem 5.8. StrQualNE and StrQualSPE are coNP-hard for deterministic two-player Rabin games.

Proof. The proof is similar to the proof of Theorem 5.6 and is accomplished by a reduction from the unsatisfiability problem for Boolean formulae in conjunctive normal form. Given a Boolean formula $\varphi = C_1 \wedge \cdots \wedge C_m$ in conjunctive normal form, where without loss of generality each clause is nonempty, we construct a deterministic two-player Rabin game $\mathcal{G}$ as follows. The arena of $\mathcal{G}$ is the same as in the proof of Theorem 5.6, depicted in Figure 5.1. However, this time player 1 wins every play of the game, and player 0's objective consists of all *Rabin* pairs of the form $(\{X\}, \{\neg X\})$ or $(\{\neg X\}, \{X\})$.

Clearly, $\mathcal{G}$ can be constructed from φ in polynomial time. We claim that the following statements are equivalent:

1. φ is unsatisfiable;
2. $(\mathcal{G}, C_1)$ has a subgame-perfect equilibrium with payoff $(0, 1)$;
3. $(\mathcal{G}, C_1)$ has a Nash equilibrium with payoff $(0, 1)$.

$(1. \Rightarrow 2.)$ Assume that φ is unsatisfiable, and consider the two-player zero-sum Rabin-Streett game $\widetilde{\mathcal{G}}$, which is derived from $\mathcal{G}$ by setting player 1's objective to the complement of player 0's objective. By Theorem 3.15, the game $(\widetilde{\mathcal{G}}, C_1)$ has a pure subgame-perfect equilibrium (σ_0, σ_1). We claim that $\sigma_1[x]$ is winning in $(\widetilde{\mathcal{G}}, v)$ for every history xv of $(\mathcal{G}, C_1)$. Consequently, (σ_0, σ_1) is also a subgame-perfect equilibrium of $(\mathcal{G}, C_1)$ with payoff $(0, 1)$. Otherwise, player 0 would have a positional winning strategy in $(\widetilde{\mathcal{G}}, v)$. But a positional strategy τ of player 0 picks for each clause a literal contained in this clause. Since φ is unsatisfiable, there must exist a variable X and clauses C and C' such that $\tau(C) = X$ and $\tau(C') = \neg X$. Player 1 could counter this strategy by playing from X to C' and from any other literal to C, a contradiction.

$(2. \Rightarrow 3.)$ Trivial.

$(3. \Rightarrow 1.)$ Let (σ_0, σ_1) be a Nash equilibrium of $(\mathcal{G}, C_1)$ with payoff $(0, 1)$, and assume that φ is satisfiable. Consider the following positional strategy τ of player 0: whenever the play reaches a clause, then τ plays to a literal that is mapped to true by the satisfying assignment. This strategy ensures that for each variable X at most one of the literals X or $\neg X$ is visited infinitely often. Since the construction of $\mathcal{G}$ ensures that, under any strategy profile, at least one literal is visited infinitely often, τ ensures a winning play for player 0. Hence, player 0 can improve her payoff by playing τ instead of σ_0, a contradiction to the fact that (σ_0, σ_1) is a Nash equilibrium. $\qquad \square$

The next result shows that StrQualNE is not only coNP-hard for Rabin games, but also NP-hard. In fact, it is even NP-hard to decide whether in
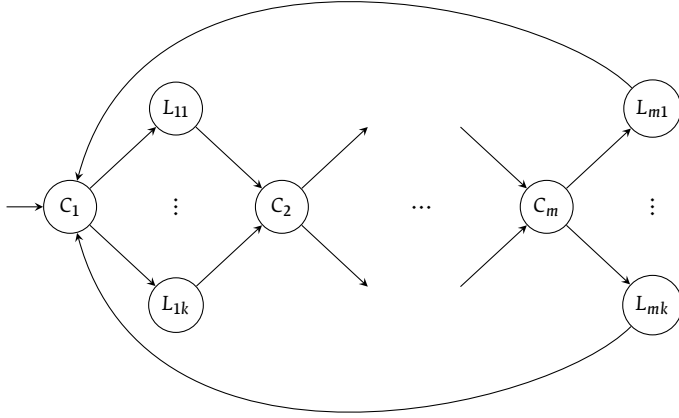
Figure 5.2. Reducing SAT to deciding the existence of a play winning for all players in a deterministic Rabin game.

a deterministic Rabin game there exists a play that fulfils the objective of each player.

Proposition 5.9. The problem of deciding, given a deterministic Rabin game, whether there exists a play that is won by each player is NP-hard.

Proof. We reduce from SAT: given a Boolean formula $\varphi = C_1 \wedge \cdots \wedge C_m$ in conjunctive normal form over propositional variables $X_1, \ldots, X_n$, where without loss of generality each clause is nonempty, we show how to construct in polynomial time a deterministic $(n + 1)$-player Rabin game $\mathcal{G}$ such that φ is satisfiable if and only if there exists a play of $\mathcal{G}$ that is won by each player. The game has vertices $C_1, \ldots, C_m$ and, for each clause C and each literal L that occurs in C, a vertex (C, L). All vertices are controlled by player 0. There are edges from a clause C_j to each vertex (C_j, L) such that L occurs in C_j and from there to $C_{(j \bmod m)+1}$. The arena of $\mathcal{G}$ is schematically depicted in Figure 5.2. The Rabin objectives are defined as follows:

- player 0 wins every play of $\mathcal{G}$;
- player $i \neq 0$ wins if each vertex of the form (C, X_i) is visited only finitely often *or* each vertex of the form $(C, \neg X_i)$ is visited only finitely often.

Clearly, $\mathcal{G}$ can be constructed from φ in polynomial time. To establish the reduction, we need to show that φ is satisfiable if and only if there exists a play of $\mathcal{G}$ that is won by each player.

($\Rightarrow$) Assume that $\alpha: \{X_1, \dots, X_n\} \to \{\text{true}, \text{false}\}$ is a satisfying assignment of φ. Clearly, the positional strategy of player 0 where from each clause C she plays to a fixed vertex (C, L) such that L is mapped to true by α induces a play that is won by each player.

($\Leftarrow$) Assume that there exists a play π of $\mathcal{G}$ that is won by each player. Obviously, it is not possible that both a vertex (C, X_i) and a vertex $(C', \neg X_i)$ are visited infinitely often in π since this would violate player i's objective. Consider the variable assignment that maps X to true if some vertex (C, X) is visited infinitely often in π. This assignment satisfies the formula because, by the construction of $\mathcal{G}$, for each clause C there exists a literal L in C such that the vertex (C, L) is visited infinitely often in π. □

It follows from Theorem 5.8 and Proposition 5.9 that, unless NP = coNP, both StrQualNE and StrQualSPE are not contained in NP $\cup$ coNP, even for deterministic Rabin games. A slightly stronger result is that these problems are hard for the class DP (see Appendix A).

Theorem 5.10. StrQualNE and StrQualSPE are DP-hard for deterministic Rabin games.

Proof. Let us focus on StrQualNE; the proof for StrQualSPE is similar. The proof proceeds by a reduction from SAT-UNSAT: we show how to construct in polynomial time from a pair (φ, ψ) of Boolean formulae in conjunctive normal form a game $(\mathcal{G}, v_0)$ such that φ is satisfiable and ψ is not if and only if $(\mathcal{G}, v_0)$ has a Nash equilibrium with payoff $(0, 1, \dots, 1)$.

By Proposition 5.9, we know that from φ we can construct in polynomial time a deterministic Rabin game $(\mathcal{G}_1, v_1)$ such that φ is satisfiable if and only if there exists a play of $(\mathcal{G}_1, v_1)$ that is won by each player. Moreover, by (the proof of) Theorem 5.6, we know that from ψ we can construct in polynomial time a two-player deterministic Rabin game $(\mathcal{G}_2, v_2)$ such that every play of $\mathcal{G}_2$ is won by player 1, and ψ is unsatisfiable if and only if $(\mathcal{G}_2, v_2)$ has a Nash equilibrium with payoff $(0, 1)$. Assume that $\mathcal{G}_1$ is played by players $1, \dots, n$ and that $\mathcal{G}_2$ is played by players 0 and 1. The game $\mathcal{G}$ is the disjoint union of $\mathcal{G}_1$ and $\mathcal{G}_2$ combined with a new vertex v_0, controlled by player 0, with transitions to v_1 and v_2. All plays that go through v_1 or v_2 are lost by player 0, respectively players $2, \dots, n$. We claim that φ is satisfiable and ψ is not if and only if $(\mathcal{G}, v_0)$ has a Nash equilibrium with payoff $(0, 1, \dots, 1)$.

($\Rightarrow$) Assume that φ is satisfiable and ψ is not. Hence, there exists a strategy profile $(\sigma_1, \dots, \sigma_n)$ of $(\mathcal{G}_1, v_1)$ where all players win and a Nash equilibrium

(τ_0, τ_1) of $(\mathcal{G}_2, v_2)$ with payoff $(0, 1)$. Define a strategy profile $\bar{\sigma}^* = (\sigma_0^*, \ldots, \sigma_n^*)$ of $(\mathcal{G}, v_0)$ by setting $\sigma_i^*[v_0] = \sigma_i$ for $i \geq 1$, $\sigma_i^*[v_0] = \tau_i$ for $i \leq 1$ and $\sigma_0^*(v_1 \mid v_0) = 1$. Clearly, $\bar{\sigma}^*$ has payoff $(0, 1, \ldots, 1)$. Moreover, $\bar{\sigma}^*$ is a Nash equilibrium because $(\sigma_1, \ldots, \sigma_n)$ and (τ_0, τ_1) are Nash equilibria with suitable payoffs. In particular, player 0 cannot improve her payoff of 0 by playing to v_2 since (τ_0, τ_1) is a Nash equilibrium of $(\mathcal{G}_2, v_2)$ where player 0 receives payoff 0.

($\Leftarrow$) Assume that $\bar{\sigma} = (\sigma_0, \ldots, \sigma_n)$ is a Nash equilibrium of $(\mathcal{G}, v_0)$ with payoff $(0, 1, \ldots, 1)$. Since all players $1, \ldots, n$ win, we must have $\sigma_0(v_1 \mid v_0) = 1$. Hence, $\bar{\sigma}$ induces a Nash equilibrium of $(\mathcal{G}_1, v_1)$ with payoff $(1, \ldots, 1)$ and, by the construction of $\mathcal{G}_1$, the formula φ is satisfiable. Why is ψ unsatisfiable? Assume the opposite. Then, by the construction of $\mathcal{G}_2$, there exists a strategy τ for player 0 such that $(\tau, \sigma_1[v_0])$ gives payoff > 0 to player 0 in $(\mathcal{G}_2, v_2)$. But then player 0 could improve her payoff in $(\mathcal{G}, v_0)$ by playing from v_0 to v_2 and applying strategy τ afterwards, a contradiction. $\qquad\square$

For stochastic Rabin games, we can show a completeness result: for these games, StrQualNE and StrQualSPE are also hard for $P^{NP[\log]}$.

Theorem 5.11. StrQualNE and StrQualSPE are $P^{NP[\log]}$-hard for Rabin SMGs.

Proof. Let us focus on StrQualNE; the proof for StrQualSPE is completely analogous. Wagner (1990) and, independently, Buss & Hay (1991) showed that $P^{NP[\log]}$ equals the closure of NP with respect to *polynomial-time Boolean formula reducibility*. The canonical complete problem for this class is to decide, given a Boolean combination α of statements of the form "φ is satisfiable", where φ ranges over all Boolean formulae, whether α evaluates to true. We claim that for every such statement α we can construct in polynomial time a Rabin SMG $(\mathcal{G}, v_0)$ such that α evaluates to true if and only if $(\mathcal{G}, v_0)$ has a Nash equilibrium with payoff $(0, 1, \ldots, 1)$. The game $\mathcal{G}$ is constructed by induction on the complexity of α; without loss of generality, we assume that negations are only applied to atoms. If α is of the form "φ is satisfiable" or "φ is not satisfiable", then the existence of a suitable game $\mathcal{G}$ follows from Proposition 5.9 or Theorem 5.8, respectively.

Now, let $\alpha = \alpha_1 \wedge \alpha_2$, and assume that we already have constructed suitable games $(\mathcal{G}_1, v_1)$ and $(\mathcal{G}_2, v_2)$, played by the same players $0, 1, \ldots, n$. The game $\mathcal{G}$ is the disjoint union of $\mathcal{G}_1$ and $\mathcal{G}_2$ combined with one new stochastic vertex v_0. From v_0, the game moves with probability $\frac{1}{2}$ each to v_1 or v_2. Obviously, $(\mathcal{G}, v_0)$ has a Nash equilibrium with payoff $(0, 1, \ldots, 1)$ if and only if both $(\mathcal{G}_1, v_1)$ and $(\mathcal{G}_2, v_2)$ have such an equilibrium.

Algorithm 5.1. Finding end components in parity SMGs.

Input: parity SMG $\mathcal{G} = (\Pi, V, (V_i)_{i\in\Pi}, \Delta, \chi, (\Omega_i)_{i\in\Pi})$, $\bar{x} = (x_i)_{i\in\Pi} \in \{0,1\}^{\Pi}$
Output: $\bigcup\{U \subseteq V : U$ is an end component of $\mathcal{G}$ with payoff $\bar{x}\}$

output FindEC(V)

procedure FindEC(X)
  $Z := \varnothing$
  compute all end components of $\mathcal{G}$ maximal in X
  for each such end component U **do**
    $P := \{i \in \Pi : \min \Omega_i(\chi(U)) \equiv x_i \bmod 2\}$
    if $P = \varnothing$ **then**
      (* U is an end component with payoff $\bar{x}$ *)
      $Z := Z \cup U$
    else
      (* U has the wrong payoff *)
      $Y := \bigcap_{i\in P}\{v \in U : \Omega_i(\chi(v)) > \min \Omega_i(\chi(U))\}$
      $Z := Z \cup$ FindEC(Y)
    end if
  end for
  return Z
end procedure

Finally, let $\alpha = \alpha_1 \vee \alpha_2$, and assume that we already have constructed suitable games $(\mathcal{G}_1, v_1)$ and $(\mathcal{G}_2, v_2)$, again played by the same players $0, 1, \ldots, n$. As in the previous case, the game $\mathcal{G}$ is the disjoint union of $\mathcal{G}_1$ and $\mathcal{G}_2$ combined with one new vertex v_0, which has transitions to both v_1 and v_2. However, this time v_0 is controlled by player 1. Obviously, $(\mathcal{G}, v_0)$ has a Nash equilibrium with payoff $(0, 1, \ldots, 1)$ if and only if at least one of the games $(\mathcal{G}_1, v_1)$ and $(\mathcal{G}_2, v_2)$ has such an equilibrium. □

To solve StrQualNE for parity SMGs, we employ Algorithm 5.1, which computes for a game $\mathcal{G}$ with priority functions $(\Omega_i)_{i\in\Pi}$ and $\bar{x} \in \{0,1\}^{\Pi}$ the union of all end components with payoff $\bar{x}$. The algorithm is a straightforward adaptation of the algorithm for computing the union of all accepting end components in a Streett MDP (Chatterjee et al. 2005). At the heart of the algorithm lies the procedure FindEC that returns on input $X \subseteq V$ the union of all end components with payoff $\bar{x}$ that are contained in X. The procedure

starts by computing all end components maximal in X. If such an end component U has payoff $\bar{x}$, all vertices in U can be added to the result of the procedure. Otherwise, there exists a player i such that either $x_i = 0$ and the least priority for player i in U is odd or $x_i = 1$ and the least priority for player i in U is even. Each end component with payoff $\bar{x}$ inside U must exclude all vertices with this least priority. Hence, we call the procedure recursively on the subset of U that results from removing these vertices.

Note that on input X, the total number of recursive calls to the procedure FindEC is bounded by $|X|$. Since, additionally, the set of all end components maximal in a set X can be computed in polynomial time (see Appendix B), this proves that Algorithm 5.1 runs in polynomial time.

Theorem 5.12. StrQualNE is in $UP \cap coUP$ for parity SMGs.

Proof. An unambiguous nondeterministic polynomial-time algorithm that decides StrQualNE for parity SMGs works as follows: On input $\mathcal{G}, v_0, \bar{x}$, the algorithm starts by guessing, for each player i with $x_i = 0$, the set Z_i of vertices v with $\mathrm{val}_i^{\mathcal{G}}(v) = 0$. Then, for each $v \in V$, the guess whether $v \in Z_i$ or $v \notin Z_i$ is verified by running the UP algorithm for the respective problem. If some guess was not correct, the algorithm rejects immediately. Otherwise, it constructs the subarena $Z := \bigcap_{i \in \Pi : x_i = 0} Z_i$ and uses Algorithm 5.1 to determine the union T of all end components with payoff $\bar{x}$. If $v_0 \notin Z$, the algorithm rejects immediately. Otherwise, it computes in polynomial time the value $\mathrm{val}^{\mathcal{G}(\bar{x})}(v_0)$ of the MDP $\mathcal{G}(\bar{x})$ from v_0. If this value equals 1, the algorithm accepts; otherwise, it rejects. Analogously, an algorithm for the complement of StrQualNE accepts if and only if $v_0 \notin Z$ or $\mathrm{val}^{\mathcal{G}(\bar{x})}(v_0) < 1$.

Obviously, both algorithms run in polynomial time. Moreover, on each input there exists at most one accepting run because the algorithms only accept if each of the sets Z_i has been guessed correctly. Finally, the correctness of both algorithms follows from Lemma 5.3. □

Recall from Section 2.5 that it is an open question whether the qualitative decision problem for parity S2Gs admits a polynomial-time algorithm. Such an algorithm would allow us compute the domain of the MDP $\mathcal{G}(\bar{x})$ efficiently, which would imply that StrQualNE is in P for parity SMGs. In fact, given a class $\mathcal{C}$ of parity S2Gs for which the qualitative decision problem is in P, we can easily derive a class of parity SMGs for which StrQualNE is in P. As in Section 3.4, denote by $\mathcal{C}^*$ the class of all parity SMGs such that for each player i the coalition game $\mathcal{G}_i$ is in $\mathcal{C}$.

Theorem 5.13. Let $\mathcal{C}$ be a class of finite parity S2Gs such that the qualitative decision problem is decidable in P for games in $\mathcal{C}$. Then StrQualNE is in P for games in $\mathcal{C}^*$.

Proof. Consider the algorithm given in the proof of Theorem 5.12. For each player i, the set Z_i can be computed in polynomial time if $\mathcal{G}_i \in \mathcal{C}$, and there is no need to guess this set. The resulting deterministic algorithm still runs in polynomial time. □

By Theorem 2.20, for each $d \in \mathbb{N}$, we can decide the qualitative decision problem for parity S2Gs with at most d priorities in polynomial time. Hence, it follows from Theorem 5.13 that StrQualNE is decidable in polynomial time for parity SMGs with at most d priorities. In particular, StrQualNE is in P for (co-)Büchi SMGs.

Corollary 5.14. For each $d \in \mathbb{N}$, StrQualNE is in P for parity SMGs with at most d priorities.

5.2 The positive-one fragment

In this section, we prove the decidability of the problem OneNE. Formally, OneNE is the following decision problem:

Given a finite SMG $(\mathcal{G}, v_0)$ and $p \in [0,1]$, decide whether $(\mathcal{G}, v_0)$ has a Nash equilibrium with payoff $\geq (p, 1, \ldots, 1)$.

Being more general than the qualitative decision problem, OneNE is PSPACE-hard for Muller SMGs. In order to put the problem into PSPACE, we describe a polynomial-space algorithm that, given a Muller SMG $\mathcal{G}$ and $p \in [0,1]$, computes the set of vertices v such that $(\mathcal{G}, v)$ has a Nash equilibrium with payoff $\geq (p, 1, \ldots, 1)$.

Algorithm 5.2 is a variant of the classical algorithm for computing in a Markov decision process the set of states from where the optimal probability of reaching a certain set of target states equals 1 (see Baier & Katoen 2008, Chapter 10). The general idea of the algorithm is to find a subarena in which the players can ensure to visit an end component with payoff $(0, 1, \ldots, 1)$ or $(1, 1, \ldots, 1)$ with probability 1; let us call such an end component *good*. Additionally, at every vertex v in this subarena the optimal probability of reaching an end component with payoff $(1, 1, \ldots, 1)$ must be at least as high

---

Algorithm 5.2. Solving OneNE for Muller SMGs.

Input: Muller SMG $\mathcal{G} = (\Pi, V, (V_i)_{i \in \Pi}, \Delta, \chi, (\mathcal{F}_i)_{i \in \Pi})$, $p \in [0, 1]$
Output: $\{v \in V : (\mathcal{G}, v) \text{ has a Nash equilibrium with payoff} \geq (p, 1 \ldots, 1)\}$

compute $z_v := \mathrm{val}_0^{\mathcal{G}}(v)$ for each $v \in V$
$X := V$
repeat
  $X' := X$
  (* identify good end components *)
  $S := \bigcup \{U \subseteq X : U \text{ is an end component of } \mathcal{G} \upharpoonright X \text{ with payoff } (0, 1, \ldots, 1)\}$
  $T := \bigcup \{U \subseteq X : U \text{ is an end component of } \mathcal{G} \upharpoonright X \text{ with payoff } (1, 1, \ldots, 1)\}$
  compute $p_v := \sup_{\bar{\sigma}} \mathrm{Pr}_v^{\bar{\sigma}}(\mathrm{Reach}(T))$ in $\mathcal{G} \upharpoonright X$ for each $v \in X$
  $X := \{v \in X : S \cup T \text{ reachable from } v \text{ inside } \mathcal{G} \upharpoonright X\} \cap \{v \in X : z_v \leq p_v\}$
  (* compute the largest subarena contained in X *)
  repeat
    $X'' := X$
    $X := \{v \in X : v\Delta \cap X \neq \varnothing\} \cap \{v \in X : v \in \bigcup_{i \in \Pi} V_i \text{ or } v\Delta \subseteq X\}$
  until $X = X''$
until $X = X'$ or $X = \varnothing$
(* $S \cup T$ reachable from all vertices in $\mathcal{G} \upharpoonright X$, and $\mathrm{val}_0^{\mathcal{G}}(v) \leq p_v$ for all $v \in X$ *)
output $\{v \in X : p \leq p_v\}$

---

as the value $\mathrm{val}_0^{\mathcal{G}}(v)$; otherwise player 0 would prefer switching to an optimal strategy at vertex v. Such a subarena can be found by an iterative process; in every iteration, the algorithm computes (inside the current subarena) the union of all good end components and the optimal probability of reaching an end component where all players win. All vertices from where a good end component is not reachable or where the latter probability is strictly less than $\mathrm{val}_0^{\mathcal{G}}(v)$ are then removed from the current subarena. If the resulting set X of vertices still forms a subarena, we can output all vertices from where the optimal probability of reaching an end component where all players win is at least p; otherwise we have to continue the process with the largest subarena contained in X.

Theorem 5.15. OneNE is PSPACE-complete for Muller SMGs.

Proof. Hardness follows from Theorem 2.18. To prove membership in PSPACE, we show that Algorithm 5.2 is a polynomial-space algorithm that solves

OneNE for Muller SMGs. First, for each vertex v, its value for player 0 can be computed in polynomial space since we are dealing with Muller objectives. Second, the sets S and T can be determined by enumerating all possible subsets, one at a time, and checking which ones are end components with the right payoff. Third, the numbers p_v can be computed in polynomial time via linear programming (see Appendix B). Fourth, the sets X, X' and X'' are subsets of V and can thus be stored using polynomial space. Finally, the algorithm terminates because in each iteration of one of the two repeat loops, the set X becomes smaller until the termination criterion is met.

Now let $(\mathcal{G}, v_0)$ be an arbitrary Muller SMG, and let $p \in [0,1]$; we claim that $(\mathcal{G}, v_0)$ has a Nash equilibrium with payoff $\geq (p, 1, \dots, 1)$ if and only if v_0 is output by the algorithm on input $\mathcal{G}, p$.

($\Rightarrow$) Assume that $(\mathcal{G}, v_0)$ has a Nash equilibrium $\bar{\sigma}$ with payoff $\geq (p, 1, \dots, 1)$, and consider the set $R := \{v \in V : \Pr_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(v)) > 0\}$. We claim that every vertex $v \in R$ remains inside the set X maintained by the algorithm. Since X is initially set to V, this is clearly true before the main loop has been entered. Now assume that $R \subseteq X$; let S and T be defined as in the algorithm, and let xv be a history ending in $v \in R$ that is consistent with $\bar{\sigma}$. By Lemma 2.3 and since $\bar{\sigma}$ has payoff $\geq (0, 1, \dots, 1)$, we have $\Pr_v^{\bar{\sigma}[x]}(\mathrm{Reach}(S \cup T)) = 1$; in particular, $S \cup T$ is reachable from v inside $\mathcal{G} \upharpoonright R$, and therefore also inside $\mathcal{G} \upharpoonright X$. Moreover, since $\bar{\sigma}$ is a Nash equilibrium, we have

$$
\begin{aligned}
z_v &= \mathrm{val}_0^{\mathcal{G}}(v) \\
&\leq \Pr_{v_0}^{\bar{\sigma}}(\mathrm{Win}_0 \mid xv \cdot V^{\omega}) &&\text{(by Lemma 3.3)} \\
&= \Pr_v^{\bar{\sigma}[x]}(\mathrm{Win}_0) &&\text{(by Lemma 2.2)} \\
&\leq \Pr_v^{\bar{\sigma}[x]}(\mathrm{Reach}(T)) &&\text{(by Lemma 2.3)} \\
&\leq \sup\{\Pr_v^{\bar{\sigma}}(\mathrm{Reach}(T)) : \bar{\sigma} \text{ strategy profile of } \mathcal{G} \upharpoonright R\} \\
&\leq \sup\{\Pr_v^{\bar{\sigma}}(\mathrm{Reach}(T)) : \bar{\sigma} \text{ strategy profile of } \mathcal{G} \upharpoonright X\} \\
&= p_v \, .
\end{aligned}
$$

Finally, no vertex $v \in R$ is removed from X in the inner repeat loop because R is a subarena of $\mathcal{G}$.

It follows that each vertex $v \in R$ is still in the set X after completing the main loop; in particular, this holds for v_0. Moreover, substituting v by v_0 and x by the empty word in the above calculation yields that $p \leq \Pr_{v_0}^{\bar{\sigma}}(\mathrm{Win}_0) \leq p_{v_0}$. Hence, v_0 is in the output of the algorithm.

($\Leftarrow$) Assume that v_0 is output by the algorithm on input $\mathcal{G}, p$, and let X, S, T and $(p_v)_{v \in V}$ have the same values as when the algorithm terminates. We claim that $\sup_{\bar{\sigma}} \Pr_v^{\bar{\sigma}}(\text{Reach}(S \cup T)) = 1$ for all $v \in X$, where the supremum ranges over all strategy profiles of $\mathcal{G} \restriction X$. To see this, consider the stationary strategy profile $\bar{\sigma}$ that for each non-stochastic vertex $v \in X$ picks a successor $w \in v\Delta \cap X$ uniformly at random. All bottom SCCs of $(\mathcal{G} \restriction X)^{\bar{\sigma}}$ contain a vertex of $S \cup T$ since otherwise the whole SCC would have been removed from X by the algorithm. Hence, $S \cup T$ is reached almost surely in $(\mathcal{G} \restriction X)^{\bar{\sigma}}$.

Since reachability MDPs admit optimal positional strategies (Theorem B.9), we can fix positional strategy profiles $\bar{\sigma}^S = (\sigma_i^S)_{i \in \Pi}$ and $\bar{\sigma}^T = (\sigma_i^T)_{i \in \Pi}$ of $\mathcal{G} \restriction X$ such that $\Pr_v^{\bar{\sigma}^S}(\text{Reach}(S \cup T)) = 1$ and $\Pr_v^{\bar{\sigma}^T}(\text{Reach}(T)) = p_v$ for all $v \in X$. Moreover, by Lemma 2.4, we can fix pure finite-state strategy profiles $\bar{\tau}^S = (\tau_i^S)_{i \in \Pi}$ and $\bar{\tau}^T = (\tau_i^T)_{i \in \Pi}$ of $\mathcal{G} \restriction S$ and $\mathcal{G} \restriction T$, respectively, such that $\Pr_v^{\bar{\tau}^S}(\bigcap_{i \in \Pi \setminus \{0\}} \text{Win}_i) = 1$ for all $v \in S$ and $\Pr_v^{\bar{\tau}^T}(\bigcap_{i \in \Pi} \text{Win}_i) = 1$ for all $v \in T$. We define a pure finite-state strategy profile $\bar{\sigma} = (\sigma_i)_{i \in \Pi}$ of $(\mathcal{G}, v_0)$ by setting

$$
\sigma_i(xv) = \begin{cases}
\sigma_i^T(v) & \text{if } v \in X \setminus T \text{ and } p_v > 0, \\
\sigma_i^S(v) & \text{if } v \in X \setminus S \text{ and } p_v = 0, \\
\tau_i^T(x_1 v) & \text{if } v \in T, \\
\tau_i^S(x_2 v) & \text{if } v \in S \text{ and } p_v = 0, \\
\text{arbitrary} & \text{otherwise,}
\end{cases}
$$

for all $xv \in V^* V_i$ and $i \in \Pi$, where x_1 and x_2 are the longest suffixes of x contained in T and S, respectively. We claim that $\bar{\sigma}$ fulfils the following three properties:

1. $\Pr_{v_0}^{\bar{\sigma}}(xv \cdot V^\omega) > 0$ only if $v \in X$;
2. $\Pr_{v_0}^{\bar{\sigma}}(\text{Win}_0 \mid xv \cdot V^\omega) \geq p_v$ if xv is consistent with $\bar{\sigma}$ and $v \in X$;
3. $\Pr_{v_0}^{\bar{\sigma}}(\text{Win}_i) = 1$ for each player $i \neq 0$.

It follows from 1., 2. and the definition of X that $\Pr_{v_0}^{\bar{\sigma}}(\text{Win}_0 \mid xv \cdot V^\omega) \geq p_v \geq \text{val}_0^{\mathcal{G}}(v)$ for each history xv consistent with $\bar{\sigma}$. Moreover, it follows from 3. that $\Pr_{v_0}^{\bar{\sigma}}(\text{Win}_i \mid xv \cdot V^\omega) = 1 \geq \text{val}_i^{\mathcal{G}}(v)$ for each player $i \neq 0$ and for each history xv consistent with $\bar{\sigma}$. Hence, by Lemma 3.8, the game $(\mathcal{G}, v_0)$ has a pure finite-state Nash equilibrium $\bar{\sigma}^*$ with $\Pr_{v_0}^{\bar{\sigma}^*} = \Pr_{v_0}^{\bar{\sigma}}$. Moreover, since $p_{v_0} \geq p$, it follows from 2. and 3. that $\bar{\sigma}^*$ has payoff $\geq (p, 1, \dots, 1)$.

To complete the proof, we need to verify properties 1.–3. Property 1. is immediate from the definition of $\bar{\sigma}$ and the fact that X is a subarena of $\mathcal{G}$

with $v_0 \in X$. For 2., assume that xv is a history of $(\mathcal{G}, v_0)$ consistent with $\bar{\sigma}$ and ending in a vertex $v \in X$. The claim holds trivially if $p_v = 0$. Otherwise, $\bar{\sigma}[x]$ guarantees to reach T with probability p_v. But once T has been reached, $\bar{\sigma}[x]$ behaves like $\bar{\tau}^T$, which guarantees to fulfil the objective of player 0 almost surely. Hence, $\mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Win}_0 \mid xv \cdot V^\omega) = \mathrm{Pr}_v^{\bar{\sigma}[x]}(\mathrm{Win}_0) \geq \mathrm{Pr}_v^{\bar{\sigma}[x]}(\mathrm{Reach}(T)) = p_v$.

In order to prove 3., define $Z := \{v \in X : p_v = 0\}$. By definition, Z is a subarena of $\mathcal{G} \upharpoonright X$ with $Z \cap T = \varnothing$. We claim that $\mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}((S \cap Z) \cup T)) = 1$. This proves 3. because, once $(S \cap Z) \cup T$ has been reached, $\bar{\sigma}$ behaves either like $\bar{\tau}^S$ or like $\bar{\tau}^T$, which guarantee to fulfil the objective of each player $i \neq 0$ almost surely. By the definition of $\bar{\sigma}$, we have $\mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(Z) \setminus \mathrm{Reach}(S)) = 0$ and therefore $\mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(Z)) = \mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(S \cap Z))$. Moreover, if $\pi \in X^\omega \setminus \mathrm{Reach}(Z)$, then $\mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(T) \mid \pi|_k \cdot X^\omega) \geq p_{\pi(k-1)} \geq \min\{p_v : v \in X \setminus Z\}$ for all $k > 0$ and therefore $\lim_{k \to \infty} \mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(T) \mid \pi|_k \cdot X^\omega) \neq 0$. Using Levy's zero-one-law (Theorem A.6), we can conclude that

$$
\begin{aligned}
1 &= \mathrm{Pr}_{v_0}^{\bar{\sigma}}(\{\pi \in X^\omega : \lim_{k \to \infty} \mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(T) \mid \pi|_k \cdot X^\omega) = 1_{\mathrm{Reach}(T)}(\pi)\}) \\
&\leq \mathrm{Pr}_{v_0}^{\bar{\sigma}}(\{\pi \in X^\omega : \lim_{k \to \infty} \mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(T) \mid \pi|_k \cdot X^\omega) = 0\} \cup \mathrm{Reach}(T)) \\
&\leq \mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(Z) \cup \mathrm{Reach}(T)) \\
&= \mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(Z)) + \mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(T)) \\
&= \mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(S \cap Z)) + \mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(T)) \\
&= \mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}((S \cap Z) \cup T)),
\end{aligned}
$$

which proves the claim. □

Note that the Nash equilibrium $\bar{\sigma}^*$ constructed in the proof of Theorem 5.15 is a pure finite-state equilibrium. It follows that, for finite SMGs with ω-regular objectives, the problem OneNE does not change when one asks for a pure finite-state equilibrium with payoff $\geq (p, 1, \ldots, 1)$ instead of an arbitrary equilibrium with such a payoff.

Of course, Algorithm 5.2 can also be used to solve OneNE for SMGs with Streett-Rabin or parity objectives in polynomial space. In fact, we can do better, and the complexity of OneNE is comparable to the complexity of the problem StrQualNE, discussed in the previous section. In particular, we can show that OneNE lies in NP for Streett SMGs. In fact, the same upper bound holds for Streett-Rabin SMGs with a Streett objective for player 0.

Theorem 5.16. OneNE is in NP for Streett-Rabin SMGs with a Streett objective for player 0.

Proof. To get a nondeterministic polynomial-time algorithm for the problem, it suffices to modify Algorithm 5.2 as follows: Instead of computing $\mathrm{val}_0^{\mathcal{G}}(v)$ explicitly, we guess a positional strategy τ of the coalition $\Pi \setminus \{0\}$ in the coalition game $\mathcal{G}_0$ and compute $z_v := 1 - \mathrm{val}^{\tau}(v)$. Moreover, instead of computing the sets S and T explicitly, we guess suitable end components (at most $|V|$ many) and take their union. Clearly, $z_v \geq \mathrm{val}_0^{\mathcal{G}}(v)$ for all $v \in V$. It follows that, in every iteration of the main loop, the sets S, T and X are subsets of the "real" sets S, T and X, and each of the numbers p_v is bounded from above by the "real" value of p_v. In particular, if v is in the output of the modified algorithm on input $\mathcal{G}, p$, then it is also in the output of Algorithm 5.2 on input $\mathcal{G}, p$. On the other hand, if τ is globally optimal (such a strategy exists by Corollary 2.12), then $z_v = \mathrm{val}_0^{\mathcal{G}}(v)$ for all $v \in V$, and the modified algorithm outputs the same set as Algorithm 5.2.

Why does the modified algorithm run in polynomial time? Since the values of a Streett MDP can be computed in polynomial time, so can the numbers z_v; whether a set U is an end component with payoff $(0, 1 \ldots, 1)$ or payoff $(1, \ldots, 1)$ can easily be checked in polynomial time; the set of states from where $S \cup T$ is reachable can be computed in polynomial time using a simple backward search procedure; all other operations are trivial set operations. Finally, the total number of iterations of both repeat loops is at most $2|V| + 1$ since at least one vertex is removed from the set X in each iteration but the last one. $\qquad \square$

For Streett-Rabin SMGs with a Rabin objective for player 0, the best upper bound we can show for the complexity of OneNE is that the problem lies in the class P^{NP} of problems decidable by a deterministic polynomial-time algorithm with access to an NP oracle. In particular, we do not know whether we can bound the number of oracle queries by $O(\log n)$; this would put OneNE into $\mathrm{P}^{\mathrm{NP}[\log]}$.

Theorem 5.17. OneNE is in P^{NP} for Streett-Rabin SMGs with a Rabin objective for player 0.

Proof. Again, the algorithm for proving membership in P^{NP} is virtually identical to the one for Muller SMGs. The only critical steps are the computation of the values z_v and the computation of the sets S and T in the main loop. Concerning the values, we can determine them by asking oracle queries of the form $\mathrm{val}_0^{\mathcal{G}}(v) \geq x$ repeatedly and closing in on the value $\mathrm{val}_0^{\mathcal{G}}(v)$ using

binary search.[2] Finally, to compute the sets S and T, we can ask the oracle for each $v \in X$ whether v lies in an end component of $\mathcal{G} \upharpoonright X$ with payoff $(0, 1, \ldots, 1)$ or payoff $(1, \ldots, 1)$, respectively. □

As for StrQualNE, the problem OneNE is NP-hard or coNP-hard for deterministic games with Streett or Rabin objectives, respectively, even if the number of players is bounded. Moreover, the same lower bounds hold for the problem OneSPE, where we look for a subgame-perfect equilibrium instead of a Nash equilibrium.

Theorem 5.18. OneNE and OneSPE are NP-hard for deterministic three-player Streett games; both problems are coNP-hard for deterministic three-player Rabin games.

Proof. From (the proofs of) Theorems 5.6 and 5.8, we know that deciding whether there exists a Nash (subgame-perfect) equilibrium with payoff $(0, 1)$ in a deterministic two-player Streett or Rabin game with a trivial objective for player 1 (meaning that she wins every play) is NP-hard or coNP-hard, respectively. We reduce from this problem: given a two-player Streett or Rabin game $(\mathcal{G}, v_0)$ in which all plays are won by player 1, we construct a three-player game $(\widetilde{\mathcal{G}}, \widetilde{v}_0)$ of the same type such that $(\mathcal{G}, v_0)$ has a Nash (subgame-perfect) equilibrium with payoff $(0, 1)$ if and only if $(\widetilde{\mathcal{G}}, \widetilde{v}_0)$ has a Nash (subgame-perfect) equilibrium with payoff $\geq (0, 1, 1)$. The arena of $\widetilde{\mathcal{G}}$ is obtained from the arena of $\mathcal{G}$ by adding two more vertices: the initial vertex $\widetilde{v}_0$, controlled by player 0, and a new terminal vertex $\bot$. From $\widetilde{v}_0$, player 0 can either play to $\bot$ or to v_0, in which case the game continues in $\mathcal{G}$. The objectives for the players are as follows:

- player 0 wins if her objective in $\mathcal{G}$ is fulfilled; if the game reaches $\bot$, she loses;
- player 1 wins every play of $\widetilde{\mathcal{G}}$;
- player 2 wins if $\bot$ is reached; otherwise she loses.

It is easy to see that these objectives can be represented as Streett or Rabin objectives if the original objectives are of this form. We need to show that $(\mathcal{G}, v_0)$ has a Nash (subgame-perfect) equilibrium with payoff $(0, 1)$ if and only if $(\widetilde{\mathcal{G}}, \widetilde{v}_0)$ has a Nash (subgame-perfect) equilibrium with payoff $\geq (0, 1, 1)$.

---

[2] Recall from Chapter 2 that the values of a Rabin-Streett S2G are rational numbers of polynomial bit complexity.

($\Rightarrow$) Assume that (σ_0, σ_1) is a Nash (subgame-perfect) equilibrium of $(\mathcal{G}, v_0)$ with payoff $(0, 1)$. We extend σ_0 and σ_1 to strategies in $\widetilde{\mathcal{G}}$ by setting $\sigma_0(\bot \mid \widetilde{v}_0) = 1$, $\sigma_0[\widetilde{v}_0] = \sigma_0$ and $\sigma_1[\widetilde{v}_0] = \sigma_1$. It is easy to see that, combined with the empty strategy for player 2, the resulting strategy profile is a Nash (subgame-perfect) equilibrium of $(\widetilde{\mathcal{G}}, \widetilde{v}_0)$ with payoff $(0, 1, 1)$.

($\Leftarrow$) Let $\bar{\sigma} = (\sigma_0, \sigma_1, \sigma_2)$ be a Nash equilibrium of $(\widetilde{\mathcal{G}}, \widetilde{v}_0)$ with payoff $\geq (0, 1, 1)$. Note that, since $\bar{\sigma}$ gives payoff 1 to player 2, it must hold that $\sigma_0(\bot \mid \widetilde{v}_0) = 1$, and therefore $\bar{\sigma}$ actually gives payoff 0 to player 0. We claim that $(\sigma_0[\widetilde{v}_0], \sigma_1[\widetilde{v}_0])$ is a Nash equilibrium of $(\mathcal{G}, v_0)$ with payoff $(0, 1)$. Otherwise, since player 1 wins every play of $\mathcal{G}$, there would exist a strategy τ of player 0 such that $(\tau, \sigma_1[\widetilde{v}_0])$ gives payoff > 0 to player 0. But then player 0 could improve her payoff in $(\widetilde{\mathcal{G}}, \widetilde{v}_0)$ by playing from $\widetilde{v}_0$ to v_0 and applying τ afterwards, a contradiction to $\bar{\sigma}$ being a Nash equilibrium. Finally, if $\bar{\sigma}$ is a subgame-perfect equilibrium, then $(\sigma_0[\widetilde{v}_0], \sigma_1[\widetilde{v}_0])$ is a subgame-perfect equilibrium of $(\mathcal{G}, v_0)$ by the definition of a subgame-perfect equilibrium. $\qquad\square$

It follows from Theorems 5.16 and 5.18 that OneNE is NP-complete for Streett SMGs. For Rabin games, Theorem 5.18 and Proposition 5.9 already give good evidence that OneNE and OneSPE cannot be put into NP $\cup$ coNP. As for StrQualNE and StrQualSPE, we can show the stronger result that OneNE and OneSPE are, in fact, DP-hard for deterministic Rabin games and $P^{NP[\log]}$-hard for Rabin SMGs. The proofs are virtually identical to the proofs of Theorems 5.10 and 5.11 and are left to the reader.

Theorem 5.19. OneNE and OneSPE are DP-hard for deterministic Rabin games and $P^{NP[\log]}$-hard for Rabin SMGs.

For parity SMGs, we can use Algorithm 5.1 to compute the union of the relevant end components in polynomial time. Deciding OneNE therefore reduces to computing the values for player 0.

Theorem 5.20. OneNE is in NP $\cap$ coNP for SMGs with parity objectives.

Proof. Membership in NP follows from Theorem 5.16. To prove membership in coNP, we modify Algorithm 5.2 as follows: Instead of computing $\mathrm{val}_0^{\mathcal{G}}(v)$, we guess a positional strategy τ for player 0 and set $z_v := \mathrm{val}^\tau(v)$. Clearly, $z_v \leq \mathrm{val}_0^{\mathcal{G}}(v)$ for all $v \in V$. Finally, we switch the output of the algorithm, i.e. instead of outputting the set $\{v \in X : p \leq p_v\}$, we output the set $\{v \in V : v \notin X \text{ or } p > p_v\}$ (since we seek an algorithm for the complement of OneNE).

The reasoning that the algorithm is correct is analogous to the reasoning in the proof of Theorem 5.16: In every iteration of the main loop, the sets S, T and X are *supersets* of the "real" sets S, T and X, and each of the numbers p_v is bounded from *below* by the "real" value of p_v. Hence, if v is in the output of the modified algorithm on input $\mathcal{G}, p$, then it is certainly not in the output of Algorithm 5.2 on input $\mathcal{G}, p$. On the other hand, if σ is an optimal strategy for player 0, then the modified algorithm outputs precisely the complement of the output of Algorithm 5.2.

The reasoning that the algorithm runs in polynomial time is identical to the proof of Theorem 5.16, except for the computation of the sets S and T inside the main loop: we can compute these sets using Algorithm 5.1 in polynomial time. $\square$

The natural question at this point is whether OneNE can be solved in polynomial time on classes of games where StrQualNE can, such as SSMGs. Clearly, if the value of an SS2G can be computed in polynomial time, then OneNE is also decidable in polynomial time for SSMGs. On the other hand, we can give a polynomial-time reduction from the quantitative decision problem for SS2Gs to OneNE for two-player SSMGs. This shows that lowering the complexity of OneNE to P would resolve a major open problem in the theory of two-player zero-sum stochastic games.

Proposition 5.21. There is a polynomial-time reduction from the quantitative decision problem for SS2Gs to OneNE for two-player SSMGs.

Proof. Let $\mathcal{G}$ be an SS2G, $v \in V$, and $p \in [0,1]$. From $\mathcal{G}$, v and p, we derive a two-player SSMG $\widetilde{\mathcal{G}}$ as follows: In $\widetilde{\mathcal{G}}$, there is a new vertex $\tilde{v}$, controlled by player 1, where player 1 can choose either to leave the game, in which case player 1 receives payoff $1 - p$ and player 0 receives payoff 1, or to move to v, from where the play continues in $\mathcal{G}$. Clearly, $\widetilde{\mathcal{G}}$ can be constructed in polynomial time from $\mathcal{G}$, v and p. To establish the reduction, we show that $\text{val}^{\mathcal{G}}(v) \geq p$ if and only if $(\widetilde{\mathcal{G}}, \tilde{v})$ has a Nash equilibrium with payoff $\geq (1, 1 - p)$.

($\Rightarrow$) Assume that $\text{val}^{\mathcal{G}}(v) \geq p$ and therefore $\text{val}_1^{\mathcal{G}}(v) \leq 1 - p$. Consider any positional strategy profile (σ_0, σ_1) of $(\widetilde{\mathcal{G}}, \tilde{v})$ where player 1 leaves the game and player 0 plays an optimal positional strategy in the game $(\mathcal{G}, v)$. Clearly, such a strategy profile has payoff $(1, 1 - p)$. We claim that (σ_0, σ_1) is also a Nash equilibrium: For player 0, it is obvious that no improvement is possible. For player 1, assume that she has a strategy τ that, together with σ_0, gives

her payoff $> 1 - p$. Then the residual strategy profile $(\sigma_0[\tilde{v}], \tau[\tilde{v}])$ must give player 1 payoff $> 1 - p$ in $(\mathcal{G}, v)$. But this is impossible since $\mathrm{val}_1^{\mathcal{G}}(v) \leq 1 - p$ and $\sigma_0[\tilde{v}] = \sigma_0$ is optimal.

($\Leftarrow$) Assume that $(\tilde{\mathcal{G}}, \tilde{v})$ has a Nash equilibrium with payoff $\geq (1, 1 - p)$, but $\mathrm{val}^{\mathcal{G}}(v) < p$. Then $\mathrm{val}_1^{\mathcal{G}}(v) > 1 - p$, and player 1's best response is to play from $\tilde{v}$ to v with probability 1 and to apply a strategy that gives her payoff $> 1 - p$ afterwards. But then player 0 receives payoff < 1, a contradiction. □

5.3 The qualitative fragment for deterministic games

The aim of this section is to prove that the qualitative fragment of NE, in the following denoted by QualNE, is decidable for deterministic games:

> Given a finite deterministic game $(\mathcal{G}, v_0)$ and $\bar{x}, \bar{y} \in \{0, 1\}^\Pi$, decide whether $(\mathcal{G}, v_0)$ has a Nash equilibrium with payoff $\geq \bar{x}$ and $\leq \bar{y}$.

As a by-product of the proof, we show that pure strategies are as powerful as arbitrary randomised strategies in this context. Since in a deterministic game every pure strategy profile has a binary payoff, this implies that the problem PureNE is decidable for deterministic games as well.

Characterisation of existence

The decidability of QualNE relies on the following characterisation of the existence of a Nash equilibrium with a qualitative constraint on the payoff in deterministic games with prefix-independent objectives, which resembles Proposition 5.1 for stochastic games. As in Section 5.1, given a deterministic game $\mathcal{G}$, we denote by W_i the set of all vertices v such that $\mathrm{val}_i^{\mathcal{G}}(v) > 0$ or, equivalently, $\mathrm{val}_i^{\mathcal{G}}(v) = 1$; we call W_i the *winning region* of player i.

Proposition 5.22. Let $(\mathcal{G}, v_0)$ be a finite deterministic game with prefix-independent objectives $\mathrm{Win}_i \subseteq V^\omega$, and let $\bar{x}, \bar{y} \in \{0, 1\}^k$. Then the following statements are equivalent:

1. there exists a Nash equilibrium of $(\mathcal{G}, v_0)$ with payoff $\geq \bar{x}$ and $\leq \bar{y}$;
2. there exists a strategy profile $\bar{\sigma}$ of $(\mathcal{G}, v_0)$ with payoff $\geq \bar{x}$ and $\leq \bar{y}$ such that $\mathrm{Pr}_{v_0}^{\bar{\sigma}}(\mathrm{Reach}(W_i) \setminus \mathrm{Win}_i) = 0$ for each player i;
3. there exists a play π of $(\mathcal{G}, v_0)$ with payoff $\geq \bar{x}$ and $\leq \bar{y}$ such that $\pi \in \mathrm{Win}_i$ or $\pi \notin \mathrm{Reach}(W_i)$ for each player i;
4. there exists a pure Nash equilibrium of $(\mathcal{G}, v_0)$ with payoff $\geq \bar{x}$ and $\leq \bar{y}$.

If additionally all objectives are ω-regular, then each of the above statements is equivalent to the following statement:

5. there exists a pure finite-state Nash equilibrium of $(\mathcal{G}, v_0)$ with payoff $\geq \bar{x}$ and $\leq \bar{y}$.

Proof. (1. $\Rightarrow$ 2.) Let $\bar{\sigma}$ be a Nash equilibrium of $(\mathcal{G}, v_0)$ with payoff $\geq \bar{x}$ and $\leq \bar{y}$. It follows from Lemma 3.3 that $\Pr^{\bar{\sigma}}_{v_0}(\mathrm{Win}_i \mid \mathrm{Reach}(W_i)) = 1$ and therefore $\Pr^{\bar{\sigma}}_{v_0}(\mathrm{Reach}(W_i) \setminus \mathrm{Win}_i) = 0$ for each player i.

(2. $\Rightarrow$ 3.) Assume that $\bar{\sigma}$ is a strategy profile of $(\mathcal{G}, v_0)$ with payoff $\geq \bar{x}$ and $\leq y$ such that $\Pr^{\bar{\sigma}}_{v_0}(\mathrm{Reach}(W_i) \setminus \mathrm{Win}_i) = 0$, i.e. $\Pr^{\bar{\sigma}}_{v_0}(\mathrm{Win}_i \cup V^\omega \setminus \mathrm{Reach}(W_i)) = 1$, for each player i. Let $X \subseteq V^\omega$ be the set of plays of $(\mathcal{G}, v_0)$ with payoff $\geq \bar{x}$ and $\leq \bar{y}$. Clearly, $\Pr^{\bar{\sigma}}_{v_0}(X) = 1$. Hence, also $\Pr^{\bar{\sigma}}_{v_0}(X \cap \bigcap_{i \in \Pi}(\mathrm{Win}_i \cup V^\omega \setminus \mathrm{Reach}(W_i))) = 1$. In particular, there exists a play π with payoff $\geq x$ and $\leq y$ such that $\pi \in \mathrm{Win}_i$ or $\pi \notin \mathrm{Reach}(W_i)$ for each player i.

(3. $\Rightarrow$ 4.) Assume that π is a play of $(\mathcal{G}, v_0)$ with payoff $\geq \bar{x}$ and $\leq \bar{y}$ such that $\pi \in \mathrm{Win}_i$ or $\pi \notin \mathrm{Reach}(W_i)$ for each player i. There exists a pure strategy profile $\bar{\sigma}$ of $(\mathcal{G}, v_0)$ such that π is the unique play compatible with $\bar{\sigma}$. For each player i and each history xv that is consistent with $\bar{\sigma}$, we either have $\pi \in \mathrm{Win}_i$ and $\Pr^{\bar{\sigma}}_{v_0}(\mathrm{Win}_i \mid xv \cdot V^\omega) = 1$ or $\pi \notin \mathrm{Reach}(W_i)$ and $\mathrm{val}^{\mathcal{G}}_i(v) = 0$. Hence, by Lemma 3.7, $(\mathcal{G}, v_0)$ has a pure Nash equilibrium $\bar{\sigma}^*$ with the same payoff as $\bar{\sigma}$. In particular, $\bar{\sigma}^*$ has payoff $\geq \bar{x}$ and $\leq \bar{y}$.

(4. $\Rightarrow$ 1.) Trivial.

In the following, assume that, additionally, all objectives are ω-regular.

(3. $\Rightarrow$ 5.) Assume that π is a play of $(\mathcal{G}, v_0)$ with payoff $\bar{x} \leq (z_i)_{i \in \Pi} \leq \bar{y}$ such that $\pi \in \mathrm{Win}_i$ or $\pi \notin \mathrm{Reach}(W_i)$ for each player i. Consider the (deterministic) MDP $\mathcal{M}$ that is obtained from $\mathcal{G}$ by removing all vertices v such that $v \in W_i$ for some player i with $z_i = 0$ and imposing the objective

$$\mathrm{Win} := \bigcap_{\substack{i \in \Pi \\ z_i = 1}} \mathrm{Win}_i \cap \bigcap_{\substack{i \in \Pi \\ z_i = 0}} V^\omega \setminus \mathrm{Win}_i \,.$$

The MDP $\mathcal{M}$ is well-defined since its arena is a subarena of $\mathcal{G}$. With each of the objectives Win_i, the objective Win is prefix-independent and ω-regular. Since π fulfils this objective, player 0 has a winning strategy in $(\mathcal{M}, v_0)$. But then, by Corollary 2.13, she must also have a pure finite-state winning strategy. The claim now follows from Lemma 3.8.

(5. $\Rightarrow$ 1.) Trivial. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$
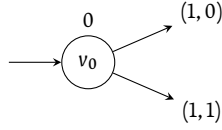
Figure 5.3. A game with only two pure but infinitely many randomised Nash equilibria.

An immediate corollary of Proposition 5.22 is that, for deterministic games with ω-regular objectives, pure finite-state strategies are as powerful as arbitrary randomised strategies as far as the existence of a Nash equilibrium with a qualitative constraint on the payoff is concerned.

Corollary 5.23. Let $(\mathcal{G}, v_0)$ be a finite deterministic game with ω-regular objectives, and let $\bar{x}, \bar{y} \in \{0,1\}^k$. Then $(\mathcal{G}, v_0)$ has a Nash equilibrium with payoff $\geq \bar{x}$ and $\leq \bar{y}$ if and only if $(\mathcal{G}, v_0)$ has a pure finite-state Nash equilibrium with payoff $\geq \bar{x}$ and $\leq \bar{y}$.

Proof. The claim follows from Proposition 5.22 and the fact that every game with ω-regular objectives can be reduced to one with parity objectives. □

As witnessed by the following simple example, Corollary 5.23 fails if the thresholds $\bar{x}$ and $\bar{y}$ are not binary.

Example 5.24. Consider the deterministic two-player game $\mathcal{G}$ depicted in Figure 5.3. Clearly, there exist precisely two pure Nash equilibria in $(\mathcal{G}, v_0)$, one with payoff $(1, 0)$ and one with payoff $(1, 1)$. However, for every $p \in [0, 1]$ there exists a randomised Nash equilibrium of $(\mathcal{G}, v_0)$ with payoff $(1, p)$.

A consequence of Corollary 5.23 is that the problems QualNE, PureNE and PureFinNE are polynomial-time equivalent for deterministic games with ω-regular objectives: deciding the existence of a pure (finite-state) Nash equilibrium with payoff $\geq \bar{x}$ and $\leq \bar{y}$ for $\bar{x}, \bar{y} \in [0,1]^\Pi$ amounts to deciding the existence a pure (finite-state) Nash equilibrium with payoff $\geq \lceil \bar{x} \rceil$ and $\leq \lfloor \bar{y} \rfloor$ (ceiling and floor applied componentwise); by Corollary 5.23, such an equilibrium exists if and only if there exists an arbitrary (possibly randomised) Nash equilibrium with this payoff. Hence, all the complexity bounds we are going to devise for the problem QualNE also apply to PureNE and PureFinNE.

Computational complexity

The decidability of QualNE for deterministic games follows from the de-
cidability of StrQualNE: by Corollary 5.23, it suffices to check for each bi-
nary payoff in between the thresholds whether there exists an equilibrium
with this payoff. However, since QualNE is more general than StrQualNE,
the complexity of QualNE is, a priori, higher. For games with Muller or
Streett objectives, this higher generality comes for free, and both problems
have the same complexity.

Theorem 5.25. QualNE is Pspace-complete for deterministic Muller games.

Proof. Hardness follows from Theorem 2.18. To prove membership in Pspace,
consider the polynomial-space algorithm for StrQualNE on the class of
Muller SMGs (Theorem 5.4). To decide QualNE, we just need to call this
algorithm for each binary payoff $\bar{x} \le \bar{z} \le \bar{y}$, which can also be done using
polynomial space. If the algorithm accepts for one such payoff, we know
that there exists a suitable Nash equilibrium. Otherwise, by Corollary 5.23,
no such equilibrium exists. □

Theorem 5.26. QualNE is NP-complete for deterministic Streett games.

Proof. NP-hardness follows from Theorem 5.6. Membership in NP follows
from Theorem 5.5 with almost the same reasoning as in the proof of Theo-
rem 5.25. Instead of enumerating all binary payoffs $\bar{x} \le \bar{z} \le \bar{y}$, the algorithm
just guesses such a payoff in the beginning. □

For deterministic Streett-Rabin games, we can show the same upper
bound as for StrQualNE: the problem lies in $P^{NP[\log]}$. However, we do not
know whether QualNE is complete for this class; we only know that the
problem is DP-hard (see Theorem 5.10).

Theorem 5.27. QualNE is in $P^{NP[\log]}$ for deterministic Streett-Rabin games.

Proof. To prove membership in $P^{NP[\log]}$, let us describe a polynomial-time
algorithm performing a logarithmic number of queries to an NP oracle for
the problem. We use the same trick as in the proof of Theorem 5.7. Given the
input $\mathcal{G}, v_0, \bar{x}, \bar{y}$, denote by R and S the set of players $i \in \Pi$ with a Rabin and
a Streett objective, respectively. The algorithm starts by determining the
number r of pairs (v, i) such that $i \in R$ and player i has a winning strategy
from vertex v. Again, it is easy to see that this number can be computed by

performing only a logarithmic number of queries to an NP oracle. Then we perform one more query; we ask whether for each player i there exists a set $Z_i \subseteq V$ as well as positional strategies $(\sigma_i)_{i \in R}$ and $(\tau_i)_{i \in S}$, where σ_i is a strategy of player i and τ_i is a strategy of the coalition $\Pi \setminus \{i\}$ in the coalition game $\mathcal{G}_i$, a binary payoff $\bar{z} = (z_i)_{i \in \Pi}$ with $\bar{x} \leq \bar{z} \leq \bar{y}$, and an end component $U \subseteq V$ with the following properties:

1. $\sum_{i \in R} |Z_i| = r$;
2. σ_i is winning from each vertex $v \in Z_i$ for each player $i \in R$;
3. τ_i is winning from each vertex $v \in V \setminus Z_i$ for each player $i \in S$;
4. $U \subseteq \bigcap_{i \in \Pi : z_i = 0} (V \setminus Z_i)$;
5. U is reachable from v_0 inside $\bigcap_{i \in \Pi : z_i = 0} (V \setminus Z_i)$;
6. U has payoff $\bar{z}$.

This query can be decided by an NP oracle by guessing suitable sets and strategies together with a suitable binary payoff and verifying 1.–7. in polynomial time. If the answer to the query is *yes*, the algorithm accepts; otherwise, it rejects.

Obviously, the algorithm runs in polynomial time. To see that the algorithm is correct, first note that for each player $i \in R$ ($i \in S$) the set Z_i is an under- (over-)approximation of the winning region W_i. In fact, for each player $i \in R$, we have $Z_i = W_i$. Otherwise, for some player $i \in R$, there would exist a vertex $v \in W_i \setminus Z_i$. But then, the number of pairs (v, i) with $i \in R$ and $v \in W_i$ would be strictly greater than r, a contradiction. Now assume that the algorithm accepts its input. Then there exists a play π of $(\mathcal{G}, v_0)$ with $\mathrm{Inf}(\pi) = U$ that stays inside $\bigcap_{i \in \Pi : z_i = 0} (V \setminus W_i)$. In particular, π has payoff $\bar{z}$. By Proposition 5.22, this play can be extended to a pure Nash equilibrium with the same payoff. On the other hand, if there exists a pure Nash equilibrium with payoff $\geq \bar{x}$ and $\leq \bar{y}$, then the query will succeed because we can set Z_i to W_i, σ_i to a globally optimal positional strategy of player i for each player $i \in R$, τ_i to a globally optimal positional strategy of the coalition $\Pi \setminus \{i\}$ for each player $i \in S$ (such strategies exist by Theorem 2.11), $\bar{z}$ to the payoff of the equilibrium, and U to the set of vertices visited infinitely often in the equilibrium; with these choices, all of the above properties are fulfilled. $\square$

An immediate consequence of Theorem 5.26 is that QualNE belongs to NP for deterministic parity games. However, in many cases, we can do better: For two payoff vectors $\bar{x}, \bar{y} \in \{0, 1\}^{\Pi}$, denote by $\mathrm{dist}(\bar{x}, \bar{y})$ the *Hamming distance* of $\bar{x}$ and $\bar{y}$, i.e. the number $\sum_{i \in \Pi} |y_i - x_i|$ of non-matching bits. Note that $\mathrm{dist}(\bar{x}, \bar{y})$

is always bounded by the number of players. We show that, if $\text{dist}(\bar{x}, \bar{y})$ is bounded by a constant, then QualNE is in UP $\cap$ coUP.

Theorem 5.28. QualNE is in UP $\cap$ coUP for deterministic parity games and bounded $\text{dist}(\bar{x}, \bar{y})$.

Proof. Assume that $\text{dist}(\bar{x}, \bar{y})$ is bounded. An unambiguous nondeterministic algorithm for QualNE works as follows: On input $\mathcal{G}, v_0, \bar{x}, \bar{y}$, the algorithm starts by guessing the winning region W_i of each player. Then, for each vertex v and each player i, the guess whether $v \in W_i$ or $v \notin W_i$ is verified by running the UP algorithm for the respective problem. If one guess was incorrect, the algorithm rejects immediately. Otherwise, the algorithm checks for each payoff $\bar{z} \in \{0,1\}^\Pi$ with $\bar{x} \le \bar{z} \le \bar{y}$ whether there exists a winning play from v_0 in the one-player Streett game with objective $\bigcap_{i \in \Pi: z_i = 1} \text{Win}_i \cap \bigcap_{i \in \Pi: z_i = 0} (C^\omega \setminus \text{Win}_i)$, played on the subarena $\bigcap_{i \in \Pi: z_i = 0} (V \setminus W_i)$ of $\mathcal{G}$. The algorithm accepts if such a play exists for at least one payoff $\bar{z}$. Analogously, a UP algorithm for the complement of QualNE accepts if there is *no* such play for all admissible payoffs $\bar{z}$.

Clearly, both algorithms run in polynomial time; they are unambiguous because they only accept if each winning region has been guessed correctly. Finally, their correctness follows from Proposition 5.22. □

If there were to exist a polynomial-time algorithm for the qualitative decision problem for deterministic parity games, then we could compute the winning region for each player efficiently, and QualNE would be decidable in polynomial time for deterministic parity games and bounded $\text{dist}(\bar{x}, \bar{y})$. In general, a polynomial-time algorithm for the qualitative decision problem on a class $\mathcal{C}$ of deterministic two-player zero-sum parity games can be extended to a polynomial-time algorithm for QualNE with bounded $\text{dist}(\bar{x}, \bar{y})$ on the class $\mathcal{C}^*$ of all deterministic multiplayer parity games where for each player i the coalition game $\mathcal{G}_i$ is in $\mathcal{C}$.

Corollary 5.29. Let $\mathcal{C}$ be a class of finite deterministic two-player zero-sum parity games such that the qualitative decision problem is decidable in P for games in $\mathcal{C}$. Then QualNE is in P for games in $\mathcal{C}^*$ and bounded $\text{dist}(\bar{x}, \bar{y})$.

Proof. By Proposition 5.22, we only need to check for every binary payoff $\bar{x} \le \bar{z} \le \bar{y}$, whether $\mathcal{G}$ has a Nash equilibrium with payoff $\bar{z}$. By Theorem 5.13, the latter property can be checked in polynomial time if $\mathcal{G} \in \mathcal{C}^*$. □

Instead of assuming that dist$(\bar{x}, \bar{y})$ is bounded by a constant, we can treat this number as a parameter for the problem QualNE. Theorem 5.13 implies that QualNE is *fixed-parameter tractable* with respect to this parameter on every suitable class $\mathcal{C}^*$, i.e. there exists a deterministic algorithm that decides QualNE for games in $\mathcal{C}^*$ in time exponential in the size of the parameter but polynomial in the size of the game. The proof is virtually identical to the proof of Corollary 5.29.

Corollary 5.30. Let $\mathcal{C}$ be a class of finite deterministic two-player zero-sum parity games such that the qualitative decision problem is decidable in P for games in $\mathcal{C}$. Then QualNE is fixed-parameter tractable with respect to the parameter dist$(\bar{x}, \bar{y})$ for games in $\mathcal{C}^*$.

In particular, it follows from Corollary 5.30 and Theorem 2.19 that QualNE is fixed-parameter tractable for deterministic parity games with a bounded number of priorities.

Corollary 5.31. For each $d \in \mathbb{N}$, QualNE is fixed-parameter tractable with respect to the parameter dist$(\bar{x}, \bar{y})$ for deterministic parity games with at most d priorities.

The natural question at this point is whether QualNE is actually decidable in polynomial time for parity games with a bounded number of priorities. As witnessed by the following theorem, this is quite unlikely: QualNE is NP-hard for deterministic co-Büchi games, and the same is true for QualSPE, the analogous problem for subgame-perfect equilibria.

Theorem 5.32. QualNE and QualSPE are NP-hard for deterministic games with co-Büchi objectives.

Proof. Again, we reduce from SAT. Given a Boolean formula $\varphi = C_1 \wedge \cdots \wedge C_m$ in conjunctive normal form over propositional variables $X_1, \ldots, X_n$, where without loss of generality $m \geq 1$ and each clause is nonempty, we build a game $\mathcal{G}$ played by players $0, 1, \ldots, n$ as follows: The game $\mathcal{G}$ has vertices $C_1, \ldots, C_m$, controlled by player 0, and for each clause C and each literal L that occurs in C a vertex (C, L), controlled by player i if $L = X_i$ or $L = \neg X_i$; additionally, the game contains a terminal vertex $\bot$. There are edges from a clause C_j to each vertex (C_j, L) such that L occurs in C_j and from there to $C_{(j \bmod m)+1}$, and there is an edge from each vertex of the form $(C, \neg X)$ to $\bot$. The arena of $\mathcal{G}$ is schematically depicted in Figure 5.4. The players' objectives are as follows:
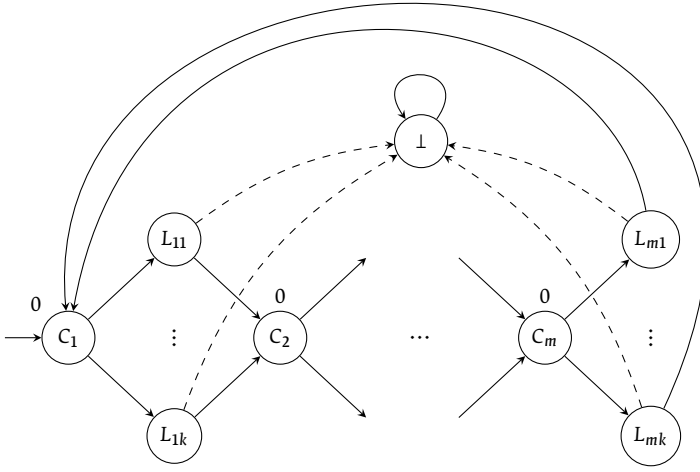
Figure 5.4. Reducing SAT to QualNE for deterministic co-Büchi games.

- player 0 wins if ⊥ is visited only finitely often (i.e. never);
- player $i \neq 0$ wins if each vertex (C, X_i) is visited only finitely often.

Clearly, $\mathcal{G}$ can be constructed from φ in polynomial time. To establish both reductions, we prove the equivalence of the following three statements:

1. φ is satisfiable;
2. $(\mathcal{G}, C_1)$ has a subgame-perfect equilibrium where player 0 wins;
3. $(\mathcal{G}, C_1)$ has a Nash equilibrium where player 0 wins.

(1. ⇒ 2.) Assume that $\alpha \colon \{X_1, \ldots, X_n\} \to \{\text{true}, \text{false}\}$ is a satisfying assignment for φ. We show that the positional strategy profile $\bar{\sigma}$ where at any time player 0 plays from a clause C to a fixed vertex (C, L) such that L is mapped to true by α and each player $i \neq 0$ plays from $\neg X_i$ to ⊥ if and only if $\alpha(X_i) = \text{true}$ is a subgame-perfect equilibrium of $(\mathcal{G}, C_1)$ where player 0 wins. First note that the induced play never reaches ⊥; hence player 0 wins. To show that $\bar{\sigma}$ is a subgame-perfect equilibrium, we only need to prove that $\bar{\sigma}$ is a Nash equilibrium of $(\mathcal{G}, v)$ for every vertex v. If $v = (C, \neg X_i)$ and $\alpha(X_i) = \text{true}$, then player i moves to ⊥ immediately. Hence, all players apart from player 0 win, but player 0 cannot improve her payoff since no clause vertex is visited. Otherwise, the induced play never reaches ⊥, and player 0 wins. Consider any player i whose objective is violated. Hence, a vertex of the form (C, X_i) is visited infinitely often. However, as player 0 plays according

to the satisfying assignment, no vertex of the form $(C', \neg X_i)$ is ever visited. Hence, player i cannot improve her payoff by playing to $\bot$.

($2. \Rightarrow 3.$) Trivial.

($3. \Rightarrow 1.$) Assume that $(\mathcal{G}, C_1)$ has a, without loss of generality pure, Nash equilibrium where player 0 wins. Since player 0 wins, the terminal vertex $\bot$ is not reached in the induced play π. Moreover, we claim that it is not the case that both a vertex (C, X_i) and a vertex $(C', \neg X_i)$ are visited infinitely often in π. Otherwise, player i would lose, but could improve her payoff by playing from $(C', \neg X_i)$ to $\bot$, a contradiction. Now consider the variable assignment α that maps X to true if some vertex (C, X) is visited infinitely often; we claim that α satisfies the formula. Consider any clause C. By the construction of $\mathcal{G}$, there exists a literal L in C such that the vertex (C, L) is visited infinitely often in π. Hence, α maps L to true and satisfies C. □

Theorem 5.32 leaves open the existence of a polynomial-time algorithm for QualNE on the class of deterministic Büchi games. In fact, we can give a polynomial-time algorithm that computes, given a deterministic game $\mathcal{G}$ with Büchi objectives given by $F_i \subseteq C$ and thresholds $\bar{x}, \bar{y} \in \{0, 1\}^{\Pi}$, the set of vertices from where there exists a Nash equilibrium with payoff $\geq \bar{x}$ and $\leq \bar{y}$.

Algorithm 5.3 is a variant of the classical algorithm for deciding the existence of a winning play in a deterministic one-player Streett game, due to Emerson & Lei (1987), and works as follows: By Proposition 5.22, the game $(\mathcal{G}, v)$ has a Nash equilibrium with payoff $\geq \bar{x}$ and $\leq \bar{y}$ if and only if there exists a play π with this payoff that stays outside the winning region W_i of each player i with $\chi(\text{Inf}(\pi)) \cap F_i = \varnothing$. Clearly, such a play exists if and only if there exists a payoff $\bar{z} \in \{0, 1\}^{\Pi}$ with $\bar{x} \leq \bar{z} \leq \bar{y}$ and an end component $U \subseteq \bigcap_{i \in \Pi : z_i = 0}(V \setminus W_i)$ with payoff $\bar{z}$ that is reachable from v inside $\bigcap_{i \in \Pi : z_i = 0}(V \setminus W_i)$. The essential part of the algorithm is the procedure SolveSubgame; on input X, its task is to find any such set contained in X.

At first, SolveSubgame computes all end components of $\mathcal{G}$ maximal in X. For each such end component U, the procedure performs the following steps: First, the set P of players i such that $\chi(U) \cap F_i = \varnothing$ is computed. If this set contains a player i with $x_i = 1$, there is no hope of finding an end component with a suitable payoff inside U, and U does not have to be explored further. Otherwise, the algorithm checks whether U does not intersect with the winning region of each player $i \in P$. If so, we have found an end component with a suitable payoff $\bar{z}$, namely $z_i = 1$ if and only if $i \in P$. Hence, the procedure

---

Algorithm 5.3. Solving QualNE for deterministic Büchi games.

Input: deterministic Büchi game $\mathcal{G} = (\Pi, V, (V_i)_{i \in \Pi}, \Delta, \chi, (F_i)_{i \in \Pi}), \bar{x}, \bar{y} \in \{0, 1\}^\Pi$
Output: $\{v \in V : (\mathcal{G}, v) \text{ has a Nash equilibrium with payoff} \geq \bar{x} \text{ and} \leq \bar{y}\}$

compute $W_i := \{v \in V : \mathrm{val}_i^{\mathcal{G}}(v) > 0\}$ for each $i \in \Pi$
$X := \bigcap_{i \in \Pi : y_i = 0}(V \setminus \chi^{-1}(F_i))$
output SolveSubgame(X)

procedure SolveSubgame(X)
  $Z := \varnothing$
  compute all end components of $\mathcal{G}$ maximal in X
  for each such end component U **do**
    $P := \{i \in \Pi : \chi(U) \cap F_i = \varnothing\}$
    if $i \notin P$ for all i with $x_i = 1$ **then**
      $Y := U \cap \bigcap_{i \in P}(V \setminus W_i)$
      if $Y = U$ **then**
        $Z := Z \cup \{v \in V : U \text{ reachable from } v \text{ in } \mathcal{G} \upharpoonright \bigcap_{i \in P}(V \setminus W_i)\}$
      else
        $Z := Z \cup$ SolveSubgame(Y)
      end if
    end if
  end for
  return Z
end procedure

---

adds U and all vertices from where U is reachable inside $\bigcap_{i \in P}(V \setminus W_i)$ to the result set. Otherwise, it removes the winning region of each player $i \in P$ from U. The resulting set of vertices may not be strongly connected any more and fewer objectives may be satisfied; hence, the procedure has to be called recursively.

As we are not interested in end components that are winning for some player i with $y_i = 0$, SolveSubgame is called initially on the subarena of $\mathcal{G}$ that results from removing all vertices v such that $\chi(v) \in F_i$ for some player i with $y_i = 0$.

Theorem 5.33. QualNE is decidable in polynomial time for deterministic Büchi games.

Proof. We claim that Algorithm 5.3 outputs the set of vertices from where there is a Nash equilibrium with payoff $\geq \bar{x}$ and $\leq \bar{y}$. Since the number of recursive calls is bounded by the size of the arena and maximal end components can be computed in polynomial time, the procedure SolveSubgame runs in polynomial time. For each player i, the set of vertices from where she has a winning strategy can also be computed in polynomial time (Theorem 2.20). Hence, the algorithm runs in polynomial time.

To prove the correctness of the algorithm, let $Z \subseteq V$ be the output of the algorithm on input $\mathcal{G}, \bar{x}.\bar{y}$. We claim that Z equals the set of vertices $v \in V$ such that $(\mathcal{G}, v)$ has a Nash equilibrium with payoff $\geq \bar{x}$ and $\leq \bar{y}$.

($\subseteq$) Assume that $v \in Z$. Hence, in some call of the procedure SolveSubgame, say on input X, the algorithm finds a maximal end component U of $\mathcal{G} \upharpoonright X$ reachable from v and contained inside $\bigcap_{i \in P}(V \setminus W_i)$ such that $\chi(U) \cap F_i = \varnothing$ only if $x_i = 0$. Let $\bar{z} \in \{0, 1\}^{\Pi}$ be defined by $z_i = 1$ if and only if $\chi(U) \cap F_i \neq \varnothing$; in particular, $\bar{x} \leq \bar{z}$. Since U is strongly connected, we can build a play π of $(\mathcal{G}, v)$ such that $\text{Inf}(\pi) = U$ and $\pi \notin \text{Reach}(W_i)$ for each player i with $z_i = 0$. Since $\text{Inf}(\pi) = U$, this play has payoff $\bar{z}$. By Proposition 5.22, π can be extended to a pure Nash equilibrium of $(\mathcal{G}, v)$ with payoff $\bar{z}$. Moreover, since the algorithm maintains the invariant that $\chi(X) \cap F_i = \varnothing$ for each $i \in \Pi$ with $y_i = 0$, we have $\bar{z} \leq \bar{y}$.

($\supseteq$) Suppose there exists a (pure) Nash equilibrium of $(\mathcal{G}, v)$ with payoff $\bar{x} \leq \bar{z} \leq \bar{y}$. Hence, by Proposition 5.22, there exists a play π of $(\mathcal{G}, v)$ with payoff $\bar{z}$ such that $\pi \notin \text{Reach}(W_i)$ for each player i with $z_i = 0$. Let X be defined as in the first call of the procedure SolveSubgame. By the definition of π, the set $\text{Inf}(\pi)$ is contained in a maximal end component U of $\mathcal{G} \upharpoonright X$. Let $P := \{i \in \Pi : \chi(U) \cap F_i = \varnothing\}$. If U is contained in $\bigcap_{i \in P}(V \setminus W_i)$, then the algorithm adds v to Z immediately. Otherwise, $\text{Inf}(\pi) \subseteq U \cap \bigcap_{i \in P}(V \setminus W_i) = Y$, and $\text{Inf}(\pi)$ is contained in a maximal end component of $\mathcal{G} \upharpoonright Y$. Hence, the procedure will eventually find a maximal end component U of $\mathcal{G} \upharpoonright X$ such that U is contained in $\bigcap_{i \in P}(V \setminus W_i)$ and $\text{Inf}(\pi) \subseteq U$. In particular, $P \subseteq \{i \in \Pi : z_i = 0\}$, and U is reachable from v inside $\mathcal{G} \upharpoonright \bigcap_{i \in P}(V \setminus W_i)$. Hence, v is added to Z. $\square$

Remark. By combining the proofs of Corollary 5.29 and Theorem 5.33, we can show that QualNE is, in fact, decidable in polynomial time for deterministic games with an arbitrary number of Büchi objectives and a bounded number of co-Büchi objectives (or even a bounded number of parity objectives with a bounded number of priorities).

Table 5.1. The complexity of StrQualNE, OneNE and QualNE.

	StrQualNE	OneNE	QualNE (det. games)
SSMGs	P-complete	$NP \cap coNP$	P-complete
Büchi	P-complete	$NP \cap coNP$	P-complete
co-Büchi	P-complete	$NP \cap coNP$	NP-complete
Parity[d]	P-complete	$NP \cap coNP$	NP-complete
Parity	$UP \cap coUP$	$NP \cap coNP$	NP-complete
Streett	NP-complete	NP-complete	NP-complete
Rabin	$P^{NP[\log]}$-complete	P^{NP} $P^{NP[\log]}$-hard	$P^{NP[\log]}$ DP-hard
Muller	Pspace-complete	Pspace-complete	Pspace-complete

5.4 Summary of results

Our main results on the complexity of StrQualNE and OneNE for SMGs as well as QualNE for deterministic games are summarised in Table 5.1; similarly to Section 2.5, Parity[d] denotes the class of all parity SMGs whose priority functions refer to at most d priorities. All lower bounds also hold for the corresponding problems for subgame-perfect equilibria, and the upper bounds for games with Rabin objectives also hold for games with Streett-Rabin objectives.

6

Conclusion

In this final chapter, we sum up the main results of this work and list some open problems. Finally, we discuss several perspectives for future work.

6.1 Summary and open problems

In this work, we have analysed the existence and computational complexity of game-theoretic equilibrium concepts in the context of turn-based stochastic games played on graphs.

In Chapter 3, we proved that every finite stochastic game with prefix-independent objectives has a Nash equilibrium in pure strategies (Theorem 3.9); if each player has an ω-regular objective, there also exists a Nash equilibrium in pure finite-state strategies (Theorem 3.10). The corresponding statements about subgame-perfect equilibria are only true for deterministic games (Theorems 3.15 and 3.16); there exists a stochastic game with Büchi objectives that has no subgame-perfect equilibrium (Proposition 3.18).

In Chapter 4, we studied the different decision problems associated to a solution concept and a strategy type in their full generality. On the positive side, we proved the problems PosNE, PosSPE, StatNE and StatSPE to be decidable; the former two problems are NP-complete, while the latter two problems are contained in PSPACE and hard for both NP and SqrtSum. What remains open is the precise complexity of StatNE and StatSPE. Given that

these problems are intimately connected to the square-root sum problem and the existential theory of the reals, such bounds seem hard to come by.

Problem 6.1. What is the precise complexity of StatNE and StatSPE?

We continued our analysis by proving that even the qualitative fragments of all other decision problems (NE, SPE, PureNE, PureSPE, FinNE, FinSPE, PureFinNE and PureFinSPE) are undecidable for SSMGs with a bounded number of players (Theorems 4.10 and 4.13). These results leave open the decidability of these problems for games with a small number of players. In particular, it is conceivable that NE or PureNE is decidable for two-player SSMGs, or even for two-player SMGs with ω-regular objectives. In fact, it follows from the decidability of OneNE (Theorem 5.15) that for two-player SMGs with ω-regular objectives the *qualitative fragments* of NE and PureNE coincide and are indeed decidable.

Problem 6.2. Is NE decidable for two-player SSMGs? Is NE decidable for two-player SMGs with ω-regular objectives?

Problem 6.3. Is PureNE decidable for two-player SSMGs? Is PureNE decidable for two-player SMGs with ω-regular objectives?

In Chapter 5, we looked at several restrictions of the original decision problems. In particular, we proved that the problems StrQualNE and OneNE are decidable for SMGs with ω-regular objectives and that QualNE is decidable for deterministic games with ω-regular objectives. There are several open questions regarding these results. For instance, the precise complexity of OneNE for Streett-Rabin SMGs with a Rabin objective for player 0 remains open: while we could prove membership in P^{NP} (Theorem 5.17), we could only show that this problem is hard for $P^{NP[\log]}$ (Theorem 5.19).

Problem 6.4. What is the precise complexity of OneNE for Streett-Rabin SMGs with a Rabin objective for player 0?

A similar open problem is to determine the precise complexity of QualNE for deterministic Streett-Rabin games; for this problem, we proved membership in $P^{NP[\log]}$ (Theorem 5.27) but hardness only for DP (Theorem 5.10).

Problem 6.5. What is the precise complexity of QualNE for deterministic Streett-Rabin games?

More fundamentally, although we could establish the decidability of QualNE for deterministic games with ω-regular objectives, the decidability of NE (where arbitrary rational payoff thresholds are allowed) for these games remains open.

Problem 6.6. Is NE decidable for deterministic simple reachability games? Is NE decidable for deterministic games with ω-regular objectives?

The decidability of QualNE for deterministic games relies on Lemma 3.7, which gives a necessary and sufficient condition for the existence of a pure Nash equilibrium with a certain payoff, and the fact that pure strategies are sufficient to realise any binary payoff. To decide NE, we cannot employ Lemma 3.7 because pure strategies are not sufficient to realise any non-binary payoff. However, *almost pure* strategies, which require randomisation only for finitely many histories, do suffice for this purpose. We conjecture that Lemma 3.7 can be extended to almost pure strategies, which would yield a positive answer to Problem 6.6.

Another fundamental open question about the decision problems we studied in Chapter 5 is whether these problems are decidable when one looks for a subgame-perfect equilibrium instead of a Nash equilibrium.

Problem 6.7. Is StrQualSPE decidable for SSMGs? Is StrQualSPE decidable for SMGs with ω-regular objectives?

Problem 6.8. Is OneSPE decidable for SSMGs? Is OneSPE decidable for SMGs with ω-regular objectives?

Problem 6.9. Is QualSPE decidable for deterministic SSMGs? Is QualSPE decidable for deterministic games with ω-regular objectives?

The only nontrivial decidable decision problem about subgame-perfect equilibria in infinite-duration games we are aware of is PureSPE for deterministic games with ω-regular objectives (Ummels 2005, 2006). However, the best known algorithm for this problem requires exponential time, even for games with Büchi objectives, for which the NP lower bound of Theorem 5.32 does not apply. A related open question is whether pure strategies are sufficient to realise any subgame-perfect equilibrium with a binary payoff in deterministic games with ω-regular objectives (as for Nash equilibria). Since PureSPE is decidable for these games, a positive answer to this question would imply the decidability of QualSPE for these games.

6.2 Perspectives

Broadly speaking, this work can be extended along two axes: one can modify the game model, or one can modify the solution concept.

Different game models

A possible extension to the game model is to add *nondeterminism*, which can be used to model behaviour that is neither controllable nor describable by a probability distribution. The easiest way to incorporate nondeterminism in the model is to add another type of vertices: vertices of this additional type are neither stochastic nor controlled by a player; when a play arrives at such a vertex, a successor is chosen nondeterministically. Formally, these vertices can be assigned to a *Byzantine player*, whose strategy is unknown. A Nash equilibrium of such a game would be a profile of strategies for the remaining players that is a Nash equilibrium (in the classical sense) for every strategy of the Byzantine player. We conjecture that many of the decidability results of Chapter 5 carry over to this setting, albeit with higher complexity.

We already pointed out in Chapter 2 that our model lacks concurrency. In a concurrent game, whenever the play reaches a vertex, all players simultaneously choose an action; the chosen profile of actions determines a probability distribution on successor vertices. The arguments for the decidability of PosNE and StatNE for SMGs also prove the decidability of these problems for concurrent SMGs. On the other hand, problems that are already undecidable for turn-based SMGs are also undecidable for concurrent SMGs. In fact, it is easy to extend our undecidability proof for NE to prove that even the existence of *any* Nash equilibrium is undecidable for concurrent SMGs. (Note that a concurrent SMG may fail to have an equilibrium.) On the other hand, Fisman et al. (2010) proved that PureNE is decidable for *deterministic* concurrent games with ω-regular objectives (see also Bouyer et al. 2010b).

Recently, much effort has been invested into extending the algorithmic results on two-player zero-sum games with finitely many states to games with a countably infinite number of states. In particular, games that are played on the configuration graphs of pushdown automata, so-called *pushdown games*, have been studied thoroughly. Walukiewicz (2001) showed that deciding the winner of a deterministic two-player zero-sum parity pushdown game is Exptime-complete, which gives hope that the decidability

results for deterministic games in Chapter 5 can be extended to pushdown games. For stochastic pushdown games, however, most problems are undecidable, even for MDPs (Etessami & Yannakakis 2005). Hence, in order to obtain decidability results, one has to consider more restricted classes such as *one-exit recursive stochastic games* (Etessami & Yannakakis 2005) or *one-counter stochastic games* (Brázdil et al. 2010).

An extension whose semantics are actually given by concurrent games with uncountably many states is the model of a *timed game*. In such a game, a set of clocks is used to measure real time: states and transitions have *guards*, which specify for which clock values resting in a state or taking a transition is legal, and clocks can be reset along transitions. For non-stochastic timed games, preliminary results on the complexity of computing Nash equilibria were recently obtained by Bouyer et al. (2010a,b).

As pointed out in Chapter 1, in the original stochastic game model, introduced by Shapley (1953), the objective of a player is not given by a set of plays, but by *rewards* that are assigned to states or transitions. There are several ways to obtain a payoff from the infinite stream of rewards a player receives; popular payoff functions include the *discounted* sum of the accumulated rewards (with respect to some discount factor $\lambda < 1$), their *limit average* and their *total* sum (which can be infinite). An SSMG can be viewed as a limit-average game, where non-zero rewards occur only on terminal vertices, or as a total-reward game (by introducing intermediate vertices on transitions to terminal vertices). Hence, all our lower bounds for SSMGs also apply to games with limit-average or total-reward objectives; we conjecture that similar bounds hold for discounted games.

Different solution concepts

After Nash (1950) introduced his equilibrium concept, a plethora of other solution concepts have been introduced to mitigate the drawbacks of this solution concept (such as the requirement for coordination). Formally, a solution concept assigns to a game a set of strategy profiles, each of which is a *solution* of the game. We say that a solution concept $\mathcal{C}$ is *stronger* (*weaker*) than a solution concept $\mathcal{D}$ if on every game the set of all $\mathcal{C}$ solutions is a subset (superset) of the set of all $\mathcal{D}$ solutions. In the literature, both solution concepts stronger and solution concepts weaker than Nash equilibrium have been explored: prominent examples in the former category are *strong* and

subgame-perfect equilibria; examples in the latter category are *correlated equilibria* and *rationalisability* (see Osborne & Rubinstein 1994).

In a Nash equilibrium, no player can gain from switching to a different strategy. It is natural to relax this requirement by allowing players to gain a small amount from deviating. Formally, for $\varepsilon > 0$, a strategy profile is an *ε-equilibrium* if each player cannot increase her probability of winning by more than ε when switching to a different strategy. ε-Equilibria are a suitable alternative to Nash equilibria when the latter solution concept is too strong. For instance, in concurrent games, the existence of ε-Nash equilibria is usually the best one can hope for (see Section 1.1). Our undecidability proof for Nash equilibria relies heavily on the fact that profitable deviations are forbidden in Nash equilibria. Hence, it is conceivable that problems such as NE and PureNE become decidable when we take ε-equilibria into account.

As argued by Abraham et al. (2006), of particular relevance for distributed computing are equilibria in which a deviation of up to a certain number k of players does not *increase* these players' payoffs and in which a deviation of up to a possibly different number t of players does not *decrease* the payoff of the other players; such an equilibrium is called (k, t)-*robust*. By definition, every Nash equilibrium is $(1, 0)$-robust, and a (k, t)-robust equilibrium is also a Nash equilibrium as long as $k \geq 1$. It seems that most of our proofs do not extend to, for instance, $(2, 0)$-robust equilibria or $(1, 1)$-robust equilibria. From an optimistic point of view, this might enable more decidability results.

Another refinement of Nash equilibrium, which can be used for *assume-guarantee reasoning*, has been introduced by Chatterjee et al. (2006) under the name *secure equilibria*. Intuitively, such an equilibrium captures rational behaviour if a player is not only interested in maximising her own payoff but also in decreasing the other players' payoffs. More precisely, a Nash equilibrium is secure if each player can only decrease another player's payoff by decreasing her own payoff. As for ε-equilibria and (k, t)-robust equilibria, we do not know whether our results carry over to this equilibrium notion.

Appendix A

Preliminaries

In this appendix, we review concepts from probability and complexity theory that are used in this work. For a thorough introduction to these topics, we recommend the textbooks (Billingsley 1995) and (Papadimitriou 1994), respectively.

A.1 Probability theory

Let Ω be an arbitrary nonempty set, called the *sample space*. An *algebra* over Ω is a collection $\mathcal{F} \subseteq \mathcal{P}(\Omega)$, whose elements are called *events*, that contains Ω and is closed under complementation and taking finite unions:

- $\Omega \in \mathcal{F}$;
- if $X \in \mathcal{F}$, then $\Omega \setminus X \in \mathcal{F}$;
- if $X, Y \in \mathcal{F}$, then $X \cup Y \in \mathcal{F}$.

A σ-*algebra* is an algebra $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ that is also closed under taking countable unions:

- if $X_0, X_1, \ldots \in \mathcal{F}$, then $\bigcup_{n \in \mathbb{N}} X_n \in \mathcal{F}$.

Given an arbitrary collection $\mathcal{F} \subseteq \mathcal{P}(\Omega)$, we denote by $\sigma(\mathcal{F})$ the σ-algebra *generated by* $\mathcal{F}$. Formally, $\sigma(\mathcal{F})$ is the intersection of all σ-algebras that contain $\mathcal{F}$. If $\mathcal{F}$ is an algebra, then $\sigma(\mathcal{F})$ can alternatively be characterised as the least *monotone class* that contains $\mathcal{F}$. Formally, we say that a collection $\mathcal{M} \subseteq \mathcal{P}(\Omega)$ is *monotone* if it is closed under taking limits of chains:

- If $X_0 \subseteq X_1 \subseteq \cdots \in \mathcal{M}$, then $\bigcup_{n \in \mathbb{N}} X_n \in \mathcal{M}$;
- If $X_0 \supseteq X_1 \supseteq \cdots \in \mathcal{M}$, then $\bigcap_{n \in \mathbb{N}} X_n \in \mathcal{M}$.

Obviously, any σ-algebra is monotone.

Theorem A.1 (Monotone class theorem). Let $\mathcal{F}$ be an algebra, and let $\mathcal{M}$ be a monotone collection of subsets of Ω. If $\mathcal{F} \subseteq \mathcal{M}$, then $\sigma(\mathcal{F}) \subseteq \mathcal{M}$. In particular, $\sigma(\mathcal{F})$ is the smallest monotone class that contains $\mathcal{F}$.

Given an algebra $\mathcal{F} \subseteq \mathcal{P}(\Omega)$, a function $P \colon \mathcal{F} \to [0,1]$ is a *probability measure* on $\mathcal{F}$ if it satisfies the following properties:

- $P(\varnothing) = 0$;
- $P(\Omega) = 1$;
- if $X_0, X_1, \ldots \in \mathcal{F}$ is a sequence of pairwise disjoint sets with $\bigcup_{n \in \mathbb{N}} X_n \in \mathcal{F}$, then $P(\bigcup_{n \in \mathbb{N}} X_n) = \sum_{n \in \mathbb{N}} P(X_n)$.

If $P(X) = 1$, we say that the event X happens *almost surely*. The following laws are proved easily from the axioms.

Proposition A.2. Let $\mathcal{F}$ be an algebra with $X, Y \in \mathcal{F}$, and let P be a probability measure on $\mathcal{F}$.

1. $P(\Omega \setminus X) = 1 - P(X)$;
2. $P(X \cup Y) = P(X) + P(Y)$ if $X \cap Y = \varnothing$;
3. $P(Y) = P(X) + P(Y \setminus X) \geq P(X)$ if $X \subseteq Y$;
4. $P(X \cap Y) = P(X)$ if $P(Y) = 1$.

Moreover, probability measures are *continuous* from below and above.

Proposition A.3. Let P be a probability measure on an algebra $\mathcal{F}$.

1. If $X_0 \subseteq X_1 \subseteq \cdots \in \mathcal{F}$ and $\bigcup_{n \in \mathbb{N}} X_n \in \mathcal{F}$, then $P(\bigcup_{n \in \mathbb{N}} X_n) = \lim_n P(X_n)$;
2. if $X_0 \supseteq X_1 \supseteq \cdots \in \mathcal{F}$ and $\bigcap_{n \in \mathbb{N}} X_n \in \mathcal{F}$, then $P(\bigcap_{n \in \mathbb{N}} X_n) = \lim_n P(X_n)$.

For an arbitrary sequence $X_0, X_1, \ldots \in \mathcal{F}$ of events, define

$$\liminf_n X_n := \bigcup_{n \in \mathbb{N}} \bigcap_{k \geq n} X_k \quad \text{and}$$
$$\limsup_n X_n := \bigcap_{n \in \mathbb{N}} \bigcup_{k \geq n} X_k \,.$$

The set $\liminf_n X_n$ consists of all elements that occur in all but finitely many of the sets X_n, and $\limsup_n X_n$ consists of all elements that occur in infinitely many of the sets X_n.

Proposition A.4. Let P be a probability measure on a σ-algebra $\mathcal{F}$, and let $X_0, X_1, \ldots \in \mathcal{F}$. Then $P(\liminf_n X_n) \leq \liminf_n P(X_n) \leq \limsup_n P(X_n) \leq P(\limsup_n X_n)$.

How can we set up a probability measure? In applications, it is often easier to define a probability measure on an algebra rather than on a σ-algebra. However, an algebra might be too small, and we would like to assign a probability to more events, e.g. to all sets in the generated σ-algebra. The following theorem allows us to do just that; in fact, the extended measure is unique.

Theorem A.5 (Carathéodory's extension theorem). Let $\mathcal{F}$ be an algebra, and let P be a probability measure on $\mathcal{F}$. Then there exists a unique probability measure P^* on $\sigma(\mathcal{F})$ such that $P^*(X) = P(X)$ for all $X \in \mathcal{F}$.

Conditional probabilities

Let P be a probability measure on an algebra $\mathcal{F}$. Given events $X, Y \in \mathcal{F}$, we define the *conditional probability* of Y given X as

$$P(Y \mid X) := \begin{cases} P(X \cap Y) / P(X) & \text{if } P(X) > 0, \\ 1 & \text{otherwise.} \end{cases}$$

From the definition, the identity $P(X \cap Y) = P(X) \cdot P(Y \mid X)$ follows immediately. Note that, unless $P(X) = 0$, the function $P(\cdot \mid X)$ that maps $Y \in \mathcal{F}$ to $P(Y \mid X)$ is also a probability measure on $\mathcal{F}$.

Random variables

Given a σ-algebra $\mathcal{F}$ over a set Ω, a (discrete) *random variable* is a mapping $\Theta \colon \Omega \to A$ into a countable set A such that

$$\Theta^{-1}(a) := \{x \in \Omega : \Theta(x) = a\} \in \mathcal{F}$$

for all $a \in A$. It is customary in probability theory to omit the argument in expressions involving a random variable. For example, we usually write $P(\Theta = a)$ instead of $P(\{x \in \Omega : \Theta(x) = a\})$.

Probability measures on infinite sequences

The sample space that arises when one deals with stochastic games is the space A^ω of infinite sequences over a countable set A. The relevant

σ-algebra $\mathcal{F}$ is generated by the *basic cylinder sets*: these are sets of the form $x \cdot A^\omega$, where $x \in A^*$, i.e. they consist of all infinite prolongations of a finite sequence x. More generally, a *cylinder set* is a finite, disjoint union of basic cylinder sets. The class $\mathcal{C}$ of all cylinder sets forms an algebra, and we can apply Carathéodory's extension theorem to extend a probability measure on $\mathcal{C}$ to a probability measure on $\mathcal{F} = \sigma(\mathcal{C})$. The σ-algebra $\mathcal{F}$ is called the *Borel σ-algebra*, and we call a set $X \in \mathcal{F}$ a *Borel set*.

The following theorem is a reformulation of *Levy's zero-one law* in the special case of the Borel σ-algebra over infinite sequences.[1] Intuitively, the theorem states that the conditional probabilities of an event X given the basic cylinder sets induced by longer and longer prefixes of an infinite sequence α approach either 1 or 0, depending on whether $\alpha \in X$ or not. More precisely, this convergence happens almost surely. We denote by $1_X : A^\omega \to \{0,1\}$ the *indicator function* of $X \subseteq A^\omega$, defined by $1_X(\alpha) = 1$ if and only if $\alpha \in X$.

Theorem A.6 (Levy's zero-one law). Let $\mathcal{F}$ be the Borel σ-algebra over A^ω, and let P be a probability measure on $\mathcal{F}$. Then

$$P(\{\alpha \in A^\omega : \lim_k P(X \mid \alpha(0) \ldots \alpha(k-1) \cdot A^\omega) = 1_X(\alpha)\}) = 1$$

for all $X \in \mathcal{F}$.

A.2 Computational complexity

We assume that the reader is familiar with (non-)deterministic Turing machines and the classes P, NP, coNP and Pspace. In the following, we give a brief description of the other complexity classes that play a role in this work.

Decision classes

Between P and NP lies the class UP of languages decidable by an *unambiguous* nondeterministic Turing machine, i.e. a nondeterministic machine that has at most one accepting run on every input. As for NP, it is neither known nor believed that UP is closed under complementation. Hence, we define coUP to be the class of problems whose complement is in UP. Obviously, we have $P \subseteq UP \subseteq NP$ and $P \subseteq coUP \subseteq coNP$; in the absence of a proof that $P \neq NP$, neither of these inclusions is known to be proper.

---

[1] For the general formulation, see (Durret 2010, Chapter 5).

Above NP and coNP lies the class DP of all languages L of the form $L = L_1 \cap L_2$ for $L_1 \in$ NP and $L_2 \in$ coNP, and its dual, the class coDP of all languages L of the form $L = L_1 \cup L_2$ for $L_1 \in$ NP and $L_2 \in$ coNP. The canonical complete problem for DP is SAT-UNSAT, the problem of deciding, given a pair (φ, ψ) of two Boolean formulae (in conjunctive normal form), whether φ is satisfiable *and* ψ is unsatisfiable. Hence, a pair (φ, ψ) belongs to SAT-UNSAT if and only if $\varphi \in$ SAT and $\psi \notin$ SAT. Any DP-complete problem is both NP-hard and coNP-hard. Hence, it is believed that NP $\cup$ coNP is *properly* contained in DP (since otherwise NP would equal coNP).

Theorem A.7. SAT-UNSAT is DP-complete.

The remaining decision classes that play a role in this work are defined via *oracle machines*: An oracle machine is a Turing machine that is equipped with one extra working tape, called the *oracle tape*. The semantics of the machine are defined with respect to a certain fixed language L, e.g. $L =$ SAT. At any time of its computation, the machine can "ask" the oracle whether the inscription of the oracle tape belongs to L or not. The omniscient oracle will return the answer immediately, and the machine may continue with its computation depending on the answer. For a complexity class $\mathcal{C}$, we denote by $P^{\mathcal{C}}$ and $NP^{\mathcal{C}}$ the classes of languages that are decidable by a deterministic, respectively non-deterministic, polynomial-time oracle machine with an oracle for a language $L \in \mathcal{C}$. Finally, $coNP^{\mathcal{C}}$ is the class of all languages whose complements are in $NP^{\mathcal{C}}$.

The *polynomial hierarchy* (PH) consists of the classes Σ_k^P, Π_k^P and Δ_k^P, where $k \in \mathbb{N}$, defined inductively by setting $\Sigma_0^P = \Pi_0^P = \Delta_0^P = P$ and

$$\Sigma_{k+1}^P = NP^{\Sigma_k^P},$$
$$\Pi_{k+1}^P = coNP^{\Sigma_k^P},$$
$$\Delta_{k+1}^P = P^{\Sigma_k^P}$$

for all $k \in \mathbb{N}$. Note that $\Sigma_1^P =$ NP, $\Pi_1^P =$ coNP and $\Delta_1^P = P$. Regarding inclusions, it is obvious that $\Delta_k^P \subseteq \Sigma_k^P \subseteq \Delta_{k+1}^P$ and $\Delta_k^P \subseteq \Pi_k^P \subseteq \Delta_{k+1}^P$ for each $k \in \mathbb{N}$. Moreover, it is easy to see that PH $\subseteq$ Pspace.

Of particular relevance to this work is the class $\Delta_2^P = P^{NP}$, which can alternatively be characterised as the class of languages decidable by a deterministic polynomial-time oracle machine with an oracle for SAT, and its subclass $P^{NP[\log]}$, the class of languages that are decidable by a deterministic

polynomial-time oracle machine that, on inputs of length n, may perform at most $O(\log n)$ queries to an oracle for SAT. Since any problem in DP can be decided with just two queries to an oracle for SAT, we have $DP \subseteq P^{NP[\log]}$.

Figure A.1 visualises the decision classes considered in this work and their relationships to each other.

Function classes

In complexity theory, a *function problem* is not merely the problem of computing the output of a function but, more generally, the problem of computing, given a binary relation R, for every input x an output y such that $(x, y) \in R$, if such an output exists; if no such output exists, the input x should be rejected.

The classical example of a function problem is FSAT, the problem of computing for a Boolean formula φ a satisfying assignment. In this case, the underlying relation is polynomial-time decidable: given a formula φ and an assignment α, we can decide in polynomial time whether α satisfies φ. Such function problems make up the class FNP. Formally, a function problem "given x, compute y such that $(x, y) \in R$" is in FNP if the relation R is *polynomially balanced* and decidable by a deterministic polynomial-time algorithm. (A relation R is polynomially balanced if there exists $k \in \mathbb{N}$ such that $|y| \le |x|^k$ for all $(x, y) \in R$).

The class $FP \subseteq FNP$ consist of all those problems in FNP for which a correct output can be computed in (deterministic) polynomial time. The problem FSAT turns out to be complete for FNP (via a suitable notion of polynomial reduction), and it is easy to see that a polynomial-time algorithm for SAT could be extended to a polynomial-time algorithm for FSAT. Hence, $P = NP$ if and only if $FP = FNP$.

Theorem A.8. $FP = FNP$ if and only if $P = NP$
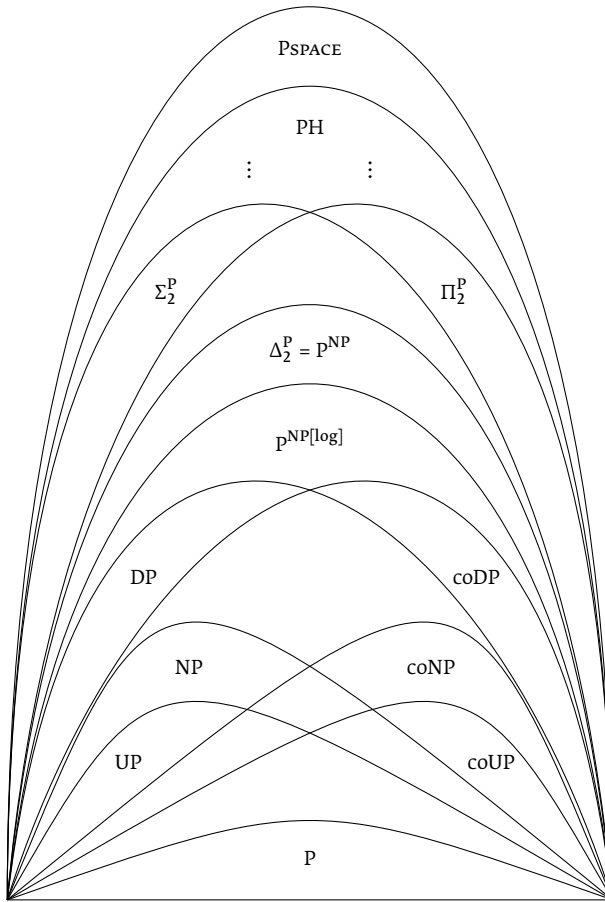
Figure A.1. A hierarchy of complexity classes.

B

Markov Chains and Markov Decision Processes

In this appendix, we review Markov chains and Markov decision processes. In particular, we discuss how to compute the (optimal) probabilities of fulfilling a given objective in these models. More details (including proofs) can be found in (Baier & Katoen 2008; Puterman 1994).

B.1 Markov chains

In a Markov chain, the system evolves solely through stochastic transitions. Moreover, the probability that the system moves to a certain successor state does only depend on the current state. Formally, a (time-homogenous) Markov chain (MC) $\mathcal{M}$ consists of:

- a nonempty, countable set S of *states*,
- a *transition function* $\delta \colon S \to \mathcal{D}(S)$, and
- a *colouring function* $\chi \colon S \to C$ into a set C of *colours*.

We denote by $\delta(t \mid s)$ the probability that $\mathcal{M}$ moves from state s to state t, i.e. $\delta(t \mid s) = \delta(s)(t)$. The *transition graph* of $\mathcal{M}$ is the directed graph (S, E) with $(s, t) \in E$ if and only if $\delta(t \mid s) > 0$.

Given an initial state $s \in S$, we define a probability measure on the Borel σ-algebra over S^ω as follows: the probability of a basic cylinder set $s_0 \ldots s_k \cdot S^\omega$

equals the product $\prod_{j=1}^{k} \delta(s_j \mid s_{j-1})$ if $s = s_0$; basic cylinder sets that start in a state different from s have probability 0. By Carathéodory's extension theorem (Theorem A.5), this definition induces a unique probability measure on the Borel σ-algebra over S^{ω}; we denote this measure by $\Pr_s^{\mathcal{M}}$. Finally, we obtain a probability measure on the Borel σ-algebra over C^{ω} by viewing the colouring function χ as a continuous function $S^{\omega} \to C^{\omega}$. We abuse notation and denote this measure also by $\Pr_s^{\mathcal{M}}$.

Remark. More generally, a Markov chain is a sequence $(\Theta_n)_{n\in\mathbb{N}}$ of discrete random variables into S such that the probability of being in a state t at time $k + 1$ only depends on the probabilities of being in states $s \in S$ at time k:

$$\Pr(\Theta_{k+1} = t \mid \Theta_k = s_k, \Theta_{k-1} = s_{k-1}, \ldots, \Theta_0 = s_0) = \Pr(\Theta_{k+1} = t \mid \Theta_k = s_k)$$

for all $k \in \mathbb{N}$ and $s_0, \ldots, s_k, t \in S$. This definition is more general since the probabilities $\Pr(\Theta_{k+1} = t \mid \Theta_k = s)$ may depend on k. If these probabilities do not depend on k, the Markov chain is called *time-homogenous*, in which case $\Pr(\Theta_{k+1} = t \mid \Theta_k = s) = \delta(t \mid s)$ for all $k \in \mathbb{N}$ and $s, t \in S$.

Reachability objectives

The basic probabilities that we wish to compute are the probabilities $\Pr_s^{\mathcal{M}}(\mathrm{Reach}(F))$ of reaching a designated subset $F \subseteq S$ of states. In terms of basic cylinder sets, we have

$$\Pr_s^{\mathcal{M}}(\mathrm{Reach}(F)) = \sum_{xt\in(S\setminus F)^*F} \Pr_s^{\mathcal{M}}(xt \cdot S^{\omega}).$$

Moreover, given a set $Z \subseteq S$ of states such that $\Pr_s^{\mathcal{M}}(\mathrm{Reach}(F)) = 0$ for all $s \in Z$, the probabilities $x_s := \Pr_s^{\mathcal{M}}(\mathrm{Reach}(F))$ satisfy the following equations:

$$
\begin{aligned}
x_s &= 1 && \text{if } s \in F; \\
x_s &= 0 && \text{if } s \in Z; \\
x_s &= \sum_{t\in S} \delta(t \mid s) \cdot x_t && \text{if } s \in S \setminus (F \cup Z).
\end{aligned}
\tag{B.1}
$$

In fact, the probabilities $\Pr_s^{\mathcal{M}}(\mathrm{Reach}(F))$ form the *least* solution of (B.1).

Theorem B.1. Let $\mathcal{M}$ be a Markov chain, and let $F, Z \subseteq S$ be sets of states such that $\Pr_s^{\mathcal{M}}(\mathrm{Reach}(F)) = 0$ for all $s \in Z$. If $(x_s)_{s\in S} \in [0,1]^S$ is a solution of (B.1), then $x_s \geq \Pr_s^{\mathcal{M}}(\mathrm{Reach}(F))$ for all $s \in S$.

By Theorem B.1 (taking $Z = \varnothing$), the probabilities $\Pr_s^{\mathcal{M}}(\mathrm{Reach}(F))$ can be computed in polynomial time by solving the following linear programme:

Minimise $\sum_{s \in S} x_s$ subject to

$$x_s \geq 0 \qquad\qquad \text{for all } s \in S,$$

$$x_s = 1 \qquad\qquad \text{for all } s \in F,$$

$$x_s = \sum_{t \in S} \delta(t \mid s) \cdot x_t \quad \text{for all } s \in S \setminus F.$$

Corollary B.2. Given a finite MC $\mathcal{M}$ (with rational transition probabilities) and a set $F \subseteq S$, the probabilities $\Pr_s^{\mathcal{M}}(\mathrm{Reach}(F))$ can be computed in polynomial time.

In practice, there is an easier way to compute reachability properties, which is supported by the following theorem.

Theorem B.3. Let $\mathcal{M}$ be a finite MC, and let $F \subseteq S$. If Z equals the set of all $s \in S$ such that $\Pr_s^{\mathcal{M}}(\mathrm{Reach}(F)) = 0$, then the probabilities $x_s := \Pr_s^{\mathcal{M}}(\mathrm{Reach}(F))$ are the *only* solution of (B.1).

Since $\Pr_s^{\mathcal{M}}(\mathrm{Reach}(F)) = 0$ if and only if there is no path from s to F in the transition graph of $\mathcal{M}$, the set Z in Theorem B.3 can be computed in linear time. To determine $\Pr_s^{\mathcal{M}}(\mathrm{Reach}(F))$ for each $s \in S$, we can then solve (B.1) using Gaussian elimination.

Infinitary objectives

The central notion for the verification of objectives that speak about the infinite behaviour of a Markov chain is that of a *bottom strongly connected component* (bottom SCC, BSCC). A BSCC of a Markov chain $\mathcal{M}$ is a maximal subset T of states that is strongly connected (i.e. in the subgraph of the transition graph induced by T every state has a path to every other state) and that has no transitions leading outside T, i.e. $\delta(s \mid t) = 0$ for all $t \in T$ and $s \in S \setminus T$. The importance of BSCCs stems from the following fact.

Lemma B.4. Let $\mathcal{M}$ be a finite MC. Then $\Pr_s^{\mathcal{M}}(\{\pi \in S^\omega : \mathrm{Inf}(\pi) \text{ is a BSCC}\}) = 1$ for all $s \in S$.

By Lemma B.4, to compute the probabilities of fulfilling a given Muller objective, it suffices to compute the probabilities of reaching a bottom SCC that corresponds to an accepting set.

Theorem B.5. Let $\mathcal{M}$ be a finite Markov chain, and let $\mathcal{F} \subseteq \mathcal{P}(C)$. Then $\Pr_s^{\mathcal{M}}(\text{Muller}(\mathcal{F})) = \Pr_s^{\mathcal{M}}(\text{Reach}(U))$ for all $s \in S$, where U is the union of all BSCCs T of $\mathcal{M}$ such that $\chi(T) \in \mathcal{F}$.

Since all SCCs of a finite graph can be identified in linear time, e.g. using Tarjan's algorithm (see Cormen et al. 2009), it follows from Theorem B.5 and Corollary B.2 that the probabilities of fulfilling a given Streett, Rabin or Muller objective can be computed in polynomial time.

Corollary B.6. Given a finite MC $\mathcal{M}$ (with rational transition probabilities) and a Streett, Rabin or Muller objective Win, the probabilities $\Pr_s^{\mathcal{M}}(\text{Win})$ can be computed in polynomial time.

B.2 Markov decision processes

Markov decision processes extend Markov chains with controlled states. Formally, a Markov decision process (MDP) $\mathcal{M}$ consists of:

- a nonempty, countable set S of *states*,
- a subset $S_0 \subseteq S$ of *controlled* states,
- a *transition relation* $\Delta \subseteq S \times ([0,1] \cup \{\bot\}) \times S$, and
- a *colouring function* $\chi \colon S \to C$ into a set C of *colours*.

We require that a transition is labelled with $\bot$ if and only if it originates in a controlled state, and that transition probabilities are unique: if $s \in S \setminus S_0$ and $t \in S$, then there exists precisely one $p \in [0,1]$ with $(s,p,t) \in \Delta$; let us denote this probability by $\Delta(t \mid s)$. Naturally, we assume that for each $s \in S \setminus S_0$ the probabilities on outgoing transitions sum up to 1: $\sum_{t \in S} \Delta(t \mid s) = 1$. For the sake of simplicity, we require additionally that for each $s \in S_0$ there exists at least one state t with $(s, \bot, t) \in \Delta$.

For a state $s \in S$, we denote by $s\Delta$ the set of all states $t \in S$ such that there exists $0 \neq p \in [0,1] \cup \{\bot\}$ with $(s,p,t) \in \Delta$. The *transition graph* of an MDP $\mathcal{M}$ is the directed graph (S, E), where $(s,t) \in E$ if and only if $t \in s\Delta$.

Remark. In the literature, MDPs are often defined using *actions*: in each state, the controller chooses an action, which determines a probability distribution on successor states. The two definitions are essentially equivalent: On the one hand, we can view states as actions. On the other hand, we can simulate actions by alternating between controlled and non-controlled states.

The behaviour of the controller is described by a *strategy* (other names in the literature include *policy* and *scheduler*). Formally, a (*randomised*) strategy in $\mathcal{M}$ is a mapping $\sigma \colon S^* S_0 \to \mathcal{D}(S)$ that assigns to each finite sequence of states that ends in a controlled state a probability distribution on states such that $\sigma(t \mid xs) := \sigma(xs)(t) > 0$ only if $(s, \bot, t) \in \Delta$. We extend σ to a mapping $S^+ \to \mathcal{D}(S)$ by setting $\sigma(t \mid xs) := \Delta(t \mid s)$ for all $x \in S^*$, $s \in S \setminus S_0$ and $t \in S$.

As for SMGs (see Section 2.2), we call a strategy σ *pure* (or *deterministic*) if $\sigma(t \mid xs) \in \{0, 1\}$ for all $xs \in S^* S_0$ and $t \in S$, and we call σ *stationary* if $\sigma(xs) = \sigma(s)$ for all $xs \in S^* S_0$. Finally, we say that a strategy is *positional* if it is both pure and stationary.

Given a strategy σ and an initial state s, we define a probability measure on the Borel σ-algebra over S^ω as for SMGs: the probability of a basic cylinder set $s_0 \ldots s_k \cdot S^\omega$ equals the product $\prod_{j=1}^{k} \sigma(s_j \mid s_0 \ldots s_{j-1})$ if $s = s_0$; basic cylinder sets that start in a different state than s have probability 0. By Carathéodory's extension theorem (Theorem A.5), this definition induces a unique probability measure on the Borel σ-algebra over S^ω; we denote both this measure and the corresponding measure on the Borel σ-algebra over C^ω (defined via the colouring χ) by Pr_s^σ.

The central problem in the analysis of Markov decision processes is computing the *optimal* probabilities of fulfilling a certain objective. Formally, given an MDP $\mathcal{M}$, a state s and an objective Win (over states or colours), we want to compute $\sup_\sigma \mathrm{Pr}_s^\sigma(\mathrm{Win})$. Note that this supremum ranges over *all* strategies in $\mathcal{M}$ and that an optimal strategy does not need to exist.

Reachability objectives

For reachability objectives, the optimal probabilities can again be characterised as the least solution of a system of equations.

Theorem B.7. Let $\mathcal{M}$ be a Markov decision process, and let $F \subseteq S$. The optimal probabilities $\sup_\sigma \mathrm{Pr}_s^\sigma(\mathrm{Reach}(F))$ form the least solution (over $[0,1]^S$) of the following system of equations:

$$
\begin{aligned}
x_s &= 1 && \text{if } s \in F; \\
x_s &= \max\{x_t : t \in s\Delta\} && \text{if } s \in S_0 \setminus F; \\
x_s &= \sum_{t \in S} \Delta(t \mid s) \cdot x_t && \text{if } s \in S \setminus (F \cup S_0).
\end{aligned}
$$

By replacing equations containing max with suitable inequalities, we obtain the following linear programme, whose optimal solution is the vector of optimal reachability probabilities:

Minimise $\sum_{s \in S} x_s$ subject to

$$x_s \geq 0 \qquad\qquad \text{for all } s \in S,$$

$$x_s = 1 \qquad\qquad \text{for all } s \in F,$$

$$x_s \geq x_t \qquad\qquad \text{for all } (s, \bot, t) \in \Delta,$$

$$x_s = \sum_{t \in S} \Delta(t \mid s) \cdot x_t \quad \text{for all } s \in S \setminus (F \cup S_0).$$

Corollary B.8. Given a finite MDP $\mathcal{M}$ (with rational transition probabilities) and a set $F \subseteq S$, the optimal probabilities $\sup_\sigma \Pr_s^\sigma(\text{Reach}(F))$ can be computed in polynomial time.

Do optimal strategies exist in MDPs with reachability objectives? For finite MDPs, the answer is positive. In fact, there always exists a globally optimal *positional* strategy. However, infinite MDPs with reachability objectives do, in general, not admit optimal strategies (see Example 2.8).

Theorem B.9. Let $\mathcal{M}$ be a finite MDP, and let $F \subseteq S$. There exists a positional strategy τ in $\mathcal{M}$ such that $\Pr_s^\tau(\text{Reach}(F)) = \sup_\sigma \Pr_s^\sigma(\text{Reach}(F))$ for all $s \in S$.

Infinitary objectives

For computing the optimal probability of fulfilling an infinitary objective, *end components* take the role that BSCCs play for Markov chains. Formally, a *sub-MDP* of an MDP $\mathcal{M}$ is a subset $T \subseteq S$ of states such that:

- $T \neq \emptyset$,
- $s\Delta \cap T \neq \emptyset$ for all $s \in T$, and
- $s\Delta \subseteq T$ for all $s \in T \setminus S_0$.

A set $T \subseteq S$ is an end component of $\mathcal{M}$ if T is a sub-MDP of $\mathcal{M}$ that is strongly connected (with respect to the transition graph of $\mathcal{M}$). Finally, we say that an end component T of $\mathcal{M}$ is *maximal* in a set $U \subseteq S$ if there exists no end component T' of $\mathcal{M}$ such that $T \subsetneq T' \subseteq U$.

Algorithm B.1 is a polynomial-time algorithm for computing all end components maximal in a given set U (for a finite MDP $\mathcal{M}$). The core of the algorithm is the procedure FindMEC, which on input $X \subseteq S$ computes all end

---

Algorithm B.1. Finding maximal end components.

Input: MDP $\mathcal{M} = (S, S_0, \Delta, \chi)$, $U \subseteq S$
Output: $\{T \subseteq S : T$ is an end component of $\mathcal{M}$ maximal in $U\}$

output $\text{FindMEC}(U)$

procedure $\text{FindMEC}(X)$
  $Z := \varnothing$
  compute all nontrivial SCCs of $(X, \{(s, t) : t \in s\Delta\})$
  for each such SCC T **do**
    $C := \{s \in T \setminus S_0 : s\Delta \not\subseteq T\}$
    if $C = \varnothing$ **then**
      $Z := Z \cup \{T\}$
    else
      $Z := Z \cup \text{FindMEC}(T \setminus C)$
    end if
  end for
  return Z
end procedure

---

components of $\mathcal{M}$ maximal in X. The procedure first computes all nontrivial SCCs in the transition graph of $\mathcal{M}$ restricted to X (i.e. all maximal strongly connected subsets of X that contain at least one edge). If such an SCC T is a sub-MDP, then T is also a maximal end component and can be added to the output of the algorithm. Otherwise, in order to find a maximal end component inside T, we can remove all non-controlled states s with $s\Delta \not\subseteq T$ from T. The resulting set might not be strongly connected any more; hence, the procedure has to be called recursively on this set.

The termination of Algorithm B.1 is guaranteed by the fact that the procedure FindMEC on input X calls itself only on proper subsets of X. Moreover, since recursive calls are limited to disjoint subsets, the total number of recursive calls is bounded by the number of states. For identifying all nontrivial SCCs of a directed graph, we can again employ Tarjan's linear-time algorithm. Hence, Algorithm B.1 computes all end components maximal in U in quadratic time.

Theorem B.10. Given a finite MDP $\mathcal{M}$ and a set $U \subseteq S$, the set of all end components of $\mathcal{M}$ maximal in U can be computed in quadratic time.

The fundamental fact about end components in finite MDPs is that, under any strategy, the set of states visited infinitely often is almost surely an end component.

Lemma B.11. Let $\mathcal{M}$ be a finite MDP, and let σ be a strategy in $\mathcal{M}$. Then $\mathrm{Pr}_s^\sigma(\{\pi \in S^\omega : \mathrm{Inf}(\pi) \text{ is an end component}\}) = 1$ for all $s \in S$.

Moreover, we can build a strategy σ that, when started in an end component T, visits almost surely all (and only) states in T infinitely often. There are two ways to construct such a strategy: First, the stationary strategy that moves from a state $s \in T \cap S_0$ to each state $t \in s\Delta \cap T$ with the same probability does the job. Second, it is not very hard to construct a pure strategy that achieves the same task.

Lemma B.12. Let $\mathcal{M}$ be a finite MDP, and let T be an end component of $\mathcal{M}$. There exists both a stationary strategy σ and a pure strategy σ such that $\mathrm{Pr}_s^\sigma(\{\pi \in S^\omega : \mathrm{Inf}(\pi) = T\}) = 1$ for all $s \in T$.

By Lemmas B.11 and B.12, computing the optimal probabilities of fulfilling a given Muller objective reduces to computing the optimal probabilities of reaching an *accepting* end component.

Theorem B.13. Let $\mathcal{M}$ be a finite Markov decision process, and let $\mathcal{F} \subseteq \mathcal{P}(C)$. Then $\sup_\sigma \mathrm{Pr}_s^\sigma(\mathrm{Muller}(\mathcal{F})) = \sup_\sigma \mathrm{Pr}_s^\sigma(\mathrm{Reach}(U))$ for all $s \in S$, where U is the union of all end components T of $\mathcal{M}$ such that $\chi(T) \in \mathcal{F}$.

Given a family $\mathcal{F} \subseteq \mathcal{P}(C)$ of accepting sets, in order to compute the union of all accepting components, we employ Algorithm B.1 to compute, for each $F \in \mathcal{F}$, all end components maximal in $\chi^{-1}(F)$. If such an end component T contains all colours $c \in F$, we can include T in the union of all accepting end component; otherwise, there is no hope of finding an accepting end component inside T (at least with respect to the accepting set F). In fact, the same idea can be used to compute the union of all accepting end components with respect to a Rabin objective. Finally, for Streett objectives, Chatterjee et al. (2005) gave an algorithm for computing the union of all accepting end components, which employs Algorithm B.1 as a subroutine (see Algorithm 5.1).

Corollary B.14. Given a finite MDP $\mathcal{M}$ (with rational transition probabilities) and a Streett, Rabin or Muller objective Win, the optimal probabilities $\sup_\sigma \mathrm{Pr}_s^\sigma(\mathrm{Win})$ can be computed in polynomial time.

Bibliography

ABADI, M., LAMPORT, L. & WOLPER, P. (1989). Realizable and unrealizable specifications of reactive systems. In *Proceedings of the 16th International Colloquium on Automata, Languages and Programming, ICALP '89*, vol. 372 of *Lecture Notes in Computer Science*, pp. 1–17. Springer-Verlag.

ABRAHAM, I., DOLEV, D., GONEN, R. & HALPERN, J. Y. (2006). Distributed computing meets game theory: Robust mechanisms for rational secret sharing and multiparty computation. In *Proceedings of the 25th Annual ACM Symposium on Principles of Distributed Computing, PODC 2006*, pp. 53–62. ACM Press.

ABRAMSKY, S., JAGADEESAN, R. & MALACARIA, P. (2000). Full abstraction for PCF. *Information and Computation*, vol. 163(2), pp. 409–470.

ALLENDER, E., BÜRGISSER, P., KJELDGAARD-PEDERSEN, J. & MILTERSEN, P. B. (2009). On the complexity of numerical analysis. *SIAM Journal on Computing*, vol. 38(5), pp. 1987–2006.

ANDERSSON, D. & MILTERSEN, P. B. (2009). The complexity of solving stochastic games on graphs. In *Proceedings of the 20th International Symposium on Algorithms and Computation, ISAAC 2009*, vol. 5878 of *Lecture Notes in Computer Science*, pp. 112–121. Springer-Verlag.

AUMANN, R. J. (1981). Survey of repeated games. In *Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern*, pp. 11–42. Bibliographisches Institut Mannheim/Wien/Zürich.

BAIER, C., BRÁZDIL, T., GRÖSSER, M. & KUČERA, A. (2007). Stochastic game logic. In *Proceedings of the Fourth International Conference on Quantitative Evaluation of Systems, QEST 2007*, pp. 227–236. IEEE Computer Society Press.

BAIER, C., GRÖSSER, M., LEUCKER, M., BOLLIG, B. & CIESINSKI, F. (2004). Probabilistic controller synthesis. In *Proceedings of the 3rd IFIP International Conference on Theoretical Computer Science, IFIP TCS 2004*, pp. 493–506. Kluwer Academic Publishers.

BAIER, C. & KATOEN, J.-P. (2008). *Principles of Model Checking.* MIT Press.

BELLMAN, R. E. (1957). *Dynamic Programming.* Princeton University Press.

BERWANGER, D., DAWAR, A., HUNTER, P. & KREUTZER, S. (2006). DAG-width and parity games. In *Proceedings of the 23rd Symposium on Theoretical Aspects of Computer Science, STACS 2006*, vol. 3884 of *Lecture Notes in Computer Science*, pp. 524–536. Springer-Verlag.

BERWANGER, D. & GRÄDEL, E. (2005). Entanglement – A measure for the complexity of directed graphs with applications to logic and games. In *Proceedings of the 11th International Conference on Logic for Programming, Artificial Intelligence and Reasoning, LPAR 2004*, vol. 3452 of *Lecture Notes in Computer Science*, pp. 209–223. Springer-Verlag.

BILLINGSLEY, P. (1995). *Probability and Measure.* John Wiley and Sons, 3rd edn.

BJÖRKLUND, H., SANDBERG, S. & VOROBYOV, S. (2003). A discrete subexponential algorithm for parity games. In *Proceedings of the 20th Annual Symposium on Theoretical Aspects of Computer Science, STACS 2003*, vol. 2607 of *Lecture Notes in Computer Science*, pp. 663–674. Springer-Verlag.

BJÖRKLUND, H. & VOROBYOV, S. (2005). Combinatorial structure and randomized subexponential algorithms for infinite games. *Theoretical Computer Science*, vol. 349, pp. 347–360.

BOREL, É. (1921). La théorie du jeu et les équations intégrales à noyau symétrique. *Comptes Rendus hebdomadaires des séances de l'Académie des Sciences*, vol. 173, pp. 1304–1308.

BOROS, E. & GURVICH, V. (2003). On Nash-solvability in pure stationary strategies of finite games with perfect information which may have cycles. *Mathematical Social Sciences*, vol. 46(2), pp. 207–241.

BOUYER, P., BRENGUIER, R. & MARKEY, N. (2010a). Computing equilibria in two-player timed games via turn-based finite games. In *Proceedings of*

the 8th International Conference on Formal Modeling and Analysis of Timed Systems, FORMATS 2010, vol. 6246 of *Lecture Notes in Computer Science*, pp. 62–76. Springer-Verlag.

Bouyer, P., Brenguier, R. & Markey, N. (2010b). Nash equilibria for reachability objectives in multi-player timed games. In *Proceedings of the 21st International Conference on Concurrency Theory, CONCUR 2010*, vol. 6269 of *Lecture Notes in Computer Science*, pp. 192–206. Springer-Verlag.

Brázdil, T., Brožek, V., Etessami, K., Kučera, A. & Wojtczak, D. (2010). One-counter Markov decision processes. In *Proceedings of the 21st ACM-SIAM Symposium on Discrete Algorithms, SODA 2010*, pp. 863–874. ACM Press.

Brázdil, T., Brožek, V., Forejt, V. & Kučera, A. (2006). Stochastic games with branching-time winning objectives. In *Proceedings of the 21st IEEE Symposium on Logic in Computer Science, LICS 2006*, pp. 349–358. IEEE Computer Society Press.

Büchi, J. R. & Landweber, L. H. (1969). Solving sequential conditions by finite-state strategies. *Transactions of the American Mathematical Society*, vol. 138, pp. 295–311.

Buss, S. R. & Hay, L. (1991). On truth-table reducibility to SAT. *Information and Computation*, vol. 91(1), pp. 86–102.

Canny, J. (1988). Some algebraic and geometric computations in PSPACE. In *Proceedings of the 20th annual ACM Symposium on Theory of Computing, STOC '88*, pp. 460–469. ACM Press.

Chandra, A. K., Kozen, D. & Stockmeyer, L. J. (1981). Alternation. *Journal of the ACM*, vol. 28(1), pp. 114–133.

Chatterjee, K. (2007). Stochastic Müller games are PSPACE-complete. In *Proceedings of the 27th International Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2007*, vol. 4855 of *Lecture Notes in Computer Science*, pp. 436–448. Springer-Verlag.

Chatterjee, K., de Alfaro, L. & Henzinger, T. A. (2005). The complexity of stochastic Rabin and Streett games. In *Proceedings of the 32nd International Colloquium on Automata, Languages and Programming, ICALP 2005*, vol. 3580 of *Lecture Notes in Computer Science*, pp. 878–890. Springer-Verlag.

CHATTERJEE, K., HENZINGER, T. A. & JURDZIŃSKI, M. (2006). Games with secure equilibria. *Theoretical Computer Science*, vol. 365(1–2), pp. 67–82.

CHATTERJEE, K., JURDZIŃSKI, M. & HENZINGER, T. A. (2003). Simple stochastic parity games. In *Proceedings of the 12th Annual Conference of the European Association for Computer Science Logic, CSL 2003*, vol. 2803 of *Lecture Notes in Computer Science*, pp. 100–113. Springer-Verlag.

CHATTERJEE, K., JURDZIŃSKI, M. & HENZINGER, T. A. (2004a). Quantitative stochastic parity games. In *Proceedings of the 15th ACM-SIAM Symposium on Discrete Algorithms, SODA 2004*, pp. 121–130. ACM Press.

CHATTERJEE, K., MAJUMDAR, R. & JURDZIŃSKI, M. (2004b). On Nash equilibria in stochastic games. In *Proceedings of the 13th Annual Conference of the European Association for Computer Science Logic, CSL 2004*, vol. 3210 of *Lecture Notes in Computer Science*, pp. 26–40. Springer-Verlag.

CHEN, X. & DENG, X. (2006). Settling the complexity of two-player Nash equilibrium. In *Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2006*, pp. 261–272. IEEE Computer Society Press.

CHURCH, A. (1957). Applications of recursive arithmetic to the problem of circuit synthesis. In *Summaries of the Summer Institute of Symbolic Logic*, vol. 1, pp. 3–50. Cornell University.

CHURCH, A. (1963). Logic, arithmetic, and automata. In *Proceedings of the International Congress of Mathematicians, 1962*, pp. 23–35. Institute Mittag-Leffler.

CLARKE, E. M., GRUMBERG, O. & PELED, D. A. (1999). *Model Checking*. MIT Press.

CODENOTTI, B. & ŠTEFANKOVIČ, D. (2005). On the computational complexity of Nash equilibria for (0, 1) bimatrix games. *Information Processing Letters*, vol. 94(3), pp. 145–150.

CONDON, A. (1992). The complexity of stochastic games. *Information and Computation*, vol. 96(2), pp. 203–224.

CONITZER, V. & SANDHOLM, T. (2003). Complexity results about Nash equilibria. In *Proceedings of the 18th International Joint Conference on Artificial Intelligence, IJCAI 2003*, pp. 765–771. Morgan Kaufmann.

Cormen, T. H., Leiserson, C. E., Rivest, R. L. & Stein, C. (2009). *Introduction to Algorithms*. MIT Press, 3rd edn.

Courcoubetis, C. A. & Yannakakis, M. (1995). The complexity of probabilistic verification. *Journal of the ACM*, vol. 42(4), pp. 857–907.

Courcoubetis, C. A. & Yannakakis, M. (1998). Markov decision processes and regular events. *IEEE Transactions on Automatic Control*, vol. 43(10), pp. 1399–1418.

Cournot, A. A. (1838). *Recherches sur les principes mathématiques de la théorie des richesses*. Hachette.

Daskalakis, C., Goldberg, P. W. & Papadimitriou, C. H. (2006). The complexity of computing a Nash equilibrium. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing, STOC 2006*, pp. 71–78. ACM Press.

Daskalakis, C., Goldberg, P. W. & Papadimitriou, C. H. (2009). The complexity of computing a Nash equilibrium. *SIAM Journal on Computing*, vol. 39(1), pp. 195–259.

de Alfaro, L. (1997). *Formal Verification of Probabilistic Systems*. Ph.D. thesis, Stanford University.

de Alfaro, L. (1998). How to specify and verify the long-run average behavior of probabilistic systems. In *Proceedings of the 13th IEEE Symposium on Logic in Computer Science, LICS '98*, pp. 454–465. IEEE Computer Society Press.

de Alfaro, L. & Henzinger, T. A. (2000). Concurrent omega-regular games. In *Proceedings of the 15th IEEE Symposium on Logic in Computer Science, LICS 2000*, pp. 141–154. IEEE Computer Society Press.

de Alfaro, L., Henzinger, T. A. & Kupferman, O. (2007). Concurrent reachability games. *Theoretical Computer Science*, vol. 386(3), pp. 188–217.

Dijkstra, E. W. (1971). Hierarchical ordering of sequential processes. *Acta Informatica*, vol. 1(2), pp. 115–138.

Durret, R. (2010). *Probability: Theory and Examples*. Cambridge University Press, 4th edn.

DZIEMBOWSKI, S., JURDZIŃSKI, M. & WALUKIEWICZ, I. (1997). How much memory is needed to win infinite games? In *Proceedings of the 12th Annual IEEE Symposium on Logic in Computer Science, LICS '97*, pp. 99–110. IEEE Computer Society Press.

EMERSON, E. A. & JUTLA, C. S. (1991). Tree automata, mu-calculus and determinacy (extended abstract). In *Proceedings of the 32nd Annual Symposium on Foundations of Computer Science, FoCS '91*, pp. 368–377. IEEE Computer Society Press.

EMERSON, E. A. & JUTLA, C. S. (1999). The complexity of tree automata and logics of programs. *SIAM Journal on Computing*, vol. 29(1), pp. 132–158.

EMERSON, E. A. & LEI, C.-L. (1987). Modalities for model checking: branching time logic strikes bsack. *Science of Computer Programming*, vol. 8(3), pp. 275–306.

ETESSAMI, K., KWIATKOWSKA, M. Z., VARDI, M. Y. & YANNAKAKIS, M. (2008). Multi-objective model checking of Markov decision processes. *Logical Methods in Computer Science*, vol. 4(4).

ETESSAMI, K. & YANNAKAKIS, M. (2005). Recursive Markov decision processes and recursive stochastic games. In *Proceedings of the 32nd International Colloquium on Automata, Languages and Programming, ICALP 2005*, vol. 3580 of *Lecture Notes in Computer Science*, pp. 891–903. Springer-Verlag.

ETESSAMI, K. & YANNAKAKIS, M. (2010). On the complexity of Nash equilibria and other fixed points. *SIAM Journal on Computing*, vol. 39(6), pp. 2531–2597.

FAELLA, M. (2009). Admissible strategies in infinite games over graphs. In *Proceedings of the 34th International Symposium on Mathematical Foundations of Computer Science, MFCS '09*, vol. 5734 of *Lecture Notes in Computer Science*, pp. 307–318. Springer-Verlag.

FILAR, J. & VRIEZE, K. (1997). *Competitive Markov decision processes*. Springer-Verlag.

FINK, A. M. (1964). Equilibrium in a stochastic *n*-person game. *Journal of Science in Hiroshima University*, vol. 28(1), pp. 89–93.

FISMAN, D., KUPFERMAN, O. & LUSTIG, Y. (2010). Rational synthesis. In *Proceedings of the 16th International Conference on Tools and Algorithms for the Construction and Analysis of Systems, TACAS 2010*, vol. 6015 of *Lecture Notes in Computer Science*, pp. 190–204. Springer-Verlag.

FLESCH, J., KUIPERS, J., SCHOENMAKERS, G. & VRIEZE, K. (2010). Subgame-perfection in stochastic games with perfect information and recursive payoffs. *Mathematics of Operations Research*, vol. 35, pp. 193–207.

FRIEDMANN, O. (2009). An exponential lower bound for the parity game strategy improvement algorithm as we know it. In *Proceedings of the 24th IEEE Symposium on Logic in Computer Science, LICS 2009*, pp. 145–156. IEEE Computer Society Press.

GALE, D. & STEWART, F. M. (1953). Infinite games with perfect information. In *Contributions to the Theory of Games II*, vol. 28 of *Annals of Mathematical Studies*, pp. 245–266. Princeton University Press.

GAREY, M. R., GRAHAM, R. L. & JOHNSON, D. S. (1976). Some NP-complete geometric problems. In *Proceedings of the 8th Annual ACM Symposium on Theory of Computing, STOC '76*, pp. 10–22. ACM Press.

GILBOA, I. & ZEMEL, E. (1989). Nash and correlated equilibria: some complexity considerations. *Games and Economic Behavior*, vol. 1(1), pp. 80–93.

GIMBERT, H. & HORN, F. (2009). Solving simple stochastic games with few random vertices. *Logical Methods in Computer Science*, vol. 5(2).

GIMBERT, H. & HORN, F. (2010). Solving simple stochastic tail games. In *Proceedings of the 21st ACM-SIAM Symposium on Discrete Algorithms, SODA 2010*, pp. 847–862. ACM Press.

GRÄDEL, E. (2007). Finite model theory and descriptive complexity. In *Finite Model Theory and Its Applications*, pp. 125–230. Springer-Verlag.

GRÄDEL, E., THOMAS, W. & WILKE, T., eds. (2002). *Automata, Logics, and Infinite Games*, vol. 2500 of *Lecture Notes in Computer Science*. Springer-Verlag.

GRÄDEL, E. & UMMELS, M. (2008). Solution concepts and algorithms for infinite multiplayer games. In *New Perspectives in Games and Interactions*, vol. 4 of *Texts in Logic and Games*, pp. 151–178. Amsterdam University Press.

HANSSON, H. & JONSSON, B. (1994). A logic for reasoning about time and reliability. *Formal Aspects of Computing*, vol. 6(5), pp. 512–535.

HEMACHANDRA, L. A. (1989). The strong exponential hierarchy collapses. *Journal of Computer and System Sciences*, vol. 39(3), pp. 299–322.

HENNESSY, M. & MILNER, R. (1985). Algebraic laws for nondeterminism and concurrency. *Journal of the ACM*, vol. 32(1), pp. 137–161.

HORN, F. (2005). Streett games on finite graphs. In *Proceedings of the 2nd Workshop on Games in Design and Verification, GDV 2005*.

HORN, F. (2008a). Explicit Muller games are PTIME. In *Proceedings of the 28th International Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2008*, vol. 2 of *Leibniz International Proceedings in Informatics*. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, Germany.

HORN, F. (2008b). *Random Games*. Ph.D. thesis, Université Paris 7.

HUNTER, P. & DAWAR, A. (2005). Complexity bounds for regular games. In *Proceedings of the 30th International Symposium on Mathematical Foundations of Computer Science, MFCS 2005*, vol. 3618 of *Lecture Notes in Computer Science*, pp. 495–506. Springer-Verlag.

HUNTER, P. & KREUTZER, S. (2007). Digraph measures: Kelly decompositions, games, and orderings. In *Proceedings of the 18th ACM-SIAM Symposium on Discrete Algorithms, SODA 2007*, pp. 637–644. ACM Press.

HYLAND, J. M. E. & ONG, C.-H. L. (2000). On full abstraction for PCF: I. models, observables and the full abstraction problem, II. dialogue games and innocent strategies, III. a fully abstract and universal game model. *Information and Computation*, vol. 163(2), pp. 285–408.

IMMERMAN, N. (1981). Number of quantifiers is better than number of tape cells. *Journal of Computer and System Sciences*, vol. 22(3), pp. 384–406.

JURDZIŃSKI, M. (1998). Deciding the winner in parity games is in UP ∩ co-UP. *Information Processing Letters*, vol. 68(3), pp. 119–124.

JURDZIŃSKI, M., PATERSON, M. & ZWICK, U. (2008). A deterministic subexponential algorithm for solving parity games. *SIAM Journal on Computing*, vol. 38(5), pp. 1519–1532.

KLARLUND, N. (1994). Progress measures, immediate determinacy, and a subset construction for tree automata. *Annals of Pure and Applied Logic*, vol. 69(2–3), pp. 243–268.

KOZEN, D. (1983). Results on the propositional mu-calculus. *Theoretical Computer Science*, vol. 27, pp. 333–354.

KUHN, H. W. (1950). Extensive games. *Proceedings of the National Academy of Sciences of the USA*, vol. 36, pp. 570–576.

KUIPERS, J., FLESCH, J., SCHOENMAKERS, G. & VRIEZE, K. (2009). Pure subgame-perfect equilibria in free transition games. *European Journal of Operations Research*, vol. 199(2), pp. 442–447.

LIGGETT, T. M. & LIPPMAN, S. A. (1969). Stochastic games with perfect information and time average payoff. *SIAM Review*, vol. 11(4), pp. 604–607.

LUCE, R. D. & RAIFFA, H. (1957). *Games and Decisions: Introduction and Critical Survey*. John Wiley and Sons.

MAITRA, A. P. & SUDDERTH, W. D. (1998). Finitely additive stochastic games with Borel measurable payoffs. *International Journal of Game Theory*, vol. 27(2), pp. 257–267.

MAITRA, A. P. & SUDDERTH, W. D. (2003). Stochastic games with Borel payoffs. In Neyman & Sorin (2003), pp. 367–373.

MARTIN, D. A. (1975). Borel determinacy. *Annals of Mathematics*, vol. 102, pp. 363–371.

MARTIN, D. A. (1985). Purely inductive proof of Borel determinacy. In *Proceedings of Symposia in Pure Mathematics*, vol. 42, pp. 303–308.

MARTIN, D. A. (1998). The determinacy of Blackwell games. *Journal of Symbolic Logic*, vol. 63(4), pp. 1565–1581.

McIVER, A. & MORGAN, C. (2002). Games, probability and the quantitative μ-calculus. In *Proceedings of the 9th International Conference on Logic for Programming, Artificial Intelligence and Reasoning, LPAR 2002*, vol. 2514 of *Lecture Notes in Computer Science*, pp. 292–310. Springer-Verlag.

McNAUGHTON, R. (1993). Infinite games played on finite graphs. *Annals of Pure and Applied Logic*, vol. 65(2), pp. 149–184.

Mostowski, A. W. (1991). Games with forbidden positions. Tech. Rep. 78, Instytut Matematyki, Uniwersytet Gdański, Poland.

Nash, Jr., J. F. (1950). Equilibrium points in N-person games. *Proceedings of the National Academy of Sciences of the USA*, vol. 36, pp. 48–49.

von Neumann, J. (1928). Zur Theorie der Gesellschaftsspiele. *Mathematische Annalen*, vol. 100(1), pp. 295–320.

von Neumann, J. & Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. John Wiley and Sons.

Neyman, A. & Sorin, S., eds. (2003). *Stochastic Games and Applications*, vol. 570 of *NATO Science Series C*. Springer-Verlag.

Nisan, N., Roughgarden, T., Tardos, É. & Vazirani, V. V., eds. (2007). *Algorithmic Game Theory*. Cambridge University Press.

Obdržálek, J. (2003). Fast mu-calculus model checking when tree-width is bounded. In *Proceedings of the 15th International Conference on Computer Aided Verification, CAV 2003*, vol. 2725 of *Lecture Notes in Computer Science*, pp. 80–92. Springer-Verlag.

Obdržálek, J. (2006). DAG-width – connectivity measure for directed graphs. In *Proceedings of the 17th ACM-SIAM Symposium on Discrete Algorithms, SODA 2006*, pp. 814–821. ACM Press.

Obdržálek, J. (2007). Clique-width and parity games. In *Proceedings of the 16th Annual Conference of the European Association for Computer Science Logic, CSL 2007*, vol. 4646 of *Lecture Notes in Computer Science*, pp. 54–68. Springer-Verlag.

Osborne, M. J. & Rubinstein, A. (1994). *A Course in Game Theory*. MIT Press.

Papadimitriou, C. H. (1994). *Complexity Theory*. Addison Wesley.

Pnueli, A. & Rosner, R. (1989). On the synthesis of a reactive module. In *Proceedings of the Sixteenth Annual ACM Symposium on Principles of Programming Languages, POPL '89*, pp. 179–190. ACM Press.

Puterman, M. L. (1994). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley and Sons.

RAMADGE, P. J. & WONHAM, W. M. (1989). The control of discrete event systems. *Proceedings of the IEEE*, vol. 77(1), pp. 81–98.

SELTEN, R. (1965). Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit. *Zeitschrift für die gesamte Staatswissenschaft*, vol. 121, pp. 301–324 and 667–689.

SHAPLEY, L. S. (1953). Stochastic games. *Proceedings of the National Academy of Sciences of the USA*, vol. 39, pp. 1095–1100.

SOLAN, E. & VIEILLE, N. (2003). Deterministic multi-player Dynkin games. *Journal of Mathematical Economics*, vol. 39(8), pp. 911–929.

THOMAS, W. (1990). Automata on infinite objects. In *Handbook of Theoretical Computer Science*, vol. B: Formal Models and Semantics, pp. 133–192. Elsevier.

UMMELS, M. (2005). Rational behaviour and strategy construction in infinite multiplayer games. Diploma Thesis, RWTH Aachen University.

UMMELS, M. (2006). Rational behaviour and strategy construction in infinite multiplayer games. In *Proceedings of the 26th International Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2006*, vol. 4337 of *Lecture Notes in Computer Science*, pp. 212–223. Springer-Verlag.

UMMELS, M. (2008). The complexity of Nash equilibria in infinite multiplayer games. In *Proceedings of the 11th International Conference on Foundations of Software Science and Computation Structures, FOSSACS 2008*, vol. 4962 of *Lecture Notes in Computer Science*, pp. 20–34. Springer-Verlag.

UMMELS, M. & WOJTCZAK, D. (2009a). The complexity of Nash equilibria in simple stochastic multiplayer games. In *Proceedings of the 36th International Colloquium on Automata, Languages and Programming, ICALP 2009*, vol. 5556 of *Lecture Notes in Computer Science*, pp. 297–308. Springer-Verlag.

UMMELS, M. & WOJTCZAK, D. (2009b). Decision problems for Nash equilibria in stochastic games. In *Proceedings of the 18th Annual Conference of the European Association for Computer Science Logic, CSL '09*, vol. 5771 of *Lecture Notes in Computer Science*, pp. 515–529. Springer-Verlag.

VIEILLE, N. (2000a). Two-player stochastic games I: A reduction. *Israel Journal of Mathematics*, vol. 119(1), pp. 55–91.

VIELLE, N. (2000b). Two-player stochastic games II: The case of recursive games. *Israel Journal of Mathematics*, vol. 119(1), pp. 93–126.

VÖGE, J. & JURDZIŃSKI, M. (2000). A discrete strategy improvement algorithm for solving parity games. In *Proceedings of the 12th International Conference on Computer Aided Verification, CAV 2000*, vol. 1855 of *Lecture Notes in Computer Science*, pp. 202–215. Springer-Verlag.

WAGNER, K. W. (1990). Bounded query classes. *SIAM Journal on Computing*, vol. 19(5), pp. 833–846.

WALUKIEWICZ, I. (2001). Pushdown processes: games and model checking. *Information and Computation*, vol. 164(2), pp. 234–263.

ZERMELO, E. (1913). Über eine Anwendung der Mengenlehre auf die Theorie des Schachspiels. In *Proceedings of the 5th International Congress of Mathematicians*, vol. II, pp. 501–504. Cambridge University Press.

ZIELONKA, W. (1998). Infinite games on finitely coloured graphs with applications to automata on infinite trees. *Theoretical Computer Science*, vol. 200(1–2), pp. 135–183.

ZIELONKA, W. (2004). Perfect-information stochastic parity games. In *Proceedings of the 7th International Conference on Foundations of Software Science and Computation Structures, FOSSACS 2004*, vol. 2987 of *Lecture Notes in Computer Science*, pp. 499–513. Springer-Verlag.

Notation

Note: See referenced pages for formal definitions.

$(\bar{\sigma}_{-i}, \tau)$ strategy profile $\bar{\sigma}$ with strategy for player i replaced by τ 37

$[0,1]$ closed interval from 0 to 1   31

$|x|$ length of x 31

$\prec$ proper prefix relation   31

$\preceq$ prefix relation   31

$\vdash$ successor relation   88

A^* finite sequences over A 31

A^+ non-empty finite sequences over A 31

A^ω infinite sequences over A 31

$\alpha|_k$ prefix of length k of α 31

$\mathrm{Bnd}(U)$ boundary of value class U 52

$\mathrm{B\ddot{u}chi}(F)$ infinite sequences hitting F infinitely often   34

C set of colours   32

χ colouring function   32

$\mathrm{coB\ddot{u}chi}(F)$ infinite sequences staying in F from some point onwards   34

$\mathcal{C}^*$ class of SMGs derived from class $\mathcal{C}$ of S2Gs   70

Δ transition relation   32

$\delta(q)$ enabled transitions in state q 88

$\Delta(w \mid v)$ transition probability   32

$\mathrm{dist}(\bar{x}, \bar{y})$ Hamming distance of $\bar{x}$ and $\bar{y}$ 126

$\mathcal{D}(A)$ probability distributions over A 31

ε empty word   31

$\mathcal{G}(\bar{x})$ MDP induced by Muller SMG $\mathcal{G}$ and payoff $\bar{x}$ 101

$\mathcal{G}[x]$ residual game after history x 40

$\mathcal{G}^{\bar{\sigma}_{-i}}$	MDP induced by strategy profile $\bar{\sigma}$ and player i	39, 40
$\mathcal{G}^{\bar{\sigma}}$	Markov chain induced by strategy profile $\bar{\sigma}$	39, 40
$\mathcal{G}_i$	coalition game against player i	43
$\mathcal{G} \upharpoonright U$	restriction of $\mathcal{G}$ to subarena U	41
$\bar{\mathcal{G}}$	S2G $\mathcal{G}$ with boundary states made absorbing	52
$\text{Inf}(\alpha)$	elements occurring infinitely often in α	31
$\lim_n X_n$	limit of chain $(X_n)_{n \in \mathbb{N}}$	57
$\text{Muller}(\mathcal{F})$	infinite sequences fulfilling Muller objective $\mathcal{F}$	35
$\mathbb{N}$	natural numbers	31
$\text{Parity}(\Omega)$	infinite sequences fulfilling parity objective Ω	34
$\text{Parity}[d]$	Parity SMGs or S2Gs with at most d priorities	49, 133
Π	set of players	32
$\text{Pr}_{v_0}^{\bar{\sigma}}$	probability measure induced by strategy profile $\bar{\sigma}$ and initial vertex v_0	38
$\mathcal{P}(A)$	power set of A	31
$\text{Rabin}(\Omega)$	infinite sequences fulfilling Rabin objective Ω	35
$\text{Reach}(F)$	infinite sequences hitting F	34
$\mathbb{R}$	real numbers	31
$\mathfrak{R}$	ordered field of real numbers	82
$\text{Safe}(F)$	infinite sequences staying in F	34
$\sigma(w \mid xv)$	transition probability with strategy σ	37
$\sigma[x]$	residual strategy after history x	40
$\bar{\sigma}(w \mid xv)$	transition probability with strategy profile $\bar{\sigma}$	37
$\bar{\sigma}[x]$	residual strategy profile after history x	40
$\text{Streett}(\Omega)$	infinite sequences fulfilling Streett objective Ω	35
V	set of vertices	32
V_i	set of vertices controlled by player i	32
$\text{val}^{\mathcal{G}}(v)$	value of S2G $\mathcal{G}$ from v	43
$\text{val}^{\tau}(v)$	value of strategy τ from v	42
$\text{val}_i^{\mathcal{G}}(v)$	value of SMG $\mathcal{G}$ for player i from v	42
$v\Delta$	Δ-successors of v	32
W_i	winning region for player i	100, 122
Win_i	objective for player i	33
$x^{-1}X$	residual language of X with respect to prefix x	31

Index