TECHNISCHE UNIVERSITÄT MÜNCHEN

Lehrstuhl für Wirtschaftsinformatik und Entscheidungstheorie

Set-Valued Solution Concepts in Social Choice and Game Theory

Axiomatic and Computational Aspects

Markus Brill

Vollständiger Abdruck der von der Fakultät für Informatik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:		Univ.-Prof. Tobias Nipkow, Ph.D.
Prüfer der Dissertation:	1.	Univ.-Prof. Dr. Felix Brandt
	2.	Prof. Dr. Jérôme Lang, Université Paris-Dauphine/Frankreich

Die Dissertation wurde am 5. Juli 2012 bei der Technischen Universität München eingereicht und durch die Fakultät für Informatik am 12. Oktober 2012 angenommen.

SET-VALUED SOLUTION CONCEPTS IN SOCIAL CHOICE AND GAME THEORY

MARKUS BRILL

Axiomatic and Computational Aspects

Für Papa

ABSTRACT

This thesis studies axiomatic and computational aspects of set-valued solution concepts in social choice and game theory. It is divided into two parts.

The first part focusses on solution concepts for normal-form games that are based on varying notions of dominance. These concepts are intuitively appealing and admit unique minimal solutions in important subclasses of games. Examples include Shapley's saddles, Harsanyi and Selten's primitive formations, Basu and Weibull's CURB sets, and Dutta and Laslier's minimal covering sets. Two generic algorithms for computing these concepts are proposed. For each of these algorithms, properties of the underlying dominance notion are identified that ensure the soundness and efficiency of the algorithm. Furthermore, it is shown that several solution concepts based on weak and very weak dominance are computationally intractable, even in two-player games.

The second part is concerned with social choice functions (SCFs), an important subclass of which is formed by tournament solutions. The winner determination problem is shown to be computationally intractable for different variants of Dodgson's rule, Young's rule, and Tideman's method of ranked pairs. For a number of tractable SCFs such as maximin and Borda's rule, the complexity of computing possible and necessary winners for partially specified tournaments is determined. Special emphasis is then put on tournament solutions that are defined via retentiveness. The axiomatic properties and the asymptotic behavior of these solutions is studied in depth, and a new attractive tournament solution is proposed. Finally, necessary and sufficient conditions for the strategyproofness of irresolute SCFs are presented.

PUBLICATIONS

This thesis is based on the following publications. Authors are listed in alphabetical order.

1. H. Aziz, M. Brill, F. Fischer, P. Harrenstein, J. Lang, and H. G. Seedig. Possible and necessary winners of partial tournaments. In V. Conitzer and M. Winikoff, editors, *Proceedings of the 11th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*. IFAAMAS, 2012.

2. F. Brandt and M. Brill. Necessary and sufficient conditions for the strategyproofness of irresolute social choice functions. In K. Apt, editor, *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge (TARK)*, pages 136–142. ACM Press, 2011.

3. F. Brandt and M. Brill. Computing dominance-based solution concepts. In B. Faltings, K. Leyton-Brown, and P. Ipeirotis, editors, Proceedings of the 13th ACM Conference on Electronic Commerce (ACM-EC), page 233. ACM Press, 2012.

4. F. Brandt, M. Brill, F. Fischer, and P. Harrenstein. Computational aspects of Shapley's saddles. In *Proceedings of the 8th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 209–216. IFAAMAS, 2009.

5. F. Brandt, M. Brill, F. Fischer, P. Harrenstein, and J. Hoffmann. Computing Shapley's saddles. *ACM SIGecom Exchanges*, 8(2), 2009.

6. F. Brandt, M. Brill, F. Fischer, and P. Harrenstein. Minimal retentive sets in tournaments. In W. van der Hoek, G. A. Kaminka, Y. Lespérance, and M. Luck, editors, *Proceedings of the 9th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 47–54. IFAAMAS, 2010.

7. F. Brandt, M. Brill, E. Hemaspaandra, and L. Hemaspaandra. Bypassing combinatorial protections: Polynomial-time algorithms for single-peaked electorates. In M. Fox and D. Poole, editors, *Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI)*, pages 715–722. AAAI Press, 2010.

8. F. Brandt, M. Brill, F. Fischer, and P. Harrenstein. On the complexity of iterated weak dominance in constant-sum games. *Theory of Computing Systems*, 49(1):162–181, 2011.

9. F. Brandt, M. Brill, F. Fischer, and J. Hoffmann. The computational complexity of weak saddles. *Theory of Computing Systems*, 49(1):139–161, 2011.

10. M. Brill and F. Fischer. The price of neutrality for the ranked pairs method. In J. Hoffmann and B. Selman, editors, *Proceedings of the 26th AAAI Conference on Artificial Intelligence (AAAI)*, pages 1299–1305. AAAI Press, 2012.

ACKNOWLEDGMENTS

I want to say *thank you* to many people.

To Felix Brandt, for being the best advisor I can imagine. I well remember the day when I first met Felix, and his support, guidance, and brilliance have been outstanding ever since. I truly hope that we will have fruitful collaborations for decades to come.

To all former and current members of the PAMAS research group. Haris Aziz, Felix Brandt, Felix Fischer, Paul Harrenstein, Keyvan Kardel, Evangelia Pyrga, Hans Georg Seedig, and Troels Bjerre Sørensen: You have been amazing colleagues and friends, and I have been very lucky to be in the right place at the right time.

To all colleagues at LMU and TUM, for creating a stimulating working environment. To all my coauthors and collaborators, for teaching me everything I know about research. Besides PAMAS members, the set of my coauthors includes Edith and Lane Hemaspaandra, Jan Hoffmann, and Jérôme Lang. Jérôme has also provided extensive and very helpful feedback to my thesis and my work in general. My work has further benefitted from helpful discussions with Craig Boutilier, Vincent Conitzer, Edith Elkind, Ulle Endriss, Piotr Faliszewski, Christian Klamler, Jean-François Laslier, Michel Le Breton, Vincent Merlin, Rolf Niedermeier, Noam Nisan, Ariel Procaccia, Jörg Rothe, Tim Roughgarden, Tuomas Sandholm, Arkadii Slinko, Toby Walsh, Gerhard Woeginger, Lirong Xia, and Bill Zwicker, among others.

To Jean-François Laslier and his research group at École Polytechnique, for hosting me in the summer of 2011. To José Rui Figueira, *the special one*, for organizing an unforgettable summer school in Estoril. To the A-Team Aarhus consisting of Matúš Mihaľák, Leo Rüst, Rahul Savani, and Alex Souza, for the fun we had in Denmark. To the writing group under the guidance of Aniko Balazs, for helpful tips and feedback. To Tobias Nipkow, for chairing my thesis committee.

To TopMath and the M9 research group, for allowing me to dive into research at an early stage. In particular, I want to thank René Brandenberg, Peter Gritzmann, and Anusch Taraz from M9, and Andrea Echtler, Ralf Franken, Christian Kredler, and Denise Lichtig from TopMath.

To my TopMath football team, for all the great memories, including winning the Elite-Cup trophy twice. Let me write them in formation: Stefan Jerg—Péter Koltai, Stephan Ritscher—Thorsten Knott, Lukas Höhndorf, Suyang Pu—Sascha Böhm.

To all my friends in academia, including Doro Baumeister, Nadja Betzler, Paul Dütting, Gábor Erdélyi, Britta Dorn, and Umberto Grandi, for making conferences so much fun.

To all my friends outside academia, including Maria Eltsova, Melanie Götz, Eva Langhauser, Nico Noppenberger, Nóra Regős, and Tea Todorova, for supporting me throughout. To Kriszta and Wolfgang Endreß, for providing me with Hungarian wines to keep me going.

To my parents and family, for all they have done for me.

Finally, to all those great people who deserve a thank you but whom I forgot.

FUNDING SOURCES My work was supported by the Deutsche Forschungsgemeinschaft under grants BR-2312/6-1 (within the European Science Foundation's EUROCORES program LogICCC) and BR-2312/10-1, by the TUM Graduate School, and by ParisTech.

CONTENTS

1 INTRODUCTION

This thesis studies axiomatic and computational aspects of set-valued solution concepts in social choice and game theory. Let me start by explaining each of these terms.

Game theory studies strategic interactions of multiple agents in situations where the well-being of a single agent depends not only on his own actions, but also on the actions of all the other agents. The term *agent* is used to refer to an autonomous decision maker that can be ascribed preferences over different states of the world. For example, an agent can be a person, an institution, or a country. Whereas early developments of the theory were mainly motivated by the analysis of parlor games such as chess and checkers, game theory has developed into an important field at the intersection of economics and mathematics that has numerous applications in the social sciences and beyond.

Social choice theory studies how a group of agents can make collective decisions based on the—possibly conflicting—preferences of the members of the group. In the most general setting, there is a set of outcomes over which each group member has preferences. A social choice mechanism aggregates these preferences to social preferences, on the basis of which a collective decision or *social choice* is made. Social choice theory is an inherently interdisciplinary field that has attracted researchers (and practitioners) from such diverse areas as mathematics, economics, political science, and psychology.

The main objects of study in both game theory and social choice theory are *solution concepts*. However, there is a considerable variation in the meaning of this term between the two disciplines. In (non-cooperative) game theory, a solution concept tries to capture rational behavior. Given the specification of an interactive decision scenario, a so-called *game*, it attempts to give recommendations to agents as to what actions maximize the agent's well-being. This prescriptive view of game-theoretic solution concepts is often complemented with a descriptive perspective that tries to predict the actions of rational agents. Social choice theory, on the other hand, is not focussed on recommending or predicting behavior of agents. Rather, it can be understood as a tool that facilitates group decision-making by providing methods to aggregate preferences. It is those preference aggregation mechanisms that are referred to as solution concepts in social choice theory. The problems that these solutions aim to solve are thus of a cooperative character. This is also reflected in fairness criteria, often referred to as *axioms*, that are used to evaluate different methods.

In the context of voting, the equal treatment of voters—"one person, one vote"—constitutes a prime example of such a fairness principle. However, the largely cooperative attitude of preference aggregation does not deter agents from gaming the system. The most common manifestation of manipulative behavior can be observed when agents lie about their preferences in order to achieve a preferred outcome. Hence, in order to be considered sensible, a solution concept in social choice theory should not only exhibit desirable fairness criteria, but also some degree of resistance to strategic manipulation.

A solution concept is *set-valued* if it may return sets of recommendations or choices instead of a single one. Whereas game theory differentiates between set-valued and *point-valued* solution concepts, the counterpart of a set-valued concept in social choice theory is a *resolute* preference aggregation mechanism. From a mathematical perspective, set-valued solution concepts are often more elegant—partly because they do not require tie-breaking in symmetric situations—but sometimes also more challenging. The interpretation of the outcome of a set-valued concept is not so clear; after all, a final choice needs to be made from the set, but this decision is left unspecified by the concept. As such, set-valued concepts might be thought of as focussing on *excluding* implausible options rather than singling out optimal ones. Although being more cautious, this approach often turns out to be suitable where more discriminatory concepts fail. Despite their intuitive appeal, set-valued solution concepts are traditionally less studied than their respective counterparts.

In recent years, the role of *computational* aspects has gained significant importance in both social choice and game theory. This development, which is often attributed to the rise of the internet as a platform where many agents interact, is in fact a quite natural one. Indeed, the internet can be seen as the predominant modern arena of conflict and cooperation. This viewpoint virtually begs for studies that incorporate concepts and techniques that haven been developed in game theory and social choice. As a consequence, two interdisciplinary research areas have emerged: *algorithmic game theory* and *computational social choice*. The flow of ideas goes in both directions. On the one hand, strategic considerations and fairness notions are taken into account when designing, say, multiagent systems. On the other hand, approaches and techniques developed in computer science are used to gain a better understanding of concepts and phenomena in social choice and game theory. A prime example for the latter direction can be found in the vast literature that studies the *computational complexity of solution concepts*. The relevance of this endeavor is obvious: the absence of efficient algorithms for computing solutions would render a solution concept virtually useless, at least for large problem instances that do not exhibit additional structure. In game theory, computational intractability of a solution concept even challenges the plausibility of the concept as a tool to predict rational behavior: if it is impossible to compute a solution in reasonable time, why should agents be expected to behave according to the solution? In social choice settings, on the other hand, computational intractability can play the beneficial role of protecting preference aggregation functions from manipulative attacks. The basic idea is that, although most functions are manipulable in principle, *finding* a manipulative action might be computationally infeasible.

Axiomatic aspects are a traditional focus of theoretical economics, a field that has played a significant role in the development of both social choice and game theory. Unlike in mathematics, the term *axiom* here refers to (more or less desirable) *properties* that a concept may or may not satisfy. The *axiomatic method*, which consists in setting up a set of axioms and exploring which concepts—if any—satisfy these axioms, is particularly popular in social choice theory. Famous impossibility results like that of Arrow (1951) identify sets of axioms that are incompatible in the sense that no solution concept can satisfy all of them. Even apart from characterizations and impossibility results, studying axiomatic aspects of solution concepts is a worthwhile task. For instance, knowing which properties are satisfied by which preference aggregation mechanisms is instrumental in making informed decisions as to which mechanism to use in a given context.

1.1 ILLUSTRATIVE EXAMPLES

To illustrate the main concepts studied in this thesis, let us consider some simple examples of interactive decision situations.

1.1.1 *Example 1: Choice of Action*

Imagine that two gladiators are set to fight against each other in an ancient arena. Before the fight begins, each gladiator has to choose one—and only one—weapon from his personal arsenal. Assume that the gladiators know their opponent's arsenal, but that the actual choice of the opponent is only observable after they have made their own choice. The result of the fight will depend, at least in a probabilistic way, on the weapons that are chosen. What makes the decision situation an interactive one is the fact that, in order to evaluate different choices, a gladiator needs to make assumptions about the choice behavior of his opponent.

The situation can be modeled as a two-player normal-form *game*. Each gladiator is a *player*, and weapon choices correspond to *actions*. Every pair of actions, one for each player, describes an *outcome* of the game. In the gladiator game, an outcome does not necessarily correspond to the result of the fight; it simply describes which weapons have been chosen.

Since the prospects of winning depend on the choices of weapons, a gladiator presumably *prefers* certain outcomes over others. In economic theory, preferences of rational decision-makers are usually assumed to be complete and transitive. In other words, every decision maker is able to rank-order the set of outcomes, allowing for indifferences. For notational convenience, preferences are usually represented by a function that maps each outcome to a real-valued *utility*. The interpretation of such a utility function is that an outcome o_1 is (strictly) preferred to another outcome o_2 if and only if the utility attached to o_1 is (strictly) greater than the utility attached to o_2. It is noteworthy that the numbers representing utilities do not bear any absolute meaning. Their sole purpose is to enable ordinal comparisons between pairs of outcomes. Thus, there is a continuum of utility functions representing a given preference relation.

Of course, utility functions of different players do not need to be identical. In the gladiator scenario, it is in fact reasonable to assume that the preferences of the opponents are diametrically opposed, in the sense that one gladiator prefers an outcome o_1 over another outcome o_2 if and only if his opponent prefers o_2 over o_1.[1]

A particular instance of a gladiator game can therefore be specified by listing the choices for each gladiator, and the preferences over outcomes for one of them (assuming that the opponent's preferences are reversed). This can conveniently be done with a matrix whose rows and columns correspond to the actions of the two players. One player chooses a row, the other player chooses a column. This uniquely defines a matrix cell. By convention, the entries in matrix cells are the utilities for the row player. The row player therefore prefers outcomes that are attached higher numbers, whereas the column player prefers lower numbers.

$$\begin{array}{c|ccc|} & b_1 & b_2 & b_3 \\ \hline a_1 & 1 & 5 & 2 \\ a_2 & 4 & 3 & 5 \\ a_3 & 2 & 2 & 1 \\ \hline \end{array}$$

Figure 1: First gladiator game

Figure 1 presents an example where the row player has to choose between actions a_1, a_2, and a_3, and the column player has to choose between b_1, b_2, and b_3. The numbers in the matrix describe the preferences of the players, and partition the set of outcomes into five indifference classes. One possible interpretation of these numbers in the gladiator context is in terms of winning probabilities: the higher

1 Games with this property are called *strictly competitive*. A particularly well-studied subclass of strictly competitive games is given by *zero-sum* games (see page 18).

the number, the more likely is a victory of the first gladiator. How exactly these numbers come about is irrelevant; they are simply taken as given.

Game-theoretic reasoning can now be applied to the game in Figure 1. The matrix reveals that choosing action a_3 is not a very good idea for the row player: the utility of a_2 is higher than that of a_3 *no matter which action the column player chooses*. In game-theoretic terminology, a_2 *dominates* a_3. It seems to be a reasonable principle to never choose a dominated action, and the row player should thus exclude action a_3 from consideration. The column player, on the other hand, cannot exclude any of his actions with the help of this argument. If he however anticipates that the row gladiator does *not* choose a_3, his action b_3 becomes dominated by b_1. This thought process, called *iterated dominance* and studied in detail in Chapter 6, can generally be used to reduce a game. Iterated dominance rarely yields a unique outcome, and often cannot eliminate any actions. For the game in Figure 1, iterated dominance at least identified two actions, one for each player, that should *not* be chosen.

In order to interpret the exclusion of actions in the gladiator context, assume that gladiators usually bring their whole arsenal to the arena, in order to keep all their options open. Dominated actions would then correspond to actions that a gladiator can safely leave at home, without any risk of ever regretting this decision.

	b_1	b_2	b_3
a_1	5	1	0
a_2	4	$\boxed{3}$	4
a_3	1	2	5

Figure 2: Second gladiator game

Now consider the game in Figure 2. No player has a dominated action, so iterated dominance has no bite. Nevertheless, the outcome (a_2, b_2) can be identified as a desirable one, as it is stable in the following sense: given the action of the other player in this outcome, no player wants to change his own action. It can be checked that this property does not hold for any other outcome in this game. A stable outcome constitutes a *saddle point* of the matrix, as the corresponding entry is maximal in its column and minimal in its row at the same time. An alternative characterization of saddle points highlights their relation to dominance: although no action in the game in Figure 2 is dominated, action a_2 dominates both a_1 and a_3 *when the column player is restricted to play* b_2, and analogously for the column player.

Saddle points need not exist, as the first game (Figure 1) witnesses. When they do exist, however, they constitute good predictions of the outcome of the game, in the sense that they are selected by most reasonable game-theoretic solution concepts. Gladiators appreciate saddle points because they simplify weapon choice tremendously. Each gladiator can simply bring the weapon corresponding to the saddle point, and no gladiator will regret doing so.

Figure 3 presents a game that has neither a saddle point nor a dominated action. In game theory, different approaches have been developed in order to deal with situations like this. Most prominent is the introduction of *mixed strategies*, i.e., probability distributions over actions, that lead to the notion of *Nash equilibrium*. This direction is not pursued here, though it will be considered later in this thesis.

	b_1	b_2	b_3	b_4
a_1	5	3	4	6
a_2	2	4	5	3
a_3	4	2	6	1
a_4	1	3	0	7

Figure 3: Third gladiator game

A lesser-known approach, introduced by Shapley (1964), consists in identifying pairs of *sets of actions* that stand in a best-response relation to each other. In the game in Figure 3, consider the pair $(\{a_1, a_2\}, \{b_1, b_2\})$, consisting of the first two actions of each player. Assuming that the column player chooses an action from his set $\{b_1, b_2\}$, it would be suboptimal for the row player to choose a_3 or a_4, as both of these actions are dominated under this assumption. Since the analogous statement holds for the column player, the sets $\{a_1, a_2\}$ and $\{b_1, b_2\}$ are stable in a generalized sense. Shapley used the term *saddle* to refer to inclusion-minimal pairs satisfying this property, and showed that every game has at least one saddle.

Like iterated dominance, saddles enable gladiators to partition their arsenal into weapons they might want to bring to the arena and weapons they can leave at home. Inclusion-minimality ensures that gladiators need to carry as few weapons as possible, while still having all "best replies" to any weapon the other gladiator brings to the arena.

Solution concepts of this kind will be the focus of the first part of this thesis. For various notions of dominance, the multiplicity and the computational complexity of saddles will be considered.

1.1.2 *Example 2: Choice Based on Pairwise Comparisons*

In the second part of this thesis, tournament solutions play an important role. The purpose of a tournament solution is to single out best alternatives based on pairwise comparisons between alternatives.

Imagine that five gladiators meet in the arena and hold a round-robin tournament, such that each pair of gladiators face each other exactly once. This makes for $\binom{5}{2} = 10$ fights in total. The results of these fights can be represented by a tournament graph like the one in Figure 4.

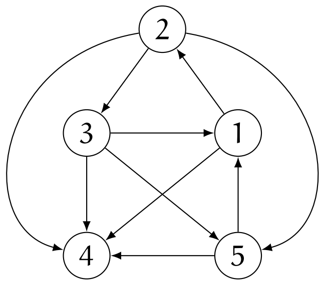


Figure 4: Graphical representation of the outcome of a round-robin tournament. An arrow from node i to node j represents that gladiator i has won the fight against gladiator j.

Given the results of the individual fights, the organizer of the tournament needs to decide which gladiator should be declared the overall winner. This decision is straightforward in tournaments where one gladiator beats all others. In the absence of such a gladiator, however, several solutions are possible. One natural idea is to choose the gladiator(s) with the highest number of wins. This corresponds to *Copeland's* tournament solution and would result in a tie between gladiators 2 and 3 in the example above. Many other tournament solutions have been proposed, and several of them will be studied in this thesis.

1.1.3 *Example 3: Choice Based on Preferences*

Besides tournament solutions, the second part of this thesis is concerned with mechanisms that aggregate preferences. Consider again the gladiators that motivated the examples above. Imagine that, after a long day of exhausting fights, nine gladiators meet for dinner. Since even a gladiator cannot eat a whole animal all by himself, they have to decide which animal they like to eat. The cook offers an antelope (a), a buffalo (b), and a caribou (c), and asks the gladiators to rank-order these alternatives. (Assume that the gladiators' total budget does not allow to purchase more than one animal.) Figure 5 shows a compiled version of the gladiators' preferences.

How should the gladiators decide based on these preferences? In principle, the situation is very similar to a political election, where *voters* entertain preferences over *alternatives* such as parties or candidates. Thus, a first approach might be to invoke familiar voting rules. Many political elections use plurality rule, which chooses the alternative that is ranked first by most voters. In the example, this would be the antelope with four first-place votes. Other voting rules, like the one used in French presidential elections, first eliminate the alterna-

4	3	2
a	c	b
b	b	c
c	a	a

Figure 5: Gladiators' preferences over animals. Four gladiators prefer a over b over c, three prefer c over b over a, and two prefer b over c over a.

tive with *fewest* first-place votes and then have a runoff between the remaining two alternatives. This would result in b being eliminated and c winning the runoff versus a, as five out of nine gladiators prefer c over a. A third approach to voting, and one that will be extensively studied in this thesis, is related to the tournament setting from the previous section. Indeed, a round-robin tournament between the alternatives a, b, and c can be constructed by looking at pairwise majority comparisons. It has already been observed that a majority of gladiators prefers c over a. Likewise, a majority of gladiators (five out of nine) that prefer b over a and another majority of gladiators (six out of nine) that prefer b over c can be found. The method of pairwise comparisons therefore suggests that b should be chosen, as b is majority-preferred to all other alternatives (see Figure 6 for a graphical illustration).

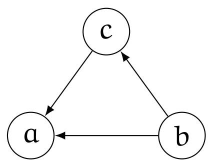


Figure 6: Graphical representation of majorities

The fact that three seemingly plausible voting rules lead to three different results in such an easy example is somewhat disconcerting. It is examples of this kind that have spurred the interest into voting rules and their properties. More generally, social choice theory studies scenarios where the preferences of individuals need to be aggregated in one form or another. As famously shown by Arrow (1951), every preference aggregation method has its drawbacks. The role of social choice theory can be seen as informing decision makers (the cook in the example above, or a country in the case of presidential elections) about the properties an aggregation mechanism does or does not satisfy. The decision maker can then choose a mechanism by weighing the pros and cons of different methods.

It should be noted that most voting rules may yield *sets* of alternatives that are tied for winner. This is easily seen for the plurality rule and for runoff rules (which coincide with plurality in the case of two alternatives). Furthermore, Condorcet's paradox (Figure 7) shows

1	1	1
a	c	b
b	a	c
c	b	a

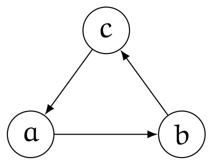


Figure 7: Condorcet's paradox. Even though all three voters have transitive preferences, the majority relation is cyclic.

that the pairwise majority relation may be cyclic and thus unable to produce a single winner.

Ties in the outcome are not necessarily a bad thing. In fact, fairness axioms require that ties be declared in certain symmetric situations (see Section 1.2.3). For practical purposes, however, ties must somehow be broken. How this is done is often left unspecified. In this sense, preference aggregation mechanism share some of the problems (and benefits) of set-valued solution concepts in game theory. The following section discusses these issues in more detail.

1.2 THE VALUE OF SET-VALUED SOLUTION CONCEPTS

While the interpretation of set-valuedness is similar in both disciplines, there are formal distinctions between set-valued solution concepts in social choice theory and set-valued solution concepts in game theory. These differences become apparent when one compares set-valued concepts with their respective traditional counterparts: *resolute* choice functions in social choice theory and *point-valued* solution concepts in game theory. In order to get a clear picture of those contrasts, let me propose a classification of solution concepts that encompasses both social choice and game theory.

1.2.1 *A Classification of Solution Concepts*

From a high-level perspective, both game theory and social choice deal with *preferences* and *objects*, and the purpose of a solution concept is to enable the choice of objects based on given preferences.[2] In game theory, every player has to choose among his *actions*, and preferences (over action profiles) are given by utility functions. In social choice theory, society wants to choose among *alternatives*, and preferences (over alternatives) are given by voters' rankings of the alternatives. A solution concept maps preferences to a nonempty set of *solutions*. Intuitively, each solution represents a way to make a choice among the objects.

I classify solution concepts according to two dimensions, namely form and multiplicity. The *form* of a solution refers to the mathemat-

2 This unifying perspective resembles, in a less formal way, Gibbard's (1973) definition of *game forms*.

		resolute	irresolute
single object	SC	Borda's rule with lexicographic tie-breaking	uncovered set (page 99), Borda's rule (page 101)
	GT		rationalizability (Pearce, 1984; Bernheim, 1984)
set of objects	SC	minimal covering set (page 100)	minimal upward covering sets (Brandt and Fischer, 2008b)
	GT	iterated strict dominance (page 25)	iterated weak dominance (page 25), S-sets (page 28)
randomization over objects	SC	random dictatorship (e.g., Zeckhauser, 1973)	maximal lotteries (Fishburn, 1984)
	GT		Nash equilibrium (page 23)

Table 1: Examples of solution concepts in social choice (SC) and game theory (GT) of varying form and multiplicity

ical shape a solution takes. I distinguish solution concepts for which every solution is *(i)* a single object, *(ii)* a nonempty set of objects, and *(iii)* a probability distribution over objects. The *multiplicity* of a solution concept refers to the number of solutions per instance. A solution concept is called *resolute* if it always outputs exactly one solution, and *irresolute* if it sometimes outputs more than one solution.[3] Table 1 illustrates this two-dimensional classification, and gives examples for each of the resulting six classes. Several comments are in order.

First, the distinction between irresolute solution concepts yielding single objects and resolute solution concepts yielding sets of objects is somewhat blurred. In fact, I will identify those two classes when I discuss social choice functions in Part II of this thesis.

Second, several prominent concepts may fail to produce any solution, and therefore do not fit my more restrictive definition of a solution concept. Examples include pure Nash equilibrium and Condorcet winner (yielding single objects), downward minimal covering sets and W-sets (yielding sets of objects), and quasi-strict Nash equilibrium (yielding randomizations over objects).

Third, a further generalization of the classification is possible for the case when one is not interested in the *choice* of objects, but rather in *rankings* of all the objects. This can be captured by allowing for additional form options like rankings of objects, sets of rankings of objects, and randomizations over rankings of objects.

3 By definition, a solution concept always outputs at least one solution.

We can now compare set-valued solution concepts with their traditional counterparts. Let us start with game theory, and then turn to social choice theory.

1.2.2 *Set-Valued Solution Concepts in Game Theory*

The traditional focus of game theory has been on solutions that give unambiguous strategic advice to players, in the form of either a single action (called *pure* strategy) or a probability distribution over actions (called *mixed* strategy).[4] Solution concepts of these forms are sometimes called *point-valued*, as every player is assigned a single point in the space of all possible strategies.

Point-valued solution concepts like Nash equilibrium have been criticized on various grounds. In particular, their need for randomization has been deemed unsuitable, impractical, or even infeasible (see, e.g., Luce and Raiffa, 1957; McKelvey and Ordeshook, 1976). Some of these drawbacks can be remedied by turning to solution concepts that recommend *sets of actions* to players. Dufwenberg et al. (2001, pp. 119–120) list several advantages of set-valued solution concepts:

> First, as argued by Basu and Weibull (1991), there is no obvious reason why recommendations should take the form of a single strategy rather than a set of strategies. Second, if one does not consider mixed strategies as reasonable objects of choice (see e.g. Ariel Rubinstein's arguments in Osborne and Rubinstein (1994, Section 3.2.1) then in many games no equilibria exist while appropriate set-valued solutions might have no such problems. Third, some notions that arise in decision-theoretic approaches to analyzing games, like the product set of rationalizable strategies (Bernheim, 1984; Pearce, 1984), fit quite nicely into the framework of set-valued solutions. Fourth, in many games some player will have no "strict" incentive to comply with a recommended profile because he has multiple optimal choices given that all others comply. If all such strategies are made part of the recommendation, this will come as a strategy set. Similar concerns presumably motivate Nash's (1951) notion of "strict solvability," and certainly motivate the work of Basu and Weibull (1991) and Hurkens (1995, see especially pp. 13–14).

I have already mentioned that the *interpretation* of set-valued solutions is not obvious. This is particularly true when the assigned sets of actions are viewed as recommendations to players, as it is not clear

4 Game-theoretic solution concepts are usually not required to be resolute. Nevertheless, it is desirable to identify classes of games for which a solution concept *is* resolute. See Sections 3.3 and 6.1 for results of this kind.

what action a player should pick after all. This apparent drawback could, however, also be seen as a benefit: the uncertainty about eventual choices of the opponents leads every player to stick himself to a set of actions rather than a single action. This line of thought is reminiscent of Nash equilibrium, and the set-valued solution concepts that are studied in Chapters 3 to 5 can be seen as selecting minimal sets of actions that stand in a best-response relation with one another.

1.2.3 *Set-Valued Solution Concepts in Social Choice Theory*

Many authors in social choice theory assume that solution concepts (called *social choice functions* or SCFs for short) are resolute. For instance, the classic result by Gibbard (1973) and Satterthwaite (1975) and Moulin's (1988b) *no show paradox* heavily rely on this assumption. However, resoluteness has been criticized for being unnatural and unreasonable (Gärdenfors, 1976; Kelly, 1977). For example, consider a situation with two agents and two alternatives such that each agent prefers a different alternative.[5] The problem is not that a resolute SCF has to select a single alternative (which is a well-motivated practical requirement), but that it has to select a single alternative *based on the individual preferences alone* (see, e.g., Kelly, 1977). As a consequence, any resolute SCF has to be biased towards an alternative or a voter (or both). Resoluteness is therefore at variance with such elementary fairness notions as *neutrality* (symmetry among the alternatives) and *anonymity* (symmetry among the voters). In a similar vein, Barberà (1977b, p. 1574) speaks of "situations where individual preferences are 'symmetric' and yet the single-valuedness of [SCFs] forces the alternatives to be treated in an 'asymmetric' manner."

It is for these reasons that this thesis almost exclusively considers *irresolute*, i.e., set-valued SCFs. It should be mentioned that this approach is not without its problems. First and foremost, there is the issue of *interpreting* set-valued outcomes. In the context of social choice, alternatives usually represent *mutually exclusive* outcomes. It is assumed—often implicitly so—, that a final, resolute choice will be made from the set of chosen alternatives, but this final resolution is not captured by the SCF. As such, an irresolute SCF can be seen as an incompletely specified choice mechanism, or as a pre-processing method that eliminates alternatives that should *not* get selected. Second, this lack of specification also has serious consequences for the study of strategyproofness, as voters' preferences over sets of alternatives crucially depend on the final choice process. In Chapter 11, different variants of preferences over sets and their implications on the manipulability of irresolute SCFs will be considered. Third, the very same fairness notions that motivate irresoluteness might lead to computational intractability, as exemplified in Section 8.1. There,

5 Barberà (1977b) refers to scenarios like this as "natural ties."

we study two natural variants of the SCF *ranked pairs* and reveal a trade-off between neutrality and tractability.

1.3 OVERVIEW OF THIS THESIS

This thesis is divided into two parts. The first part studies set-valued solution concepts for normal-form games. These solution concepts are based on varying notions of dominance. After introducing basic game-theoretic concepts in Chapter 2, formal definitions of dominance-based solutions are presented in Chapter 3, together with results on the multiplicity of solutions. Then, the complexity of computing solutions is considered. Two generic algorithms, a greedy and a sophisticated one, are proposed in Chapter 4. For each algorithm, properties of the underlying dominance notion are identified that ensure that the algorithms are sound and efficient. It is furthermore shown that certain problems associated with concepts based on weak and very weak dominance are computationally intractable (Chapter 5). Finally, Chapter 6 considers the classic notion of iterated dominance.

The second part of this thesis is concerned with irresolute SCFs, an important subclass of which is formed by tournament solutions. After introducing basic concepts in Chapter 7, the computational complexity of winner determination is studied for different variants of ranked pairs, Young's rule, and Dodgson's rule in Chapter 8. For a number of tractable SCFs such as maximin and Borda's rule, the complexity of computing possible and necessary winners for partially specified tournaments is determined in Chapter 9. Moreover, Chapter 10 studies axiomatic aspects of tournament solutions that are defined via retentiveness, and Chapter 11 investigates the strategyproofness of irresolute SCFs according to three well-known preference extensions.

1.3.1 *Publications on Which This Thesis is Based*

The work presented in this thesis is based on a number of papers that have been published in journals and conference proceedings.[6] In particular, Chapters 3 and 4 are based on [3,4,5], Chapter 5 is based on [9], and Chapter 6 is based on [8]. As for the second part, Chapter 8 is partly based on [7,10], Chapter 9 is based on [1], Chapter 10 is based on [6], and Chapter 11 is based on [2]. Since all these papers have coauthors, I will use the personal pronoun "we" throughout this thesis.

6 Numbers in brackets refer to the list on pages ix–x.

1.3.2 *Work Excluded From This Thesis*

Apart from the papers mentioned above, my work has resulted in a couple of other publications that will not be discussed in this thesis for reasons of coherence. I list them here for completeness.

- F. Brandt, M. Brill, and H. G. Seedig. On the fixed-parameter tractability of composition-consistent tournament solutions. In T. Walsh, editor, *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI)*, pages 85–90. AAAI Press, 2011.
- H. Aziz, M. Brill, and P. Harrenstein. Testing substitutability of weak preferences. In *Proceedings of the 2nd International Workshop on Matching Under Preferences (MATCHUP)*, 2012.

1.4 PREREQUISITES

In order to appreciate the concepts presented in this thesis, only a basic knowledge of mathematical objects such as sets, relations, functions, graphs etc. is required. For the computational results, I assume the reader to be familiar with the basic notions of complexity theory, such as polynomial-time many-one reductions, Turing reductions, hardness and completeness, and with standard complexity classes such as P, NP, and coNP (see, e.g., Papadimitriou, 1994). I will further use the complexity classes Σ_2^p and Θ_2^p. $\Sigma_2^p = \mathrm{NP}^{\mathrm{NP}}$ forms part of the second level of the polynomial hierarchy and consists of all problems that can be solved in polynomial time by a non-deterministic Turing machine with access to an NP oracle. $\Theta_2^p = \mathrm{P}_{||}^{\mathrm{NP}}$ consists of all problems that can be solved in polynomial time by a deterministic Turing machine with parallel (i.e., non-adaptive) access to an NP oracle.

Part I

GAME THEORY

2 GAMES, DOMINANCE, AND SOLUTIONS

In this chapter, we introduce game-theoretic concepts and notations that will be used throughout the first part of this thesis. We refer to the textbooks by Owen (1982), Myerson (1991), and Osborne and Rubinstein (1994) for more detailed accounts.

2.1 STRATEGIC GAMES

A finite *game in normal form* is a tuple $\Gamma = (N, (A_i)_{i \in N}, (u_i)_{i \in N})$, where $N = \{1, \ldots, n\}$ is a nonempty finite set of *players*. For each player $i \in N$, A_i is a nonempty finite set of *actions* available to player i, and $u_i : (\prod_{i \in N} A_i) \to \mathbb{R}$ is a *utility function*. A utility functions represents the preferences of a player over *action profiles* (i.e., combination of actions) by mapping each action profile $a \in \prod_{i \in N} A_i$ to a real number $u_i(a)$, with the interpretation that player i (weakly) prefers an action profile a to another action profile a' if and only if $u_i(a) \geqslant u_i(a')$. Since we are only concerned with finite normal-form games, we drop these qualifications and by 'game' understand finite game in normal form. We assume throughout this thesis that games are given explicitly, i.e., as tables containing the utilities for every possible action profile.

normal-form game

action profiles

Let $\Delta(M)$ denote the set of all probability distributions over a finite set M. A (mixed) *strategy* of a player $i \in N$ is an element of $\Delta(A_i)$. Whenever we are concerned with mixed strategies, we assume that preferences of players satisfy the axioms of von Neumann and Morgenstern (1944), such that utility functions can be extended to strategy profiles by taking *expected* utilities: for $s = (s_1, \ldots, s_n) \in \prod_{j \in N} \Delta(A_j)$ and $i \in N$,

strategy

expected utility

$$u_i(s) = \sum_{a \in \prod_{k \in N} A_k} u_i(a) \left(\prod_{j \in N} s_j(a_j) \right),$$

where $s_i(a_i)$ denotes the probability that player i assigns to action a_i in strategy s_i. A strategy that assigns probability one to a single action is called *pure* and usually identified with the corresponding action.

pure strategy

The *support* of a strategy s_i, denoted by *supp*(s_i), is the set of all actions that are assigned a positive probability, i.e., *supp*$(s_i) = \{a_i \in A_i : s_i(a_i) > 0\}$. For a strategy profile $s = (s_1, \ldots, s_n)$, we furthermore let *supp*(s) denote the tuple $(\textit{supp}(s_1), \ldots, \textit{supp}(s_n))$.

support

	b_1	b_2
a_1	$(2,5)$	$(-1,0)$
a_2	$(-2,1)$	$(5,2)$
a_3	$(-1,0)$	$(-1,0)$

(a) a bimatrix game

	b_1	b_2	b_3
a_1	1	0	-1
a_2	-2	2	4
a_3	0	2	1

(b) a matrix game

	b_1	b_2
a_1	$(2,2)$	$(-1,5)$
a_2	$(-2,6)$	$(4,0)$
a_3	$(0,4)$	$(1,3)$

(c) a constant-sum game

	a_1	a_2	a_3
a_1	0	2	0
a_2	-2	0	-4
a_3	0	4	0

(d) a symmetric matrix game

	a_1	a_2	a_3
a_1	0	2	-1
a_2	-2	0	-4
a_3	1	4	0

(e) a confrontation game

	a_1	a_2	a_3
a_1	0	1	-1
a_2	-1	0	-1
a_3	1	1	0

(f) a tournament game

Figure 8: Examples of two-player games

bimatrix

A two-player game $\Gamma = (\{1,2\}, (A_1, A_2), (u_1, u_2))$ is often called a *bimatrix* game, as it can conveniently be represented as an $|A_1| \times |A_2|$ bimatrix M, i.e., a matrix with rows indexed by A_1, columns indexed by A_2, and $M(a_1, a_2) = (u_1(a_1, a_2), u_2(a_1, a_2))$ for every action profile $(a_1, a_2) \in A_1 \times A_2$. We will commonly refer to actions of players 1 and 2 by the rows and columns of this matrix, respectively. A two-player game is a *matrix game* (or *zero-sum game*) if $u_1(a_1, a_2) + u_2(a_1, a_2) = 0$ for all $(a_1, a_2) \in A_1 \times A_2$. Such a game can conveniently be represented by a single matrix that only contains the utilities of player 1. Since all of the solution concepts considered in this thesis are invariant under positive affine transformations of the utility functions, our results on matrix games also hold for the larger class of two-player *constant-sum games*, in which the utilities of both players add up to the same value in every action profile. In fact, and mainly due to reasons of notational convenience, the results in Chapter 6 are presented for constant-sum games rather than matrix games.

matrix game

constant-sum games

We let A denote the set of all actions of a game, i.e., $A = \bigcup_{i \in N} A_i$. We usually assume that the actions sets of different players are pairwise disjoint, and thus have $|A| = \sum_{i \in N} |A_i|$. An exception are symmetric games, in which all players have the same action set $A_i = A$. We are only interested in symmetry in the context of matrix games.

symmetric matrix games

A matrix game is *symmetric* if the corresponding matrix is skew-symmetric. For symmetric matrix games, we often simplify notation and write $\Gamma = (A, u)$, where A is the set of actions and u denotes the utility function of player 1. See Figure 8 for example games.

confrontation games

Confrontation games are symmetric matrix games characterized by the fact that the two players get the same utility if and only if they play the same action (Duggan and Le Breton, 1996a). Formally, a symmetric matrix game $\Gamma = (A, u)$ is called confrontation game if it satisfies the *off-diagonal property*:

off-diagonal property

$$\text{for all } a, b \in A,\ u(a, b) = 0 \text{ if and only if } a = b.$$

If moreover $u(a, b) \in \{-1, 0, 1\}$ for all $a, b \in A$, we have a *tournament game*.[1]

tournament game

A game is called *binary* if $u_i(a) \in \{0, 1\}$ for all $i \in N$ and $a \in \prod_{i \in N} A_i$. A binary two-player game is called a win-lose game. The set of *outcomes* of a game is given by $\{(u_1(a), \dots, u_n(a)) : a \in \prod_{i \in N} A_i\}$. Thus, a win-lose game only allows the outcomes $(0, 0)$, $(0, 1)$, $(1, 0)$, and $(1, 1)$.

outcomes

The following notation will be useful for reasoning about set-valued solution concepts. Let A_N denote the n-tuple $(A_1, \dots, A_n)$ containing the action sets of all players.[2] An n-tuple $X = (X_1, \dots, X_n)$ is said to be *nonempty* if $X_i \neq \emptyset$ for all $i \in N$. For two n-tuples $X = (X_1, \dots, X_n)$ and $Y = (Y_1, \dots, Y_n)$, we write $X \subseteq Y$ if X is nonempty and $X_i \subseteq Y_i$ for all $i \in N$. To simplify the exposition, we will frequently abuse terminology and refer to an n-tuple $X \subseteq A_N$ as a "set." For every player i, we furthermore let X_{-i} denote the set $\prod_{j \in N \setminus \{i\}} X_j$ of all opponent action profiles where each opponent $j \in N \setminus \{i\}$ is restricted to play only actions from $X_j \subseteq A_j$.

$X \subseteq A_N$

A *subgame* of a game $\Gamma = (N, (A_i)_{i \in N}, (u_i)_{i \in N})$ is a game $\Gamma' = (N, (X_i)_{i \in N}, (u'_i)_{i \in N})$ where $X \subseteq A_N$ and $u'_i(a) = u_i(a)$ for all $a \in \prod_{i=1}^n X_i$. Γ is then called a *supergame* of Γ'. For a symmetric matrix game $\Gamma = (A, u)$ and a nonempty subset $X \subseteq A$ of actions, $\Gamma|_X$ denotes the subgame of Γ restricted to X, i.e., $\Gamma|_X$ is the symmetric matrix game $(X, u|_{X \times X})$. We often identify a tuple $X \subseteq A_N$ with the corresponding induced subgame.

subgame

2.2 DOMINANCE

In this section, we define the dominance notions on which the main solution concepts in this thesis will be based. Furthermore, we introduce a number of properties that will be critical for the results in the following chapters.

2.2.1 *Dominance Structures*

Roughly speaking, an action a dominates another action b if a yields a higher payoff than b. We will see that there are several differ-

1 The term *tournament game* refers to the fact that such a game $\Gamma = (A, u)$ can be represented by a tournament graph with vertex set A and edge set $\{(a, b) : u(a, b) = 1\}$. In a similar fashion, a confrontation game can be represented by a *weighted* tournament graph.

2 For a symmetric matrix game $\Gamma = (A, u)$, we thus have $A_N = (A, A)$.

ent ways to formalize this intuitive idea. Consider a player $i \in N$. Whether an action (or a combination of actions) in A_i dominates another action in A_i naturally depends on which actions the other players have at their disposal. This is reflected in the following definition, in which a dominance structure is defined as a mapping from opponent action profiles to dominance relations on A_i. In order to accommodate for Börgers dominance (Börgers, 1993) and mixed dominance structures, we furthermore have dominance relations on A_i relate *sets of actions* to individual actions.

Definition 2.1. *Let $\Gamma = (N, (A_i)_{i\in N}, (u_i)_{i\in N})$ be a game in normal form and $X \subseteq A_N$. For each player $i \in N$, a* dominance structure D *maps X_{-i} to a subset of $2^{A_i} \times A_i$ such that $(X_i, a_i) \in D(X_{-i})$ implies $(Y_i, a_i) \in D(X_{-i})$ for all Y_i with $X_i \subseteq Y_i \subseteq A_i$.*

For $X_i \subseteq A_i$ and $a_i \in A_i$, we write $X_i \; D(X_{-i}) \; a_i$ if $(X_i, a_i) \in D(X_{-i})$. If X_i consists of a single action x_i, we write $x_i \; D(X_{-i}) \; a_i$ instead of $\{x_i\} \; D(X_{-i}) \; a_i$ to avoid cluttered notation. If $X_i \; D(X_{-i}) \; a_i$, we say that a_i *is* D-*dominated by* X_i *with respect to* X_{-i}. We furthermore say that a_i is D-*dominated* if there exists $X_i \subseteq A_i$ such that a_i is D-dominated by X_i with respect to A_{-i}, and we call a game D-*irreducible* if no action is D-dominated.

D-*dominated*

D-*irreducible*

We go on to define the main dominance structures considered in this thesis, together with their mixed counterparts that allow for randomized strategies (consult Figure 9 for examples).

Definition 2.2. *Let $\Gamma = (N, (A_i)_{i\in N}, (u_i)_{i\in N})$ be a game in normal form and $i \in N$. Furthermore, let $X \subseteq A_N$ and $a_i \in A_i$.*

- strict dominance (S): $X_i \; S(X_{-i}) \; a_i$ *if there exists $x_i \in X_i$ with $u_i(x_i, x_{-i}) > u_i(a_i, x_{-i})$ for all $x_{-i} \in X_{-i}$.*
- weak dominance (W)*: $X_i \; W(X_{-i}) \; a_i$ if there exists $x_i \in X_i$ with $u_i(x_i, x_{-i}) \geqslant u_i(a_i, x_{-i})$ for all $x_{-i} \in X_{-i}$ and the inequality is strict for at least one $x_{-i} \in X_{-i}$.*
- very weak dominance (V)*: $X_i \; V(X_{-i}) \; a_i$ if there exists $x_i \in X_i$ with $u_i(x_i, x_{-i}) \geqslant u_i(a_i, x_{-i})$ for all $x_{-i} \in X_{-i}$.*
- Börgers[3] dominance (B)*: $X_i \; B(X_{-i}) \; a_i$ if $X_i \; W(Y_{-i}) \; a_i$ for all $Y_{-i} \subseteq X_{-i}$.*
- mixed strict dominance (S^*): $X_i \; S^*(X_{-i}) \; a_i$ *if there exists $s_i \in \Delta(X_i)$ with $u_i(s_i, x_{-i}) > u_i(a_i, x_{-i})$ for all $x_{-i} \in X_{-i}$.*
- mixed weak dominance (W^*)*: $X_i \; W^*(X_{-i}) \; a_i$ if there exists $s_i \in \Delta(X_i)$ with $u_i(s_i, x_{-i}) \geqslant u_i(a_i, x_{-i})$ for all $x_{-i} \in X_{-i}$ and the inequality is strict for at least one $x_{-i} \in X_{-i}$.*

3 Börgers dominance is referred to as *inherent* dominance by Apt (2004, 2011).

	b_1	b_2	b_3	b_4
a_1	1	2	3	0
a_2	2	1	3	0
a_3	0	1	0	1
a_4	1	1	2	2

Figure 9: Example matrix game with $A_1 = \{a_1, a_2, a_3, a_4\}$ and $A_2 = \{b_1, b_2, b_3, b_4\}$. For $X_2 = \{b_1, b_2, b_3\}$, the following dominance relations hold, among others: $a_1\ S(X_2)\ a_3$, $a_2\ W(X_2)\ a_3$, and $\{a_1, a_2\}\, B(X_2)\ a_4$.

- mixed very weak dominance (V^*)*:* $X_i\ V^*(X_{-i})\ a_i$ *if there exists* $s_i \in \Delta(X_i)$ *with* $u_i(s_i, x_{-i}) \geqslant u_i(a_i, x_{-i})$ *for all* $x_{-i} \in X_{-i}$.

Börgers dominance has a mixed counterpart as well, requiring that $X_i\ W^*(Y_{-i})\ a_i$ for all $Y_{-i} \subseteq X_{-i}$. However, mixed Börgers dominance coincides with mixed strict dominance (Duggan and Le Breton, 1996a).

The following dominance structures are only well-defined in symmetric matrix games, and we will refer to them as *symmetric dominance structures*.

Definition 2.3. *Let* $\Gamma = (A, u)$ *be a symmetric matrix game,* $X, Y \subseteq A$, *and* $a \in A$.

- covering (C_M): $X\ C_M(Y)\ a$ *if there exists* $x \in X \cap Y$ *with*
  - $u(x, a) > 0$ *and*
  - $u(x, y) \geqslant u(a, y)$ *for all* $y \in Y$.
- deep covering (C_D): $X\ C_D(Y)\ a$ *if there exists* $x \in X \cap Y$ *with*
  - $u(x, a) > 0$,
  - $u(x, y) \geqslant u(a, y)$ *for all* $y \in Y$, *and*
  - $u(x, y) > u(a, y)$ *for all* $y \in Y$ *with* $u(a, y) = 0$.

Covering was introduced by McKelvey (1986) and later generalized by Dutta and Laslier (1999), and deep covering is a generalization of a notion by Duggan (2012).

For two dominance structures D and D', we write $D \subseteq D'$ if $D(X_{-i}) \subseteq D'(X_{-i})$ for all $X \subseteq A_N$. The following relations follow immediately from the respective definitions: $S \subseteq B \subseteq W \subseteq V$, $C_D \subseteq C_M$, and $D \subseteq D^*$ for all $D \in \{S, W, V\}$. Since mixed Börgers dominance coincides with S^*, we also have $B \subseteq S^*$.

2.2.2 *Properties of Dominance Structures*

An action $a_i \in A_i$ is said to be D-maximal with respect to X_{-i} if it is not D-dominated by A_i.

Definition 2.4. *Let* D *be a dominance structure and* $X \subseteq A_N$*. The* D-*maximal elements of* A_i *with respect to* X_{-i} *are defined as*

$$\max(D(X_{-i})) = A_i \setminus \{a_i \in A_i : A_i \; D(X_{-i}) \; a_i\}.$$

We now define a number of properties of dominance structures.

Definition 2.5. *Let* $X \subseteq A_N$ *and* $a_i \in A_i$*. A dominance structure* D *satisfies*

- monotonicity (MON) *if* $X_i \; D(X_{-i}) \; a_i$ *implies* $X_i \; D(Y_{-i}) \; a_i$ *for all* $Y_{-i} \subseteq X_{-i}$*,*
- computational tractability (COM) *if* $X_i \; D(X_{-i}) \; a_i$ *can be checked in polynomial time,*
- maximal domination (MAX) *if* $A_i \; D(X_{-i}) \; a_i$ *implies*

  $$\max(D(X_{-i})) \; D(X_{-i}) \; a_i,$$

  and
- singularity (SING) *if* $X_i \; D(X_{-i}) \; a_i$ *implies the existence of an action* $x_i \in X_i$ *with* $x_i \; D(X_{-i}) \; a_i$*.*

It is easily seen that S, B, and V are monotonic, and that W is not. S and W satisfy maximal domination because the relations $S(X_{-i})$ and $W(X_{-i})$—restricted to pairs of singletons—are transitive and irreflexive. On the other hand, V violates MAX because $\max(V(X_{-i}))$ may be empty. It directly follows from the definitions that S, W, V, C_M, and C_D are singular. All the dominance structures introduced in this section are computationally tractable.[4]

The following properties are defined for symmetric dominance structures.

Definition 2.6. *Let* $X, X' \subseteq A$ *and* $a, b, c \in A$*. A symmetric dominance structure* D *satisfies*

- weak monotonicity (weak MON) *if* $a \; D(X) \; b$ *implies* $a \; D(Y) \; b$ *for all* $Y \subseteq X$ *with* $a \in Y$*,*
- transitivity (TRA) *if* $a \; D(X) \; b$*,* $b \; D(X') \; c$*, and* $a \in X \cap X'$ *imply* $a \; D(X \cap X') \; c$*,*
- computational tractability of finding subsets (SUB-COM) *if a nonempty subset of a* D*-set can be computed in polynomial time, and*

4 See Algorithm 4 on page 46 for checking Börgers dominance.

- uniqueness (UNI) *if every symmetric matrix game has a unique* D-*set.*

See Sections A.1 and A.2 in the appendix for complete lists of dominance structures and their properties.

2.3 SOLUTION CONCEPTS

Solution concepts in game theory try to capture rational behavior. The solutions they produce can be seen as recommendations to players as well as predictions of the actual outcome. Formally, a (game-theoretic) *solution concept* S associates with each game Γ a set $S(\Gamma)$ of *solutions*. We distinguish point-valued and set-valued solution concepts. For *point-valued* solution concepts, each solution $X \in S(\Gamma)$ is a strategy profile, i.e., $X \in \prod_{i \in N} \Delta(A_i)$. For *set-valued* solution concepts, on the other hand, each solution is a product set in A, i.e., a set $X = \prod_{i \in N} X_i$, where $\emptyset \neq X_i \subseteq A_i$ for all $i \in N$. For notational convenience, we usually identify a product set $X = \prod_{i \in N} X_i$ with the tuple $(X_1, \ldots, X_n) \subseteq A_N$.

solution concept

point-valued

set-valued

2.3.1 *Nash Equilibrium*

Nash equilibrium is undoubtedly the most important point-valued solution concept in game theory. A combination of strategies is a Nash equilibrium if no player can benefit by unilaterally changing his strategy. For a strategy profile s, let s_{-i} denote the profile of all strategies in s except s_i.

Definition 2.7. *A strategy profile* $s^* \in \prod_{i \in N} \Delta(A_i)$ *is a* Nash equilibrium *if for each player* $i \in N$ *and every strategy* $s_i \in \Delta(A_i)$*,*

$$u_i(s_i^*, s_{-i}^*) \geqslant u_i(s_i, s_{-i}^*).$$

The set of all Nash equilibria of a game Γ *is denoted by* $N(\Gamma)$*.*

Nash equilibria can also be characterized in terms of best responses. Given a profile s_{-i} of opponent strategies, the set $BR_i(s_{-i})$ of *best responses* of player i is given by the set of all actions that maximize the utility of player i against s_{-i}, i.e.,

best responses

$$BR_i(s_{-i}) = \arg\max_{s_i \in \Delta(A_i)} u_i(s_i, s_{-i}).$$

A strategy profile s is a Nash equilibrium if and only if $s_i \in BR_i(s_{-i})$ for all $i \in N$.

An action is called *essential* if it is played with positive probability in *some* Nash equilibrium. The *essential set* $ES(\Gamma)$ is an n-tuple containing

essential set

	b_1	b_2	b_3
a_1	1	1	2
a_2	1	1	3
a_3	0	1	1

Figure 10: Matrix game with a continuum of Nash equilibria

the essential actions of each player. Formally,

$$ES_i(\Gamma) = \{a_i \in A_i : \exists s \in N(\Gamma) \text{ with } a_i \in supp(s_i)\}$$

and $ES(\Gamma) = (ES_1(\Gamma), \dots, ES_n(\Gamma))$. Nash (1951) has shown that every normal-form game has a Nash equilibrium. As a consequence, the essential set is never empty.

Since a game may have many Nash equilibria, several refinements have been proposed to single out particularly desirable ones (see van Damme, 1983, for an overview). A *quasi-strict equilibrium* (Harsanyi, 1973) is a Nash equilibrium in which all actions played with positive probability yield *strictly* more utility than actions played with probability zero.

Definition 2.8. *A Nash equilibrium* $s \in N(\Gamma)$ *is called* quasi-strict *if for all* $i \in N$ *and all* $a, b \in A_i$ *with* $s_i(a) > 0$ *and* $s_i(b) = 0$, $u_i(a, s_{-i}) > u_i(b, s_{-i})$.

In a quasi-strict equilibrium, every player assigns a positive probability to each of his best responses. Norde (1999) has shown that every bimatrix game has a quasi-strict equilibrium.

Let us illustrate these concepts with the help of an easy example. The matrix game Γ in Figure 10 has a continuum of Nash equilibria in *mixed* strategies, namely

$$N(\Gamma) = \{\,(s_1, s_2) : \exists\, p, q \in [0,1] \text{ such that } s_1 = (p, 1-p, 0) \\ \text{and } s_2 = (q, 1-q, 0)\,\}.$$

The essential set is thus given by $ES(\Gamma) = (\{a_1, a_2\}, \{b_1, b_2\})$. Observe that $N(\Gamma)$ is convex and that the strategies constituting Nash equilibria are "interchangeable," in the sense that any combination of equilibrium strategies of player 1 and equilibrium strategies of player 2 constitutes an equilibrium. This is not a coincidence, as the following well-known fact shows.

Fact 2.9. *Let* Γ *be a matrix game. The following properties hold:*

(i) Convexity: *If* $s, s' \in N(\Gamma)$, *then* $\lambda s + (1-\lambda)s' \in N(\Gamma)$ *for every* $\lambda \in [0, 1]$.

(ii) Interchangeability: *If* $s, s' \in N(\Gamma)$, *then* $(s_1, s_2') \in N(\Gamma)$ *and* $(s_1', s_2) \in N(\Gamma)$.

The convexity of Nash equilibria in matrix games implies that every matrix game Γ has a quasi-strict Nash equilibrium s with $supp(s) = ES(\Gamma)$. Brandt and Fischer (2008a) have shown that $supp(s) = ES(\Gamma)$ is in fact necessary for any quasi-strict equilibrium s of Γ. As a consequence, all quasi-strict equilibria of a matrix game have the same support. In the game in Figure 10, the set of quasi-strict equilibria is given by the set of all combinations of strategies in which both players play both their essential actions with positive probability. Thus, almost all Nash equilibria of this game are quasi-strict.

2.3.2 *Iterated Dominance*

A simple conviction in game theory is that a player need not bother to consider an action that is dominated. Of course, the validity of such an argument depends on the notion of dominance one employs, and different dominance structures will easily lead to different conclusion about the plausibility of certain actions. We need the following definitions.

Definition 2.10. *An* elimination sequence *of a game is a finite sequence $\Sigma = (\Sigma_1, \dots, \Sigma_n)$ of pairwise disjoint subsets of actions in A. For a game $\Gamma = (N, (A_i)_{i \in N}, (u_i)_{i \in N})$ and an elimination sequence $\Sigma = (\Sigma_1, \dots, \Sigma_n)$ of Γ we have $\Gamma(\Sigma)$ denote the subgame where the actions in $\Sigma_1 \cup \dots \cup \Sigma_n$ have been removed, i.e., $\Gamma(\Sigma) = (N, (A'_i)_{i \in N}, (u'_i)_{i \in N})$ where $A'_i = A_i \setminus (\Sigma_1 \cup \dots \cup \Sigma_n)$ and u' is the restriction of u to $\prod_{i \in N} A'_i$.*

In the special case of symmetric matrix games, $A_1 = A_2 = A$ implies that rows and columns corresponding to the same actions can only be eliminated simultaneously.

The *validity* of an elimination sequence with respect to a dominance structure D is defined inductively.

Definition 2.11. *An elimination sequence $\Sigma = (\Sigma_1, \dots, \Sigma_m)$ is* valid for Γ with respect to D *if either it is the empty sequence, or if $(\Sigma_1, \dots, \Sigma_{m-1})$ is valid for Γ with respect to D and every action $a_i \in \Sigma_m$ is D-dominated in $\Gamma(\Sigma_1, \dots, \Sigma_{m-1})$.*

For each dominance structure D, we can define a set-valued solution concept. Call a tuple $X \subseteq A_N$ D-*irreducible* if the corresponding subgame is D-irreducible.

Definition 2.12. *The solution concept* iterated D-dominance *returns all D-irreducible tuples $X \subseteq A_N$ that can be reached via an elimination sequence that is valid with respect to D.*

Iterated S-dominance and iterated W-dominance are well-established solution concepts, which have a long history and appear in virtually every textbook on game theory. It is well-known that the resulting subgame is independent of the order in which S-dominated *order-independence* actions are eliminated (see Section 6.1 for details). Thus, iterated S-dominance always returns a single tuple. By contrast, the game in Figure 26 on page 76 demonstrates that this is not necessarily the case for iterated W-dominance. The latter phenomenon gives rise to a number of computational issues, some of which will be studied in Chapter 6.

2.4 SUMMARY

We have introduced basic game-theoretic concepts and terminology. Table 2 summarizes some notation for later reference.

$\Gamma = (N, (A_i)_{i\in N}, (u_i)_{i\in N})$	game in normal form
$N = \{1, \dots, n\}$	set of players
A_i	set of actions of player i
$\Delta(A_i)$	set of (mixed) strategies of player i
$a \in \prod_{i=1}^{n} A_i$	action profile
$s \in \prod_{i=1}^{n} \Delta(A_i)$	strategy profile
$\mathit{supp}(s_i)$	support of a strategy s_i
$u_i : \prod_{i=1}^{n} \Delta(A_i) \to \mathbb{R}$	utility function of player i
$A = \bigcup_{i=1}^{n} A_i$	set of all actions
$A_N = (A_1, \dots, A_n)$	tuple of action sets
$X \subseteq A_N$	tuple of action subsets
$\Gamma = (A, u)$	symmetric matrix game
$\Gamma\|_X = (X, u\|_{X\times X})$	subgame of a symmetric matrix game
$N(\Gamma)$	Nash equilibria of Γ
$\mathit{BR}_i(s_{-i})$	best responses of player i to s_{-i}
$\mathit{ES}(\Gamma)$	essential set of Γ

Table 2: Notation for game-theoretic concepts

3 DOMINANCE-BASED SOLUTION CONCEPTS

In this chapter, we introduce dominance-based solution concepts and study the existence and multiplicity of solutions. Thereby, we prepare the ground for the computational results in the following chapters.

3.1 MOTIVATION

saddle point

Saddle points, i.e., combinations of actions such that no player can gain by deviating, are among the earliest solution concepts considered in game theory (see, e.g., von Neumann and Morgenstern, 1944). In matrix games, every saddle point happens to coincide with the optimal outcome both players can guarantee in the worst case and thus enjoys a very strong normative foundation. Unfortunately, however, not every matrix game possesses a saddle point. In order to remedy this situation, Borel (1921) introduced *mixed*—i.e., randomized—strategies and von Neumann (1928) proved that every matrix game contains a mixed saddle point (or equilibrium) that moreover maintains the appealing normative properties of saddle points. The existence result was later generalized to arbitrary normal-form games by Nash (1951), at the expense of its normative foundation. Since then, Nash equilibrium has commonly been criticized for resting on very demanding epistemic assumptions such as the common knowledge of von Neumann-Morgenstern utilities.[1]

generalized saddle point (GSP)

saddle

Shapley (1953a,b) showed that the existence of saddle points (and even uniqueness in the case of matrix games) can also be guaranteed by moving to *minimal sets* of actions rather than randomizations over them.[2] Shapley defines a *generalized saddle point (GSP)* to be a tuple of subsets of actions for each player that satisfies a simple external stability condition: every action not contained in a player's subset is dominated by some action in the set, given that the remaining players choose actions from their respective sets. A GSP is minimal if it does not contain another GSP. Minimal GSPs, which Shapley calls *saddles*, also satisfy internal stability in the sense that no two actions within a set dominate each other, given that the remaining players choose actions from their respective sets. While Shapley was the first to conceive GSPs, he was not the only one. Apparently unaware of Shapley's work, Samuelson (1992) uses the very related concept of a *consistent pair* to point out epistemic inconsistencies in the concept of

1 See, e.g., Luce and Raiffa (1957, pp. 74–76), Fishburn (1978), Bernheim (1984), Pearce (1984), Myerson (1991, pp. 88–91), Börgers (1993), Aumann and Brandenburger (1995), Perea (2007), Jungbauer and Ritzberger (2011).

2 The main results of the 1953 reports later reappeared in revised form (Shapley, 1964).

iterated W-dominance. Also, *weakly admissible sets* as defined by McKelvey and Ordeshook (1976) in the context of spatial voting games and the *minimal covering set* as defined by Dutta (1988) in the context of majority tournaments are GSPs (Duggan and Le Breton, 1996a).

In a regrettably unpublished paper, Duggan and Le Breton (1996b) extend Shapley's approach to normal-form games and other dominance structures. Their framework—which is very similar to the one used in this thesis—is rich enough to cover common set-valued solution concepts such as *rationalizability* (Bernheim, 1984; Pearce, 1984) and *CURB sets* (Basu and Weibull, 1991).

rationalizability

3.2 D-SOLUTIONS AND D-SETS

Generalizing a classic cooperative solution concept by von Neumann and Morgenstern (1944), a set of actions X can be said to be *stable* if it consists precisely of those alternatives not dominated by X (see also Wilson, 1970). This fixed-point characterization can be split into two conditions of internal and external stability: first, there should be no reason to restrict the selection by excluding some action from it; second, there should be an argument against each proposal to include an outside action into the selection.

stable sets

Definition 3.1. *Let $\Gamma = (N, (A_i)_{i \in N}, (u_i)_{i \in N})$ be a normal-form game and D a dominance structure. A tuple $X \subseteq A_N$ is a* D-solution *in Γ if for every $i \in N$,*

$$X_i \setminus \{x_i\} \; D(X_{-i}) \; x_i \text{ for no } x_i \in X_i \text{, and} \tag{1}$$

$$X_i \; D(X_{-i}) \; a_i \text{ for all } a_i \in A_i \setminus X_i. \tag{2}$$

We refer to (1) and (2) as *internal* and *external* D*-stability*, respectively. Similar to Nash equilibria, D-solutions can be characterized in terms of best responses. To see this, define $BR_i^D(X_{-i})$ as the set of subsets of A_i that satisfy conditions (1) and (2). Then, X is a D-solution if and only if $X_i \in BR_i^D(X_{-i})$ for all $i \in N$.

We are mainly interested in *inclusion-minimal* D-solutions. Following Duggan and Le Breton (1996b), we call them D*-sets*.

Definition 3.2. *A* D-set *is a D-solution X such that there does not exist a D-solution Y with $Y \subseteq X$ and $Y \neq X$.*

Figure 11 contains examples of D-sets for all dominance structures considered in this thesis.

Various set-valued solution concepts that have been proposed in the literature can be characterized as D-sets for some dominance structure D. Shapley's (1964) *saddles* and *weak saddles* for matrix games correspond to S- and V-sets, respectively. Dutta and Laslier's (1999) *minimal covering sets* for symmetric matrix games correspond to C_M-sets and Duggan's (2012) *minimal deep covering sets* for binary symmet-

Shapley's saddles

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
a_1	0	2	−2	1	2			
a_2	−2	0	2	1	2			
a_3	2	−2	0	1	1		3	
a_4	−1	−1	−1	0	1			
a_5	−2	−2	−1	−1	0			
a_6						0	3	−3
a_7			−3			−3	0	3
a_8						3	−3	0

	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8
b_1	0	0	0	0	1	1	1	1
b_2	0	0	1	−1	0	1	1	−1
b_3	0	−1	0	1	1	1	−1	1
b_4	0	1	−1	0	0	−1	1	1
b_5	−1	0	−1	0	0	1	1	1
b_6	−1	−1	−1	1	−1	0	−1	1
b_7	−1	−1	1	−1	−1	1	0	−1
b_8	−1	1	−1	−1	−1	−1	1	0

D	D-set
S	$(\{a_1, a_2, a_3, a_4, a_5\}, \{a_1, a_2, a_3, a_4, a_5\})$
W, V, B	$(\{a_1, a_2, a_3, a_4\}, \{a_1, a_2, a_3, a_4\})$
S^*, W^*, V^*	$(\{a_1, a_2, a_3\}, \{a_1, a_2, a_3\})$

D	D-set
C_D	$(\{b_1, b_2, b_3, b_4, b_5\}, \{b_1, b_2, b_3, b_4, b_5\})$
C_M	$(\{b_1, b_2, b_3, b_4\}, \{b_1, b_2, b_3, b_4\})$
V	$(\{b_1\}, \{b_1\})$

Figure 11: Example games with unique D-sets for several dominance structures D. The game on the left is a confrontation game and the game on the right is a symmetric matrix game.

ric matrix games correspond to C_D-sets. Furthermore, mixed refinements of Shapley's saddles, as proposed by Duggan and Le Breton (2001) for binary symmetric matrix games, correspond to S^*- and W^*-sets.

Two further solutions that fit into our framework are Harsanyi and Selten's (1988) *formations* and Basu and Weibull's (1991) *CURB sets*. The respective dominance structures are defined in terms of best response sets. An action a_i is *rationally dominated* with respect to a set X_{-i} of opponent action profiles if it is *not* a best response to any mixed opponent strategy with support in X_{-i}. A subtle difference occurs if there are more than two players (and therefore more than one opponent). While in *correlated rational dominance* (R_c), opponents are allowed to play joint, i.e., correlated, mixtures (and thus to act like a single opponent), *uncorrelated rational dominance* (R_u) restricts opponents to independent mixtures. Formally, $X_i \; R_c(X_{-i}) \; a_i$ if and only if

rational dominance

$$a_i \notin \arg\max_{a_i' \in X_i} u_i(a_i', x_{-i}) \quad \text{for all } x_{-i} \in \Delta(X_{-i}),$$

and $X_i \; R_u(X_{-i}) \; a_i$ if and only if

$$a_i \notin \arg\max_{a_i' \in X_i} u_i(a_i', x_{-i}) \quad \text{for all } x_{-i} \in \prod_{j \neq i} \Delta(X_j).$$

minimal CURB sets

primitive formations

A tuple of sets is a CURB set if and only if it is externally R_u-stable, and *minimal CURB sets* coincide with R_u-sets. Similarly, a tuple of sets is a formation if and only if it is externally R_c-stable, and *primitive formations* are R_c-sets. Since it is well known that an action is not a best response to some correlated opponent strategy if and only if it is dominated by a mixed strategy (see, e.g., Pearce, 1984, Lemma 3), the dominance structures R_c and S^* are identical.

Fact 3.3. *The dominance structures S^* and R_c coincide.*

As a consequence, all our results concerning S^*-sets directly apply to primitive formations as well. The same is true for minimal CURB sets in two-player games, due to the equivalence of R_c and R_u for $n = 2$.[3]

existence of solutions

Monotonicity turns out to be sufficient for the *existence of solutions*. If a dominance structure D satisfies MON, a D-solution can be constructed via iterated D-dominance: when the elimination process comes to an end, the resulting set is not only internally D-stable, but—due to MON—also externally D-stable. Note, however, that these solutions need not be minimal (see, e.g., Figure 11).[4] The same argument applies to symmetric dominance structures satisfying weak monotonicity. As a consequence, D-sets are guaranteed to exist for the dominance structures S, S^*, B, V, V^*, C_M, and C_D. Weak dominance (W) and mixed weak dominance (W^*), on the other hand, are not monotonic, so the above argument does not apply to those dominance structures. In fact, there are games without any W- or W^*-solutions (see Figure 12 for an example). For this reason, W-sets and W^*-sets do not qualify as solution concepts. Nevertheless, the concept of a W-set will prove useful to establish a connection between C_M-sets and V-sets in confrontation games (see Section 3.3).

weak saddles

The fact that W-solutions may fail to exist was first observed by Samuelson (1992). There are at least three approaches to restore the existence of W-solutions. First, one can ignore internal stability and consider minimal *externally* W-stable sets, so-called *weak saddles*. This approach, introduced by Duggan and Le Breton (2001), is followed in Chapter 5. Second, one can look for restricted classes of games in which W-solutions are guaranteed to exist. One such class is the class of confrontation games, where the W-set is unique and coincides with the V-set. Third, one can consider the so-called *monotonic kernel* of the dominance structure W, which turns out to be identical to Börgers dominance (see Duggan and Le Breton, 1996b).

internal stability

Another beneficial property of (weakly) monotonic dominance structures is that *minimal* externally stable sets also happen to be inter-

3 It was recently shown that CURB sets are computationally intractable in games with more than two players (Hansen et al., 2010). In fact, even checking uncorrelated rational dominance is coNP-hard.

4 Under fairly general conditions, D-solutions obtained via iterated D-dominance are *maximal* (Duggan and Le Breton, 1996b). The maximal S^*-solution of a two-player game, for instance, consists of all *rationalizable* actions.

	a_1	a_2	a_3
a_1	0	1	0
a_2	-1	0	1
a_3	0	-1	0

Figure 12: Symmetric matrix game without W- and W^*-solutions

nally stable. This is again due to the fact that the iterative elimination of dominated actions preserves external stability.

Proposition 3.4. *Let* D *be a dominance structure satisfying* MON *or weak* MON*. Then, a set* $X \subseteq A_N$ *is a* D*-set if and only if it is a minimal externally* D*-stable set.*

Our proofs will frequently exploit this equivalence of D-sets and minimal externally D-stable sets. In particular, we will make use of the following easy corollary.

Corollary 3.5. *Let* D *be a dominance structure satisfying* MON *or weak* MON *and let* $X \subseteq A_N$ *be externally* D*-stable. Then, there exists a* D*-set* Y *with* $Y \subseteq X$*.*

3.3 UNIQUENESS RESULTS FOR D-SETS

Resoluteness is often considered a desirable property for a solution concept. It is however unrealistic to expect unique solutions for games that are not zero-sum. For example, coordination games like the one in Figure 13 have multiple Nash equilibria in pure strategies, all of which should be selected by any reasonable solution concept—at least as long as they do not Pareto-dominate each other. In this section, we present uniqueness proofs for D-sets in (subclasses of) matrix games. *coordination games*

	b_1	b_2
a_1	$(1,1)$	$(0,0)$
a_2	$(0,0)$	$(1,1)$

Figure 13: A simple coordination game with two pure Nash equilibria

We start by showing that the number of S-sets, S^*-sets, and B-sets is polynomially bounded by the size of the game. This follows from a more general statement about the intersection of externally stable sets: for monotonic dominance structures satisfying maximal domination, these sets are closed under intersection.[5]

5 $X \cap Y$ is to be read componentwise. Hence, $X \cap Y \neq \emptyset$ if and only if $X_i \cap Y_i \neq \emptyset$ for all $i \in N$.

Proposition 3.6. *Let* D *be a dominance structure satisfying MON and MAX. If* X *and* Y *are externally* D*-stable and* $X \cap Y \neq \emptyset$*, then* $X \cap Y$ *is externally* D*-stable.*

Proof. Suppose that X and Y are externally D-stable and $X \cap Y \neq \emptyset$. In order to show that $X \cap Y$ is externally D-stable, fix $i \in N$ and consider $a_i \in A_i \setminus (X_i \cap Y_i)$. Without loss of generality, assume that $a_i \notin X_i$. As X is an externally D-stable, $X_i\ D(X_{-i})\ a_i$, and thus $A_i\ D(X_{-i})\ a_i$. Now MON implies $A_i\ \ D(X_{-i} \cap Y_{-i})\ \ a_i$. Since $a_i \in A_i \setminus (X_i \cap Y_i)$ was chosen arbitrarily, $\max(D(X_{-i} \cap Y_{-i})) \subseteq X_i \cap Y_i$. Moreover, maximal domination implies $\max(D(X_{-i} \cap Y_{-i}))\ D(X_{-i} \cap Y_{-i})\ a_i$, which finally yields $(X_i \cap Y_i)\ D(X_{-i} \cap Y_{-i})\ a_i$. □

A direct corollary is the fact that any two distinct D-sets are disjoint.

Corollary 3.7. *Let* D *be a dominance structure satisfying MON and MAX. If* X *and* Y *are distinct* D*-sets, then* $X \cap Y = \emptyset$*.*

Proof. Assume for contradiction that $X \cap Y \neq \emptyset$. Then Proposition 3.6 implies that $X \cap Y$ is externally D-stable. It follows from Corollary 3.5 that there exists a D-set Z with $Z \subseteq X \cap Y$, violating the minimality of X or Y (or both). □

It is shown in Section 4.3 that S, S^*, and B satisfy MON and MAX. As a consequence, the number of S-sets, S^*-sets, and B-sets can never exceed the number of action profiles of the game.

Corollary 3.8. *The number of* S*-sets,* S^**-sets, and* B*-sets of a normal-form game is bounded by a polynomial in the size of the game.*

Proposition 3.6 can be utilized to prove uniqueness results for matrix games. This was first observed by Shapley (1964), who has shown that every matrix game has a unique S-set. The result was later generalized by Duggan and Le Breton (1996a) to S^*-sets and B-sets.

Theorem 3.9 (Shapley, 1964; Duggan and Le Breton, 1996a)**.** *Every matrix game has a unique* S*-set, a unique* S^**-set, and a unique* B*-set.*

Proof. We first prove uniqueness of S^*-sets. Let Γ be a matrix game and s^* a quasi-strict equilibrium of Γ. Consider an arbitrary S^*-set $X = (X_1, X_2)$ and let s be a Nash equilibrium of the subgame induced by X.

We claim that s is a Nash equilibrium of Γ. To see this, observe that X is a R_c-set (Fact 3.3). This implies that the set of all best responses to s_2 is contained in $\Delta(X_1)$. Since both s and s^* are Nash equilibria of Γ, we can apply interchangeability (Fact 2.9(*ii*)) and get that $(s_1^*, s_2) \in N(\Gamma)$. In particular, s_1^* is a best response to s_2. We thus have $supp(s_1^*) \subseteq X_1$, and an analogous statement shows $supp(s_2^*) \subseteq X_2$.

Since $supp(s^*) = ES(\Gamma)$ and X was chosen arbitrarily, we have shown that every S^*-set contains $ES(\Gamma)$. In other words, any two S^*-sets have a nonempty intersection. The uniqueness of S^*-sets now follows from Corollary 3.7.

We now show that S-sets and B-sets are also unique in matrix games. Recall that $S \subseteq B \subseteq S^*$. Assume for contradiction that there exists a matrix game with two distinct S-sets X and Y. Since $S \subseteq S^*$, X and Y are externally S^*-stable. Corollary 3.5 implies the existence of two disjoint S^*-sets $X' \subseteq X$ and $Y' \subseteq Y$, contradicting the uniqueness of S^*-sets. The argument for B-sets is analogous. □

In tournament games, S^*-sets and B-sets coincide.

Proposition 3.10. *For every tournament game, the unique S^*-set coincides with the unique B-set.*

Proof. First, observe that uniqueness of a D-set X in a symmetric matrix game implies symmetry of X, i.e., $X_1 = X_2$. (If $X_1 \neq X_2$, (X_2, X_1) would be another D-set.)

uniqueness implies symmetry

Second, Proposition 3.5 reduces the problem of showing the coincidence of S^*-sets and B-sets to the problem of showing the coincidence of externally S^*-stable sets and externally B-stable sets.

Combining these two insights, it is sufficient to prove the following statement for any tournament game $\Gamma = (A, u)$ and any $X \subseteq A$:

$$(X, X) \text{ is externally } S^*\text{-stable} \Leftrightarrow (X, X) \text{ is externally } B\text{-stable.}$$

The direction from right to left is trivial since $B \subseteq S^*$. For the direction from left to right, consider an action $b \in A \setminus X$ with $X \; S^*(X) \; b$. We need to show that $X \; B(X) \; b$. Since Γ is a tournament game and $b \notin X$, $u(b, x) \neq 0$ for all $x \in X$. Moreover, $u(b, x)$ cannot equal 1 since $X \; S^*(X) \; b$ and $u(x, y) \leqslant 1$ for all $x, y \in A$. It follows that $u(b, x) = -1$ for all $x \in X$.

It is now easy to show that $X \; B(X) \; b$, as for all nonempty $Y \subseteq X$ and all $y \in Y$, $y \; W(Y) \; b$. The reason is that the strict inequality $u(y, y) > u(b, y)$ is sufficient for y weakly dominating b with respect to Y. □

The game on the left-hand side of Figure 11 shows that Proposition 3.10 cannot be generalized to confrontation games.

We now turn to symmetric dominance structures and prove a uniqueness result that applies to a whole class of such structures.

Proposition 3.11. *Let D be a symmetric dominance structure satisfying weak MON. If $D \subseteq C_M$, then every symmetric matrix game has a unique D-set.*

It is unknown whether $D \subseteq C_M$ is *necessary* for the uniqueness of D-sets in symmetric matrix games. Various symmetric dominance structures D with $C_M \subseteq D$ have been shown to admit disjoint minimal solutions, sometimes involving rather elaborate combinatorial arguments. Examples include *unidirectional covering* (Brandt and Fischer, 2008b) and *extending* (Brandt, 2011b; Brandt et al., 2013).

In order to prove Proposition 3.11, we need the following two lemmata.[6]

Lemma 3.12. *Let $D \subseteq C_M$ be a symmetric dominance structure satisfying weak MON. Let furthermore $\Gamma = (A, u)$ be a symmetric matrix game and $X, Y \subseteq A$. If (X, X) and (Y, Y) are externally D-stable, then $(X \cap Y, X \cap Y)$ is externally D-stable.*

Proof. We first show that $X \cap Y \neq \emptyset$. Assume for contradiction that $X \cap Y = \emptyset$ and let $a_0 \in Y$. As (X, X) is externally D-stable and $D \subseteq C_M$, there exists $a_1 \in X$ with $a_1 \; C_M(X) \; a_0$. As (Y, Y) is externally D-stable and $D \subseteq C_M$, there exists $a_2 \in Y$ with $a_2 \; C_M(X) \; a_1$. Repeatedly applying these arguments yields an infinite sequence $(a_0, a_1, a_2, ...)$ such that

- for all *even* $i \geqslant 0$, $a_i \in Y$ and $a_{i+1} \; C_M(X) \; a_i$, and
- for all *odd* $i \geqslant 1$, $a_i \in X$ and $a_{i+1} \; C_M(Y) \; a_i$.

Since A is finite, this sequence must contain repetitions. Without loss of generality, assume that $a_k = a_0$ for some even $k > 0$. We can therefore construct the following chain of inequalities (consult Figure 14 for an example with $k = 6$):

$$\begin{aligned} u(a_0, a_1) &\leqslant u(a_1, a_1) \leqslant u(a_1, a_0) \leqslant u(a_2, a_0) \leqslant u(a_2, a_{k-1}) \\ &\leqslant u(a_3, a_{k-1}) \leqslant u(a_3, a_{k-2}) \leqslant u(a_4, a_{k-2}) \leqslant \dots \\ &\leqslant u(a_0, a_2) \leqslant u(a_0, a_1) \end{aligned} \quad (3)$$

It follows that all utilities in this chain of inequalities are equal. Since Γ is a symmetric matrix game, $u(a_1, a_1) = 0$. Hence, all utilities in this chain are zero. In particular, $u(a_1, a_0) = 0$, which contradicts the assumption that $a_1 \; C_M(Y) \; a_0$. This proves that $X \cap Y \neq \emptyset$.

In order to show that $(X \cap Y, X \cap Y)$ is externally D-stable, take an arbitrary $a_0 \notin X \cap Y$. Without loss of generality, assume that $a_0 \notin X$. As (X, X) is externally D-stable and $D \subseteq C_M$, there exists $a_1 \in X$ with $a_1 \; C_M(X) \; a_0$. If $a_1 \notin Y$, there exists $a_2 \in Y$ with $a_2 \; C_M(X) \; a_1$. This construction finally yields some $a_k \in X \cap Y$, for otherwise we have a contradiction as in the first part of the proof. Repeated application of weak MON now yields $a_k \; D(X \cap Y) \; a_i$ for all $i < k$. In particular, $a_k \; D(X \cap Y) \; a_0$, as desired. □

The proof of the following lemma is similar to that of Lemma 3.12, and we omit it.

Lemma 3.13. *Under the assumptions of Lemma 3.12, if (X, Y) is externally D-stable, then $(X \cap Y, X \cap Y)$ is externally D-stable.*

We are now ready to prove Proposition 3.11.

6 These lemmata are adapted from Duggan and Le Breton (1996a), who proved the analogous statements for *W*-sets in confrontation games.

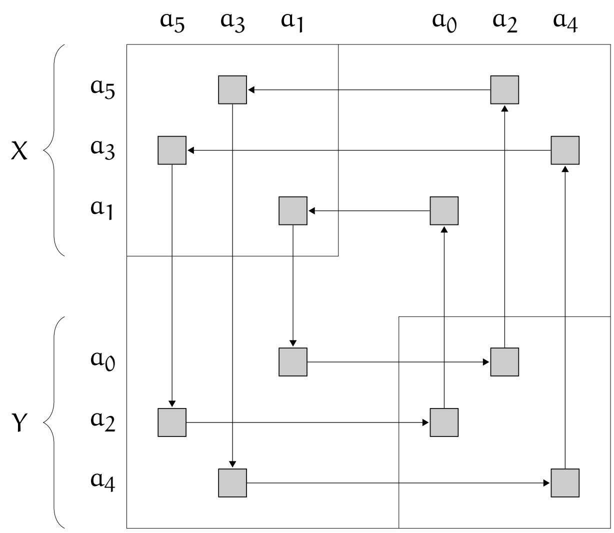


Figure 14: Construction used in the proof of Lemma 3.12. An arrow from an action profile (a_i, a_j) to another action profile $(a_{i'}, a_{j'})$ indicates that $u(a_i, a_j) \geqslant u(a_{i'}, a_{j'})$.

Proof of Proposition 3.11. Lemma 3.13 implies that every D-set (X, Y) satisfies $X = Y$, as otherwise $(X \cap Y, X \cap Y)$ would be a smaller externally D-stable set. Similarly, Lemma 3.12 implies that there cannot exist two D-sets (X, X) and (Y, Y) with $X \neq Y$. □

Since $C_D \subseteq C_M$ and both C_D and C_M are weakly monotonic, it immediately follows that C_D and C_M satisfy uniqueness.

Corollary 3.14. *Every symmetric matrix game has a unique C_D-set and a unique C_M-set.*

We now show how the results on symmetric dominance structures can be utilized for studying V-sets. In contrast to S-sets, V-sets are not unique in matrix games.[7] In confrontation games, the picture is different. Duggan and Le Breton (1996a) have shown that these games have a unique W-set, which moreover coincides with the (unique) C_M-set. As already mentioned in Footnote 6, the proof of the latter statement proceeds along the same lines as the one of Proposition 3.11. In particular, the chain of weak inequalities (3) is used to show that $u(a_1, a_0) = 0$, which violates the off-diagonal property of confrontation games. It can easily be checked that all arguments carry over to very weak dominance. As a consequence, V-sets, W-sets, and C_M-sets all coincide in confrontation games.

7 It is in fact easily seen that even a symmetric matrix game can have multiple V-sets: if all action profiles yield the same utility, then every single profile constitutes a V-set. Moreover, we will in Section 3.4 construct a class of symmetric matrix games with an exponential number of V-sets. And we will show in Section 5.7 that several natural problems like *finding* V-sets, checking whether a given action is contained in *some* V-set, or deciding whether there is a unique V-set are computationally intractable.

Proposition 3.15. *Every confrontation game has a unique* V*-set, and this* V*-set coincides with the unique* C_M*-set.*

Our final result on confrontation games states that the unique C_D-set coincides with the C_M-set as well.

Proposition 3.16. *For every confrontation game, the unique* C_D*-set coincides with the unique* C_M*-set.*

Proof. Following a similar argument as in the proof of Proposition 3.10, it is sufficient to prove the following statement for any confrontation game $\Gamma = (A, u)$ and $X \subsetneq A$:

$$(X, X) \text{ is externally } C_D\text{-stable} \Leftrightarrow (X, X) \text{ is externally } C_M\text{-stable.}$$

This equivalence is easily verified by inspecting the definitions of C_M and C_D (page 21): assuming $X = Y$ and $a \notin X$, the dominance relations $C_M(X)$ and $C_D(X)$ only differ in games in which some off-diagonal matrix entry equals 0. Obviously, this cannot happen in confrontation games. □

V^*-sets, on the other hand, are not even unique in confrontation games (see Figure 15 on page 39 for an example). In order to guarantee the uniqueness of V^*-sets, we have to restrict the class of games even further. Again, we utilize a result by Duggan and Le Breton (2001), who have shown that tournament games have a unique W^*-set, which moreover coincides with the W-set (and hence with the V-set, the C_M-set, and the C_D-set). The proof is similar to the one of Proposition 3.10 and shows that a pair (X_1, X_2) is externally W^*-stable if and only if it is externally W-stable. Observing that all the arguments can be adapted to very weak dominance yields the final result of this section.

Proposition 3.17. *Every tournament game has a unique* V^**-set, and this* V^**-set coincides with the unique* V*-set.*

3.4 GAMES WITH AN EXPONENTIAL NUMBER OF D-SETS

In this section we show that there are games with an exponential number of V-sets and V^*-sets. We need the following lemma.

Lemma 3.18. *Let* $\Gamma = (A, u)$ *be a symmetric matrix game and* $D \in \{V, V^*\}$*. Define* Γ' *as the matrix game that is identical to* Γ *except that player* 1 *has an additional action* $\hat{a}$ *that always yields a utility of* 1*. That is,* $\Gamma' = (\{1,2\}, (A \cup \{\hat{a}\}, A), (u_1, u_2))$ *with* $u_1(a, b) = u(a, b)$ *for all* $a, b \in A$*,* $u_1(\hat{a}, a) = 1$ *for all* $a \in A$*, and* $u_2 = -u_1$*. Then, there exists no subset* $X \subseteq A$ *of actions of player* 1 *such that* $X\ D(X)\ \hat{a}$*.*

Proof. Assume for contradiction that $X\ V^*(X)\ \hat{a}$ for some $X \subseteq A$, and let $s_1 \in \Delta(X)$ be a strategy with $u_1(s_1, x_{-i}) \geqslant 1$ for all $x_{-i} \in X$. In the subgame $\Gamma|_X$, the strategy s_1 guarantees a utility of at least 1 to the row player. This contradicts the fact that $\Gamma|_X$, being a symmetric zero-sum game, has a value of 0. Since $V \subseteq V^*$, the statement for V follows immediately. □

A D-set (X_1, X_2) is *symmetric* if $X_1 = X_2$. It is straightforward to verify that every symmetric matrix game has a symmetric D-set for all dominance structures D. For $D \in \{V, V^*\}$, any symmetric matrix game with multiple symmetric D-sets can be used to show that the number of D-sets may be exponential in general.

symmetric D-sets

Proposition 3.19. *Let* $D \in \{V, V^*\}$. *If there exists a symmetric matrix game with at least two symmetric* D*-sets, then there exists a family of symmetric matrix games such that the number of* D*-sets is exponential in the size of the game.*

Proof. Let $D \in \{V, V^*\}$ and consider a symmetric matrix game $\Gamma = (A, u)$ with $k \geqslant 2$ symmetric D-sets. We construct a family $(\Gamma_i)_{i \in \mathbb{N}}$ such that $\Gamma_i = (A^i, u^i)$ is a symmetric matrix game and the number of symmetric D-sets in Γ_i is $k^{|A^i|/|A|}$.

Let $\Gamma_1 = \Gamma$. For $i \geqslant 1$, $\Gamma_{i+1} = (A^{i+1}, u^{i+1})$ is defined inductively as follows.

$$A^{i+1} = A^{i,0} \cup A^{i,1} \cup A^{i,2},$$

where for each $\ell \in \{0, 1, 2\}$, $A^{i,\ell}$ is a copy of A^i. For $a \in A^{i,\ell}$ and $b \in A^{i,\ell'}$, the utility function u^{i+1} is defined by

$$u^{i+1}(a, b) = \begin{cases} u^i(a, b) & \text{if } \ell = \ell', \\ -1 & \text{if } \ell' = \ell + 1, \\ 1 & \text{if } \ell' = \ell + 2, \end{cases}$$

where $\ell + c$ should be understood to mean $\ell + c \mod 3$. If M_i is the matrix representing Γ_i, $\mathbb{1}$ is the $|A^i| \times |A^i|$ matrix containing only ones, and $-\mathbb{1}$ is $(-1) \cdot \mathbb{1}$, then the game Γ_{i+1} is represented by the block matrix

$$M_{i+1} = \begin{pmatrix} M_i & -\mathbb{1} & \mathbb{1} \\ \mathbb{1} & M_i & -\mathbb{1} \\ -\mathbb{1} & \mathbb{1} & M_i \end{pmatrix}.$$

We will show that, for each $i \geqslant 1$, the symmetric D-sets of Γ_{i+1} can be characterized in terms of the symmetric D-sets of Γ_i. The following notation will be useful. For $X \subseteq A^{i+1}$ and $\ell \in \{0, 1, 2\}$, let $X_\ell = X \cap A^{i,\ell}$ denote the part of X that lies in block ℓ. We claim that for each $i \geqslant 1$,

$$\begin{aligned} &(X, X) \text{ is a symmetric D-set in } \Gamma_{i+1} \quad \text{if and only if} \\ &(X_\ell, X_\ell) \text{ is a symmetric D-set in } \Gamma_i \text{ for all } \ell \in \{0, 1, 2\}. \end{aligned} \tag{4}$$

Before proving this equivalence, we make the following observation.

$$\text{If } (X,X) \text{ is a D-set in } \Gamma_{i+1}, \text{ then } X_\ell \neq \emptyset \text{ for all } \ell \in \{0,1,2\}. \tag{5}$$

To see this, let $x \in X$ be an arbitrary action and choose ℓ such that $a \in X_\ell$. Consider the game where the actions of player 2 are restricted to X_ℓ. As $u^{i+1}(a,b) = 1$ for all $a \in X_{\ell+1}$ and $b \in X_\ell$, Lemma 3.18 implies that no action in $X_{\ell+1}$ is D-dominated by X_ℓ. Therefore, at least one of the actions in $X_{\ell+1}$ has to be contained in X, i.e., $X_{\ell+1} \neq \emptyset$. Repeating the argument, $X_{\ell+1} \neq \emptyset$ implies $X_{\ell+2} \neq \emptyset$, which proves (5).

We are now ready to prove the equivalence (4). For the direction from left to right, assume that (X,X) is a D-set in Γ_{i+1} and let $\ell \in \{0,1,2\}$. We need to show that (X_ℓ, X_ℓ) is a D-set in Γ_i. By (5), we know that $X_\ell \neq \emptyset$. To show that (X_ℓ, X_ℓ) is externally D-stable, consider some $a \in A^{i,\ell} \setminus X_\ell$. As (X,X) is externally D-stable in Γ_{i+1}, $X\ D(X)\ a$. However, the definition of u^{i+1} ensures that none of the actions in $X_{\ell+1} \cup X_{\ell+2}$ is essential for this dominance relation to hold. It thus follows that $X_\ell\ D(X)\ a$. Monotonicity of D finally yields $X_\ell\ D(X_\ell)\ a$. For minimality of (X_ℓ, X_ℓ), note that the existence of an externally D-stable set $(X', Y') \neq (X_\ell, X_\ell)$ in Γ_i with $X', Y' \subseteq X_\ell$ would contradict the minimality of (X,X) in Γ_{i+1}.

For the direction from right to left, (X,X) is externally D-stable in Γ_{i+1} because each (X_ℓ, X_ℓ) is externally D-stable in Γ_i. Furthermore (X,X) is minimal, as a proper subset of (X,X) that is externally D-stable in Γ_{i+1} would yield an externally D-stable subset of (X_ℓ, X_ℓ) for some $\ell \in \{0,1,2\}$, contradicting the minimality of (X_ℓ, X_ℓ) in Γ_i.

Let k_i denote the number of symmetric D-sets in Γ_i. It follows from (4) that $k_{i+1} = k_i^3$ for all $i \geqslant 1$. As $k_1 = k$, this yields $k_i = k^{3^{i-1}}$. As $|A^i| = 3^{i-1}|A|$, the number of D-sets in $\Gamma_i = (A^i, u^i)$ equals $k_i = k^{|A^i|/|A|}$. In particular, k_i is exponential in $|A^i|$. □

We will now use Proposition 3.19 to show that games can have an exponential number of V-sets and V*-sets. All we have to do is construct symmetric matrix games with more than one solution, and this is easily achieved. For instance, take a 2×2 game where all utilities are zero. It follows that the number of V-sets and V*-sets may be exponential in the size of the game. An immediate consequence is that no polynomial-time algorithm can compute all of these sets.

Corollary 3.20. *Computing all V-sets or all V*-sets of a game requires exponential time in the worst case, even for symmetric matrix games.*

For mixed very weak dominance, the result can be strengthened. The game in Figure 15 proves that, unlike V-sets, V*-sets are *not* unique in confrontation games. Applying Proposition 3.19 again, we get the following.[8]

8 It is easily seen that the games $(\Gamma_i)_{i \in \mathbb{N}}$ constructed in the proof of Proposition 3.19 are confrontation games whenever Γ is a confrontation game.

	a_1	a_2	a_3	a_4	a_5	a_6
a_1	0	3	−3	−1	−1	2
a_2	−3	0	3	−1	2	−1
a_3	3	−3	0	2	−1	−1
a_4	1	1	−2	0	3	−3
a_5	1	−2	1	−3	0	3
a_6	−2	1	1	3	−3	0

Figure 15: A confrontation game with two symmetric V^*-sets: $(\{a_1, a_2, a_3\}, \{a_1, a_2, a_3\})$ and $(\{a_4, a_5, a_6\}, \{a_4, a_5, a_6\})$. In the former V^*-set, a_4 is dominated by $\frac{2}{3}a_1 + \frac{1}{3}a_3$, a_5 is dominated by $\frac{1}{3}a_2 + \frac{2}{3}a_3$, and a_6 is dominated by $\frac{1}{3}a_1 + \frac{2}{3}a_2$. In the latter V^*-set, a_1 is dominated by $\frac{2}{3}a_5 + \frac{1}{3}a_6$, a_2 is dominated by $\frac{2}{3}a_4 + \frac{1}{3}a_5$, and a_3 is dominated by $\frac{1}{3}a_4 + \frac{2}{3}a_6$.

Corollary 3.21. *Computing all V^*-sets of a game requires exponential time in the worst case, even for confrontation games.*

3.5 SUMMARY

We have studied the existence and multiplicity of D-set for various dominance structures D. With the exception of W and W^*, all dominance structures yield solutions that are guaranteed to exist. Table 3 summarizes the results on the multiplicity of solutions. For a given dominance structure D and a class of games (ordered by set inclusion), the table shows bounds on the *asymptotic number* of D-sets (unique, polynomial, or exponential). If a cell spans several columns, the corresponding D-sets coincide within the respective class of games. For each table cell that is *not* labelled with 'exp,' the next chapter will provide an algorithm that computes all D-sets for games in the respective class.

<table>
<tr><th></th><th>S</th><th>B</th><th>S^*</th><th>C_D</th><th>C_M</th><th>V</th><th>V^*</th></tr>
<tr><td>normal-form games</td><td>poly</td><td>poly</td><td>poly</td><td colspan="2" rowspan="2"></td><td rowspan="3">exp</td><td rowspan="4">exp</td></tr>
<tr><td>matrix games</td><td rowspan="4">unique</td><td rowspan="3">unique</td><td rowspan="3">unique</td></tr>
<tr><td>symmetric matrix games</td><td>unique</td><td>unique</td></tr>
<tr><td>confrontation games</td><td colspan="3">unique</td></tr>
<tr><td>tournament games</td><td colspan="2">unique</td><td colspan="4">unique</td></tr>
</table>

Table 3: The multiplicity of D-sets

4

ALGORITHMS FOR DOMINANCE-BASED SOLUTIONS

The goal of this chapter is to efficiently compute D-sets for different dominance structures D. We introduce two generic algorithms, a *greedy* and a *sophisticated* one. In principle, these algorithms can be applied to any dominance structure. For each algorithm, we identify properties of a dominance structure that ensure soundness and efficiency of the algorithm.

4.1 GENERIC GREEDY ALGORITHM

Shapley (1964) did not only show that every matrix game possesses a unique S-set, he also described an algorithm—attributed to Harlan Mills—to compute this set. The idea behind this algorithm is that given a subset of the S-set, the S-set itself can be computed by iteratively adding actions that are maximal, i.e., not dominated with respect to the current subset of actions of the other player. We generalize Mills' algorithm in two directions. First, we identify general conditions on a dominance structure D that ensure that this greedy approach works. Second, we consider arbitrary n-player normal-form games, thereby losing uniqueness of D-sets, and devise an algorithm that computes *all* D-sets of such games in polynomial time. We first observe a particularly useful consequence of Proposition 3.6 (page 32).

Mills' algorithm

Corollary 4.1. *Let $X^0 \subseteq A_N$. Under the assumptions of Proposition 3.6, the minimal externally D-stable set containing X^0 is unique: if Y and Z are externally D-stable with $X^0 \subseteq Y$ and $X^0 \subseteq Z$, then $Y \subseteq Z$ or $Z \subseteq Y$.*

Proof. Suppose that both Y and Z are minimal among all externally D-stable sets containing X^0, and that neither $Y \subseteq Z$ nor $Z \subseteq Y$. As both Y and Z contain X^0, $Y \cap Z$ is nonempty and Proposition 3.6 implies that $Y \cap Z$ is externally D-stable. This contradicts minimality of both Y and Z. □

If D moreover satisfies computational tractability, the minimal externally D-stable set containing X^0 can be computed efficiently by greedily adding undominated actions.

Proposition 4.2. *Let $X^0 \subseteq A_N$. If D satisfies MON, MAX, and COM, the minimal externally D-stable set containing X^0 can be computed in polynomial time.*

Algorithm 1 Minimal externally D-stable set containing X^0

procedure `min_ext`$(\Gamma, (X_1^0, \dots, X_n^0))$
  for all $i \in N$ **do**
    $X_i \leftarrow X_i^0$
  end for
  repeat
    for all $i \in N$ **do**
      $Y_i \leftarrow \max(D(X_{-i})) \setminus X_i$
      $X_i \leftarrow X_i \cup Y_i$
    end for
  until $\bigcup_{i=0}^{n} Y_i = \emptyset$
  return $(X_1, \dots, X_n)$

Algorithm 2 Generic greedy algorithm

procedure `D_set`(Γ)
  $C \leftarrow \emptyset$
  for all $(a_1, \dots, a_n) \in \prod_{i \in N} A_i$ **do**
    $C \leftarrow C \cup$ `min_ext`$(\Gamma, (\{a_1\}, \dots, \{a_n\}))$
  end for
  return $\{X \in C : X \text{ is inclusion-minimal in } C\}$

Proof. We show that Algorithm 1 computes the minimal externally D-stable set containing X^0 and runs in polynomial time. Algorithm 1 starts with X^0 and iteratively adds all actions that are not yet dominated. As D satisfies COM, these actions can be computed efficiently. Moreover, the number of loops is bounded by $\sum_{i=1}^{n} |A_i|$.

Let X^{min} be the minimal externally D-stable set containing X^0. We show that during the execution of Algorithm 1, the set X is always a subset of X^{min}. At the end of the algorithm, $\bigcup_{i=0}^{n} Y_i = \emptyset$ implies that $\max(D(X_{-i})) \subseteq X_i$ for all $i \in N$. As D satisfies MAX, this shows that X is externally D-stable.

We prove $X \subseteq X^{min}$ by induction on $|X| = \sum_{i=1}^{n} |X_i|$. At the beginning of the algorithm, $X = X^0 \subseteq X^{min}$ by definition of X^{min}. Now assume that $X \subseteq X^{min}$ at the beginning of a particular iteration. We have to show that for all $i \in N$, $Y_i \subseteq X_i^{min}$. Let $a_i \in Y_i = \max(D(X_{-i})) \setminus X_i$, and assume for contradiction that $a_i \notin X_i^{min}$. Since X^{min} is externally D-stable, X_i^{min} $D(X_{-i}^{min})$ a_i. By the induction hypothesis, $X_{-i} \subseteq X_{-i}^{min}$, which together with MON implies X_i^{min} $D(X_{-i})$ a_i. It follows that A_i $D(X_{-i})$ a_i, contradicting the assumption that $a \in \max(D(X_{-i}))$. □

Whenever X^0 is contained in a D-set, Algorithm 1 returns this D-set. This property can be used to construct an algorithm to compute all D-sets of a game: call Algorithm 1 for every possible combination of singleton sets of actions of the different players. The result is a collection of externally D-stable sets, and the D-sets of the game are

	a_1	a_2	a_3	a_4	a_5	a_6
a_1	0	1	-1	1	1	-1
a_2	-1	0	1	1	-1	1
a_3	1	-1	0	-1	1	1
a_4	-1	-1	1	0	1	-1
a_5	-1	1	-1	-1	0	1
a_6	1	-1	-1	1	-1	0

Figure 16: Tournament game with unique V-set $(\{a_1, a_2, a_3\}, \{a_1, a_2, a_3\})$. Initiating Algorithm 1 with a pair $(\{a_i\}, \{a_j\})$ results in a proper superset of the V-set for every choice of i and j.

the inclusion-minimal elements of this collection. This idea is made precise in Algorithm 2.

Theorem 4.3. *If* D *satisfies MON, MAX, and COM, all* D*-sets can be computed in polynomial time.*

Proof. We show that Algorithm 2 computes all D-sets of Γ and runs in polynomial time. Polynomial running time follows immediately because Algorithm 1 is invoked $|A|$ times, and inclusion-maximality can be checked easily.

As for correctness, we first show that every D-set X is an element of the set C. To see this, note that Proposition 3.6 implies that X is the minimal externally D-stable set containing $(\{x_1\}, \dots, \{x_n\})$ for every $(x_1, \dots, x_n) \in \prod_{i=1}^{n} X_i$. By definition, each D-set is inclusion-minimal.

To show that all inclusion-minimal elements of C are D-sets, it is sufficient to observe that all elements of C are externally D-stable. $\square$

4.2 GENERIC SOPHISTICATED ALGORITHM

Algorithm 2 is not sound for all dominance structures. For instance, very weak dominance violates maximal domination and therefore does not satisfy the conditions of Theorem 4.3. The example given in Figure 16 shows that, even in tournament games, Algorithm 2 fails to find the unique V-set.

We will identify conditions on dominance structures D that allow for the following sophisticated method: instead of adding all D-undominated actions, merely add actions contained in a D-set of the *subgame induced by the* D*-undominated actions*. This immediately gives rise to a recursive algorithm, whose running time may however be exponential. If a nonempty subset of a D-set can be found efficiently, an efficient algorithm can be constructed.

Algorithm 3 Generic sophisticated algorithm

procedure `D_set_symm`(Γ)
  $X \leftarrow$ subset of $\mathcal{S}_D(\Gamma)$
  repeat
    $\Gamma_X \leftarrow \Gamma|_{\max(D(X))\setminus X}$
    $X' \leftarrow$ subset of $\mathcal{S}_D(\Gamma_X)$
    $X \leftarrow X \cup X'$
  until $\max(D(X)) \setminus X = \emptyset$
  return (X, X)

In this section, we will only be concerned with symmetric dominance structures D satisfying uniqueness. As we have already observed on page 33, uniqueness of D-sets implies that the D-set is symmetric, i.e., $X_1 = X_2$. In this case, the set $X_1 = X_2$ will be denoted by $\mathcal{S}_D(\Gamma)$. The following lemma is the key ingredient for the sophisticated algorithm.

Lemma 4.4. *Let Γ be a symmetric matrix game and D a symmetric dominance structure satisfying weak MON, TRA, SING, and UNI. Let furthermore $X \subseteq \mathcal{S}_D(\Gamma)$ and define $\Gamma_X = \Gamma|_{\max(D(X))\setminus X}$. Then, $\mathcal{S}_D(\Gamma_X) \subseteq \mathcal{S}_D(\Gamma)$.*

Proof. Let $A' = \max(D(X)) \setminus X$. We can assume that A' is nonempty, as otherwise $\mathcal{S}_D(\Gamma_X) = \mathcal{S}_D(\Gamma|_{A'})$ is empty and there is nothing to prove.

Now partition the set A' of undominated actions into two sets $C = A' \cap \mathcal{S}_D(\Gamma)$ and $C' = A' \setminus \mathcal{S}_D(\Gamma)$ of actions contained in $\mathcal{S}_D(\Gamma)$ and actions not contained in $\mathcal{S}_D(\Gamma)$. We will show that (C, C) is externally D-stable in Γ_X. Then, Corollary 3.5 and UNI imply that $\mathcal{S}_D(\Gamma_X) \subseteq C$ and, therefore, $\mathcal{S}_D(\Gamma_X) \subseteq \mathcal{S}_D(\Gamma)$.

In order to show that (C, C) is externally D-stable in Γ_X, consider some $z \in C'$. Since $z \notin \mathcal{S}_D(\Gamma)$, singularity of D implies the existence of a $y \in \mathcal{S}_D(\Gamma)$ with $y\ D(\mathcal{S}_D(\Gamma))\ z$. It is easy to see that $y \notin X$, since otherwise weak MON would imply that $y\ D(X)\ z$, contradicting the assumption that $z \in A'$. On the other hand, assume that $y \in \mathcal{S}_D(\Gamma) \setminus (X \cup C)$. Then there is some $x \in X$ such that $x\ D(X)\ y$. However, according to TRA, $x\ D(X)\ y$ and $y\ D(\mathcal{S}_D(\Gamma))\ z$ imply $x\ D(X)\ z$, again contradicting the assumption that $z \in A'$. Thus $y \in C$, and using weak MON again, $y\ D(\mathcal{S}_D(\Gamma))\ z$ and $z \in A'$ imply $y\ D(A')\ z$. Hence C is externally D-stable in Γ_X. □

Two further properties are required to turn the insight of Lemma 4.4 into an efficient algorithm: first, we need a polynomial-time subroutine to compute a nonempty subset of the unique D-set; second, the dominance structure D itself must be computationally tractable.

Theorem 4.5. *If D satisfies weak MON, TRA, SING, UNI, SUB-COM, and COM, the D-set of a symmetric matrix game can be computed in polynomial time.*

Proof. We show that Algorithm 3 computes $S_D(\Gamma)$ and runs in polynomial time. In each iteration, at least one action is added to the set X, so the algorithm is guaranteed to terminate after at most $|A|$ iterations. Each iteration consists of *(i)* computing the set $\max(D(X)) \setminus X$ of undominated actions and *(ii)* finding a subset of $S_D(\Gamma_X)$. Since D satisfies COM and SUB-COM, both tasks can be performed in polynomial time.

As for correctness, we show by induction on the number of iterations that $X \subseteq S_D(\Gamma)$ holds at any time. When the algorithm terminates, X is externally D-stable, which together with the induction hypothesis implies that $X = S_D(\Gamma)$.

The base case is trivial. Now assume that $X \subseteq S_D(\Gamma)$ at the beginning of a particular iteration. Then $X \cup X' \subseteq X \cup S_D(\Gamma_X) \subseteq S_D(\Gamma)$, where the first inclusion is due to $X' \subseteq S_D(\Gamma_X)$ and the second inclusion follows from Lemma 4.4 and the induction hypothesis. □

4.3 GREEDY ALGORITHMS

In this section, we investigate the consequences of Theorem 4.3 on S-, B-, and S^*-sets.

Corollary 4.6. *All S-sets of a normal-form game can be computed in polynomial time.*

Proof. According to Theorem 4.3, it suffices to show that S satisfies MON, MAX, and COM. It can easily be verified that S satisfies MON and MAX. Furthermore, S satisfies COM because $x_i \; S(X_{-i}) \; a_i$ can be checked efficiently by simply comparing $u_i(x_i, x_{-i})$ and $u_i(a_i, x_{-i})$ for each $x_{-i} \in X_{-i}$. □

The same is true for Börgers dominance.

Corollary 4.7. *All* B*-sets of a normal-form game can be computed in polynomial time.*

Proof. According to Theorem 4.3, it suffices to show that B satisfies MON, MAX, and COM. As was the case for S, it can easily be checked that B satisfies MON and MAX.

It remains to be shown that B satisfies COM. Consider the following procedure, formalized in Algorithm 4, which checks whether $X_i \; B(X_{-i}) \; a_i$ holds. First check whether X_i weakly dominates a_i. If no, then X_i does not Börgers-dominate a_i either. If yes, we can find an action $x_i \in X_i$ with $u_i(x_i, x_{-i}) \geqslant u_i(a_i, x_{-i})$ for all $x_{-i} \in X_{-i}$. Define $C(x_i)$ as the set of all tuples $x_{-i} \in X_{-i}$ for which the latter inequality is strict. $C(x_i)$ is nonempty by definition of W. It follows that $x_i \; W(Y_{-i}) \; a_i$ for all Y_{-i} with $Y_{-i} \cap C(x_i) \neq \emptyset$. We can therefore restrict attention to subsets of $Y_{-i} \setminus C(x_i)$ and "recursively" check whether $X_i \; B(Y_{-i} \setminus C(x_i)) \; a_i$. It is easily verified that this procedure correctly checks Börgers dominance and runs in polynomial time. □

Algorithm 4 Checking Börgers dominance

procedure Boergers_dom($\Gamma, (X_i)_{i \in N}, a_i$)
  $Y_{-i} \leftarrow X_{-i}$
  repeat
    if $X_i\ W(Y_{-i})\ a_i$ **then**
      Choose $x_i \in X_i$ such that $x_i\ W(Y_{-i})\ a_i$
      $C(x_i) \leftarrow \{y_{-i} \in Y_{-i} : u_i(x_i, y_{-i}) > u_i(a_i, y_{-i})\}$
      $Y_{-i} \leftarrow Y_{-i} \setminus C(x_i)$
    else
      return "no"
    end if
  until $Y_{-i} = \emptyset$
  return "yes"

The requirements for the greedy algorithm are also met by mixed strict dominance.

Corollary 4.8. *All S^*-sets of a normal-form game can be computed in polynomial time.*

Proof. According to Theorem 4.3, it suffices to show that S^* satisfies MON, MAX, and COM. It is easily verified that S^* satisfies MON. Furthermore, S^* satisfies COM because $X_i\ S^*(X_{-i})\ a_i$ can be checked efficiently with the help of a linear program (see Proposition 1 by Conitzer and Sandholm, 2005).

We now show that S^* satisfies MAX. Without loss of generality, assume that $u_i(a) \geqslant 0$ for all $i \in N$ and $a \in \prod_{i=1}^n A_i$. The following geometric interpretation will be useful. For an action a_i of player $i \in N$, define $u_i(a_i, X_{-i}) = (u_i(a_i, x_{-i}))_{x_{-i} \in X_{-i}}$ as the vector of possible utilities for player i if he plays a_i and the other players play some $x_{-i} \in X_{-i}$. For a set $Y_i \subseteq A_i$ of actions of player i, denote by $u_i(Y_i, X_{-i}) = \cup_{y_i \in Y_i} u_i(y_i, X_{-i})$ the union of all such vectors, and write $m = |X_{-i}|$ for their dimension. For a set of vectors $V \subseteq \mathbb{R}^m_{\geqslant 0}$, define $L[V]$ to be the *lower contour set* of $conv(V)$, i.e.,

$$L[V] = \bigcup \{x \in \mathbb{R}^m_{\geqslant 0} :\ \exists v \in conv(V) \text{ with } v \geqslant x\},$$

where $v \geqslant x$ is to be read componentwise.

The underlying intuition is that each action whose vector of utilities lies in the *interior* of $L[V]$ is strictly dominated by some strategy in $\Delta(V)$. More formally, $X_i\ S^*(X_{-i})\ a_i$ if and only if $u_i(a_i, X_{-i}) \in int(L[u_i(X_i, X_{-i})])$.

Suppose $A_i\ S^*(X_{-i})\ a_i$. We have to show that

$$\max(S^*(X_{-i}))\ S^*(X_{-i})\ a_i\,,$$

i.e., there exists $s_i \in \Delta(\max(S^*(X_{-i})))$ with $u_i(s_i, x_{-i}) > u_i(a_i, x_{-i})$ for all $x_{-i} \in X_{-i}$. Since $A_i\ S^*(X_{-i})\ a_i$, we know that there must be some $s_i \in \Delta(A_i)$ with $u_i(s_i, x_{-i}) > u_i(a_i, x_{-i})$ for all $x_{-i} \in X_{-i}$. It thus suffices to show that

$$L[u_i(\max(S^*(X_{-i})), X_{-i})] = L[u_i(A_i, X_{-i})]\,,$$

such that the set of actions strictly dominated by some strategy in $\Delta(\max(S^*(X_{-i})))$ coincides with the set of actions strictly dominated by some strategy in $\Delta(A_i)$.

The inclusion from left to right is trivial since $\max(S^*(X_{-i})) \subseteq A_i$. For the inclusion from right to left, recall that a convex and compact set in $\mathbb{R}^m$ is equal to the convex hull of the set of its extreme points. As both $L[u_i(A_i, X_{-i})]$ and $L[u_i(\max(S^*(X_{-i})), X_{-i})]$ are compact and convex, it remains to be shown that no point in $u_i(A_i \setminus \max(S^*(X_{-i})), X_{-i})$ is an extreme point of $L[u_i(A_i, X_{-i})]$. This follows from the fact that any such point is strictly dominated by some $a_i^* \in \Delta(A_i)$. Indeed, the definition of $\max(S^*(X_{-i}))$ ensures that for each $a_i \in A_i \setminus \max(S^*(X_{-i}))$, there exists $a_i^* \in \Delta(A_i)$ with $u_i(a_i^*, x_{-i}) > u_i(a_i, x_{-i})$ for all $x_{-i} \in X_{-i}$. □

4.4 SOPHISTICATED ALGORITHMS

In this section, we investigate the consequences of Theorem 4.5 on C_D-, C_M-, V-, and V*-sets.

Let us first consider C_M and C_D, which are only defined in symmetric matrix games. We have already seen (Corollary 3.14) that both dominance structures yield unique minimal solutions, and it can easily be shown that the other requirements for the sophisticated algorithm are satisfied as well.

Corollary 4.9. *The C_D-set and the C_M-set of a symmetric matrix game can be computed in polynomial time.*

Proof. According to Theorem 4.5, it is sufficient to show that both C_M and C_D satisfy weak MON, TRA, SING, UNI, SUB-COM, and COM.

It is easily verified that both C_M and C_D satisfy weak MON, TRA, SING, and COM. UNI was shown in Corollary 3.14. Finally, Dutta and Laslier (1999) have shown that the essential set $ES(\Gamma)$ is a (nonempty) subset of $S_{C_M}(\Gamma)$ and $ES(\Gamma)$ can be computed in polynomial time using linear programming (see Brandt and Fischer, 2008b). This proves that C_M satisfies SUB-COM. The same is true for C_D because $C_D \subseteq C_M$ implies $S_{C_M}(\Gamma) \subseteq S_{C_D}(\Gamma)$. □

The following positive results now follow from Propositions 3.15 and 3.17.

Corollary 4.10. *The unique V-set of a confrontation game can be computed in polynomial time.*

Corollary 4.11. *The unique V*-set of a tournament game can be computed in polynomial time.*

4.5 SUMMARY

We proposed two generic algorithms for computing D-sets and studied their soundness and efficiency for various dominance structures D and subclasses of games. Our results yielded greedy algorithms for computing all S-sets, S^*-sets, and B-sets of a given normal-form game, and sophisticated algorithms for computing the unique C_M-set and the unique C_D-set of a given symmetric matrix game. Within the subclass of confrontation games, these algorithms coincide and also yield the V-set. Our algorithms subsume existing algorithms for computing saddles in matrix games (Shapley, 1964), minimal covering sets in *binary* symmetric matrix games (Brandt and Fischer, 2008b), and CURB sets in two-player games (Benisch et al., 2010). Interestingly, the sophisticated algorithms rely on the repeated computation of Nash equilibria via linear programming, even though the corresponding solution concepts are purely ordinal. Whether V-sets and V^*-sets can be computed efficiently in matrix games remains an interesting open problem.

In order to summarize the results, we revisit the table from the end of the previous chapter and highlight the cells to which our algorithms apply. If a cell is highlighted in dark gray, the greedy algorithm finds all D-sets in the given class in polynomial time. If it is highlighted in light gray, the analogous statement holds for the sophisticated algorithm.

<table>
<tr><th></th><th>S</th><th>B</th><th>S^*</th><th>C_D</th><th>C_M</th><th>V</th><th>V^*</th></tr>
<tr><td>normal-form games</td><td>poly</td><td>poly</td><td>poly</td><td></td><td></td><td rowspan="3">exp</td><td rowspan="4">exp</td></tr>
<tr><td>matrix games</td><td rowspan="4">unique</td><td rowspan="3">unique</td><td rowspan="3">unique</td><td></td><td></td></tr>
<tr><td>symmetric matrix games</td><td>unique</td><td>unique</td></tr>
<tr><td>confrontation games</td><td colspan="3">unique</td></tr>
<tr><td>tournament games</td><td colspan="2">unique</td><td colspan="4">unique</td></tr>
</table>

Table 4: Algorithmic results for D-sets

5 HARDNESS RESULTS FOR DOMINANCE-BASED SOLUTIONS

The algorithms from the previous chapter do not apply to all dominance structures considered in this thesis. In fact, we have seen in Section 3.4 that some dominance structures give rise to an *exponential* number of D-sets, even in symmetric matrix games. While this immediately implies that no polynomial-time algorithm can compute *all* D-sets, a number of natural computational questions remain open. For example, such questions concern the complexity of *finding* a D-set, checking whether a given action is contained in *some* D-set, or deciding whether there is a *unique* D-set. In this chapter, we resolve this question for weak (W) and very weak (V) dominance. Since W-sets may fail to exist (see Figure 12 on page 31), we consider *weak saddles*, a closely related solution concept whose existence is guaranteed. Weak saddles are defined as inclusion-minimal externally W-stable sets. We show that finding weak saddles of bimatrix games is NP-hard, under polynomial-time Turing reductions, and that recognizing weak saddles is coNP-complete. We moreover prove that deciding whether a given action is contained in some weak saddle is hard for parallel access to NP and thus not even in NP unless the polynomial hierarchy collapses. Finally, we extend most of these intractability results to V-sets.

5.1 WEAK SADDLES IN TWO-PLAYER GAMES

As mentioned in Section 3.2, the existence of W-solutions can be restored by ignoring internal stability. Weak saddles are defined as minimal externally W-stable sets.

Definition 5.1. *Let* $\Gamma = (N, (A_i)_{i\in N}, (u_i)_{i\in N})$ *be a game. A tuple* $S \subseteq A_N$ *is a* weak generalized saddle point (WGSP) *of* Γ *if for every player* $i \in N$, S_i $W(S_{-i})$ a_i *for all* $a_i \in A_i \setminus S_i$. *A* weak saddle *is a WGSP that contains no other WGSP.*

A WGSP thus is a tuple that is externally W-stable. Observe that the tuple A_N of all actions is always a WGSP, thereby guaranteeing existence of a weak saddle in every game. Weak saddles do not have to be unique, as shown in the example in Figure 17.

This section is mostly concerned with (non-symmetric) bimatrix games. For such games, we can simplify notation and write $\Gamma = (A, B, u)$, where A is the set of actions of player 1, B is the set of actions of player 2, and $u : A \times B \to \mathbb{R} \times \mathbb{R}$ is the utility function on the

	b_1	b_2	b_3
a_1	0 2	1 1	0 1
a_2	1 1	1 1	1 2

Figure 17: Example bimatrix game with two weak saddles $(\{a_1\}, \{b_1, b_2\})$ and $(\{a_1, a_2\}, \{b_2\})$. In this chapter, we follow the convention to write player 1's utility in the lower left corner and player 2's utility in the upper right corner of the corresponding matrix cell.

understanding that $u(a, b) = (u_1(a, b), u_2(a, b))$ for all $(a, b) \in A \times B$.

$x \rightsquigarrow y$

For an action a and a weak saddle $S = (S_1, S_2)$, we will sometimes slightly abuse notation and write $a \in S$ if $a \in (S_1 \cup S_2)$. In such cases, whether a is a row or a column should be either clear from the context or irrelevant for the argument. This partial identification of S and $S_1 \cup S_2$ is also reflected in referring to S as a "set" rather than a "pair" or "tuple." When reasoning about the structure of the saddles of a game, the following notation will be useful. For two actions $x, y \in A \cup B$, we write $x \rightsquigarrow y$ if every weak saddle containing x also contains y. Observe that $\rightsquigarrow$ as a relation on $(A \cup B) \times (A \cup B)$ is transitive. We now identify two sufficient conditions for $x \rightsquigarrow y$ to hold.

Fact 5.2. *Let $\Gamma = (A, B, u)$ be a two-player game, $b \in B$ an action of player 2, and $a \in A$ an action of player 1. Then $b \rightsquigarrow a$ if one of the following two conditions holds:*[1]

(i) a *is the unique action maximizing* $u_1(\cdot, b)$, *i.e.,*

$$\{a\} = \arg\max_{a' \in A} u_1(a', b).$$

(ii) a *maximizes* $u_1(\cdot, b)$, *and all actions maximizing* $u_1(\cdot, b)$ *yield identical utilities for* all *opponent actions, i.e.,* $a \in \arg\max_{a' \in A} u_1(a', b)$ *and* $u_1(a_1, b') = u_1(a_2, b')$ *for all* $a_1, a_2 \in \arg\max_{a' \in A} u_1(a', b)$ *and all* $b' \in B$.

Part *(i)* of the statement above can be generalized in the following way. An action a is in the weak saddle if it is a unique best response to a subset of saddle actions: if $\{b_1, \dots, b_t\} \subseteq S$ and there is no $a' \in A \setminus \{a\}$ with $u_1(a', b_i) \geqslant u_1(a, b_i)$ for all $i \in [t]$, then $a \in S$.[2] In this case, we write $\{b_1, \dots, b_t\} \rightsquigarrow a$. Moreover, for two sets of actions X and Y, we write $X \rightsquigarrow Y$ if $X \rightsquigarrow y$ for all $y \in Y$. For example, in the game of Figure 17, $b_1 \rightsquigarrow a_1 \rightsquigarrow b_2$, $\{b_2, b_3\} \rightsquigarrow a_2$, and $\{b_1, b_3\} \rightsquigarrow \{a_1, a_2\}$.

1 The statement remains true if the roles of the two players are reversed.
2 For $n \in \mathbb{N}$, we denote $[n] = \{1, \dots, n\}$.

In [6], we constructed a class of games that established a strong relationship between weak saddles and inclusion-maximal cliques in undirected graphs. Based on this construction and a reduction from the NP-complete problem CLIQUE, we showed that deciding whether there exists a weak saddle with a *certain number of actions* is NP-hard. This construction, however, did not permit any statements about the more important problems of *finding* a weak saddle, *recognizing* a weak saddle, or *deciding whether a certain action is contained* in some weak saddle. In particular, we will be interested in the following computational problems for a given game Γ.

- FINDWEAKSADDLE: Find a weak saddle of Γ.
- ISWEAKSADDLE: Is a given tuple $(S_1, \dots, S_n)$ a weak saddle of Γ?
- UNIQUEWEAKSADDLE: Does Γ contain exactly one weak saddle?
- INWEAKSADDLE: Is a given action a contained in a weak saddle of Γ?
- INALLWEAKSADDLES: Is a given action a contained in *every* weak saddle of Γ?
- NONTRIVIALWEAKSADDLE: Does Γ contain a weak saddle that does *not* consist of all actions?

5.2 A GENERAL CONSTRUCTION

We will now derive various hardness results for weak saddles. We begin by presenting a general construction that transforms a Boolean formula φ into a bimatrix game Γ_φ, such that the existence of certain weak saddles in Γ_φ depends on the satisfiability of φ. This construction will be instrumental for each of the hardness proofs given in this section.

Let $\varphi = C_1 \wedge \dots \wedge C_m$ be a Boolean formula in conjunctive normal form (CNF) over a finite set $V = \{v_1, \dots, v_n\}$ of variables. Denote by $L = \{v_1, \bar{v}_1, \dots, v_n, \bar{v}_n\}$ the set of all *literals*, where a literal is either a variable or its negation. Each *clause* C_j is a set of literals. An *assignment* $\alpha \subseteq L$ is a subset of the literals with the interpretation that all literals in α are set to "true." Assignment α is *valid* if $\ell \in \alpha$ implies $\bar{\ell} \notin \alpha$ for all $\ell \in L$.[3] We say that α *satisfies* a clause C_j if α is valid and $C_j \cap \alpha \neq \emptyset$. An assignment that satisfies all clauses of φ will be called a *satisfying assignment* for φ. A satisfying assignment α will be called *minimal* if there does not exist a satisfying assignment α' with $\alpha' \subset \alpha$. A formula that has a satisfying assignment will be called *satisfiable*. Clearly, every satisfiable formula has at least one minimal satisfying assignment.

Boolean formula

literal

clause

minimal satisfying assignment

3 If $\ell = \bar{v}_i$, then $\bar{\ell} = v_i$.

	b^*	v_1	$\bar{v}_1$	v_2	$\bar{v}_2$	$\cdots$	v_n	$\bar{v}_n$
a^*	1 1	0 0	0 0	0 0	0 0	$\cdots$	0 0	0 0
d^*	0 0	1 1	1 1	1 1	1 1	$\cdots$	1 1	1 1
C_1	1 0	0 0	0 1	0 1	0 0	$\cdots$	0 1	0 1
C_2	1 0	0 1	0 0	0 1	0 1	$\cdots$	0 0	0 1
$\vdots$					$\vdots$			
C_m	1 0	0 1	0 1	0 1	0 1	$\cdots$	0 1	0 1

Figure 18: Subgame of Γ_φ for a formula $\varphi = C_1 \wedge \cdots \wedge C_m$ with $v_1, \bar{v}_2 \in C_1$ and $\bar{v}_1, v_n \in C_2$.

We assume without loss of generality that φ does not contain any *trivial* clauses, i.e., clauses that contain both a variable v and its negation $\bar{v}$, and that no literal is contained in every clause. The game $\Gamma_\varphi = (A, B, u)$ is defined in three steps.

Step 1. Player 1 has actions $\{a^*, d^*\} \cup C$, where $C = \{C_1, \dots, C_m\}$ is the set of clauses of φ. Player 2 has actions $B = \{b^*\} \cup L$, where L is the set of literals.[4] The utility function is given by

- $u(a^*, b^*) = (1, 1)$,
- $u(d^*, \ell) = (1, 1)$ for all $\ell \in L$,
- $u(C_j, b^*) = (0, 1)$ for all $j \in [m]$,
- $u(C_j, \ell) = (1, 0)$ for all $j \in [m]$ and $\ell \in L \setminus C_j$, and
- $u(a, b) = (0, 0)$ otherwise.

An example of such a game is shown in Figure 18. Observe that $(\{a^*\}, \{b^*\})$ is a weak saddle, and thus no strict superset can be a weak saddle. Furthermore, row d^* dominates row C_j with respect to a set of columns $\{\ell_1, \dots, \ell_t\} \subseteq L$ if and only if $\ell_i \in C_j$ for some $i \in [t]$. In particular, for a valid assignment α it holds that $d^*\ W(\alpha)\ C_j$ if and only if α satisfies C_j. Another noteworthy property of this game is that no weak saddle contains any of the rows C_j, because $C_j \rightsquigarrow b^* \rightsquigarrow a^*$ for each $j \in [m]$.

The basic idea behind this construction is the following. The game Γ_φ will have a weak saddle containing row d^* if and only if φ is satisfiable. More precisely, we will show that whenever a weak saddle (S_1, S_2) contains d^*, the set S_2 of saddle columns is a minimal satisfying assignment. Such a saddle will be called an *assignment*

assignment saddle

4 There shall be no confusion by identifying literals with corresponding actions of player 2, which will henceforth be called "literal columns" (or "literal columns").

saddle. In order to prove that assignment saddles only exist if φ is satisfiable, we need to ensure that a pair (S_1, S_2) with $d^* \in S_1$ and $S_2 = \alpha$ cannot be a weak saddle if α does not satisfy φ or if α is not a valid assignment. This is achieved by means of additional actions (see step 2 below), for which the utilities are defined in such a way that every "wrong" (i.e., unsatisfying or invalid) assignment yields a set containing both a^* and b^*. Obviously, such a set can never be a weak saddle because it contains the weak saddle $(\{a^*\},\{b^*\})$ as a proper subset. In fact, $(\{a^*\},\{b^*\})$ will be the unique weak saddle in cases where there is no satisfying assignment.

Step 2. We augment the action sets of both players. Player 1 has one additional row ℓ' for each literal $\ell \in L$.[5] Player 2 has one additional column y_i for each variable $v_i \in V$. Utilities for profiles involving new actions are defined as follows (for an overview, refer to Figure 19):

- $u(a^*, y_i) = (1, 0)$ for all $i \in [n]$,
- $u(\ell', \ell) = (2, 1)$ for all $\ell \in L$,
- $u(\ell', y_i) = (0, 1)$ for all $i \in [n]$ and $\ell' \in \{v_i', \overline{v}_i'\}$, and
- $u(a, b) = (0, 0)$ otherwise.

Observe that, by Fact 5.2 and the discussion following it, $\ell \rightsquigarrow \ell'$, $\{v_i', \overline{v}_i'\} \rightsquigarrow y_i$, and $y_i \rightsquigarrow a^* \rightsquigarrow b^*$ for each $\ell \in L$ and each $i \in [n]$. This means that no assignment saddle can contain both v_i as well as its negation $\overline{v}_i$.

There only remains one subtlety to be dealt with. In the game defined so far, there are weak saddles containing row d^*, whose existence is independent of the satisfiability of φ, namely $(\{d^*, \ell'\},\{\ell\})$ for each $\ell \in L$. We destroy these saddles by using additional rows.

Step 3. We introduce new rows $r_1, \overline{r}_1, \dots, r_n, \overline{r}_n$, one for each literal, with the property that $r_i \rightsquigarrow b^*$, and that r_i and $\overline{r}_i$ can only be weakly dominated (by v_i and $\overline{v}_i$, respectively) if at least one literal column other than v_i or $\overline{v}_i$ is in the saddle. For this, we define

- $u(r_i, b^*) = u(\overline{r}_i, b^*) = (0, 1)$ for all $i \in [n]$,
- $u(r_i, v_i) = u(\overline{r}_i, \overline{v}_i) = (2, 0)$ for all $i \in [n]$,
- $u(r_i, \ell) = u(\overline{r}_i, \ell) = (-1, 0)$ for all $\ell \in \{v_{i \bmod n+1}, \overline{v}_{i \bmod n+1}\}$ and $i \in [n]$, and
- $u(a, b) = (0, 0)$ otherwise.

The game Γ_φ now has action sets $A = \{a^*, d^*\} \cup C \cup L \cup \{r_1, \dots, \overline{r}_n\}$ for player 1 and $B = \{b^*\} \cup L \cup \{y_1, \dots, y_n\}$ for player 2. The size of Γ_φ thus is clearly polynomial in the size of φ. A complete example of such a game is given in Figure 19.

5 Action ℓ' of player 1 and action ℓ of player 2 refer to the same literal, but we name them differently to avoid confusion.

	b^*	v_1	$\bar{v}_1$	v_2	$\bar{v}_2$	$\cdots$	v_n	$\bar{v}_n$	y_1	y_2	$\cdots$	y_n
a^*	1, 1					$\cdots$			1, 0	1, 0	$\cdots$	1, 0
d^*		1, 1	1, 1	1, 1	1, 1	$\cdots$	1, 1	1, 1			$\cdots$	
C_1	0, 1		1, 0	1, 0		$\cdots$	1, 0	1, 0			$\cdots$	
C_2	0, 1	1, 0		1, 0	1, 0	$\cdots$		1, 0			$\cdots$	
$\vdots$						$\vdots$						
C_m	0, 1	1, 0	1, 0	1, 0	1, 0	$\cdots$	1, 0	1, 0			$\cdots$	
v'_1		2, 1				$\cdots$			0, 1		$\cdots$	
$\bar{v}'_1$			2, 1			$\cdots$			0, 1		$\cdots$	
v'_2				2, 1		$\cdots$				0, 1	$\cdots$	
$\bar{v}'_2$					2, 1	$\cdots$				0, 1	$\cdots$	
$\vdots$						$\vdots$						
v'_n						$\cdots$	2, 1				$\cdots$	0, 1
$\bar{v}'_n$						$\cdots$		2, 1			$\cdots$	0, 1
r_1	0, 1	2, 0		-1, 0	-1, 0	$\cdots$					$\cdots$	
$\bar{r}_1$	0, 1		2, 0	-1, 0	-1, 0	$\cdots$					$\cdots$	
r_2	0, 1			2, 0		$\cdots$					$\cdots$	
$\bar{r}_2$	0, 1				2, 0	$\cdots$					$\cdots$	
$\vdots$						$\vdots$						
r_n	0, 1	-1, 0	-1, 0			$\cdots$	2, 0				$\cdots$	
$\bar{r}_n$	0, 1	-1, 0	-1, 0			$\cdots$		2, 0			$\cdots$	

Figure 19: Game Γ_φ used in the proof of Proposition 5.3. Utilities equal $(0, 0)$ unless specified otherwise. $S^\alpha = (\{d^*\} \cup \alpha, \alpha)$ is a weak generalized saddle point of Γ_φ if and only if α satisfies φ. For improved readability, thick lines are used to separate different types of actions.

For a valid assignment α, define $S^\alpha = (\{d^*\} \cup \alpha, \alpha)$. It should be clear from the argumentation above that S^α is a weak generalized saddle point of Γ_φ if and only if α satisfies φ. In particular, S^α is a weak saddle if and only if α is a *minimal* satisfying assignment. To show that membership of a given action in a weak saddle is NP-hard, it suffices to show that there are no other weak saddles containing row d^*. We do so in the following section.

5.3 MEMBERSHIP IS NP-HARD

We now show that it is NP-hard to decide whether a given action is contained in some weak saddle.

Proposition 5.3. InWeakSaddle *is NP-hard, even for two-player games.*

Proof. We give a reduction from SAT. For a CNF formula φ, we show that the game Γ_φ, defined in Section 5.2, has a weak saddle that contains action d^* if and only if φ is satisfiable. The direction from right to left is straightforward. If α is a minimal satisfying assignment for φ, then S^α is a weak saddle that contains d^*.

For the other direction, we will show that all weak saddles containing d^* are (essentially) assignment saddles. Let $S = (S_1, S_2)$ be a weak saddle of Γ_φ such that $d^* \in S_1$. We can assume that $S_2 \subseteq L$. If this was not the case, i.e., if there was a column $c \in \{b^*, y_1, \dots, y_n\}$ with $c \in S_2$, then $c \rightsquigarrow a^* \rightsquigarrow b^*$, and $(\{a^*\}, \{b^*\})$ would be a smaller saddle contained in S, a contradiction. We will now show that

(i) $|S_2| \geqslant 2$,

(ii) $|\{v_i, \bar{v}_i\} \cap S_2| \leqslant 1$ for all $i \in [n]$, and

(iii) $C \cap S_1 = \emptyset$.

For *(i)*, suppose that $|S_2| = 1$. Without loss of generality, $S_2 = \{v_i\}$. Then, both v_i' and r_i have to be in S_1, as they are maximal with respect to $\{v_i\}$. Together with $r_i \rightsquigarrow b^*$, this however contradicts the fact that $b^* \notin S_2$.

For *(ii)*, suppose that there exists $i \in [n]$ with $\{v_i, \bar{v}_i\} \subseteq S_2$. Then at least one of the rows v_i' or r_i *and* at least one of the rows $\bar{v}_i'$ or $\bar{r}_i$ is in the set S_1. Since $r_i \rightsquigarrow b^*$ as well as $\bar{r}_i \rightsquigarrow b^*$, and since $b^* \notin S_2$, we deduce that $\{v_i', \bar{v}_i'\} \subseteq S_1$. On the other hand, $\{v_i', \bar{v}_i'\} \rightsquigarrow y_i$, again contradicting $S_2 \subseteq L$.

For *(iii)*, merely observe that $C_j \rightsquigarrow b^*$ for all $j \in [m]$.

We now show that $d^* \; W(S_2) \; C_j$ for all $j \in [m]$. Consider some $j \in [m]$. From *(iii)* we know that there exists a row $s \in S_1$ with $s \; W(S_2) \; C_j$. We consider two cases. First, assume that $|\{\ell \in S_2 : u_1(C_j, \ell) = 1\}| \geqslant 2$. It follows from our assumption and from the definition of u_1 that d^* is the only row that can weakly dominate C_j with respect to S_2. If, on the other hand, $|\{\ell \in S_2 : u_1(C_j, \ell) = 1\}| \leqslant 1$, $d^* \; W(S_2) \; C_j$ follows immediately from $S_2 \subseteq L$ and *(i)*.

Define the assignment $\alpha = S_2$ and note that by *(ii)*, α is valid. The fact that $d^* \; W(\alpha) \; C_j$ implies that there exists $\ell \in \alpha$ with $u_1(C_j, \ell) = 0$, which means that $\ell \in C_j$. Thus α satisfies C_j for all $j \in [m]$. In other words, φ is satisfiable. □

5.4 MEMBERSHIP IS CONP-HARD

We have just seen that it is NP-hard to decide whether there exists a weak saddle containing a given action. In order to prove that this problem is also coNP-hard, we first show the following: given a game and an action c, it is possible to augment the game with additional actions such that every weak saddle of the augmented game that contains c contains all actions of this game.

Lemma 5.4. *Let $\Gamma = (A, B, u)$ be a two-player game and $c \in A \cup B$ an action of Γ. Then there exists a supergame $\Gamma^c = (A', B', u')$ of Γ with the following properties:*

(i) If S is a weak saddle of Γ^c containing c, then $S = (A', B')$.

(ii) If S is a weak saddle of Γ that does not *contain c, then S is a weak saddle of Γ^c.*

(iii) The size of Γ^c is polynomial in the size of Γ.

Proof. Let $n = |A|$ and $m = |B|$. Without loss of generality, we may assume that all utilities in Γ are positive and that c is a column, i.e., $u_\ell(a, b) > 0$ for all $(a, b) \in A \times B$, $\ell \in [2]$, and $c \in B$. Define

$$\lambda = \max_{a \in A} u_1(a, c) + 1\,,$$

such that λ is greater than the maximum utility to player 1 in column c. Now, let Γ^c be a supergame of Γ with $n + m - 1$ additional rows and n additional columns, i.e., $\Gamma^c = (A', B', u')$, where $A' = A \cup \{a'_1, \dots, a'_{n+m-1}\}$, $B' = B \cup \{b'_1, \dots, b'_n\}$ and $u'|_{A \times B} = u$. Utilities for action profiles not in $A \times B$ are shown in Figure 20.

For *(i)*, let $S = (S_1, S_2)$ be a weak saddle of Γ^c with $c \in S_2$. Using the second part of Fact 5.2, we get $c \rightsquigarrow A' \setminus A \rightsquigarrow B' \setminus \{c\} \rightsquigarrow A$. For *(ii)*, observe that our assumption about the utilities in Γ implies that each additional action is dominated by each original action as long as c is not contained in the weak saddle. Finally, *(iii)* is immediate from the definition of Γ^c. □

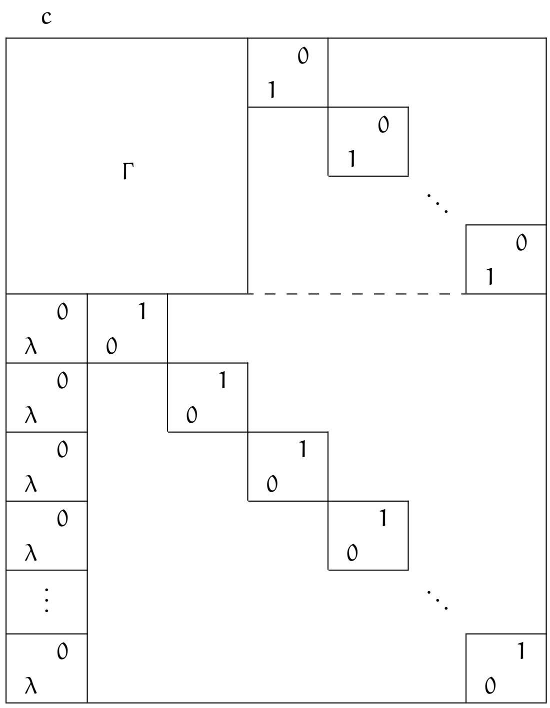


Figure 20: Construction used in the proof of Lemma 5.4. Utilities for new action profiles are $(0, 0)$ unless specified otherwise, and λ is chosen so as to maximize $u_1'(\cdot, c)$. Every weak saddle containing column c equals the set of all actions.

We are now ready to show that InWeakSaddle is coNP-hard.

Proposition 5.5. InWeakSaddle *is coNP-hard, even for two-player games.*

Proof. We give a reduction from UNSAT. For a given CNF formula φ, consider the game $\Gamma_\varphi^{b^*}$ obtained by augmenting the game Γ_φ defined in Section 5.2 in such a way that every weak saddle containing action b^* in fact contains all actions. We show that $\Gamma_\varphi^{b^*}$ has a weak saddle containing b^* if and only if φ is unsatisfiable.

For the direction from left to right, assume that there exists a weak saddle S with $b^* \in S$. By Lemma 5.4, S is trivial, i.e., equals the set of all actions. Furthermore, S must be the unique weak saddle of $\Gamma_\varphi^{b^*}$, because any other weak saddle would violate minimality of S. In particular, S^α cannot be a saddle for any assignment α, which by the discussion in Section 5.2 means that φ is unsatisfiable.

For the direction from right to left, assume that φ is unsatisfiable. Similar reasoning as in the proof of Proposition 5.3 shows that every weak saddle $S = (S_1, S_2)$ satisfies $S_2 \not\subseteq L$, i.e., S contains at least one column not corresponding to a literal. However, since $b \rightsquigarrow a^*$ for every column $b \in B \setminus L$ and $a^* \rightsquigarrow b^*$, we have that $b^* \in S_2$ for every weak saddle of $\Gamma_\varphi^{b^*}$. $\square$

The proof of Proposition 5.5 implies several other hardness results.

Corollary 5.6. *The following hold:*

- IsWeakSaddle *is coNP-complete.*
- InAllWeakSaddles *is coNP-complete.*
- UniqueWeakSaddle *is coNP-hard.*

All hardness results hold even for two-player games.

Proof. Let φ be a Boolean formula, which without loss of generality we can assume to have either no satisfying assignment or more than one. (For any Boolean formula, this property can for example be achieved by adding a clause with two new variables, thereby multiplying the number of satisfying assignments by three.)

Recall the definition of the game $\Gamma_\varphi^{b^*}$ used in the proof of Proposition 5.5. It is easily verified that the following statements are equivalent: formula φ is unsatisfiable, $\Gamma_\varphi^{b^*}$ has a trivial weak saddle, $\Gamma_\varphi^{b^*}$ has a unique weak saddle, and b^* is contained in all weak saddles of $\Gamma_\varphi^{b^*}$. This provides a reduction from UNSAT to each of the problems above.

Membership of InAllWeakSaddles in coNP holds because any externally stable set that does not contain the action in question serves as a witness that this action is *not* contained in every weak saddle. For membership of IsWeakSaddle, consider a tuple S of actions that is *not* a weak saddle. Then either S is not externally stable, or there exists a proper subset of S that is externally stable. In both cases there is a witness of polynomial size. □

5.5 FINDING A SADDLE IS NP-HARD

A particularly interesting consequence of Proposition 5.5 concerns the existence of a nontrivial weak saddle. As we will see, the hardness of deciding the latter can be used to obtain a result about the complexity of the search problem.

Corollary 5.7. NontrivialWeakSaddle *is NP-complete. Hardness holds even for two-player games.*

Proof. For membership in NP, observe that proving the existence of a nontrivial weak saddle is tantamount to finding a proper subset of the set of all actions that is externally stable. By definition, every such subset is guaranteed to contain a weak saddle. Obviously, external stability can be checked in polynomial time.

Hardness is again straightforward from the proof of Proposition 5.5, since the game $\Gamma_\varphi^{b^*}$ has a nontrivial weak saddle if and only if formula φ is satisfiable. □

Corollary 5.8. FINDWEAKSADDLE *is NP-hard under polynomial-time Turing reductions, even for two-player games.*

Proof. Suppose there exists an algorithm that computes some weak saddle of a game in time polynomial in the size of the game. Such an algorithm could obviously be used to solve the NP-hard problem NONTRIVIALWEAKSADDLE in polynomial time. Just run the algorithm once. If it returns a nontrivial saddle, the answer is "yes." Otherwise the set of all actions must be the unique weak saddle of the game, and the answer is "no." □

5.6 MEMBERSHIP IS Θ_2^p-HARD

Now that we have established that INWEAKSADDLE is both NP-hard and coNP-hard, we will raise the lower bound to Θ_2^p. Wagner provided a sufficient condition for Θ_2^p-hardness that turned out to be very useful (see, e.g., Hemaspaandra et al., 1997).

Lemma 5.9 (Wagner (1987))**.** *Let* S *be an NP-complete set, and let* T *be an arbitrary set. If there exists a polynomial-time computable function* f *such that*

$$\|\{i : x_i \in S\}\| \text{ is odd } \iff f(x_1, \dots, x_{2k}) \in T \tag{6}$$

for all $k \geqslant 1$ *and all strings* $x_1, \dots, x_{2k}$ *satisfying* $x_{j-1} \in S$ *whenever* $x_j \in S$ *for every* j *with* $1 < j \leqslant 2k$*, then* T *is* Θ_2^p*-hard.*

We now apply Wagner's Lemma to show Θ_2^p-hardness of INWEAKSADDLE.

Theorem 5.10. INWEAKSADDLE *is* Θ_2^p*-hard, even for two-player games.*

Proof. We apply Lemma 5.9 with $S = \text{SAT}$ and $T = \text{INWEAKSADDLE}$. Fix an arbitrary $k \geqslant 1$ and let $\varphi_1, \dots, \varphi_{2k}$ be $2k$ Boolean formulas such that satisfiability of φ_j implies satisfiability of φ_{j-1}, for each j, $1 < j \leqslant 2k$.

We will now define a polynomial-time computable function f which maps the given $2k$ Boolean formulas to an instance of INWEAKSADDLE such that (6) is satisfied. For *odd* $i \in [2k]$, let $\Gamma_i = (A_i, B_i, u_i)$ be the game Γ_{φ_i} as defined in the proof of Proposition 5.3, with decision row d^* renamed as d_i. Recall that this game has a weak saddle containing d_i if and only if φ_i is satisfiable. Analogously, for *even* $i \in [2k]$, let $\Gamma_i = (A_i, B_i, u_i)$ be the game $\Gamma_{\varphi_i}^{d_i}$ as defined in the proof of Proposition 5.5, with decision column b^* renamed as d_i. Thus, Γ_i has a weak saddle containing d_i if and only if φ_i is *unsatisfiable*. For all $i \in [2k]$, we may without loss of generality assume that all utilities in Γ_i are positive and strictly smaller than some $K \in \mathbb{N}$, and that the decision action d_i of game Γ_i is a row, i.e., $0 < u_\ell(a,b) < K$ for all $(a,b) \in A_i \times B_i$ and $\ell \in [2]$, and $d_i \in A_i$.[6]

Now define the game Γ by combining the games Γ_i, $i \in [2k]$, with one additional row z_i and two additional columns c_i^1 and c_i^2 for each $i \in [2k]$, as well as a decision row d^*, i.e., $\Gamma = (A, B, u)$ where $A = \bigcup_{i=1}^{2k} A_i \cup \{z_1, \dots, z_{2k}\} \cup \{d^*\}$ and $B = \bigcup_{i=1}^{2k} B_i \cup \bigcup_{i=1}^{2k} \{c_i^1, c_i^2\}$. For $a \in A_i$ and $b \in B_j$, utilities are defined as $u(a,b) = u_i(a,b)$ if $i = j$ and $u(a,b) = (0,0)$ otherwise. Furthermore, for $b \in \bigcup B_j$, let $u(z_i, b) = (0,1)$ for all $i \in [2k]$ and $u(d^*, b) = (0,1)$. The definition of u on profiles containing a new column c_i^ℓ, $i \in [2k]$, $\ell \in [2]$ is quite complicated, and we recommend consulting Figure 21 for an overview. Player 2 has only two distinct utilities for these columns: for $a \in A$ and $\ell \in [2]$,

$$u_2(a, c_i^\ell) = \begin{cases} K & \text{if } a = d_i \\ 0 & \text{otherwise.} \end{cases}$$

Recall that all utilities in the games Γ_i are smaller than K, such that the utility for player 2 in the profiles (d_i, c_i^1) and (d_i, c_i^2) is maximal in Γ.

The utilities for player 1 are defined in order to connect the games Γ_{2i-1} and Γ_{2i}, for each $i \in [k]$. We need some notation. For $i \in [2k]$, let i° be $i+1$ if i is odd and $i-1$ if i is even. Thus, each pair $\{i, i^\circ\}$ is of the form $\{2j-1, 2j\}$ for some j. For $a \in \bigcup A_j$, define

$$u_1(a, c_i^\ell) = \begin{cases} 1 & \text{if } \ell = 1 \text{ and } a \in A_i \\ 2 & \text{if } \ell = 1 \text{ and } a \in A_{i^\circ} \\ 0 & \text{otherwise.} \end{cases}$$

Furthermore, for $i, j \in [2k]$, let

$$u_1(z_j, c_i^1) = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{otherwise,} \end{cases} \qquad \text{and} \qquad u_1(z_j, c_i^2) = \begin{cases} 0 & \text{if } j = i^\circ \\ 1 & \text{otherwise.} \end{cases}$$

Finally, let $u_1(d^*, c_i^1) = 0$ and $u_1(d^*, c_i^2) = 1$ for all $i \in [2k]$.

An example of the game Γ for the case $k = 2$ is depicted in Figure 21, where we assume without loss of generality that each d_i is the first row of Γ_i.

The following facts are readily appreciated.

Fact 5.11. *If S is a weak saddle of Γ_i and $d_i \notin S$, then S is a weak saddle of Γ.*

6 Adding a positive number to every utility does not change the dominance relation between the actions. As the minimum utility in Γ_i is -1, adding a number greater than 1 suffices. If d_i is a column, as in the proof of Proposition 5.5, we can simply transpose the game by exchanging the two players.

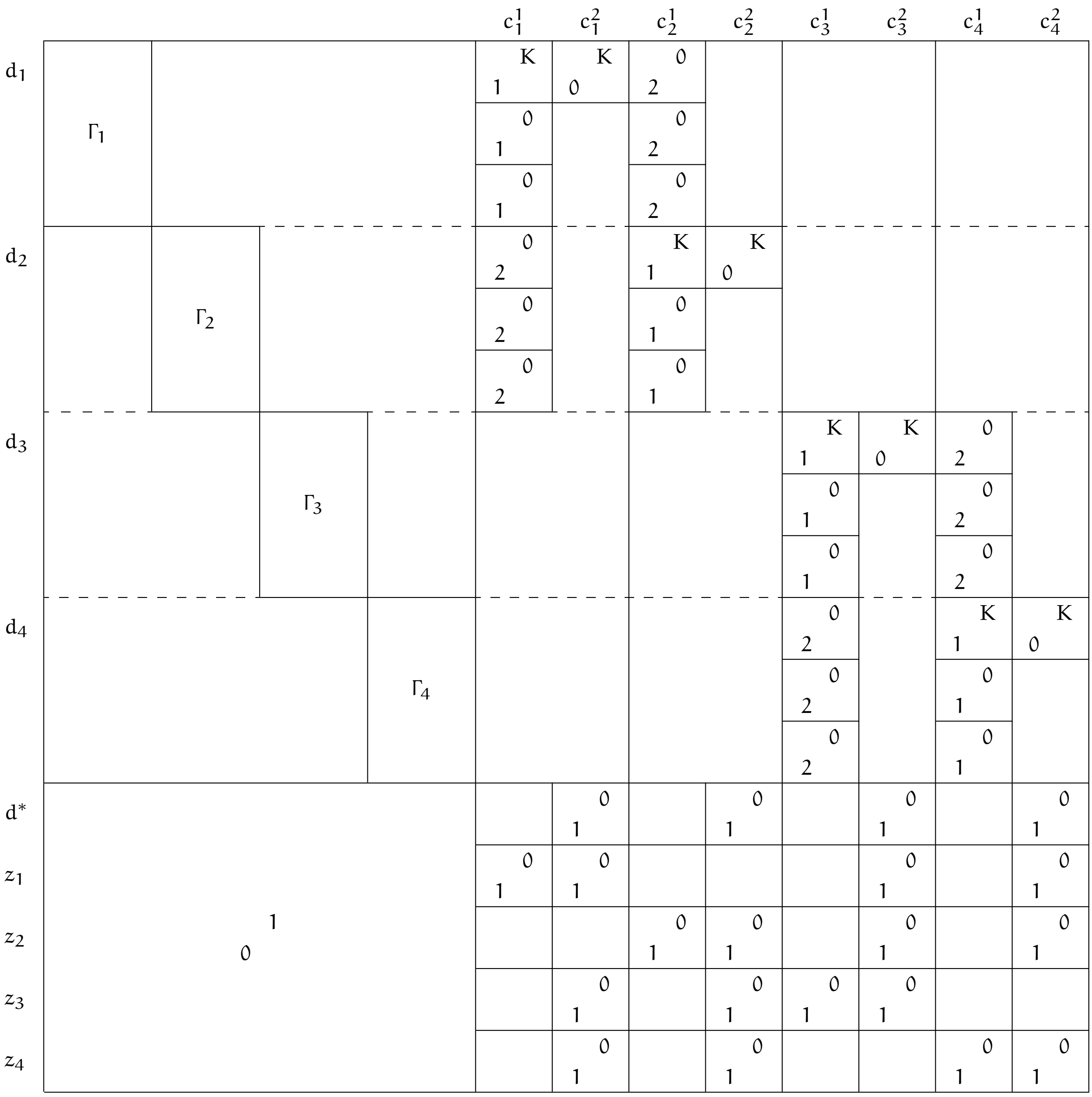


Figure 21: Game Γ used in the proof of Theorem 5.10. Utilities are $(0,0)$ unless specified otherwise. Γ has a weak saddle containing row d^* if and only if both Γ_1 and Γ_2 *or* both Γ_3 and Γ_4 have a weak saddle containing their respective decision row d_i.

For a weak saddle $S = (S_1, S_2)$ of Γ and $i \in [2k]$, define $S^i = (S_1 \cap A_i, S_2 \cap B_i)$ as the intersection of S with Γ_i.

Fact 5.12. *If S is a weak saddle of Γ, then S^i is either a weak saddle of Γ_i or empty.*

For Fact 5.11 it suffices to check external stability. For Fact 5.12, observe that our assumption that $u_\ell(a,b) > 0$ implies that weak domination with respect to a subset of $A_i \cup B_i$ can only occur among actions belonging to $A_i \cup B_i$. Therefore, if some action profile in $A_i \times B_i$ is contained in a weak saddle, all actions of Γ_i *not* contained in the saddle must be dominated by some saddle action of the same subgame Γ_i.

In order to be able to apply Lemma 5.9, we now prove (6), which here amounts to showing the following equivalence:

$$\|\{i : \varphi_i \in \text{SAT}\}\| \text{ is odd } \Leftrightarrow \Gamma \text{ has a weak saddle } S \text{ with } d^* \in S. \quad (7)$$

For the direction from left to right, assume that there is an odd number i such that φ_i is satisfiable and $\varphi_{i^\circ} = \varphi_{i+1}$ is not. Then, there exist weak saddles S^i and S^{i° of the games Γ_i and Γ_{i°, respectively, such that $d_i \in S^i$ and $d_{i^\circ} \in S^{i^\circ}$. Define $S = S^i \cup S^{i^\circ} \cup \{d^*, z_1, \dots, z_{2k}\} \cup \{c_i^1, c_i^2, c_{i^\circ}^1, c_{i^\circ}^2\}$. We claim that S is a weak saddle of Γ. The proof consists of two parts.

First, we have to show that S is externally stable, i.e., that all actions not in the saddle are weakly dominated by saddle actions. To see this, let $a \in A_j$ be a row that is not in S. If $j \notin \{i, i^\circ\}$, then a is weakly dominated by *every* saddle row because it yields utility 0 to player 1 against any saddle column. If, on the other hand, $j \in \{i, i^\circ\}$, then a is weakly dominated by the same row that weakly dominates it in the subgame Γ_j. The argument for non-saddle columns $b \in \bigcup_j B_j$ is analogous. Moreover, every column c_j^ℓ with $j \notin \{i, i^\circ\}$ is weakly dominated by the saddle columns $c_i^1, c_i^2, c_{i^\circ}^1$, and $c_{i^\circ}^2$.

Second, we have to show that S is inclusion-minimal, i.e., that no proper subset of S is a weak saddle of Γ. Let $\tilde{S} \subseteq S$ be a weak saddle. By Fact 5.12 and the observation that $\tilde{S}^i$ cannot be empty, we know that $\tilde{S}^i = S^i$, as otherwise inclusion-minimality of S^i in Γ_i would be violated. In particular, $d_i \in \tilde{S}^i$, which implies that $\{c_i^1, c_i^2\} \subseteq \tilde{S}$. The same reasoning for i° shows that $\tilde{S}^{i^\circ} = S^{i^\circ}$ and $\{c_{i^\circ}^1, c_{i^\circ}^2\} \subseteq \tilde{S}$. Now, $\{c_i^1, c_i^2\} \rightsquigarrow z_i$ and $\{c_{i^\circ}^1, c_{i^\circ}^2\} \rightsquigarrow z_{i^\circ}$. Furthermore, all rows z_j with $j \notin \{i, i^\circ\}$, as well as d^*, are in $\tilde{S}$, because they are all maximal and identical with respect to S. Here, maximality is due to the fact that they are the only rows that yield a positive utility to player 1 against both saddle columns c_i^2 and $c_{i^\circ}^2$. Thus $\tilde{S} = S$, meaning that S is indeed inclusion-minimal.

For the direction from right to left, let S be a weak saddle of Γ with $d^* \in S$. From the definition of $u_2(d^*, \cdot)$, we infer that $S \cap \bigcup_j B_j \neq \emptyset$, which in turn implies that $S \cap \bigcup_j A_j \neq \emptyset$. We can now deduce that there is at least one column $c_i^\ell \in S$, as otherwise row d^* would always yield 0 against all saddle actions and $S \setminus \{d^*\}$ would be externally stable. Now observe that for any $i \in [2k]$, the definition of $u_2(\cdot, c_i^\ell)$ implies that every weak saddle of Γ contains either none or both of the columns c_i^1 and c_i^2. We thus have $\{c_i^1, c_i^2\} \subseteq S$. Furthermore, $z_i \in S$ because $\{c_i^1, c_i^2\} \rightsquigarrow z_i$. However, z_i must not weakly dominate d^* with respect to S, because otherwise $S \setminus \{d^*\}$ would be externally stable. This means there has to be a saddle column $c \in S$ with $u_1(z_i, c) < u_1(d^*, c)$. The only column satisfying this property is

$c^2_{i^\circ}$, which means that both $c^2_{i^\circ}$ and, by the same argument as above, $c^1_{i^\circ}$ are contained in S. Now that both c^1_i and $c^1_{i^\circ}$ are in S, at least one row from each of the games Γ_i and Γ_{i° has to be a saddle action, i.e., $S^i \neq \emptyset$ and $S^{i^\circ} \neq \emptyset$. By Fact 5.12, we conclude that S^i and S^{i° are weak saddles of Γ_i and Γ_{i°, respectively.

It remains to be shown that $d_i \in S^i$ and $d_{i^\circ} \in S^{i^\circ}$. If $d_i \notin S^i$, then by Fact 5.11, $S^i \subset S$ would be a weak saddle of Γ, contradicting inclusion-minimality of S. The argument for S^{i° is analogous. It finally follows from the construction that φ_i is satisfiable and φ_{i° is unsatisfiable,[7] which completes the proof of (7). By Lemma 5.9, InWeakSaddle is Θ^p_2-hard. □

We conclude this section by showing that Σ^p_2 is an upper bound for the membership problem.

Proposition 5.13. InWeakSaddle *is in* Σ^p_2.

Proof. Let $\Gamma = (N, (A_i)_{i \in N}, (u_i)_{i \in N})$ be a game and $d^* \in \bigcup_i A_i$ a designated action. First observe that we can verify in polynomial time whether a subset of A_N is externally stable. We can guess a weak saddle S containing d^* in nondeterministic polynomial time and verify its minimality by checking that none of its subsets are externally stable. This places InWeakSaddle in $\mathrm{NP}^{\mathrm{coNP}} = \mathrm{NP}^{\mathrm{NP}}$ and thus in Σ^p_2. □

5.7 HARDNESS RESULTS FOR V-SETS

In this section, we show that most of our results for weak saddles also hold for V-sets. It is worth noting, however, that the results for V-sets do not follow in an obvious way from those for weak saddles, or vice versa. While the proofs are based on the same general idea, and again on one core construction, there are some significant technical differences.

Recall Proposition 3.4 (page 31), which states that D-sets coincide with minimal externally D-stable sets whenever D is monotonic. Since very weak dominance is monotonic, we can work with the following characterization of V-sets.

Fact 5.14. *A tuple $S \subseteq A_N$ is a V-set if and only if S is a minimal externally V-stable set.*

As in the case of weak saddles we define, for each Boolean formula φ, a two-player game Γ_φ that admits certain types of V-sets if and only if φ is satisfiable. Let $\varphi = C_1 \wedge \cdots \wedge C_m$ be a 3-CNF

7 Here we have assumed without loss of generality that $i < i^\circ$, i.e., i is odd and $i^\circ = i + 1$ is even.

formula[8] over variables $v_1, \dots, v_n$, where $C_i = \{\ell_{i,1}, \ell_{i,2}, \ell_{i,3}\}$. Call a pair $\{\ell_{i,j}, \ell_{i',j'}\}$ of literal occurrences *conflicting*, and write $[\ell_{i,j}, \ell_{i',j'}]$, if $i \neq i'$ and $\ell_{i,j} = \overline{\ell_{i',j'}}$.

Define the bimatrix game $\Gamma_\varphi = (A, B, u)$ as follows. The set A of actions of player 1 comprises the set $C = \{C_1, \dots, C_m\}$ of clauses of φ as well as one additional action for each conflicting pair $[\ell_{i,j}, \ell_{i',j'}]$ of literals.[9] The set B of actions of player 2 is the set of all literal occurrences, i.e., $B = \bigcup_{j=1}^{m} \{\ell_{j,1}, \ell_{j,2}, \ell_{j,3}\}$. The utility function is given by

$$u(C_i, \ell_{p,q}) = \begin{cases} (0,1) & \text{if } p = i, \\ (1,0) & \text{if } p = i \bmod m + 1, \\ (0,0) & \text{otherwise,} \end{cases}$$

and

$$u([\ell_{i,j}, \ell_{i',j'}], \ell_{p,q}) = \begin{cases} (1,0) & \text{if } i = p \text{ and } j = q, \\ (1,0) & \text{if } i' = p \text{ and } j' = q, \\ (0,0) & \text{otherwise.} \end{cases}$$

An example of a game Γ_φ is shown in Figure 22.

Consider a V-set (S_1, S_2) of Γ_φ. We will exploit the following three properties, which are easy consequences of the definition of Γ_φ:

(I) If $C_i \in S_1$ for some $i \in [m]$, then $\ell_{i,j} \in S_2$ for some $j \in [3]$.

(II) If $\ell_{i,j} \in S_2$ for some $i \in [m]$ and $j \in [3]$, then $C_{i \bmod m+1} \in S_1$ or $[\ell_{i,j}, \ell_{i',j'}] \in S_1$ for some $i' \in [m]$ and $j' \in [3]$.

(III) For two conflicting literals $\ell_{i,j}$ and $\ell_{i',j'}$, we have $\{\ell_{i,j}, \ell_{i',j'}\} \rightsquigarrow [\ell_{i,j}, \ell_{i',j'}]$.

The idea underlying the definition of Γ_φ is formalized in the following lemma.

Lemma 5.15. *Γ_φ has a V-set (S_1, S_2) with $S_1 = C$ if and only if φ is satisfiable.*

Proof. For the direction from left to right, consider a saddle (S_1, S_2) with $S_1 = C$. By *(III)*, S_2 does not include any conflicting literals and thus defines a valid assignment for φ. Moreover, *(I)* ensures that $|\{\ell_{i,1}, \ell_{i,2}, \ell_{i,3}\} \cap S_2| \geqslant 1$ for each $i \in [m]$, which means that this assignment satisfies φ.

8 A formula in 3-CNF is a CNF formula where every clause consists of exactly three literals. The problems SAT and UNSAT remain NP-hard and coNP-hard, respectively, even for this restricted class of formulas. While the construction works for arbitrary CNF formulas, we employ 3-CNFs for ease of notation.

9 We identify $[\ell_{i,j}, \ell_{i',j'}]$ and $[\ell_{i',j'}, \ell_{i,j}]$ and thus have only one action per conflicting pair.

	$\ell_{1,1}$	$\ell_{1,2}$	$\ell_{1,3}$	$\ell_{2,1}$	$\ell_{2,2}$	$\ell_{2,3}$	$\cdots$	$\ell_{m,1}$	$\ell_{m,2}$	$\ell_{m,3}$
C_1	1 0	1 0	1 0				$\cdots$	0 1	0 1	0 1
C_2	0 1	0 1	0 1	1 0	1 0	1 0	$\cdots$			
C_3				0 1	0 1	0 1	$\cdots$			
$\vdots$					$\vdots$					
C_m							$\cdots$	1 0	1 0	1 0
$[\ell_{1,1}, \ell_{2,3}]$	0 1					0 1	$\cdots$			
$[\ell_{1,2}, \ell_{2,1}]$		0 1		0 1			$\cdots$			
$[\ell_{1,1}, \ell_{m,1}]$	0 1						$\cdots$	0 1		
$\vdots$					$\vdots$					

Figure 22: Game Γ_φ for a formula φ with $C_1 = v_1 \vee \bar{v}_2 \vee v_3$, $C_2 = \bar{v}_1 \vee v_2 \vee v_4$, and $C_m = \bar{v}_1 \vee \bar{v}_2 \vee v_4$. Utilities are $(0, 0)$ unless specified otherwise.

For the direction from right to left, let α be a satisfying assignment of φ and $f : [m] \to [3]$ a function such that $\ell_{i,f(i)} \in \alpha$ for all $i \in [m]$. It is then easily verified that $(C, \bigcup_{i=1}^{m} \{\ell_{i,f(i)}\})$ is a V-set of Γ_φ. $\square$

In the following we define two bimatrix games Γ'_φ and Γ''_φ that extend Γ_φ with new actions in such a way that properties *(I)*, *(II)*, and *(III)* still hold. In particular, statements similar to Lemma 5.15 will hold for Γ'_φ and Γ''_φ. The game Γ'_φ is then used to prove the NP-hardness of InVset, while Γ''_φ is used in the proofs of all other hardness results.

The game Γ'_φ, shown in Figure 23, is defined by adding a column d to Γ_φ. Utilities for the new action profiles are defined as $u(C_i, d) = (0, 0)$ for all $i \in [m]$, and $u([\ell_{i,j}, \ell_{i',j'}], d) = (1, 1)$ for each conflicting pair.

Lemma 5.16. *Γ'_φ has a* V-set (S_1, S_2) *with* $C_1 \in S_1$ *if and only if* φ *is satisfiable.*

Proof. By Lemma 5.15, Γ_φ has a V-set (S_1, S_2) with $S_1 = C$ if and only if φ is satisfiable. Since $u_2(C_i, d) = 0$ for all $i \in [m]$, this property still holds for Γ'_φ.

It remains to be shown that if (S_1, S_2) is a V-set of Γ'_φ with $C_1 \in S_1$, then $S_1 = C$. If $C_1 \in S_1$, properties *(I)* and *(II)* imply that $C_i \in S_1$ for all $i \in [m]$. On the other hand, observe that $[\ell_{i,j}, \ell_{i',j'}] \rightsquigarrow d$ for every conflicting pair $[\ell_{i,j}, \ell_{i',j'}]$ and that $(\{[\ell_{i,j}, \ell_{i',j'}]\}, \{d\})$ is a V-set. Obviously, this is the only V-set containing $[\ell_{i,j}, \ell_{i',j'}]$. Therefore, a

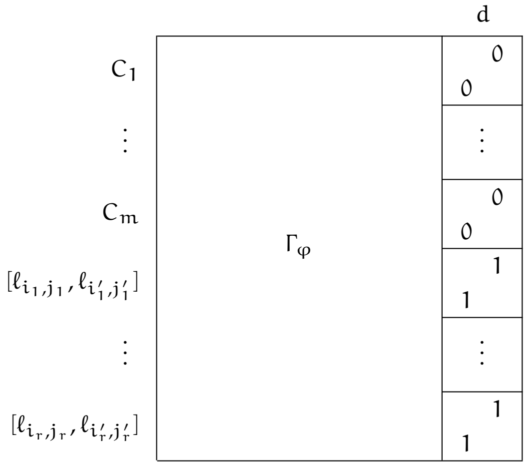


Figure 23: Game Γ'_φ used in the proof of Lemma 5.16

saddle containing C_1 does not contain any row that corresponds to a conflicting pair. □

In order to obtain further hardness results, we define the bimatrix game Γ''_φ as another supergame of Γ_φ. In addition to the properties *(I)*, *(II)*, and *(III)*, Γ''_φ will have the following new property:

(IV) For every row $[\ell_{i,j}, \ell_{i,j'}]$ that corresponds to a conflicting pair, it is true that $[\ell_{i,j}, \ell_{i,j'}] \rightsquigarrow a$ for every action a of Γ''_φ.

Let r denote the number of conflicting pairs of φ and rename the actions of $\Gamma_\varphi = (A, B, u)$ in such a way that $A = \{a_1, \dots, a_{m+r}\}$ with $C_i = a_i$ for all $i \in [m]$ and $B = \{b_1, \dots, b_{3m}\}$. To obtain the game Γ''_φ shown in Figure 24, we augment Γ_φ by s additional columns $d_1, \dots, d_s$ and s additional rows $e_1, \dots, e_s$, where $s = \max(|A|, |B|) + 1$. Utilities for new action profiles are defined as follows:

- $u(e_i, d_j) = (2, 0)$ if $j = i$,
- $u(e_i, d_j) = (0, 2)$ if $j = i \bmod s + 1$,
- $u(e_i, b_j) = (0, 1)$ if $i \in \{j, j+1\}$,
- $u(a_i, d_j) = (1, 0)$ if $j \in \{i, i+1\}$,
- $u([\ell_{i,j}, \ell_{i',j'}], d_1) = (0, 1)$ for all conflicting pairs, and
- $u(a, b) = (0, 0)$ otherwise.

Note that Γ''_φ satisfies properties *(I)*, *(II)*, and *(III)*, since $u_2(C_i, d_j) = 0$ for all $i \in [m]$ and $u_1(e_i, b) = 0$ for all $b \in B$. We can thus prove the following lemma analogously to Lemma 5.15.

Lemma 5.17. *Γ''_φ has a V-set (S_1, S_2) with $S_1 = C$ if and only if φ is satisfiable.*

	b_1	b_2	$\cdots$	b_{3m}	d_1	d_2	d_3	d_4	$\cdots$	d_{s-1}	d_s
$a_1 = C_1$	0, 1	0, 1	$\cdots$	1, 0	1, 0	1, 0			$\cdots$		
$a_2 = C_2$	1, 0	1, 0	$\cdots$			1, 0	1, 0		$\cdots$		
$a_3 = C_3$			$\cdots$				1, 0	1, 0	$\cdots$		
$\vdots$						$\vdots$					
$a_m = C_m$			$\cdots$	0, 1					$\cdots$		
$a_{m+1} = [\ell_{i_1,j_1}, \ell_{i'_1,j'_1}]$		1, 0	$\cdots$		0, 1				$\cdots$		
$a_{m+2} = [\ell_{i_2,j_2}, \ell_{i'_2,j'_2}]$	1, 0		$\cdots$		0, 1				$\cdots$		
$\vdots$						$\vdots$					
$a_{m+r} = [\ell_{i_r,j_r}, \ell_{i'_r,j'_r}]$			$\cdots$		0, 1				$\cdots$	1, 0	1, 0
e_1	0, 1		$\cdots$		2, 0	0, 2			$\cdots$		
e_2	0, 1	0, 1	$\cdots$			2, 0	0, 2		$\cdots$		
e_3		0, 1	$\cdots$				2, 0	0, 2	$\cdots$		
e_4			$\cdots$					2, 0	$\cdots$		
$\vdots$						$\vdots$					
e_s			$\cdots$		0, 2				$\cdots$		2, 0

Figure 24: Game Γ''_φ used in the proof of Theorem 5.19. Utilities are $(0,0)$ unless specified otherwise. Row labels have been moved to the right for improved readability.

In order to prove that Γ''_φ satisfies property *(IV)*, note that $[\ell_{i,j}, \ell_{i',j'}] \rightsquigarrow d_1$ for every conflicting pair $[\ell_{i,j}, \ell_{i',j'}]$. Furthermore, we have $d_i \rightsquigarrow e_i$ for every $i \in [s]$, and $e_i \rightsquigarrow d_{i+1}$ for every $i \in [s-1]$. It therefore follows from the transitivity of $\rightsquigarrow$ that $[\ell_{i,j}, \ell_{i',j'}] \rightsquigarrow d_k$ and $[\ell_{i,j}, \ell_{i',j'}] \rightsquigarrow e_k$ for every $[\ell_{i,j}, \ell_{i',j'}] \in A$ and every $k \in [s]$. Finally, by construction, $\{d_i, d_{i+1}\} \rightsquigarrow a_i$ for all $i \in \{1, \dots, |A|\}$, and $\{e_i, e_{i+1}\} \rightsquigarrow b_i$ for all $i \in \{1, \dots, |B|\}$. Since $s > \max(|A|, |B|)$, this implies *(IV)*.

Lemma 5.18. *Γ''_φ has a nontrivial V-set if and only if φ is satisfiable.*

Proof. If φ is satisfiable, then by Lemma 5.17 there exists a nontrivial V-set.

Conversely assume that φ is not satisfiable. Then by *(IV)* there is no nontrivial saddle (S_1, S_2) with $[\ell_{i,j}, \ell_{i',j'}] \in S_1$ for a conflicting pair

$[\ell_{i,j}, \ell_{i',j'}]$. By Lemma 5.17, there is no saddle (S_1, S_2) with $S_1 = C$. Furthermore, it follows from *(I)*, *(II)*, and *(IV)* that there cannot be a saddle (S_1, S_2) with $S_1 \subset C$. It remains to show that a nontrivial V-set cannot contain any of the new actions e_i or d_j. As mentioned above, $d_i \rightsquigarrow e_i$ and $e_i \rightsquigarrow d_{i \bmod s+1}$ for all $i \in [s]$. Hence we can conclude—analogously to the proof of *(IV)*—that $d_i \rightsquigarrow a$ and $e_i \rightsquigarrow a$ for every action a. Thus, d_i and e_i cannot be part of a nontrivial saddle for any $i \in [s]$. □

Computational problems for V-sets are defined analogously to their counterparts for weak saddles. Combining the four lemmata above, we get the following.

Theorem 5.19. *The following hold:*

(i) InVset *is NP-hard.*

(ii) InVset *is coNP-hard.*

(iii) IsVset *is coNP-complete.*

(iv) InAllVsets *is coNP-complete.*

(v) UniqueVset *is coNP-hard.*

(vi) NontrivialVset *is NP-complete.*

(vii) FindVset *is NP-hard under Turing reductions.*

All hardness results hold even for two-player games.

Proof. Let φ be a Boolean formula and let Γ'_φ and Γ''_φ be the games defined above.

(i) NP-hardness of InVset can be shown by a reduction from 3-SAT. Lemma 5.16 shows that Γ'_φ has a V-set containing C_1 if and only if φ is satisfiable.

(ii) coNP-hardness of InVset can be shown by a reduction from 3-UNSAT. Consider the game Γ''_φ and assume without loss of generality that φ has at least one pair of conflicting literals. It follows from property *(IV)* and Lemma 5.18 that each row that corresponds to a conflicting pair is contained in a V-set of Γ''_φ, namely the trivial one, if and only if φ is unsatisfiable.

(iii) A minor modification of the coNP algorithm for IsWeakSaddle shows that IsVset is in coNP. We show coNP-hardness by a reduction from 3-UNSAT. It follows from Lemma 5.18 that the set of all actions of Γ''_φ is a V-set if and only if φ is unsatisfiable.

(iv) The proof of coNP-membership of InAllVsets is similar to the proof of coNP-membership of InAllWeakSaddles. Hardness follows from the same argument as in *(ii)*.

(v) coNP-hardness of UNIQUEVSET can be shown by a reduction from 3-UNSAT. Consider the game Γ''_φ and assume without loss of generality that φ has either none or more than one satisfying assignment. Then, if φ is satisfiable, Γ''_φ has multiple V-sets, each of them corresponding to a particular satisfying assignment. If on the other hand φ is unsatisfiable, Γ''_φ has only the trivial V-set.

(vi) The proof of NP-membership of NONTRIVIALVSET is similar to the proof of NP-membership of NONTRIVIALWEAKSADDLE. NP-hardness of the problem follows from a reduction from 3-SAT. Lemma 5.18 shows that Γ''_φ has a nontrivial V-set if and only if φ is satisfiable.

(vii) NP-hardness of FINDVSET can be shown in the same way as that of FINDWEAKSADDLE.

□

An argument analogous to that for INWEAKSADDLE shows that INVSET is in Σ_2^p. On the other hand, Θ_2^p-hardness of INVSET appears much harder to obtain. In particular, the construction in the proof of Theorem 5.10 uses pairs of actions c_i^1 and c_i^2 that are identical from the point of view of player 2, and argues that every weak saddle must contain either none or both of them. This argument no longer goes through for V-sets, because c_i^1 and c_i^2 very weakly dominate each other, and indeed there are V-sets that contain only one of the two actions. Additional insights will therefore be required to raise the lower bound for INVSET.

5.8 SUMMARY

We have shown that weak saddles are computationally intractable even in bimatrix games. As it turned out, not only *finding* but also *recognizing* weak saddles is computationally hard. Most of the hardness results were shown to carry over to V-sets. Open problems concern the complexity of weak saddles and V-sets in matrix games, the gap between Θ_2^p and Σ_2^p for INWEAKSADDLE, and complete characterizations of the complexity of the search problem. It would also be worthwhile to investigate whether the hardness results of this chapter can be generalized to *mixed weak saddles* (minimal externally W^*-stable sets) and V^*-sets.

6 COMPLEXITY OF ITERATED DOMINANCE

In this final chapter of Part I, we turn to solution concepts that are defined via the iterated elimination of dominated actions. In particular, we investigate the computational complexity of iterated W-dominance (IWD) in two-player constant-sum games. It turns out that deciding whether an action is eliminable via IWD is feasible in polynomial time, whereas deciding whether a given subgame is reachable via IWD is NP-complete. The latter result is quite surprising, as we are not aware of other natural computational problems that are intractable in constant-sum normal-form games. Furthermore, we slightly improve on a result of Conitzer and Sandholm (2005) by showing that typical problems associated with IWD in win-lose games with at most one winner are NP-complete.

This chapter is organized as follows. We first review some classic results on order-independence (Section 6.1). Then, we focus on (pure and mixed) *weak* dominance, for which we introduce some basic terminology in Section 6.2. We propose the auxiliary concept of a regionalized game in Section 6.3 and show that this concept may be used as a convenient tool in the proofs of our hardness results. In Section 6.4 we deal with the computational complexity of reachability and eliminability problems in two-player constant-sum games. Finally, in Section 6.5, we address the same problems for win-lose games that allow at most one winner.

6.1 ORDER-INDEPENDENCE

An important property in the context of iterated dominance is whether the resulting subgame depends on the order in which dominated actions are eliminated. A dominance structure D is *order-independent* if for every game Γ and for any two D-irreducible subgames Γ' and Γ'' of Γ that are obtained by iteratively eliminating D-dominated actions from Γ, we have $\Gamma' = \Gamma''$. In other words, D is order-independent if the solution concept *iterated* D-*dominance* (Definition 2.12) always selects a unique tuple of subsets of actions. Table 5 summarizes order-independence results from the literature.[1] See Apt (2004, 2011) for unified proofs of most of these results.

Order-independence is desirable not only from a normative, but also from a computational point of view. Indeed, there is a straightfor-

1 Note that these results do not necessarily extend to infinite games (Dufwenberg and Stegeman, 2002). The fact that C_M and C_D are order-independent is easily established and crucially depends on the assumption that rows and columns corresponding to the same action can only be deleted simultaneously (see Section 2.3.2).

	order-independent	
strict dominance (S)	✓	(Gilboa et al., 1990)
weak dominance (W)	–	
very weak dominance (V)	–	
Börgers dominance (B)	✓	(Börgers, 1993)
mixed strict dominance (S^*)	✓	(Osborne and Rubinstein, 1994)
mixed weak dominance (W^*)	–	
mixed very weak dominance (V^*)	–	
covering (C_M)	✓	
deep covering (C_D)	✓	

Table 5: Order-independence

ward polynomial-time algorithm that computes the unique solution for any order-independent dominance structure D: simply delete, in each iteration, *all* D-dominated actions, until no more deletions are possible.[2]

On the other hand, dominance structures that are order-*dependent* give rise to a number of challenging computational problems such as: given a game, can a specified subgame (or a subgame of a given size) be reached via iterated elimination of dominated actions? The remainder of this chapter is devoted to studying the computational complexity of questions of this kind.

6.2 TERMINOLOGY FOR ITERATED WEAK DOMINANCE

We use the simplified notation for bimatrix games that was introduced in Chapter 5 and denote a two-player game by a triple $\Gamma = (A_1, A_2, u)$. In particular, the function $u : A_1 \times A_2 \to \mathbb{R} \times \mathbb{R}$ represents the utility functions of both players by letting $u(a_1, a_2) = (u_1(a_1, a_2), u_2(a_1, a_2))$.

Let $\Gamma = (A_1, A_2, u)$ be a two-player game and $a, b \in A_1$ two actions of player 1. Then, a is said to *weakly dominate* b *at* $c \in A_2$ *in* Γ if $u_1(a, c) > u_1(b, c)$ and for all $d \in A_2$, $u_1(a, d) \geqslant u_1(b, d)$. Thus, $a\ W(A_2)\ b$ if and only if a weakly dominates b at c for some $c \in A_2$. For a game $\Gamma' = (A'_1, A'_2, u)$ with $A'_1 \subseteq A_1$ and $A'_2 \subseteq A_2$, we say further that an action $c \in A_2$ *backs* the elimination of $b \in A_1$ by $a \in A_1$ in Γ' if $a, b, c \in A'_1 \cup A'_2$ and $u_1(a, c) > u_1(b, c)$, and *blocks* the elimination of b by a in Γ' if $a, b, c \in A'_1 \cup A'_2$ and $u_1(a, c) < u_1(b, c)$. Dominance, backing, and blocking for actions of player 2 are defined analogously. Note that an action is dominated by another action of the same player if some action of the other player backs the elimination

backing
blocking

2 The number of iterations is obviously bounded by the number of actions in the game. Furthermore, we have seen in Sections 4.3 and 4.4 that dominated actions can be identified efficiently for all dominance structures defined in Section 3.2.

and none of them block it. As the remainder of this chapter only concerns (iterated) weak dominance, we will drop the qualification 'weak' and by 'dominance' understand weak dominance.[3]

The following notions are based on elimination sequences as defined in Section 2.3.2. Recall that an elimination sequence is a given by a finite sequence $\Sigma = (\Sigma_1, \dots, \Sigma_n)$ of pairwise disjoint subsets of actions in A. If every Σ_i is a singleton, we say the elimination sequence $(\Sigma_1, \dots, \Sigma_n)$ is *simple*. Simple elimination sequences we usually write as sequences $\sigma = (\sigma_1, \dots, \sigma_n)$ of actions in A.

simple elimination sequence

An action a is called *eliminable by* b *at* c in a game Γ if there exists a valid elimination sequence Σ such that a is dominated by b at c in $\Gamma(\Sigma)$. Action a is *eliminable* in Γ if there are actions b and c such that a is eliminable by b at c. A subgame Γ' of Γ is *reachable* from Γ if there exists a valid elimination sequence Σ such that $\Gamma(\Sigma) = \Gamma'$. Furthermore Γ is called *solvable* if some subgame $\Gamma' = (A'_1, A'_2, u')$ with $|A'_1| = |A'_2| = 1$ is reachable from Γ.

eliminable

reachable

solvable

6.3 REGIONS AND REGIONALIZED GAMES

An essential building block of our hardness proofs are *regionalized games*, i.e., games in which the action set A_i of each player i is divided up into *regions*. Intuitively, the regions prevent eliminations of actions by actions from other regions. We assume for each player i that the regions constitute a partition of A_i, i.e., a set of nonempty and pairwise disjoint subsets of A_i the union of which exhausts A_i. More formally, a *regionalized two-player game* is a tuple (Γ, X_1, X_2) consisting of a two-player game $\Gamma = (A_1, A_2, u)$, a partition X_1 of A_1, and a partition X_2 of A_2. The elements of X_1 and X_2 are called *regions*.

regionalized games

regions

For regionalized games, the concept of a valid elimination sequence is modified so as to allow only eliminations of actions that are dominated by other actions in the same region. A *valid elimination sequence* for a regionalized game (Γ, X_1, X_2) is a sequence $\Sigma = (\Sigma_1, \dots, \Sigma_n)$ for Γ such that for each i with $1 \leqslant i \leqslant n$ and each $a \in \Sigma_i$, there is some action b and some $x \in X_1 \cup X_2$ such that $a, b \in x$ and b dominates a in $\Gamma(\Sigma_1, \dots, \Sigma_{i-1})$.

valid elimination sequence

With a slight abuse of notation we will use $(\Gamma(\Sigma_1, \dots, \Sigma_n), X_1, X_2)$ to refer to the regionalized game $(\Gamma(\Sigma_1, \dots, \Sigma_n), X'_1, X'_2)$ where $X'_1 = \{x \setminus (\Sigma_1 \cup \dots \cup \Sigma_n) : x \in X_1\} \setminus \{\emptyset\}$ and $X'_2 = \{x \setminus (\Sigma_1 \cup \dots \cup \Sigma_n) : x \in X_2\} \setminus \{\emptyset\}$.

We now prove a useful lemma: any regionalized two-player game can be transformed in polynomial time into a non-regionalized two-player game with the same valid elimination sequences. It follows that for reachability and eliminability problems, we can restrict ourselves to regionalized games, which are often more practical for and

3 The same comment applies to other notions such as *elimination sequences* and *irreducibility*.

afford more insight into the constructions used in our hardness proofs than games without regions.

Lemma 6.1. *For each regionalized game* (Γ, X_1, X_2) *with* $\Gamma = (A_1, A_2, u)$, *there is a game* $\Gamma' = (A_1', A_2', u')$ *computable in polynomial time such that the valid elimination sequences for* Γ' *and* (Γ, X_1, X_2) *coincide. Moreover,* $u'(a,b) \in \{(0,1),(1,0)\}$ *for all* $(a,b) \in (A_1' \times A_2') \setminus (A_1 \times A_2)$.

$$
\begin{array}{cc|ccc|ccc|cccc|}
 & & a_2^1 & \cdots & a_2^m & x_2^1 & \cdots & x_2^k & y_2^1 & y_2^2 & y_2^3 & y_2^4 \\
 & & & & & \multicolumn{3}{c}{\overbrace{\qquad\qquad}^{X_1}} & & & & \\
\hline
 & a_1^1 & & & & \cdot & \cdots & \cdot & (0,1) & (1,0) & (0,1) & (1,0) \\
 & \vdots & & \Gamma & & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 & a_1^\ell & & & & \cdot & \cdots & \cdot & (0,1) & (1,0) & (0,1) & (1,0) \\
\cline{3-5}
\multirow{3}{*}{$X_2 \left\{ \right.$} & x_1^1 & \cdot & \cdots & \cdot & (1,0) & \cdots & (0,1) & (1,0) & (0,1) & (1,0) & (0,1) \\
 & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 & x_1^k & \cdot & \cdots & \cdot & (0,1) & \cdots & (1,0) & (1,0) & (0,1) & (1,0) & (0,1) \\
 & y_1^1 & (1,0) & \cdots & (1,0) & (0,1) & \cdots & (0,1) & (1,0) & (1,0) & (0,1) & (0,1) \\
 & y_1^2 & (0,1) & \cdots & (0,1) & (1,0) & \cdots & (1,0) & (1,0) & (1,0) & (0,1) & (0,1) \\
 & y_1^3 & (1,0) & \cdots & (1,0) & (0,1) & \cdots & (0,1) & (0,1) & (0,1) & (1,0) & (1,0) \\
 & y_1^4 & (0,1) & \cdots & (0,1) & (1,0) & \cdots & (1,0) & (0,1) & (0,1) & (1,0) & (1,0) \\
\hline
\end{array}
$$

Figure 25: Game Γ' used in the proof of Lemma 6.1

Proof. The game Γ' is constructed from Γ by adding actions that impose the same restrictions on the elimination of actions as the regions did in (Γ, X_1, X_2). More auxiliary actions are then added to ensure that all elimination sequences that are valid for (Γ, X_1, X_2) are still valid for Γ' while no new valid elimination sequences are created.

Formally, let $\Gamma' = (A_1', A_2', u')$ with $A_1' = A_1 \cup X_2 \cup Y_1$ and $A_2' = A_2 \cup X_1 \cup Y_2$, where $Y_1 = \{y_1^1, y_1^2, y_1^3, y_1^4\}$ and $Y_2 = \{y_2^1, y_2^2, y_2^3, y_2^4\}$ are sets of actions disjoint from $A_1 \cup X_2$ and $A_2 \cup X_1$. Observe that the regions in (Γ, X_1, X_2) correspond to actions of the respective other player in Γ'. Further define the utility function u' such that $u'(a,b) = u(a,b)$ for all $(a,b) \in A_1 \times A_2$. For every $(a,x) \in A_1 \times X_1$ and $(x,b) \in X_2 \times A_2$, let

$$
u'(a,x) = \begin{cases} (1,0) & \text{if } a \in x, \\ (0,1) & \text{otherwise,} \end{cases} \quad \text{and} \quad u'(x,b) = \begin{cases} (0,1) & \text{if } b \in x, \\ (1,0) & \text{otherwise.} \end{cases}
$$

Without loss of generality we may assume that $|X_1| = |X_2| = k$ for some index $k \geqslant 0$, as we can always introduce copies of actions to the game. Thus, let $X_1 = \{x_2^1, \ldots, x_2^k\}$ and $X_2 = \{x_1^1, \ldots, x_1^k\}$ and define for all indices i and j with $1 \leqslant i, j \leqslant k$,

$$u'(x_1^i, x_2^j) = \begin{cases} (1,0) & \text{if } i = j, \\ (0,1) & \text{otherwise.} \end{cases}$$

The payoffs for the remaining action profiles are depicted in Figure 25. Obviously, Γ' can be obtained from (Γ, X_1, X_2) in polynomial time.

Before we show that the valid elimination sequences for (Γ, X_1, X_2) and Γ' coincide, we note that the utility function u' is chosen so as to ensure that none of the first player's actions in $X_2 \cup Y_1$ nor any of the second player's actions in $X_1 \cup Y_2$ appear in any valid elimination sequence for Γ'. To see this, observe that for each action $a \in X_2 \cup Y_1$ and each action $b \in A_1'$ there is some action $X_1 \cup Y_2$ that blocks the elimination of a by b in Γ'. For instance, x_2^1 blocks the elimination of y_1^2 by y_1^1. Moreover, for $a \in A_1$ and $b \in X_2 \cup Y_1$, there is some action in $X_1 \cup Y_2$ blocking the elimination of a by b in Γ'. It follows that for every valid elimination sequence Σ for Γ', if $a \in A_1'$ is dominated by $b_1 \in A_1'$ in $\Gamma'(\Sigma)$, then $a, b \in A_1$. By symmetrical arguments, an analogous statement holds for actions $a, b \in A_2'$.

Now consider a valid elimination sequence $\Sigma = (\Sigma_1, \ldots, \Sigma_n)$ for Γ'. Then, $\Sigma_1 \cup \cdots \cup \Sigma_n \subseteq A_1 \cup A_2$. Also consider an arbitrary index i with $1 \leqslant i \leqslant n$ and an arbitrary action $a \in \Sigma_i$. Without loss of generality we may assume that $a \in A_1$ and that there are actions $b \in A_1$ and $c \in A_2'$ such that a is dominated by b at c in $\Gamma'(\Sigma_1, \ldots, \Sigma_{i-1})$. Let $x \in X_1$ be the region of (Γ, X_1, X_2) with $b \in x$. It follows that $a \in x$ as well, otherwise the elimination of a by b would be blocked by x in $\Gamma'(\Sigma_1, \ldots, \Sigma_{i-1})$. With this being the case, observe that $u_1'(a, z) = u_1'(b, z)$ for all $z \in X_1 \cup Y_2$, i.e., no $z \in X_1 \cup Y_2$ backs the elimination of a by b. Hence, $c \in A_2 \setminus (\Sigma_1 \cup \cdots \cup \Sigma_{i-1})$. It follows that a is dominated by b at c in $\Gamma(\Sigma_1, \ldots, \Sigma_{i-1})$ and, because a and b are in the same region $x \in X_1$, a is dominated by b at c in $(\Gamma(\Sigma_1, \ldots, \Sigma_{i-1}), X_1, X_2)$ as well. We may conclude that Σ is also a valid elimination sequence for (Γ, X_2, X_2).

Finally, consider a valid elimination sequence $\Sigma = (\Sigma_1, \ldots, \Sigma_n)$ for (Γ, X_1, X_2), an index i with $1 \leqslant i \leqslant n$, and an action $a \in \Sigma_i$. Without loss of generality we may assume that $a \in A_1$. Suppose that a is dominated by b at c in $(\Gamma(\Sigma_1, \ldots, \Sigma_{i-1}), X_1, X_2)$ for some actions $b \in A_1$ and $c \in A_2$. Obviously, a is dominated by b at c in $\Gamma(\Sigma_1, \ldots, \Sigma_{i-1})$ as well. Moreover, a and b belong to the same region $x \in X_1$. Accordingly, no action $z \in X_1 \cup Y_2$ blocks the elimination of a by b in $\Gamma'(\Sigma_1, \ldots, \Sigma_{i-1})$. It follows that a is dominated by b at c in $\Gamma'(\Sigma_1, \ldots, \Sigma_{i-1})$ and that Σ is a valid elimination sequence for Γ'. $\square$

$$\begin{array}{c|cc|} & a_2 & b_2 \\ \hline a_1 & (1,0) & (1,0) \\ b_1 & (0,1) & (1,0) \\ c_1 & (1,0) & (0,1) \\ \hline \end{array}$$

Figure 26: IWD is order dependent

6.4 TWO-PLAYER CONSTANT-SUM GAMES

We will now show that subgame reachability is NP-complete even in games that only allow the outcomes $(0, 1)$ and $(1, 0)$. This may be attributed to the order dependence of IWD. For example, (b_1, a_2) is a valid elimination sequence for the game in Figure 26. However, if one eliminates row c_1 first, column a_2 is no longer eliminable.

In Section 6.4.2 we will find that for two-player constant-sum games a weak form of order independence can be salvaged, which allows us to formulate an efficient algorithm for the eliminability problem. Our first observation is that in the case of two-player constant-sum games we can restrict our attention to *simple* elimination sequences.

Lemma 6.2. *Let $\Gamma = (A_1, A_2, u)$ be a two-player constant-sum game and $\Sigma = (\Sigma_1, \dots, \Sigma_m)$ a valid elimination sequence. Then, there exists a simple elimination sequence $\sigma = (\sigma_1, \dots, \sigma_n)$ with $\{\sigma_1, \dots, \sigma_n\} = \Sigma_1 \cup \dots \cup \Sigma_m$ that is valid for Γ.*

Proof. Let X be a nonempty subset of A. It suffices to show that validity of the one-element sequence (X) for Γ implies the existence of some $x \in X$ such that the sequence $(X \setminus \{x\}, \{x\})$ is valid for Γ as well.

Assume for contradiction that (X) is valid but, for any $x \in X$, $(X \setminus \{x\}, \{x\})$ is not valid. Consider an arbitrary $x \in X$ and assume without loss of generality that $x \in A_1$. Then, x is dominated by some $x' \in A_1$ at some $y \in A_2$, i.e., $u_1(x', y) > u_1(x, y)$. Note that the dominance relation is asymmetric and transitive and that X is finite. Hence, without loss of generality, we may assume that $x' \notin X$.[4] By contrast, $y \in X$. To see this, observe that $(X \setminus \{x\})$ is valid for Γ. Moreover, as no action blocks the elimination of x by x' in Γ, neither is this the case for $\Gamma(X \setminus \{x\})$. If $y \notin X$, then $y \notin X \setminus \{x\}$, and x' would dominate x at y for $\Gamma(X \setminus \{x\})$. Consequently, $(X \setminus \{x\}, \{x\})$ would be valid for Γ, a contradiction.

Now, since $y \in X$, there must be some $y' \in A_2$ dominating y in Γ. By asymmetry and transitivity of the dominance relation, we may

4 As X is finite, by asymmetry and transitivity of the dominance relation there is a maximal sequence $x_1, \dots, x_k$ of pairwise distinct actions in X, such that $x = x_1$ and x_{i+1} dominates x_i for each i with $1 \leqslant i < k$. By assumption there also has to be some action x' in A that dominates x_k. By maximality of $x_1, \dots, x_k$, we have $x' \notin X$. Finally, by transitivity of the dominance relation, it follows that x' also dominates x.

$$
\begin{array}{c|ccc|}
 & y & & y' \\
\hline
x' & u_1(x',y) & \leqslant & u_1(x',y') \\
 & \vee & & \rotatebox{-90}{$\leqslant$} \\
x & u_1(x,y) & & u_1(x,y') \\
\hline
\end{array}
$$

Figure 27: Diagram illustrating the proof of Lemma 6.2

assume that $y' \notin X$. Moreover, there are no actions blocking the elimination of y by y' in Γ. Having assumed, however, that $(X \setminus \{y\}, \{y\})$ is not valid for Γ, it follows that x' does not back the elimination of y by y' in Γ, i.e., $u_2(x', y') \leqslant u_2(x', y)$. As Γ is a constant-sum game, $u_1(x', y') \geqslant u_1(x', y)$. Similarly, there is no action blocking the elimination of x by x' in Γ, whereas $(X \setminus \{x\}, \{x\})$ is not valid for Γ. Hence, y' does not back the elimination of x by x' in Γ, i.e., $u_1(x, y') \geqslant u_1(x', y')$. This situation is illustrated in Figure 27. It now follows that $u_1(x, y') > u_1(x, y)$ and, since Γ is a constant-sum game, $u_2(x, y') < u_2(x, y)$, contradicting the assumption that y' dominates y in Γ. □

As a corollary of Lemma 6.2 we find that a subgame of a two-player constant-sum game is reachable if and only if it is reachable by a simple elimination sequence. Analogous statements also hold for eliminability and solvability. Lemma 6.2 however does not hold for general strategic games. In fact, it already fails for games with outcomes in $\{(0,0), (0,1), (1,0)\}$, as Figure 28 illustrates.

	y	y'
x	$(0,0)$	$(0,1)$
x'	$(1,0)$	$(0,0)$

Figure 28: Game with weakly dominated actions x and y and a valid elimination sequence $(\{x, y\})$. The simple elimination sequences (x, y) and (y, x) are not valid.

6.4.1 *Reachability*

We are now ready to show that subgame reachability in constant-sum games is computationally intractable.

Theorem 6.3. *Given constant-sum games Γ and Γ', deciding whether Γ' is reachable from Γ is NP-complete, even if Γ only has outcomes $(0,1)$ and $(1,0)$ and Γ' is irreducible.*

Proof. For membership in NP consider arbitrary constant-sum games Γ and Γ'. Given an elimination sequence σ, it can clearly be decided in polynomial time whether Σ is a valid elimination sequence for (Γ, X_1, X_2) such that $\Gamma(\Sigma) = \Gamma'$.

The proof of hardness proceeds by a reduction from *3SAT*. By virtue of Lemma 6.1 it suffices to give a reduction for regionalized games. Consider an arbitrary *3CNF* $\varphi = C_1 \wedge \cdots \wedge C_k$, where each $C_i = (\lambda_i^1 \vee \lambda_i^2 \vee \lambda_i^3)$ is a clause and each λ_i^j is a literal, for $1 \leqslant i \leqslant k$ and $1 \leqslant j \leqslant 3$. Without loss of generality, we may assume that all clauses in φ are distinct. Define a regionalized game $(\Gamma_\varphi, X_1, X_2)$, with $\Gamma_\varphi = (A_1, A_2, u)$ as follows.

$$\begin{aligned}
A_1 &= \{p, \neg p, p{\downarrow} : p \text{ a variable in } \varphi\} \\
&\quad \cup \{C_i, (\lambda_i^1, i), (\lambda_i^2, i), (\lambda_i^3, i) : C_i \text{ a clause in } \varphi\} \\
&\quad \cup \{e\} \\
A_2 &= \{p, \neg p : p \text{ a variable in } \varphi\} \cup \{a, b\} \\
X_1 &= \{\{p, \neg p, p{\downarrow}\} : p \text{ a variable in } \varphi\} \\
&\quad \cup \{\{C_i, (\lambda_i^1, i), (\lambda_i^2, i), (\lambda_i^3, i)\} : C_i \text{ a clause in } \varphi\} \\
&\quad \cup \{\{e\}\} \\
X_2 &= \{\{p, \neg p : p \text{ a variable in } \varphi\} \cup \{a, b\}\} = \{A_2\}
\end{aligned}$$

For each variable p in φ, the payoffs in rows p, $\neg p$ and $p{\downarrow}$ are defined as in the following table, where q is a typical variable in φ distinct from p.

	p	$\neg p$		q	$\neg q$		a	b
			⋯			⋯		
p	$(1,0)$	$(0,1)$		$(0,1)$	$(0,1)$		$(1,0)$	$(0,1)$
$\neg p$	$(0,1)$	$(1,0)$		$(0,1)$	$(0,1)$		$(1,0)$	$(0,1)$
$p{\downarrow}$	$(1,0)$	$(1,0)$		$(0,1)$	$(0,1)$		$(0,1)$	$(0,1)$
			⋯			⋯		

Due to the regionalization, row $p{\downarrow}$ can be eliminated only by row p or row $\neg p$. Column a is the only action backing such an elimination. Intuitively, removing column p means setting variable p to false, removing column $\neg p$ setting variable p to true, thus choosing an assignment. Row $p{\downarrow}$ can thus be eliminated only after one of these columns has been removed, i.e., after an assignment for p has been chosen.

For each i with $1 \leqslant i \leqslant k$, the payoffs in rows C_i, (λ_i^1, i), (λ_i^2, i), (λ_i^3, i) depend on the literals occurring in C_i. In the following table, $\overline{\lambda}_i^j = \neg p$ if $\lambda_i^j = p$, and $\overline{\lambda}_i^j = p$ if $\lambda_i^j = \neg p$. We further assume $i \neq m$.

	λ_i^1	$\overline{\lambda}_i^1$	λ_i^2	$\overline{\lambda}_i^2$	λ_i^3	$\overline{\lambda}_i^3$	λ_m^j	$\overline{\lambda}_m^j$	a	b
(λ_i^1, i)	$(1,0)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$
(λ_i^2, i)	$(0,1)$	$(0,1)$	$(1,0)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$
(λ_i^3, i)	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(1,0)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$
C_i	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(1,0)$	$(0,1)$

Thus, the only columns backing the elimination of C_i are λ_i^1, λ_i^2, and λ_i^3. Also note that column a blocks the elimination of C_i. On the other hand, as we saw above, column a is essential to the elimination of the rows $p{\downarrow}$. Intuitively, this means that an assignment needs to be chosen before any of the rows C_i is eliminated.

Finally, we let $u(e, y) = (1, 0)$ if $y \neq b$, and $u(e, b) = (0, 1)$:

	λ_1^1	$\overline{\lambda}_1^1$	$\cdots$	λ_k^3	$\overline{\lambda}_k^3$	a	b
e	$(1,0)$	$(1,0)$	$\cdots$	$(1,0)$	$(1,0)$	$(1,0)$	$(0,1)$

Observe that row e is the only action in its region and as such cannot be eliminated, and that it backs the elimination of every column by b.

Now define $(\Gamma'_\varphi, X'_1, X'_2)$ with $\Gamma'_\varphi = (A'_1, A'_2, u')$ such that

$$\begin{aligned} A'_1 &= \{p, \neg p : p \text{ a variable in } \varphi\} \\ &\quad \cup \{(\lambda_i^1, i), (\lambda_i^2, i), (\lambda_i^3, i) : C_i \text{ a clause in } \varphi\} \\ &\quad \cup \{e\}, \\ A'_2 &= \{b\}, \end{aligned}$$

and the utility function u' and the partitions X'_1 and X'_2 are restricted appropriately to A'_1 and A'_2, i.e., $u' = u|_{A'_1 \times A'_2}$, $X'_1 = \{x \cap A'_1 : x \in X_1\} \setminus \{\emptyset\}$ and $X'_2 = \{x \cap A'_2 : x \in X_2\} \setminus \{\emptyset\}$. It is readily appreciated that no actions can be eliminated in $(\Gamma'_\varphi, X'_1, X'_2)$, i.e., that $(\Gamma'_\varphi, X'_1, X'_2)$ is irreducible.

We now prove that φ is satisfiable if and only if $(\Gamma'_\varphi, X'_1, X'_2)$ is reachable from $(\Gamma_\varphi, X_1, X_2)$.

For the direction from left to right, assume that φ is satisfiable and consider a satisfying assignment v. Start by eliminating, using column b, each column corresponding to a literal that is set to false by v. Subsequently, for each variable p, eliminate row $p{\downarrow}$ by row p or row $\neg p$. This is possible since either column p or column $\neg p$ have been eliminated in the first step. Next eliminate column a by column b. Since v is a satisfying assignment, there remains for each clause $C_i = (\lambda_i^1 \vee \lambda_i^2 \vee \lambda_i^3)$ a column λ_i^j, which now backs the elimi-

	c	d	u	v	w
a	$(0,2)$	$(2,0)$	$(1,1)$	$(2,0)$	$(0,2)$
b	$(1,1)$	$(2,0)$	$(0,2)$	$(2,0)$	$(0,2)$
x	$(0,2)$	$(0,2)$	$(0,2)$	$(2,0)$	$(1,1)$
y	$(0,2)$	$(1,1)$	$(0,2)$	$(2,0)$	$(1,1)$
z	$(0,2)$	$(2,0)$	$(2,0)$	$(1,1)$	$(1,1)$

Figure 29: Constant-sum game Γ illustrating that an elimination sequence need not remain valid if an action is eliminated. The elimination sequence (x, v, y, u, a) is valid for Γ, but the elimination sequence (d, x, v, y, u, a) is not.

nation of row C_i by row (λ_i^j, i). Finally eliminating by column b all other remaining columns, we reach subgame $(\Gamma'_\varphi, X'_1, X'_2)$.

For the direction from right to left, assume that $(\Gamma'_\varphi, X'_1, X'_2)$ is reachable from (Γ, X_1, X_2). Observe that this specifically requires the elimination of row $p{\downarrow}$ for each variable p occurring in φ, and recall that for this to be possible at least one of the columns p and $\neg p$ needs to be eliminated while column a is still present to back the elimination. Furthermore, row C_i must be eliminated for each $1 \leqslant i \leqslant k$, which can only take place by some row (λ_i^j, i) and backed by column λ_i^j, and only when column a is no longer present to block the elimination. We can thus define an assignment v^* that satisfies exactly those literals λ_i^j corresponding to columns present when column a is eliminated. It is readily appreciated that v^* is well-defined and satisfies φ. $\square$

Solvability is a special case of subgame reachability, and is tractable for two-player single-winner games, i.e., for constant-sum games which only allow outcomes $(0,1)$ and $(1,0)$ (Brandt et al., 2009b). Whether solvability is tractable in general constant-sum games remains an open question.

6.4.2 *Eliminability*

As we have seen, the iterated elimination of weakly dominated actions may depend on the order in which actions are eliminated. If an elimination sequence $\sigma = (\sigma_1, \ldots, \sigma_n)$ is valid for a game Γ, it does not automatically follow that σ remains valid if some dominated action d different from σ_1 is removed first. Consider for example the game Γ depicted in Figure 29 and the elimination sequence (x, v, y, u, a), which is valid for this game. Action d, which is itself dominated by action c, is the only action backing the elimination of x in Γ. Thus the elimination sequence (x, v, y, u, a) is no longer valid when d is eliminated first.

It turns out, however, that by delaying the elimination of x until y has been eliminated one can obtain an elimination sequence, viz. (v, y, x, u, a), that is valid for $\Gamma(d)$. We will see presently that this is just an example of a more general property of elimination sequences in two-player constant-sum games: given a valid elimination sequence σ and a dominated action d, one can carry out the elimination of d early and still find a valid elimination sequence that eliminates all the actions in σ, provided that one is prepared to postpone the elimination of some of these actions. This insight will be instrumental to the proof of Theorem 6.7, which states that the eliminability problem for two-player constant-sum games can be solved efficiently.

We need some auxiliary terminology and notation. Fix a game $\Gamma = (A_1, A_2, u)$, and let $\sigma = (\sigma_1, \dots, \sigma_n)$ be a sequence of actions. For σ to be a valid elimination sequence, there has to exist, for each action σ_i, an action δ_i of the same player and an action γ_i of the other player, both of which have not yet been eliminated, such that δ_i dominates σ_i at γ_i. Let $\delta = (\delta_1, \dots, \delta_n)$ and $\gamma = (\gamma_1, \dots, \gamma_n)$ be sequences of actions of Γ. We say that σ *is valid for* Γ *with respect to* δ *and* γ if, for each i with $1 \leqslant i \leqslant n$, action δ_i dominates σ_i at γ_i in $\Gamma(\sigma_1, \dots, \sigma_{i-1})$. We call an action σ_i an *obstacle* in σ with respect to δ and γ in Γ if δ_i does not dominate σ_i at γ_i in $\Gamma(\sigma_1, \dots, \sigma_{i-1})$. Obviously, there are no obstacles in σ with respect to δ and γ if and only if σ is valid with respect to δ and γ. An elimination sequence $\sigma = (\sigma_1, \dots, \sigma_n)$ will be called *weakly valid* with respect to an action sequence $\delta = (\delta_1, \dots, \delta_n)$ if, for all i with $1 \leqslant i \leqslant n$, it is the case that $\delta_i \in A \setminus \{\sigma_1, \dots, \sigma_{i-1}\}$ and no action in $A \setminus \{\sigma_1, \dots, \sigma_{i-1}\}$ blocks the elimination of σ_i by δ_i in $\Gamma(\sigma_1, \dots, \sigma_{i-1})$.

obstacle

weakly valid

We will show that for any constant-sum game Γ every elimination sequence $(\sigma_1, \dots, \sigma_n)$ that is weakly valid with respect to $(\delta_1, \dots, \delta_n)$ can be transformed into a valid elimination sequence, provided that the last action is not an obstacle, i.e., that there is an action actually backing the elimination of σ_n by δ_n in $\Gamma(\sigma_1, \dots, \sigma_{n-1})$. As a first step, the following lemma specifies a sufficient condition for the removal of an action from a weakly valid elimination sequence such that the sequence remains weakly valid and no new obstacles are created. Intuitively, this condition requires that if not eliminated, the action in question does not block any eliminations appearing later in the sequence.

Lemma 6.4. *Let $\Gamma = (A_1, A_2, u)$ be a two-player game, $\sigma = (\sigma_1, \dots, \sigma_n)$ and $\delta = (\delta_1, \dots, \delta_n)$ two action sequences such that σ is weakly valid for Γ with respect to δ. Let i be an index with $1 \leqslant i \leqslant n$ such that σ_i does not block the elimination of σ_j in $\Gamma(\sigma_1, \dots, \sigma_{i-1}, \sigma_{i+1}, \dots, \sigma_{j-1})$ for any j with $i < j \leqslant n$. Then, $(\sigma_1, \dots, \sigma_{i-1}, \sigma_{i+1}, \dots, \sigma_n)$ is weakly valid with respect to $(\delta_1, \dots, \delta_{i-1}, \delta_{i+1}, \dots, \delta_n)$. Moreover, the following holds for every index k with $1 \leqslant k \leqslant n$ and $k \neq i$, and for every action sequence $\gamma = (\gamma_1, \dots, \gamma_n)$: if σ_k is not an obstacle in σ with respect to δ and γ,* *then* σ_k *is not an obstacle in* $(\sigma_1, \ldots, \sigma_{i-1}, \sigma_{i+1}, \ldots, \sigma_n)$ *with respect to* $(\delta_1, \ldots, \delta_{i-1}, \delta_{i+1}, \ldots, \delta_n)$ *and* $(\gamma_1, \ldots, \gamma_{i-1}, \gamma_{i+1}, \ldots, \gamma_n)$ *either.*

Proof. Consider an arbitrary index m with $1 \leqslant m \leqslant n$ and $m \neq i$. First consider the case when $m < i$. Then, since σ is weakly valid with respect to δ, it follows immediately that no action in $A \setminus \{\sigma_1, \ldots, \sigma_{m-1}\}$ blocks the elimination of σ_m by δ_m. Now assume that $m > i$. Then, $\delta_m \in A \setminus \{\sigma_1, \ldots, \sigma_{m-1}\}$, and thus

$$\delta_m \in A \setminus \{\sigma_1, \ldots, \sigma_{i-1}, \sigma_{i+1}, \ldots, \sigma_{m-1}\}.$$

Moreover, since $m > i$, $\delta_m \neq \sigma_i$. It follows that no action in $A \setminus \{\sigma_1, \ldots, \sigma_{m-1}\}$ blocks the elimination of σ_m by δ_m in

$$\Gamma(\sigma_1, \ldots, \sigma_{i-1}, \sigma_{i+1}, \ldots \sigma_{m-1}).$$

By assumption, σ_i does not block this elimination either, and we may conclude that $(\sigma_1, \ldots, \sigma_{i-1}, \sigma_{i+1}, \ldots, \sigma_n)$ is weakly valid with respect to $(\delta_1, \ldots, \delta_{i-1}, \delta_{i+1}, \ldots, \delta_n)$.

For the second part of the claim, assume that σ_k is not an obstacle in σ with respect to δ and γ, i.e., δ_k dominates σ_k at γ_k in $\Gamma(\sigma_1, \ldots, \sigma_{k-1})$. Observe that $\gamma_k \in A \setminus \{\sigma_1, \ldots, \sigma_{k-1}\}$. The case when $k < i$ is trivial, so assume that $k > i$. We have already seen that

$$(\sigma_1, \ldots, \sigma_{i-1}, \sigma_{i+1}, \ldots, \sigma_n)$$

is weakly valid with respect to $(\delta_1, \ldots, \delta_{i-1}, \delta_{i+1}, \ldots, \delta_n)$. Moreover, $A \setminus \{\sigma_1, \ldots, \sigma_{k-1}\} \subseteq A \setminus \{\sigma_1, \ldots, \sigma_{i-1}, \sigma_{i+1}, \ldots, \sigma_{k-1}\}$. With action γ_k still available, δ_k dominates σ_k at γ_k in $\Gamma(\sigma_1, \ldots, \sigma_{i-1}, \sigma_{i+1}, \ldots, \sigma_{k-1})$. Thus, σ_k is not an obstacle. □

A corollary of Lemma 6.4 is that a valid elimination sequence remains valid after the removal of an action that blocks no other elimination if it remains in the game. Moreover, if an obstacle of an elimination sequence is moved to a position where it blocks no additional eliminations but where it can itself be eliminated, the number of obstacles in the sequence strictly decreases. As we will see next, this can be used to transform a weakly valid elimination sequence into a valid one, given that the last element of the former is not an obstacle.

Lemma 6.5. *Let* $\Gamma = (A_1, A_2, u)$ *be a constant-sum game. Let* a, b, *and* c *be distinct actions in* $A_1 \cup A_2$, *and* $\sigma = (\sigma_1, \ldots, \sigma_n)$ *and* $\delta = (\delta_1, \ldots, \delta_n)$ *action sequences with* $\sigma_n = a$ *and* $\delta_n = b$. *If* σ *is weakly valid with respect to* δ *in* Γ *and* b *dominates* a *at* c *in* $\Gamma(\sigma_1, \ldots, \sigma_{n-1})$, *then* a *is eliminable by* b *at* c *in* Γ.

Proof. Assume that σ is weakly valid with respect to δ and that b dominates a at c in $\Gamma(\sigma_1, \ldots, \sigma_{n-1})$. Let $\gamma = (\gamma_1, \ldots, \gamma_n)$ be an arbitrary action sequence with $\gamma_n = c$, and assume for contradiction that a

is not eliminable by b at c in Γ. Note that we may assume without loss of generality that σ, δ, and γ minimize the number of obstacles among all triples of action sequences with the above properties. We will derive a contradiction by showing that there exists a triple with strictly fewer obstacles.

Clearly, σ cannot be valid for Γ with respect to δ and γ, so there exists a smallest index i with $1 \leqslant i \leqslant n$ such that σ_i is an obstacle in σ with respect to δ and γ. By assumption, σ_n is not an obstacle with respect to δ and γ, and thus $i \neq n$. We distinguish two cases.

First assume that there is no index j with $i < j \leqslant n$ such that σ_i blocks the elimination of σ_j by δ_j in $\Gamma(\sigma_1, \dots, \sigma_{i-1}, \sigma_{i+1}, \dots, \sigma_{j-1})$. Then, by Lemma 6.4, $(\sigma_1, \dots, \sigma_{i-1}, \sigma_{i+1}, \dots, \sigma_n)$ is weakly valid with respect to $(\delta_1, \dots, \delta_{i-1}, \delta_{i+1}, \dots, \delta_n)$ and contains fewer obstacles with respect to $(\delta_1, \dots, \delta_{i-1}, \delta_{i+1}, \dots, \delta_n)$ and $(\gamma_1, \dots, \gamma_{i-1}, \gamma_{i+1}, \dots, \gamma_n)$ than σ does with respect to δ and γ. Moreover, since $i \neq n$, a is still dominated by b at c in $\Gamma(\sigma_1, \dots, \sigma_{i-1}, \sigma_{i+1}, \dots, \sigma_{n-1})$, a contradiction.

For the remainder of the proof we will thus assume that σ_i blocks the elimination of σ_j by δ_j in $\Gamma(\sigma_1, \dots, \sigma_{i-1}, \sigma_{i+1}, \dots, \sigma_{j-1})$ for some index j with $i < j \leqslant n$. Without loss of generality we may also assume that j is the smallest such index, and that $\sigma_i \in A_1$. Accordingly, $\delta_j, \sigma_j \in A_2$ and $u_2(\sigma_i, \delta_j) < u_2(\sigma_i, \sigma_j)$. It also holds that $\sigma_i, \sigma_j \in A \setminus \{\sigma_1, \dots, \sigma_{i-1}, \sigma_{i+1}, \dots, \sigma_{j-1}\}$, otherwise σ_i could not block the elimination of σ_j by δ_j in $\Gamma(\sigma_1, \dots, \sigma_{i-1}, \sigma_{i+1}, \dots, \sigma_{j-1})$. As σ is weakly valid with respect to δ, it follows that γ_i does not back the elimination of σ_i by δ_i in $\Gamma(\sigma_1, \dots, \sigma_{i-1})$. We will see, however, that there exists an index k with $i \leqslant k < j$ such that δ_k dominates σ_i at σ_j in $\Gamma(\sigma_1, \dots, \sigma_k)$, and that by delaying the elimination of σ_i until σ_k has been removed, σ_i ceases to be an obstacle while no additional ones are being created.

Define B as the smallest subset of $A \setminus \{\sigma_1, \dots, \sigma_{i-1}\}$ such that *(i)* $\sigma_i \in B$, and *(ii)* $\delta_k \in B$ whenever $\sigma_k \in B$ and δ_k blocks the elimination of σ_j by δ_j in $\Gamma(\sigma_1, \dots, \sigma_{k-1})$. Obviously, B is nonempty and finite. We may also assume that $B = \{\sigma_{i_1}, \dots, \sigma_{i_m}\}$, where $\sigma_{i_1} = \sigma_i$ and $\sigma_{i_k} = \delta_{i_{k-1}}$, for all k with $1 \leqslant k \leqslant m$. Further observe that by weak validity of σ with respect to δ, all actions in B must be eliminated before σ_j is, i.e., $i_m < j$.

Now consider the sequences

$$(\sigma'_1, \dots, \sigma'_n) = (\sigma_1, \dots, \sigma_{i-1}, \sigma_{i+1}, \dots, \sigma_{i_m}, \sigma_i, \sigma_{i_m+1}, \dots, \sigma_n),$$
$$(\delta'_1, \dots, \delta'_n) = (\delta_1, \dots, \delta_{i-1}, \delta_{i+1}, \dots, \delta_{i_m}, \delta_{i_m}, \delta_{i_m+1}, \dots, \delta_n),$$
$$(\gamma'_1, \dots, \gamma'_n) = (\gamma_1, \dots, \gamma_{i-1}, \gamma_{i+1}, \dots, \gamma_{i_m}, \sigma_j, \gamma_{i_m+1}, \dots, \gamma_n).$$

We will show that $(\sigma'_1, \dots, \sigma'_n)$ is weakly valid w.r.t. $(\delta'_1, \dots, \delta'_n)$ and, moreover, contains fewer obstacles in Γ with respect to $(\delta'_1, \dots, \delta'_n)$ and $(\gamma'_1, \dots, \gamma'_n)$ than σ does with respect to δ and γ.

$$
\begin{array}{r|ccc|}
 & \delta_j & & \sigma_j \\
\hline
\delta_{i_m} & u_1(\delta_{i_m},\delta_j) & \leqslant & u_1(\delta_{i_m},\sigma_j) \\
 & \rotatebox{-90}{$\geqslant$} & & \rotatebox{-90}{$\geqslant$} \\
\delta_{i_{m-1}} = \sigma_{i_m} & u_1(\delta_{i_{m-1}},\delta_j) & > & u_1(\delta_{i_{m-1}},\delta_j) \\
 & \rotatebox{-90}{$\geqslant$} & & \rotatebox{-90}{$\geqslant$} \\
\vdots & \vdots & \vdots & \vdots \\
 & \rotatebox{-90}{$\geqslant$} & & \rotatebox{-90}{$\geqslant$} \\
\delta_i = \delta_{i_1} = \sigma_{i_2} & u_1(\delta_i,\delta_j) & > & u_1(\delta_i,\sigma_j) \\
 & \rotatebox{-90}{$\geqslant$} & & \rotatebox{-90}{$\geqslant$} \\
\sigma_i = \sigma_{i_1} & u_1(\sigma_i,\delta_j) & > & u_1(\sigma_i,\sigma_j) \\
\hline
\end{array}
$$

Figure 30: Diagram illustrating the proof of Lemma 6.6

This yields a contradiction, because b also dominates a at c in $\Gamma(\sigma'_1,\dots,\sigma'_{n-1})$. To appreciate the latter, simply observe that the games $\Gamma(\sigma'_1,\dots,\sigma'_{n-1})$ and $\Gamma(\sigma_1,\dots,\sigma_{n-1})$ are identical and, since $i_m < j \leqslant n$, $\sigma'_n = \sigma_n = a$, $\delta'_n = \delta_n = b$, and $\gamma'_n = \gamma_n = c$.

Lemma 6.4 and the assumptions about σ_i imply that the sequence $(\sigma_1,\dots,\sigma_{i-1},\sigma_{i+1},\dots,\sigma_{i_m})$ is a weakly valid elimination sequence with respect to $(\delta_1,\dots,\delta_{i-1},\delta_{i+1},\dots,\delta_{i_m})$ in Γ. Moreover, for every index k with $i_m < k \leqslant n$, $\Gamma(\sigma_1,\dots,\sigma_{k-1})$ and

$$\Gamma(\sigma_1,\dots,\sigma_{i-1},\sigma_{i+1},\dots,\sigma_{i_m},\sigma_i,\sigma_{i_m+1},\dots,\sigma_{k-1})$$

are the same game, and in this game no elimination of σ_k by δ_k is blocked. To show that $(\sigma'_1,\dots,\sigma'_n)$ is weakly valid with respect to $(\delta'_1,\dots,\delta'_n)$ and contains fewer obstacles in Γ w.r.t. $(\delta'_1,\dots,\delta'_n)$ and $(\gamma'_1,\dots,\gamma'_n)$ than σ with respect to δ and γ, it thus suffices to show that σ_i is dominated by δ_{i_m} at σ_j in $\Gamma(\sigma_1,\dots,\sigma_{i-1},\sigma_{i+1},\dots,\sigma_{i_m})$.

Since Γ is a constant-sum game, $u_2(\sigma_i,\delta_j) < u_2(\sigma_i,\sigma_j)$ implies that $u_1(\sigma_i,\delta_j) > u_1(\sigma_i,\sigma_j)$. Furthermore, by definition of B, $u_1(\delta_{i_m},\delta_j) \geqslant u_1(\sigma_{i_m},\delta_j)$ and $u_1(\sigma_{i_{k+1}},\delta_j) \geqslant u_1(\sigma_{i_k},\delta_j)$ for every k with $1 \leqslant k < m$. Since i_m is the largest index for which $\sigma_{i_m} \in B$, it follows that δ_{i_m} does not block the elimination of σ_j by δ_j. Now recall that σ is weakly valid with respect to δ. Thus, $\delta_{i_m} \in A \setminus \{\sigma_1,\dots,\sigma_{i-1},\sigma_{i+1},\dots,\sigma_{i_m}\}$ and $\delta_j \in A \setminus \{\sigma_1,\dots,\sigma_{i-1},\sigma_{i+1},\dots,\sigma_j\}$. This implies $u_2(\delta_{i_m},\delta_j) \geqslant u_2(\delta_{i_m},\sigma_j)$ and, since Γ is constant-sum, $u_1(\delta_{i_m},\delta_j) \leqslant u_1(\delta_{i_m},\sigma_j)$. The resulting situation is depicted in Figure 30, from which it can easily be read off that

$$u_1(\sigma_i, \sigma_j) < u_1(\sigma_i, \delta_j) \leqslant u_1(\delta_i, \delta_j) \leqslant \cdots \leqslant u_1(\delta_{i_m}, \delta_j) \leqslant u_1(\delta_{i_m}, \sigma_j).$$

In particular, $u_1(\delta_{i_m}, \sigma_j) > u_1(\sigma_i, \sigma_j)$, i.e., σ_j backs the elimination of σ_i by δ_{i_m} in $\Gamma(\sigma_1, \dots, \sigma_{i-1}, \sigma_{i+1}, \dots, \sigma_{i_m})$. Since σ is weakly valid with respect to δ, none of the actions in $A \setminus \{\sigma_1, \dots, \sigma_{i_m}\}$ block the elimination σ_i by δ_i in $\Gamma(\sigma_1, \dots, \sigma_{i-1})$. By transitivity of the dominance relation, the same is true for the elimination of σ_i by δ_{i_k} in $\Gamma(\sigma_1, \dots, \sigma_{i-1}, \sigma_{i+1}, \dots, \sigma_{i_k})$ for any k with $1 \leqslant k \leqslant i_m$. It follows that σ_i is dominated by δ_{i_m} at σ_j in $\Gamma(\sigma_1, \dots, \sigma_{i-1}, \sigma_{i+1}, \dots, \sigma_{i_m})$, which completes the proof. □

We have seen in the beginning of this section that the elimination of an action can turn a valid elimination sequence into one that is only weakly valid. Using Lemma 6.5, we will now show that the *existence* of an elimination sequence ending with a particular action a is not affected by such an earlier elimination, given that the eliminated action is not directly involved in the elimination of a.

Lemma 6.6. *Let $\Gamma = (A_1, A_2, u)$ be a constant-sum game. Let a, b, and c be distinct actions in $A_1 \cup A_2$, and σ a valid elimination sequence for Γ not containing a, b, or c. Then, if a is eliminable by b at c in Γ, a is still eliminable by b at c in $\Gamma(\sigma)$.*

Proof. Assume that a is eliminable by b at c in Γ. Then there exist action sequences $\sigma' = (\sigma'_1, \dots, \sigma'_n)$, $\delta' = (\delta'_1, \dots, \delta'_n)$, and $\gamma' = (\gamma'_1, \dots, \gamma'_n)$ with $\sigma'_n = a$, $\delta'_n = b$, and $\gamma'_n = c$ such that σ' is valid with respect to δ' and γ'. Let $\sigma = (\sigma_1, \dots, \sigma_m)$, and let $\delta = (\delta_1, \dots, \delta_m)$ and $\gamma = (\gamma_1, \dots, \gamma_m)$ be action sequences such that σ is valid with respect to δ and γ. By transitivity of the dominance relation, we may further assume for each i with $1 \leqslant i \leqslant m$ that $\sigma_i \notin \{\delta'_1, \dots, \delta'_n\}$.[5] Since σ' is valid with respect to δ' and γ' and σ is valid with respect to δ and γ, it follows that $(\sigma_1, \dots, \sigma_m, \sigma'_1, \dots, \sigma'_n)$ is weakly valid with respect to $(\delta_1, \dots, \delta_m, \delta'_1, \dots, \delta'_n)$. Moreover, by the assumption that $a, b, c \notin \{\sigma_1, \dots, \sigma_m\}$, action a is still dominated by b at c in $\Gamma(\sigma_1, \dots, \sigma_m, \sigma'_1, \dots, \sigma'_{n-1})$. Lemma 6.5 now gives us the desired result. □

5 To appreciate this, suppose that $\sigma_i = \delta'_j$ for some j with $1 \leqslant j \leqslant n$, i.e., σ_i dominates σ'_j in $\Gamma(\sigma'_1, \dots, \sigma'_{j-1})$. Define for each k with $1 \leqslant k \leqslant n$ action δ''_k as follows:

$$\delta''_k = \begin{cases} \delta_i & \text{if } k = 1, \\ \delta'_{k-1} & \text{if } \delta''_{k-1} = \sigma_{k-1}, \\ \delta''_{k-1} & \text{otherwise.} \end{cases}$$

Now observe that generally $\delta''_k \in A \setminus \{\sigma'_1, \dots, \sigma'_{k-1}\}$. Moreover, by transitivity of the dominance relation, σ'_j is also dominated by δ''_j in $\Gamma(\sigma'_1, \dots, \sigma'_{j-1})$ and can act as a proxy for δ_i.

Intuitively, Lemma 6.6 says the following: to eliminate a particular action a by b backed by c, one can eliminate dominated actions more or less in an arbitrary way; one just has to be careful not to eliminate actions b and c. On the basis of this observation, we obtain the main result of this section.

Theorem 6.7. *The problem of deciding whether a given action of a constant-sum game is eliminable can be solved in polynomial time.*

Proof. Let a be the action to be eliminated, and assume without loss of generality that $a \in A_1$. Consider the algorithm that performs the following steps:

1. Compose a list $(b^1, c^1), \ldots, (b^k, c^k)$ of all pairs $(b^i, c^i) \in A_1 \times A_2$ such that c^i backs the elimination of a by b^i.
2. For each i with $1 \leqslant i \leqslant k$, arbitrarily eliminate actions *distinct from* b^i *and* c^i until no more eliminations are possible. Let $\sigma^i = (\sigma^i_1, \ldots, \sigma^i_{m_i})$ denote the resulting valid elimination sequence.
3. If for some i with $1 \leqslant i \leqslant k$, action a is eliminated in σ^i, i.e., $a \in \{\sigma^i_1, \ldots, \sigma^i_{m_i}\}$, output "yes," otherwise "no."

Obviously, this algorithm runs in polynomial time. If action a is not eliminable, the algorithm cannot find a valid elimination sequence and will always output "no." If, on the other hand, a is eliminable by b at c for some actions b and c, the algorithm will check this at some point. If it does so, it will make sure not to eliminate actions b and c. Thus, by Lemma 6.6, a will remain eliminable by b at c as more and more actions are eliminated. Since the overal number of actions is finite, a will at some point become dominated by b at c and can subsequently be eliminated. □

6.5 WIN-LOSE GAMES

win-lose games

win-lose games with at most one winner

Conitzer and Sandholm (2005) have shown that both subgame reachability and eliminability are NP-complete in *win-lose games*, i.e., games which only allow outcomes $(0, 0)$, $(0, 1)$, $(1, 0)$, and $(1, 1)$. As both win-lose and constant-sum games generalize single-winner games, it is interesting to compare these results with those for constant-sum games in the previous section. It turns out that the results of Conitzer and Sandholm even hold for *win-lose games with at most one winner*, i.e., for games with outcomes $(0, 0)$, $(0, 1)$, and $(1, 0)$. For subgame reachability, this follows from Theorem 6.3, which shows NP-completeness even for games with outcomes in $\{(0, 1), (1, 0)\}$. For eliminability, we modify the construction used in the proof of Theorem 6.3 to provide a reduction from *3SAT*.

Theorem 6.8. *Deciding whether a given action of a two-player game with outcomes in* $\{(0, 0), (0, 1), (1, 0)\}$ *is eliminable is NP-complete.*

Proof. Membership in NP is obvious.

Hardness is shown using a reduction from *3SAT*. By Lemma 6.1, it suffices to give a reduction for regionalized games. Consider a *3CNF* φ, and recall the regionalized game $(\Gamma_\varphi, X_1, X_2)$ with $\Gamma_\varphi = (A_1, A_2, u)$ defined in the proof of Theorem 6.3. This game only involved the outcomes $(0,1)$ and $(1,0)$. Define a regionalized game $(\Gamma'_\varphi, X'_1, X'_2)$ such that $A'_1 = A_1$, $A'_2 = A_2 \cup \{c, d^*\}$, $X'_1 = X_1$, and $X'_2 = X_2 \cup \{\{c, d^*\}\}$. The utility function u' extends u, i.e., $u'(a,b) = u(a,b)$ for all $a \in A_1$ and $b \in A_2$. Payoffs for columns c and d^* are as follows:[6]

	c	d^*
p	$(0,0)$	$(0,0)$
$\neg p$	$(0,0)$	$(0,0)$
$p{\downarrow}$	$(0,0)$	$(0,1)$

	c	d^*
(λ_i^1, i)	$(0,0)$	$(0,0)$
(λ_i^2, i)	$(0,0)$	$(0,0)$
(λ_i^3, i)	$(0,0)$	$(0,0)$
C_i	$(0,0)$	$(0,1)$

	c	d^*
e	$(0,1)$	$(0,0)$

An example of the resulting game is given in Figure 31.

Observe that for all actions $x \in A'_1$, $u_1(x,c) = u_1(x,d^*) = 0$. The additional actions c and d^* thus do not back or block any eliminations. Furthermore, c and d^* constitute a separate region and can therefore neither eliminate nor be eliminated by any of the actions in A_2. Finally, column d^* is dominated by c at e if and only if action $p{\downarrow}$ for each variable p and action C_i for each clause C_i have been eliminated. By virtue of an argument analogous to the one used in the proof of Theorem 6.3, we find that action d^* is eliminable if and only if φ is satisfiable. This completes the proof. □

Conitzer and Sandholm (2005) use a reduction from eliminability to solvability to show intractability of the latter in win-lose games. Their construction, however, hinges on the presence of the outcome $(1,1)$. For the more restricted class of games without $(1,1)$ as an outcome we instead reduce directly from *3SAT* and exploit the internal structure of the construction used in the proof of Theorem 6.8.

Theorem 6.9. *Deciding whether a two-player game with outcomes* $(0,0)$, $(0,1)$, $(1,0)$ *is solvable is NP-complete.*

Proof. Membership in NP is straightforward.

Hardness is shown using a reduction from *3SAT*. Consider a *3CNF* formula φ, and let $(\Gamma'_\varphi, X'_1, X'_2)$ with $\Gamma'_\varphi = (A'_1, A'_2, u')$ be the regionalized game with outcomes in $\{(0,0),(0,1),(1,0)\}$ defined in Theorem 6.8, with additional copies f and g^* of the actions c and d^* such

6 By setting $u'_1(x,c) = u'_1(x,d^*) = 1$ instead, one obtains a construction proving the intractability of the eliminability problem for games with outcomes in $(0,1)$, $(1,0)$ and $(1,1)$.

	p	$\neg p$	q	$\neg q$	r	$\neg r$	a	b	c	d^*
p	$(1,0)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(1,0)$	$(0,1)$	$(0,0)$	$(0,0)$
$\neg p$	$(0,1)$	$(1,0)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(1,0)$	$(0,1)$	$(0,0)$	$(0,0)$
$p{\downarrow}$	$(1,0)$	$(1,0)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,0)$	$(0,1)$
q	$(0,1)$	$(0,1)$	$(1,0)$	$(0,1)$	$(0,1)$	$(0,1)$	$(1,0)$	$(0,1)$	$(0,0)$	$(0,0)$
$\neg q$	$(0,1)$	$(0,1)$	$(0,1)$	$(1,0)$	$(0,1)$	$(0,1)$	$(1,0)$	$(0,1)$	$(0,0)$	$(0,0)$
$q{\downarrow}$	$(0,1)$	$(0,1)$	$(1,0)$	$(1,0)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,0)$	$(0,1)$
r	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(1,0)$	$(0,1)$	$(1,0)$	$(0,1)$	$(0,0)$	$(0,0)$
$\neg r$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(1,0)$	$(1,0)$	$(0,1)$	$(0,0)$	$(0,0)$
$r{\downarrow}$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(1,0)$	$(1,0)$	$(0,1)$	$(0,1)$	$(0,0)$	$(0,1)$
p	$(1,0)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,0)$	$(0,0)$
q	$(0,1)$	$(0,1)$	$(1,0)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,0)$	$(0,0)$
$\neg r$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(1,0)$	$(0,1)$	$(0,1)$	$(0,0)$	$(0,0)$
$p \vee q \vee \neg r$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(1,0)$	$(0,1)$	$(0,0)$	$(0,1)$
$\neg p$	$(0,1)$	$(1,0)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,0)$	$(0,0)$
q	$(0,1)$	$(0,1)$	$(1,0)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,0)$	$(0,0)$
r	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(1,0)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,0)$	$(0,0)$
$\neg p \vee q \vee r$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(1,0)$	$(0,1)$	$(0,0)$	$(0,1)$
$\neg p$	$(0,1)$	$(1,0)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,0)$	$(0,0)$
$\neg q$	$(0,1)$	$(0,1)$	$(0,1)$	$(1,0)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,0)$	$(0,0)$
$\neg r$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(1,0)$	$(0,1)$	$(0,1)$	$(0,0)$	$(0,0)$
$\neg p \vee \neg q \vee \neg r$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(1,0)$	$(0,1)$	$(0,0)$	$(0,1)$
e	$(1,0)$	$(1,0)$	$(1,0)$	$(1,0)$	$(1,0)$	$(1,0)$	$(1,0)$	$(0,1)$	$(0,1)$	$(0,0)$

Figure 31: Construction used in the proof of Theorem 6.8. Example for the formula $(p \vee q \vee \neg r) \wedge (\neg p \vee q \vee r) \wedge (\neg p \vee \neg q \vee \neg r)$.

	a_2^1	$\cdots$	a_2^m	c	d^*	f	g^*	x_2^1	$\cdots$	y_2^4	z_2^1	z_2^2	z_2^3
a_1^1	$\cdot$	$\cdots$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdots$	$\cdot$	$(0,1)$	$(0,1)$	$(0,1)$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
a_1^n	$\cdot$	$\cdots$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdots$	$\cdot$	$(0,1)$	$(0,1)$	$(0,1)$
e	$\cdot$	$\cdots$	$\cdot$	$(0,1)$	$(0,0)$	$(0,1)$	$(0,0)$	$\cdot$	$\cdots$	$\cdot$	$(0,1)$	$(0,1)$	$(0,1)$
x_1^1	$\cdot$	$\cdots$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdots$	$\cdot$	$(0,1)$	$(0,1)$	$(0,1)$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
y_1^4	$\cdot$	$\cdots$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdots$	$\cdot$	$(0,1)$	$(0,1)$	$(0,1)$
z_1^1	$(0,0)$	$(0,0)$	$(0,0)$	$(0,0)$	$(0,0)$	$(0,0)$	$(0,0)$	$(0,0)$	$(0,0)$	$(0,0)$	$(0,1)$	$(1,0)$	$(1,0)$
z_1^2	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(1,0)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(1,0)$
z_1^3	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(1,0)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,0)$	$(1,0)$	$(0,1)$
z_1^4	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,1)$	$(1,0)$	$(0,1)$	$(0,1)$	$(0,1)$	$(0,0)$	$(0,1)$	$(1,0)$

Figure 32: Construction used in the proof of Theorem 6.9

that $\{f, g^*\}$ constitutes a separate region. Thus, A_1' and X_1' are as before, while for the column player we have

$$A_2' = \{p, \neg p : p \text{ a variable in } \varphi\} \cup \{a, b, c, d^*, f, g^*\},$$
$$X_2' = \{\{p, \neg p : p \text{ a variable in } \varphi\} \cup \{a, b\}, \{c, d^*\}, \{f, g^*\}\},$$

and $u'(x, f) = u'(x, c)$ and $u'(x, g^*) = u'(x, d^*)$ for each $x \in A_1'$. By the same reasoning as in the proof of Theorem 6.8, both d^* and g^* are eliminable if and only if φ is satisfiable.

Now consider the game $\Gamma_\varphi'' = (A_1'', A_2'', u'')$ without regions corresponding to $(\Gamma_\varphi', X_1', X_2')$ as defined in the proof of Lemma 6.1, and define $\Gamma_\varphi''' = (A_1''', A_2''', u''')$ with $A_1''' = A_1'' \cup \{z_1^1, z_1^2, z_1^3, z_1^4\}$ and $A_2''' = A_2'' \cup \{z_2^1, z_2^2, z_2^3\}$. Let $u'''(x, y) = u''(x, y)$ for all $(x, y) \in A_1'' \times A_2''$, the payoffs for the remaining action profiles in $A_1''' \times A_2'''$ are shown in Figure 32.

We make the following observations about the game Γ_φ'''.

1. As long as columns d^* and g^* are not eliminated, the actions in $\{z_1^1, z_1^2, z_1^3, z_1^4\} \cup \{z_2^1, z_2^2, z_2^3\}$ do not dominate and are not dominated by any action in the game.

2. Actions z_1^2 and z_1^3 back the elimination of d^* by c, and z_1^4 backs the elimination of g^* by f. However, since action e also backs the same eliminations, and since action e itself is not eliminable in Γ_φ'', this does not make any additional eliminations possible as long as d^* and g^* have not been eliminated.

We now claim that Γ'''_φ can be solved, with (z_1^1, z_2^1) as the remaining action profile, if and only if φ is satisfiable.

For the direction from left to right, assume that φ is unsatisfiable. Then, using the same arguments as in the proof of Theorem 6.8 actions d^* and g^* cannot be eliminated. Hence, by 1, the game Γ'''_φ is not solvable.

For the direction from right to left, assume that φ is satisfiable. Again by the same arguments as in the proof of Theorem 6.8, columns d^* and g^* can be eliminated. Then, rows z_1^2, z_1^3 and z_1^4 can be eliminated by row z_1^1, followed by the elimination of columns a_2^1 through y_2^4 and z_2^3 by column z_2^1. Finally, row z_1^1 can eliminate all other remaining rows, and the elimination of column z_2^2 solves the game. □

6.6 RELATED WORK

Marx and Swinkels (1997) identify a condition under which all irreducible subgames that are reachable via iterated weak dominance are equivalent in terms of the payoff profiles that can be obtained, i.e., differ only by the addition or removal of identical actions and the renaming of actions. Since the condition is satisfied by constant-sum games, we can decide in polynomial time which payoff profiles of a constant sum game can still be obtained after the iterated removal of weakly dominated actions, by simply eliminating dominated actions arbitrarily.

This, however, does not imply any of our results, because it does not discriminate between actions that yield identical payoffs for *some* reachable subgame. In fact, Theorem 6.3 tells us that reachability of a given subgame is NP-hard to decide even in constant-sum games. The conceptual difference between our work and that of Marx and Swinkels is thus tightly linked to the question whether one is interested in *action profiles* or *payoff profiles* as "solutions" of a game. It may be argued that the computational gap between both concepts is of particular interest in this context.

6.7 SUMMARY

We have investigated the computational complexity of iterated weak dominance in two-player constant-sum games. In particular, we have shown that eliminability of an action can be decided in polynomial time, whereas deciding reachability of a given subgame is NP-complete. We have further shown the NP-completeness of typical problems associated with iterated dominance in win-lose games *with at most one winner*. Table 6 provides an overview of our results, and related results obtained earlier. In win-lose games an action is dominated by a mixed strategy if and only if it is dominated by a pure strategy (Conitzer and Sandholm, 2005). All of our results apart from Theorem 6.7 thus immediately extend to iterated W^*-dominance.

	Subgame reachability	Eliminability	Solvability
$\{(0,1),(1,0)\}$	NP-complete [a]	in P [b]	in P [d]
Constant-Sum	NP-complete [a]	in P [b]	?
$\{(0,0),(0,1),(1,0)\}$	NP-complete [a]	NP-complete [c]	NP-complete [e]
Win-Lose	NP-complete [f]	NP-complete [f]	NP-complete [f]
General	NP-complete [g]	NP-complete [g]	NP-complete [g]

[a] Theorem 6.3
[b] Theorem 6.7
[c] Theorem 6.8
[d] Brandt et al. (2009b)
[e] Theorem 6.9
[f] Conitzer and Sandholm (2005)
[g] Gilboa et al. (1993)

Table 6: Computational complexity of IWD in two-player games

Part II

SOCIAL CHOICE

7 PREFERENCE AGGREGATION

In this chapter, we introduce a general framework of preference aggregation and define most of the social choice functions that will be considered throughout the second part of this thesis. For more detailed accounts, we refer to the textbooks by Moulin (1988a, Chapters 9–11), Austen-Smith and Banks (2000), and Gaertner (2006). Furthermore, the excellent survey by Plott (1976) and Chapter 1 of Taylor's (2005) book are highly recommended.

7.1 PREFERENCES

According to Riker (1986, p. xi), social choice theory is "the description and analysis of the way that the preferences of individual members of a group are amalgamated into a decision of a group as a whole." Functions that map individual preferences to group decisions are called *social choice functions*, and we will see several examples in Section 7.2. As a natural first step, however, we need to deal with the question how individual preferences can be represented.

In the following, we use the term *voter* to refer to a an individual agent that has preferences over a finite set A of *alternatives*. The set A can be thought of as a collection of mutually exclusive outcomes, from which elements are to be chosen in accordance with the preferences of the voters. The usual assumption in social choice theory, and indeed the one we follow here, is that preferences of voter i are given by a *preference relation*, i.e., a complete binary relation $R_i \subseteq A \times A$. The interpretation of $(a, b) \in R_i$, denoted by $a \; R_i \; b$, is that voter i values alternative a at least as much as alternative b.[1] Completeness means that a voter is able to compare any pair of alternatives: for all $a, b \in A$, either $a \; R_i \; b$ or $b \; R_i \; b$ or both. In the latter case, the voter is said to be *indifferent* between a and b.

voter

alternatives

preference relation

completeness

Apart from completeness, two further restrictions are often imposed on preference relations. *Transitivity* requires that for all $a, b, c \in A$, if $a \; R_i \; b$ and $b \; R_i \; c$ then $a \; R_i \; c$. *Antisymmetry* requires that for all *distinct* a and b in A, $a \; R_i \; b$ implies *not* $b \; R_i \; a$, i.e., $(b, a) \notin R_i$. While transitivity of preferences is a standard assumption in economic the-

transitivity

antisymmetry

1 Modeling preferences as binary relations precludes the expression of preference intensities as well as interpersonal welfare comparisons. These issues are the subject of a long and ongoing debate in economic theory. We refer to Elster and Roemer (1991) and Binmore (2009) for general treatments, and to Plott (1976, pp. 539–543) for illuminating remarks in the context of social choice theory. For (im)possibility results when preference intensities and/or interpersonal welfare comparisons are possible, see Schwartz (1970) and Roberts (1980).

ory,[2] antisymmetry is a technical condition that often facilitates definitions and proofs. Since we mainly deal with social choice functions that are based on pairwise comparisons, transitivity is usually not required. On the other hand, we often impose antisymmetry for the sake of exposition. We use the term *strict preferences* to refer to antisymmetric (but not necessarily transitive) preference relations. If preference relations are both antisymmetric and transitive, we speak of *linear preferences*. At the beginning of each of the following chapters, we will specify which—if any—restrictions on preferences need to be imposed for the results in the respective chapter to hold.

strict preferences

linear preferences

preference profile

We let $N = \{1, \dots, n\}$ denote the set of voters. A *preference profile* $R = (R_1, \dots, R_n)$ is an n-tuple containing a preference relation R_i for every voter $i \in N$. If preferences are linear, a preference profile can conveniently be represented by a table whose i*th* column corresponds to the preferences of voter i. The profile in Figure 33 will be the running example of this chapter.

R_1	R_2	R_3	R_4	R_5
e	b	b	e	d
d	c	c	a	c
a	e	e	b	a
b	a	a	d	b
c	d	d	c	e

Figure 33: A preference profile $R = (R_1, R_2, R_3, R_4, R_5)$ with $n = 5$ voters. The set of alternatives is given by $A = \{a, b, c, d, e\}$. Columns correspond to preferences of voters. For instance, the first voter has linear preferences $e \; R_1 \; d \; R_1 \; a \; R_1 \; b \; R_1 \; c$.

A preference profile completely describes the preferences of the voters, and serves as the input of a social choice function. In the next section we classify social choice functions with respect to the amount of information they require. The reason is that, more often than not, a social choice function uses only part of the information that is encoded in a preference profile. For instance, knowing each voter's most preferred candidate is sufficient to determine the *plurality winner*. For most of the social choice functions considered in this thesis, different notions of *pairwise comparisons* are instrumental. Let us introduce some important notation.

plurality

pairwise comparisons

For a given preference profile $R = (R_1, \dots, R_n)$ and two distinct alternatives $a, b \in A$, define

$$n_R(a, b) = |\{i \in N : a \; R_i \; b\}| \quad \text{and} \quad m_R(a, b) = n_R(a, b) - n_R(b, a).$$

majority relation

The *majority relation* $R_M \subseteq A \times A$ is then defined by

$$a \; R_M \; b \quad \text{if and only if} \quad m_R(a, b) \geqslant 0.$$

2 For instance, see Mas-Colell et al. (1995, Definition 1.B.1).

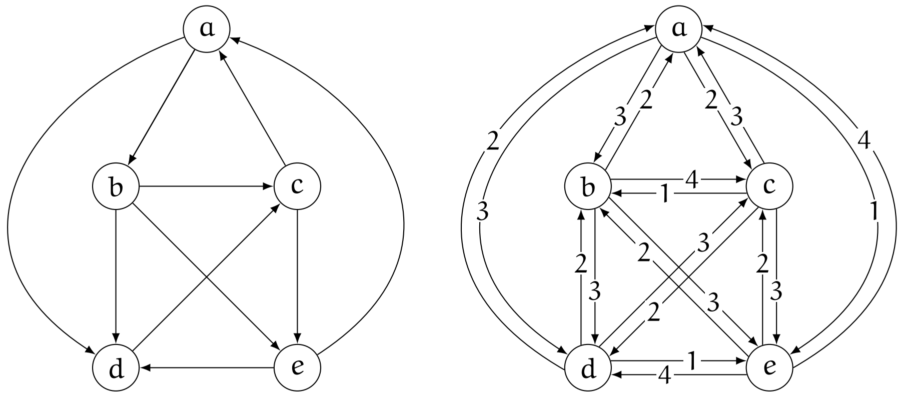


Figure 34: The majority graph (left) and the weighted tournament (right) for the preference profile R specified in Figure 33

Condorcet winner

Let P_M denote the strict part of R_M, i.e., $P_M = \{(a, b) : m_R(a, b) > 0\}$. A *Condorcet winner* is an alternative a that is preferred to any other alternative in A by a strict majority of voters, i.e., $a \; P_M \; b$ for all alternatives $b \in A \setminus \{a\}$.

tournament

weighted tournament

We frequently employ concepts from graph theory to conveniently represent pairwise comparisons between alternatives. A *tournament* is a complete and asymmetric directed graph. By a *weighted tournament* we here understand a pair (V, w) where V is a finite set and $w : V \times V \to \mathbb{N}$ is a weight function such that for all $a, b \in V$ with $a \neq b$ we have $w(a, b) + w(b, a) = n$ for some $n \in \mathbb{N}$.

majority graph

The *majority graph* $G(R)$ of a preference profile R is the asymmetric directed graph $G(R) = (A, P_M)$ on A whose edges are induced by the strict majority relation. Whenever there are no majority ties (as, for example, in the case of an odd number of voters with strict preferences), $G(R)$ is complete and therefore a tournament. Likewise, a weighted tournament (A, w) can be used to represent weighted pairwise comparisons within A by letting $w(a, b) = n_R(a, b)$. The weighted tournament representation works for any number of voters, but requires strict preferences: otherwise, $n_R(a, b)$ and $n_R(b, a)$ might not add up to the same number for all pairs $a, b \in A$. Figure 34 presents the majority graph and the weighted tournament for the example preference profile given above.

ranking

top element

A couple of further notations will be used in the following chapters, and we introduce them here. A *ranking* of a finite set X is a transitive relation $L \subseteq X \times X$ such that for each pair of distinct alternatives $x, y \in X$ either $x \; L \; y$ or $x \; L \; y$.[3] $\mathscr{L}(X)$ denotes the set of all rankings of X. The *top element* of a ranking $L \in \mathscr{L}(X)$, denoted by *top*(L), is the unique element $x \in X$ such that $x \; L \; y$ for all $y \in X \setminus \{x\}$.

---

3 For example, the strict part of a linear preference relation is a ranking of the set A of alternatives.

For a preference relation R_i, we let $R_i^{\leftarrow}$ denote the preference relation where all preferences are reversed, i.e., $a\ R_i^{\leftarrow}\ b$ if and only if $b\ R_i\ a$.

distance Furthermore, the *distance* $\delta(R_i, R_i')$ between two preference relations R_i and R_i' is defined as the number of (unordered) pairs of alternatives on which they disagree. If preferences are strict, $\delta(R_i, R_i') = |R_i \setminus R_i'| = |R_i' \setminus R_i|$. The distance between two preference profiles R and R' is defined as $\delta(R, R') = \sum_{i=1}^{n} \delta(R_i, R_i')$.

7.2 SOCIAL CHOICE FUNCTIONS

A social choice function maps a preference profile to a set of socially preferred alternatives.

Definition 7.1. *A* social choice function (SCF) *is a function* f *that maps a preference profile* R *to a nonempty subset of alternatives* $f(R) \subseteq A$.

This definition implicitly assumes that the set A of alternatives is fixed. Since this constitutes a departure from the classic social choice literature, let us relate our definition to the standard one. In the classic setting, there is a (possibly infinite) *universe* U of alternatives over which voters entertain preferences, and the input of an SCF consists of a preference profile (defined on U) *and* a finite subset $A \subseteq U$, usually called the *feasible set* or *agenda*. Most of the SCFs we consider in this thesis have been formulated in that more general setting. There are two reasons why we use the simplified definition with a fixed agenda.[4] First, we do not study any consistency conditions that relate choices from different agendas. Second, all SCFs in this thesis—if formulated in the more general setting—satisfy a condition that Fishburn (1973, p. 6) refers to as *independence of infeasible alternatives* and that resembles Arrow's IIA condition for social welfare functions. The condition requires that the choice from a feasible set $A \subseteq U$ only depends on preferences of voters among alternatives in A. For the purposes of this thesis, it does therefore not make a difference which setting we adopt, and we use the simpler one. In the context of tournament solutions, however, we *do* consider properties that relate different agendas. Consequently, tournament solutions are defined for a variable agenda in Chapter 10.

agenda

IIA

We go on to classify SCFs according to the amount of information that is required in order to determine the choice set.

7.2.1 *C1 Functions*

An SCF is called *Condorcet-consistent* if it uniquely selects the Condorcet winner whenever it exists. Fishburn (1977) classified

4 Some authors like Taylor (2005) and Moulin (2003) use the term *voting rule* to refer to SCFs with a fixed agenda.

Condorcet-consistent SCFs with respect to their informational requirements. We adopt Fishburn's terminology, but generalize the classification to also include SCFs that are not Condorcet-consistent. Three classes will be considered: C1, C2, and C3. The first class consists of all SCFs that can be evaluated by only looking at the majority relation.

Fishburn's classification

Definition 7.2. *An SCF* f *is a* C1 function *if* $f(R) = f(R')$ *for all preference profiles* R *and* R' *such that* $R_M = R'_M$.

We will now introduce a number of well-known C1 functions. Throughout this thesis, we will denote SCFs with capital letter abbreviations. In the margin, we will specify the output of the respective SCF for the preference profile R specified in Figure 33.

choice sets for the running example

TRIVIAL RULE (TRIV) The function *TRIV* returns the set A of all alternatives.

$TRIV(R) = \{a, b, c, d, e\}$

COPELAND (CO) The *Copeland score* of an alternative a is defined as the number of alternatives that a beats in pairwise majority comparisons, i.e., $s_{CO}(a) = |\{b \in A : a \; P_M \; b\}|$. The function *CO* returns the alternatives with the highest Copeland score.[5]

$CO(R) = \{b\}$

CONDORCET RULE (COND) The function *COND* returns the Condorcet winner if it exists, and the set A of all alternatives otherwise.

$COND(R) = \{a, b, c, d, e\}$

Let R^*_M denote the transitive closure of the majority relation, i.e., $a \; R^*_M \; b$ if and only if there exists $k \in \mathbb{N}$ and $a_1, \dots, a_k \in A$ with $a_1 = a$ and $a_k = b$ such that $a_i \; R_M \; a_{i+1}$ for all $i < k$.

TOP CYCLE (TC) The top cycle rule *TC* returns the maximal elements of R^*_M, i.e., $TC(R) = \{a \in A \mid a \; R^*_M \; b \text{ for all } b \in A\}$.[6]

$TC(R) = \{a, b, c, d, e\}$

The following two SCFs are defined via a subrelation of the majority relation, the *covering* relation. Let $B \subseteq A$ and $a, b \in B$. We say that a *covers* b *in* B if the following three conditions are satisfied:

covering

(*i*) $a \; P_M \; b$,

(*ii*) $b \; P_M \; c$ implies $a \; P_M \; c$ for all $c \in B \setminus \{a, b\}$, and

(*iii*) $c \; P_M \; a$ implies $c \; P_M \; b$ for all $c \in B \setminus \{a, b\}$.

UNCOVERED SET (UC) The function *UC* returns the set of all uncovered alternatives, i.e., the set of alternatives a such that there exists no alternative b that covers a in A.

$UC(R) = \{a, b, c, e\}$

5 The Copeland score of an alternative a is sometimes defined as the difference between $|\{b \in A : a \; P_M \; b\}|$ and $|\{b \in A : b \; P_M \; a\}|$. Whenever there are no majority ties, both definitions yield identical choice sets.

6 The top cycle rule is also known as *weak closure maximality*, *GETCHA*, or the *Smith set* (Good, 1971; Smith, 1973; Schwartz, 1986).

$MC(R) = \{a, b, c\}$

MINIMAL COVERING SET (MC) A subset $C \subseteq A$ is called a *covering set* if for all alternatives $b \in A \setminus C$, there exists $a \in C$ such that a covers b in $C \cup \{b\}$. Dutta (1988) and Dutta and Laslier (1999) have shown that there always exists a unique *minimal* covering set. The function *MC* returns exactly this set.

majority game

Interestingly, one can also use game-theoretic concepts to construct tournament solutions. Define the *majority game* Γ_R of a preference profile R as the matrix game in which the set of actions for both players is given by A and payoffs are defined as follows. If the first player chooses a and the second player chooses b, the payoff for the first player is 1 if $a\ P_M\ b$, -1 if $b\ P_M\ a$, and 0 otherwise. Figure 35 shows the majority game for our running example. Recall that the *essential set* (page 24) contains all actions that are played with positive probability in some Nash equilibrium of the game.

$BP(R) = \{a, b, c\}$

BIPARTISAN SET (BP) The function *BP* returns $ES_1(\Gamma_R)$ (Laffond et al., 1993b; Dutta and Laslier, 1999).

	a	b	c	d	e
a	0	1	−1	1	−1
b	−1	0	1	1	1
c	1	−1	0	−1	1
d	−1	−1	1	0	−1
e	1	−1	−1	1	0

Figure 35: Majority game Γ_R for the preference profile R specified in Figure 33

It is well-known that most of the SCFs defined above can be ordered with respect to set-inclusion (see, e.g., Laslier, 1997). For two SCFs f and f', write $f \subseteq f'$ if $f(R) \subseteq f'(R)$ for all preference profiles R.

Fact 7.3. *$BP \subseteq MC \subseteq UC \subseteq TC \subseteq COND \subseteq TRIV$.*

7.2.2 *C2 Functions*

In order to evaluate C2 functions, it is sufficient to consider the *weighted* tournament corresponding to a preference profile.

Definition 7.4. *An SCF f is a* C2 function *if f is not a C1 function and $f(R) = f(R')$ for all preference profiles R and R' such that $n_R(a, b) = n_{R'}(a, b)$ for all $a, b \in A$.*

Examples of common C2 functions include the Pareto rule, maximin, and Borda's rule. In order for the Pareto rule to always return a nonempty choice set, transitive preferences are required.

PARETO RULE (PAR) An alternative a is *Pareto-dominated* if there exists an alternative b such that $n_R(a, b) = 0$ and $n_R(b, a) > 0$. The function *PAR* returns all alternatives that are *not* Pareto-dominated.

PAR$(R) = \{a, b, c, d, e\}$

MAXIMIN (MM) The *maximin score* of an alternative a is given by its worst pairwise comparison, i.e., $s_{MM}(a) = \min_{b \in A \setminus \{a\}} n_R(a, b)$. The function *maximin*, also known as *Simpson's method* and denoted by *MM*, returns the set of all alternatives with the highest maximin score.

MM$(R) = \{b, e\}$

We go on to define Borda's rule, which is typically defined for linear preferences: each alternative receives $|A| - 1$ points for each time it is ranked first, $|A| - 2$ points for each time it is ranked second, and so forth; the alternatives with the highest total number of points are chosen. We generalize Borda's rule to (possibly intransitive) strict preferences, and formulate it in terms of $n_R(\cdot, \cdot)$ instead of ranks.

BORDA'S RULE (BO) The *Borda score* of an alternative a is defined as $s_{BO}(a) = \sum_{b \in A \setminus \{a\}} n_R(a, b)$. The function *BO* returns the alternatives with the highest Borda score.

BO$(R) = \{b, e\}$

A subclass of C2 functions that will play an important role in Chapter 11 is defined by *pairwise* SCFs (see, e.g., Young, 1974; Zwicker, 1991).

Definition 7.5. *An SCF* f *is* pairwise *if* $f(R) = f(R')$ *for all preference profiles* R *and* R' *such that* $m_R(a, b) = m_{R'}(a, b)$ *for all* $a, b \in A$.

In particular, pairwise SCFs satisfy $f(R) = f((R_1, \ldots, R_n, R_{n+1}, R_{n+1}^{\leftarrow}))$ for any preference profile $R = (R_1, \ldots, R_n)$ and any preference relation R_{n+1}. Whereas *MM* and *BO* are pairwise, *PAR* is not. In Chapter 8, we focus on a pairwise C2 function known as *ranked pairs*.

7.2.3 *C3 Functions*

All remaining SCFs are C3 functions.

Definition 7.6. *An SCF is a* C3 function *if it is neither a C1 function nor a C2 function.*

Most C3 functions explicitly use ranking information, and are only well-defined for linear preferences. The omninomination rule is one example.

OMNINOMINATION RULE (OMNI) The function *OMNI* returns all alternatives that are ranked first by at least one voter.

OMNI$(R) = \{b, d, e\}$

Two further C3 functions are *Young's rule* and *Dodgson's rule*, and we will study different variants of them in Chapter 8.

7.2.4 *Tournament Solutions*

tournament solution

Chapters 9 and 10 deal with tournament solutions, so let us clarify their relation to SCFs. An (unweighted) *tournament solution* S associates with each tournament $T = (V, E)$ a subset $S(T) \subseteq V$ of vertices. Therefore, each C1 function corresponds to a tournament solution. On the other hand, every tournament solution induces a C1 function whenever there are no majority ties. In fact, many C1 functions were first suggested as tournament solutions and have only later been generalized to incomplete tournaments in order to make them applicable to general majority graphs (see, e.g., Dutta and Laslier, 1999; Peris and Subiza, 1999). For some tournament solutions, however, no satisfactory generalization to incomplete tournaments is known. Examples from the latter category are the *tournament equilibrium set* (*TEQ*), to be defined in Chapter 10, and the *Banks set* (*BA*). In order to define the Banks set, consider a tournament (A, E) and note that a *transitive subtournament* (B, E) with $B \subseteq A$ corresponds to a ranking of B.

$BA(G(R)) = \{a, b, c, e\}$

BANKS SET (BA) The function *BA* returns the top elements of all inclusion-maximal transitive subtournaments (Banks, 1985).

A *weighted tournament solution* associates with each weighted tournament a subset of its vertices. As a consequence, there is a one-to-one correspondence between weighted tournament solution and C2 functions whenever all voters have strict preferences.

We sometimes use the term ‘tournament solution’ to refer to both weighted and unweighted tournament solutions (see Chapter 9). In Chapter 10, however, we are exclusively concerned with unweighted tournament solutions. Like SCFs, tournament solutions can be defined for a *fixed* or a *variable* agenda. While the former variant is sufficient for Chapter 9, the latter variant will be used in Chapter 10.

7.2.5 *Relations to Dominance-Based Solution Concepts*

The bipartisan set (*BP*) is explicitly defined in game-theoretic terms. Further connections to solution concepts from the first part of this thesis can be established by applying dominance-based solution concepts to the majority game Γ_R. These connections can be seen as alternative justifications for some of the C1 function defined above.

Assume that the preference profile R has no majority ties. Under this assumption, the majority graph is a tournament and the majority game Γ_R is a tournament game as defined on page 19. It was shown in Chapter 3 that D-sets are unique in tournament games for all dominance structures considered in this thesis.

Fact 7.7. *Let R be preference profile without majority ties.*

(i) The S-set of Γ_R is $(COND(R), COND(R))$.

(ii) The S^**-set and the* B*-set of* Γ_R *are equal to* $(TC(R), TC(R))$.

(iii) If $D \in \{C_M, C_D, V, V^*, W, W^*\}$*, the* D*-set of* Γ_R *is* $(MC(R), MC(R))$.

7.3 PROPERTIES OF SOCIAL CHOICE FUNCTIONS

In order to evaluate and compare different SCFs, a variety of desirable properties have been proposed (see Richelson, 1977, for an early survey). Various impossibility results like Arrow's (1951) and Gibbard (1973) and Satterthwaite's (1975) state that there exists no SCF that simultaneously satisfies a small number of natural properties. As a consequence, every SCF fails to satisfy some of these properties and compromises have to be made when choosing an SCF. The application at hand may suggest which properties are dispensable and which are not. In this section, we review a few basic properties of SCFs. More properties will be defined and studied in the following chapters.

fairness criteria

Neutrality and *anonymity* are basic fairness criteria which require, loosely speaking, that all alternatives and all voters are treated equally. In order to formally define these properties, let $R = (R_1, \dots, R_n)$ be a preference profile and $\pi : A \to A$ a permutation of A. For $i \in N$, define $\pi(R_i)$ as the preference relation given by $a\ \pi(R_i)\ b$ if and only if $\pi(a)\ R_i\ \pi(b)$. An SCF is neutral if permuting alternatives in all individual preference relations also permutes the set of chosen alternatives in the exact same way.

Definition 7.8. *An SCF* f *is* neutral *if* $f((\pi(R_1), \dots, \pi(R_n))) = \pi(f(R))$ *for all preference profiles* R *and permutations* π *of* A.

An SCF is anonymous if the set of chosen alternatives does not change when the voters are permuted.

Definition 7.9. *An SCF* f *is* anonymous *if* $f(R) = f((R_{\pi(1)}, \dots, R_{\pi(n)}))$ *for all preference profiles* R *and permutations* π *of* N.

By definition, all C1 and C2 functions are neutral and anonymous. On the other hand, C3 functions that fail neutrality or anonymity are easily constructed.

computational complexity

Another criterion we will be interested in is the *computational effort* required to evaluate an SCF. Computational tractability of the winner determination problem is obviously a significant property of any SCF: the inability to efficiently compute winners would render the method virtually useless, at least for large preference profiles that do not exhibit additional structure. The winner determination problem is usually formulated as a decision problem: given an SCF f, a preference profile R and an alternative a, does it hold that $a \in f(R)$? In Chapter 8, we study the winner determination problem for ranked pairs, Dodgson's rule, and Young's rule.

7.4 SUMMARY

We have introduced concepts and terminology related to preference aggregation mechanisms. Table 7 summarizes some notation for later reference.

A	finite set of alternatives
$N = \{1, \ldots, n\}$	set of voters
$R_i \subseteq A \times A$	complete preference relation of voter i
$R = (R_1, \ldots, R_n)$	preference profile
$n_R(a, b)$	number of voters preferring a over b
$m_R(a, b)$	majority margin between a and b
R_M	majority relation
$G(R) = (A, P_M)$	majority graph
Γ_R	majority game
$s_{CO}(a)$	Copeland score of a
$s_{BO}(a)$	Borda score of a
$s_{MM}(a)$	maximin score of a

Table 7: Notation for preference aggregation

8 COMPLEXITY OF WINNER DETERMINATION

The complexity of the winner determination problem has been studied for almost all common SCFs. A notable exception, possibly caused by some confusion regarding its exact definition, is the method of ranked pairs. We show in Section 8.1 that computing ranked pairs winners is NP-complete. We then focus on two C3 functions that are based on similar ideas: Young's rule and Dodgson's rule. Coincidentally, the resemblance of these two SCFs extends to the state of their winner determination problems. Both problems have been claimed to be Θ_2^p-complete (see Rothe et al., 2003, for Young's rule and Hemaspaandra et al., 1997, for Dodgson's rule); however, the alleged proofs are not fully satisfactory. For Young's rule, Rothe et al. (2003) used a variant of the rule that is not equivalent to Young's original definition. In Section 8.2, we show how their proof can be adapted to the original version of Young's rule. For Dodgson's rule, the proof of Hemaspaandra et al. (1997) is not entirely correct. In Section 8.3, we point out an error in the construction and show how the hardness proof can be repaired. We furthermore show that the corrected proof can be extended to cover a variant of Dodgson's rule.

8.1 RANKED PAIRS

In this section, we study the computational complexity of the *ranked pairs* method (Tideman, 1987). To the best of our knowledge, this question has not been considered before, which is particularly surprising given the extensive literature that is concerned with computational aspects of ranked pairs.[1] A possible reason for this gap might be the confusion of two variants of the method, only one of which satisfies neutrality. In Section 8.1.1, we address this confusion and describe both variants. After introducing some notation in Section 8.1.2, we show in Section 8.1.3 that deciding whether a given alternative is a ranked pairs winner is NP-complete for the neutral variant. By contrast, it can be checked efficiently whether a given ranking is a ranked pairs ranking. Finally, Section 8.1.4 discusses variants of the ranked pairs method that are not anonymous.

1 Typical problems include the hardness of manipulation (Betzler et al., 2009; Xia et al., 2009; Parkes and Xia, 2012) and the complexity of computing possible and necessary winners (Xia and Conitzer, 2011; Obraztsova and Elkind, 2011).

8.1.1 *Two Variants of the Ranked Pairs Method*

In this section we address the difference between two variants of the ranked pairs method that are commonly studied in the literature. Both variants are anonymous, i.e., treat all voters equally. Non-anonymous variants of the ranked pairs method have been suggested by Tideman (1987) and Zavist and Tideman (1989), and will be discussed in Section 8.1.4.

The easiest way to describe the ranked pairs method is to formulate it as a procedure. The procedure first defines a priority ordering over the set of all (ordered) *pairs* (a, b) of alternatives by giving priority to pairs (a, b) with a larger majority margin $m_R(a, b)$. Then, it constructs a ranking of the alternatives by starting with the empty ranking and iteratively considering pairs in order of priority. When a pair (a, b) is considered, the ranking is extended by fixing that a precedes b—unless fixing this pairwise comparison would create a cycle together with the previously fixed pairs, in which case the pair (a, b) is discarded. This procedure is guaranteed to terminate with a complete ranking of all alternatives, the top element of which is declared the winner.

What is missing from the above description is a *tie-breaking rule* for cases where two or more pairwise comparisons have the same support from the voters. This turns out to be a rather intricate issue. In principle, it is possible to employ an arbitrary tie-breaking rule. However, each fixed tie-breaking rule biases the method in favor of some alternative and thereby destroys neutrality.[2] In order to avoid this problem, Tideman (1987) originally defined the ranked pairs method to return the set of all those alternatives that result from the above procedure for *some* tie-breaking rule.[3] We will henceforth denote this variant by *RP*.

In a subsequent paper, Zavist and Tideman (1989) showed that a tie-breaking rule is in fact necessary in order to achieve the property of *independence of clones*, which was the main motivation for introducing the ranked pairs method. While Zavist and Tideman (1989) proposed a way to define a tie-breaking rule based on the preferences of a distinguished voter (see Section 8.1.4 for details), the variant that is most commonly studied in the literature considers the tie-breaking rule to be exogenously given and fixed for all profiles. This variant of ranked pairs will be denoted by RP_τ, where τ is a tie-breaking rule. Whereas *RP* is an irresolute SCF, RP_τ is a resolute one. It is straightforward to see that *RP* is neutral, i.e., treats all alternatives equally, and that RP_τ is not. An easy example for the latter statement is the case of two alternatives and two voters who each prefer a different alternative.

2 Neutrality can be maintained if the tie-breaking rule varies with the individual preferences (see Section 8.1.4).

3 This definition, sometimes called *parallel universes tie-breaking (PUT)*, can also be used to "neutralize" other SCFs that involve tie-breaking (Conitzer et al., 2009).

From a computational perspective, RP_τ is easy: constructing the ranking for a given tie-breaking rule takes time polynomial in the size of the input (see Proposition 8.4). For RP, however, the picture is different: as the number of tie-breaking rules is exponential, executing the iterative procedure for every single tie-breaking rule is infeasible. Of course, this does not preclude the existence of a clever algorithm that efficiently computes the set of all alternatives that are the top element of some chosen ranking. Our main result states that such an algorithm does not exist unless P equals NP.[4]

8.1.2 *Formal Definitions*

The resolute variant of the ranked pairs method takes as input a preference profile R and a *tie-breaking rule* $\tau \in \mathscr{L}(A \times A)$. A ranking $\succ_\tau^R$ of $A \times A$ is constructed by ordering all pairs in accordance with $m_R(\cdot,\cdot)$, using τ to break ties: $(a,b) \succ_\tau^R (c,d)$ if and only if $m_R(a,b) > m_R(c,d)$ or ($m_R(a,b) = m_R(c,d)$ and $(a,b)\ \tau\ (c,d)$). *tie-breaking rule*

The relation L_τ^R on A is constructed by starting with the empty relation and iteratively considering the pair ranked highest by $\succ_\tau^R$ among all pairs that have not been considered so far. The pair is then added to the relation L_τ^R *unless* this addition would create an L_τ^R-cycle with the pairs that have been added before.[5] After all pairs in $A \times A$ have been considered, L_τ^R is a ranking of A. The resolute variant of ranked pairs returns the top element of L_τ^R.

Definition 8.1. $RP_\tau(R) = \{top(L_\tau^R)\}$.

The outcome of RP_τ depends on the choice of τ, and RP_τ is not neutral. An irresolute and neutral variant can be defined by choosing all alternatives that are at the top of L_τ^R for *some* tie-breaking rule τ. This corresponds to Tideman's (1987) original variant RP.

Definition 8.2. $RP(R) = \{top(L_\tau^R) : \tau \in \mathscr{L}(A \times A)\}$

The alternatives in $RP(R)$ are called *ranked pairs winners for* R, and the rankings in $\{L_\tau^R : \tau \in \mathscr{L}(A \times A)\}$ are called *ranked pairs rankings for* R. *ranked pairs winner* *ranked pairs ranking*

We will work with an alternative characterization of ranked pairs rankings that was introduced by Zavist and Tideman (1989). Given a preference profile R, a ranking L of A, and two alternatives a and b, we say that a *attains* b *through* L if there exists a sequence of distinct alternatives $a_1, a_2, \ldots, a_t$, where $t \geqslant 2$, such that $a_1 = a$, $a_t = b$, $a_i\ L\ a_{i+1}$, and $m_R(a_i, a_{i+1}) \geqslant m_R(b,a)$ for all i with $1 \leqslant i < t$. In this case, we will say that a attains b *via* $(a_1, a_2, \ldots, a_t)$. A ranking L is called a *stack* if for any pair of alternatives a and b it holds that *stack*

4 A similar discrepancy can be observed for the Banks set (see page 102). Whereas Woeginger (2003) has proven that computing Banks winners is NP-complete, Hudry (2004) has shown that an arbitrary Banks winner can be found efficiently.

5 A P-cycle of a relation $P \subseteq A \times A$ is a set of pairs $\{(a_1,a_2),(a_2,a_3),\ldots,(a_{\ell-1},a_\ell)\}$ with $a_\ell = a_1$ and $a_i\ P\ a_{i+1}$ for all i with $1 \leqslant i < \ell$. Pairs of the form (a,a) are considered cycles of length 1, and are therefore never added to L_τ^R.

$a \mathrel{L} b$ implies that a attains b through L.

Lemma 8.3 (Zavist and Tideman, 1989). *A ranking of A is a ranked pairs ranking if and only if it is a stack.*

It follows that an alternative is a ranked pairs winner if and only if it is the top element of a stack.

8.1.3 *Complexity of Ranked Pairs Winners*

We are now ready to study the computational complexity of *RP*. We first observe that finding and checking ranked pairs rankings is easy. This also provides an efficient way to find *some* ranked pairs winner. The problem of deciding whether a *particular* alternative is a ranked pairs winner, however, turns out to be NP-complete. The hardness result can be adapted to a variant of the winner determination problem that asks for unique winners.

It can easily be seen that an arbitrary ranked pairs ranking can be found efficiently.

Proposition 8.4. *Finding a ranked pairs ranking is in P.*

Proof. We fix some arbitrary tie-breaking rule τ and compute L^R_τ, which, by definition, is a ranked pairs ranking. When constructing L^R_τ, in each round we have to check whether the addition of a pair (a, b) to the relation L^R_τ creates a cycle. This can efficiently be done, for example, by invoking Tarjan's (1972) algorithm for finding strongly connected components: a binary relation is acyclic if and only if there are no nontrivial strongly connected components. □

Deciding whether a given ranking is a ranked pairs ranking is also feasible in polynomial time, by checking whether the given ranking is a stack.

Proposition 8.5. *Deciding whether a given ranking is a ranked pairs ranking is in P.*

Proof. By Lemma 8.3, it suffices to check whether the given ranking L is a stack. This reduces to checking, for every pair (a, b) with $a \mathrel{L} b$, whether a attains b through L. Let a and b with $a \mathrel{L} b$ be given, and define $w = m_R(b, a)$. We construct a directed graph with vertex set A as follows. For all $x, y \in A$, there is an edge from x to y if and only if $x \mathrel{L} y$ and $m_R(x, y) \geqslant w$. It is easily verified that a attains b through L if and only if there exists a path from a to b in this graph. The latter property can be checked with a simple breadth-first search. □

It is worth noting that Proposition 8.5 can also be shown directly, without referring to stacks. For a given ranking L, define a tie-breaking rule τ_L such that $(a, b) \mathrel{\tau_L} (c, d)$ for all $(a, b) \in L$ and $(c, d) \notin L$. It can be shown that L is a ranked pairs ranking if and only if $L = L^R_{\tau_L}$. The advantage of this alternative proof is that for each "yes" instance it constructs a witnessing tie-breaking rule.

As every ranked pairs ranking yields a ranked pairs winner, Proposition 8.4 immediately implies that an arbitrary element of $RP(R)$ can be found efficiently.

Proposition 8.6. *Finding a ranked pairs winner is in P.*

Deciding whether a given alternative is a ranked pairs winner, on the other hand, turns out to be NP-complete.

Theorem 8.7. *Deciding whether a given alternative is a ranked pairs winner is NP-complete.*

Membership in NP follows from Proposition 8.5. For hardness, we give a reduction from SAT. For a Boolean formula $\varphi = C_1 \wedge \cdots \wedge C_k$ over a set $V = \{v_1, \ldots, v_m\}$ of variables, we will construct a preference profile R_φ over a set A_φ of alternatives such that a particular alternative $d \in A_\varphi$ is a ranked pairs winner for R_φ if and only if φ is satisfiable.

Let us first define the set A_φ of alternatives. For each variable $v_i \in V$, $1 \leqslant i \leqslant m$, there are four alternatives v_i, $\bar{v}_i$, v'_i, and $\bar{v}'_i$. For each clause C_j, $1 \leqslant j \leqslant k$, there is one alternative y_j. Finally, there is one alternative d for which we want to decide membership in $RP(R_\varphi)$.

Instead of constructing R_φ explicitly, we will specify a number $m(a,b)$ for each pair $(a,b) \in A_\varphi \times A_\varphi$. Debord (1987) has shown that there exists a preference profile R such that $m_R(a,b) = m(a,b)$ for all a, b, as long as $m(a,b) = -m(b,a)$ for all a, b and all the numbers $m(a,b)$ have the same parity.[6] In order to conveniently define $m(\cdot,\cdot)$, the following notation will be useful: for a natural number w, $a \succ^w b$ denotes setting $m(a,b) = w$ and $m(b,a) = -w$. *Debord's theorem* $a \succ^w b$

For each variable $v_i \in V$, $1 \leqslant i \leqslant m$, let $v_i \succ^4 \bar{v}'_i \succ^2 \bar{v}_i \succ^4 v'_i \succ^2 v_i$. For each clause C_j, $1 \leqslant j \leqslant k$, let $v_i \succ^2 y_j$ if variable $v_i \in V$ appears in clause C_j as a positive literal, and $\bar{v}_i \succ^2 y_j$ if variable v_i appears in clause c_j as a negative literal. Finally let $y_j \succ^2 d$ for $1 \leqslant j \leqslant k$ and $d \succ^2 v'_i$ and $d \succ^2 \bar{v}'_i$ for $1 \leqslant i \leqslant m$. For all pairs (a,b) for which $m(a,b)$ has not been specified so far, let $m(a,b) = m(b,a) = 0$. An example is shown in Figure 36.

As $m(a,b) \in \{-4,-2,0,2,4\}$ for all $a, b \in A_\varphi$, Debord's theorem guarantees the existence of a preference profile R_φ with $m_{R_\varphi}(a,b) = m(a,b)$ for all $a, b \in A_\varphi$, and such a profile can in fact be constructed efficiently.

The following two lemmata show that alternative d is a ranked pairs winner for R_φ if and only if the formula φ is satisfiable.

Lemma 8.8. *If $d \in RP(R_\varphi)$, then φ is satisfiable.*

6 See also Le Breton (2005).

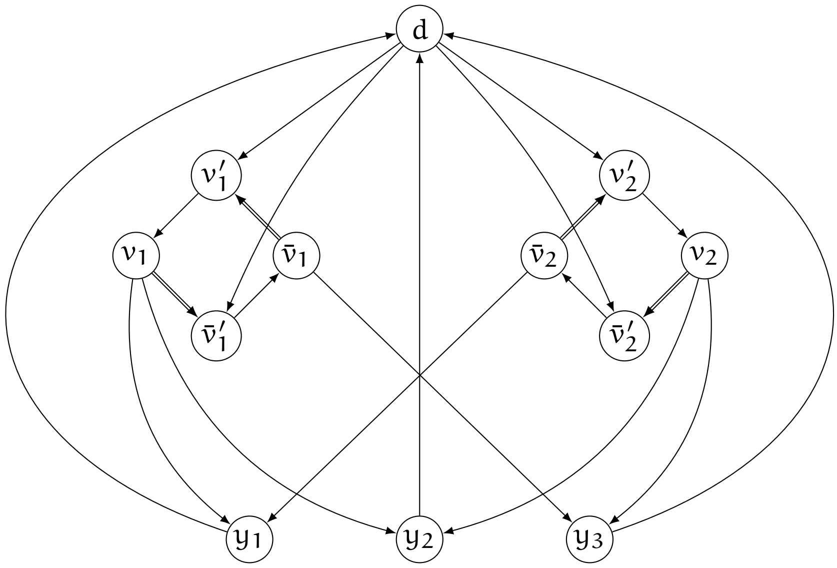


Figure 36: Graphical representation of $m_{R_\varphi}(\cdot,\cdot)$ for the Boolean formula $\varphi = \{v_1, \bar{v}_2\} \wedge \{v_1, v_2\} \wedge \{\bar{v}_1, v_2\}$. The relation $\succ^2$ is represented by arrows, and $\succ^4$ is represented by double-shafted arrows. For all pairs (a, b) that are not connected by an arrow, we have $m_{R_\varphi}(a, b) = m_{R_\varphi}(b, a) = 0$.

Proof. Assume that d is a ranked pairs winner for R_φ and let L be a stack with *top*$(L) = d$. Consider an arbitrary j with $1 \leqslant j \leqslant k$. As L is a stack and $d\ L\ y_j$, d attains y_j through L, i.e., there exists a sequence $P_j = (a_1, a_2, \ldots, a_t)$ with $a_1 = d$ and $a_t = y_j$ such that $a_i\ L\ a_{i+1}$ and $m(a_i, a_{i+1}) \geqslant 2$ for all i with $1 \leqslant i < t$. If d attains y_j via several sequences, fix one of them arbitrarily.

The definition of $m(\cdot,\cdot)$ implies that

$$P_j = (d, \bar{\ell}', \bar{\ell}, \ell', \ell, y_j) \quad \text{or}$$
$$P_j = (d, \ell', \ell, y_j),$$

where ℓ is some literal. The former is in fact not possible because $\bar{\ell}'$ does not attain ℓ through L. Therefore, each P_j is of the form $P_j = (d, \ell', \ell, y_j)$ for some $\ell \in X$.

Now define assignment α as the set of all literals that are contained in one of the sequences P_j, $1 \leqslant j \leqslant k$, i.e., $\alpha = X \cap (\bigcup_{j=1}^{k} P_j)$. We claim that α is a satisfying assignment for φ.

In order to show that α is valid, suppose there exists a literal $\ell \in X$ such that both ℓ and $\bar{\ell}$ are contained in α. This implies that there exist i and j such that d attains y_i via $P_i = (d, \ell', \ell, y_i)$ and d attains y_j via $P_i = (d, \bar{\ell}', \bar{\ell}, y_j)$. In particular, $\ell'\ L\ \ell$ and $\bar{\ell}'\ L\ \bar{\ell}$. As L is transitive and asymmetric, it follows that either $\ell'\ L\ \bar{\ell}$ or $\bar{\ell}'\ L\ \ell$. However, neither does ℓ' attain $\bar{\ell}$ through L, nor does $\bar{\ell}'$ attain ℓ through L, a contradiction.

In order to see that α satisfies φ, consider an arbitrary clause C_j. As d attains y_j via $P_j = (d, \ell', \ell, y_j)$ and $m(y_j, d) = 2$, we have that $m(\ell, y_j) \geqslant 2$. By definition of $m(\cdot,\cdot)$, this implies that $\ell \in C_j$. $\square$

Lemma 8.9. *If φ is satisfiable, then $d \in RP(R_\varphi)$.*

Proof. Assume that φ is satisfiable and let α be a satisfying assignment. Let $V_i = \{v_i, \bar{v}_i, v_i', \bar{v}_i'\}$, $1 \leqslant i \leqslant m$, and $Y = \{y_1, y_2, \dots, y_k\}$. We define a ranking L of A_φ as follows, using $B \; L \; C$ as shorthand for $b \; L \; c$ for all $b \in B$ and $c \in C$.

- For all $1 \leqslant i \leqslant m$, let $d \; L \; V_i$ and $V_i \; L \; Y$.
- For all $1 \leqslant i < j \leqslant m$, let $V_i \; L \; V_j$.
- For the definition of L within V_i, we distinguish two cases. If $v_i \in \alpha$, i.e., if v_i is set to "true" under α, let $\bar{v}_i \; L \; v_i' \; L \; v_i \; L \; \bar{v}_i'$. If, on the other hand, $v_i \notin \alpha$, let $v_i \; L \; \bar{v}_i' \; L \; \bar{v}_i \; L \; v_i'$.
- Within Y, define L arbitrarily.

We now prove that L is a stack. For each pair (a, b) with $a \; L \; b$, we need to verify that a attains b through L. If $m(b, a) \leqslant 0$, it is easily seen that a attains b through L. We can therefore assume that $m(b, a) > 0$. By definition of L and $m(\cdot,\cdot)$, a particular such pair (a, b) satisfies either

$a = d$ and $b \in Y$, or

$a, b \in V_i$ for some i with $1 \leqslant i \leqslant m$.

First consider a pair of the former type, i.e., $(a, b) = (d, y_j)$ for some j with $1 \leqslant j \leqslant k$. As α satisfies C_j, there exists $\ell \in C_j$ with $\ell \in \alpha$. Consider the sequence $P_j = (d, \ell', \ell, y_j)$. As $m(y_j, d) = 2$ and $d \succ^2 \ell' \succ^2 \ell \succ^2 y_j$, d attains y_j via P_j.

Now consider a pair of the latter type, i.e., $a, b \in V_i$ for some i with $1 \leqslant i \leqslant m$. Assume that $v_i \in \alpha$ and, therefore, $\bar{v}_i \; L \; v_i' \; L \; v_i \; L \; \bar{v}_i'$. The only nontrivial case is the pair $(\bar{v}_i, \bar{v}_i')$ with $\bar{v}_i \; L \; \bar{v}_i'$ and $m(\bar{v}_i', \bar{v}_i) = 2$. But $\bar{v}_i$ attains $\bar{v}_i'$ via $(\bar{v}_i, v_i', v_i, \bar{v}_i')$ because $\bar{v}_i \succ^4 v_i' \succ^2 v_i \succ^4 \bar{v}_i'$. The case $v_i \notin \alpha$ is analogous.

We have shown that L is a stack. Lemma 8.3 now implies that $d \in RP(R_\varphi)$, which completes the proof. $\square$

Combining Lemma 8.8 and Lemma 8.9, and observing that both A_φ and R_φ can be constructed efficiently, completes the proof of Theorem 8.7.

An interesting variant of the winner determination problem concerns the question whether a given alternative is the *unique* winner for a given preference profile. Despite its similarity to the original winner determination problem, this problem is sometimes considerably easier.[7] For *RP*, the picture is different: verifying unique winners is not feasible in polynomial time, unless P equals coNP.

Theorem 8.10. *Deciding whether a given alternative is the unique ranked pairs winner is coNP-complete.*

Proof. Membership in coNP follows from the observation that for every "no" instance there is a stack whose top element is different from the alternative in question.

For hardness, we modify the construction from the proof of Theorem 8.7 to obtain a reduction from the problem UNSAT, which asks whether a given Boolean formula is *not* satisfiable. For a Boolean formula φ, define $A'_\varphi = A_\varphi \cup \{d^*\}$, where d^* is a new alternative and A_φ is defined as in the proof of Theorem 8.7. R'_φ is defined such that $d \succ^2 d^*$ and $d^* \succ^4 a$ for all $a \in A_\varphi \setminus \{d\}$. Within A_φ, R'_φ coincides with R_φ. We show that $RP(R'_\varphi) = \{d^*\}$ if and only if φ is unsatisfiable.

For the direction from left to right, assume for contradiction that $RP(R'_\varphi) = \{d^*\}$ and φ is satisfiable. Consider a satisfying assignment α and let L be the ranking of A_φ defined in the proof of Lemma 8.9. Define the ranking L' of A'_φ by $L' = L \cup \{(d, d^*)\} \cup \{(d^*, a) : a \in A_\varphi \setminus \{d\}\}$. That is, L' extends L by inserting the new alternative d^* in the second position. As in the proof of Lemma 8.9, it can be shown that L' is a stack. It follows that $top(L') = d \in RP(R'_\varphi)$, contradicting the assumption that $RP(R'_\varphi) = \{d^*\}$.

For the direction from right to left, assume for contradiction that φ is unsatisfiable and $RP(R'_\varphi) \neq \{d^*\}$. Then there exists a tie-breaking rule τ such that $top(L_\tau^{R'_\varphi}) = a \neq d^*$. From the definition of R'_φ it follows that $a = d$, as $d^* \succ^4 b$ for all $b \in A_\varphi \setminus \{d\}$ and there are no $\succ^4$-cycles. By the same argument as in the proof of Lemma 8.8, it can be shown that φ is satisfiable, contradicting our assumption. □

8.1.4 *Non-Anonymous Variants*

As mentioned in Section 8.1.1, Tideman (1987) and Zavist and Tideman (1989) suggested ways to use the preferences of a distinguished voter, say, a chairperson, to render the ranked pairs method resolute. There are essentially two ways to achieve this, which differ in the point in time when ties are broken.

a priori The *a priori* variant uses the preferences of the chairperson in order to construct a tie-breaking rule $\tau \in \mathscr{L}(A \times A)$, which is then *a posteriori* used to compute RP_τ. The *a posteriori* variant first computes $RP(\cdot)$ and then chooses the alternative from this set that is most preferred

[7] The Banks set constitutes an example: although deciding membership is NP-complete in general, it can be checked in polynomial time whether an alternative is the *unique* Banks winner. The reason for the latter is that an alternative is the unique Banks winner if and only if it is a Condorcet winner.

by the chairperson. Both variants are neutral: if the alternatives are permuted in each ranking, including the ranking of the chairperson, the tie-breaking rule and thus the chosen alternative will change accordingly.

Whereas the *a priori* variant is a special case of RP_τ and therefore efficiently computable, the *a posteriori* variant is intractable by the results in Section 8.1.3. It follows that neutrality and tractability can be reconciled at the expense of anonymity. By moving to *non-deterministic* SCFs, one can even regain anonymity: choosing the chairperson for the *a priori* variant uniformly at random results in a procedure that is neutral, anonymous, and tractable, for appropriate generalizations of anonymity and neutrality to the case of non-deterministic SCFs. The winner determination problem for the *a posteriori* variant remains intractable when the chairperson is chosen randomly.

8.2 YOUNG'S RULE

Young's rule and Dodgsons's rule are based on the same general idea. Whenever there is a Condorcet winner, this alternative should be chosen. If there is no Condorcet winner, choose those alternatives that are "closest" to being a Condorcet winner. The rules differ in the measure of closeness. For Dodgson's rule, distance is measured in the number of *switches* that have to be preformed in voter's preferences in order to turn a specified alternative into a Condorcet winner (see Section 8.3 for details). Young's rule, on the other hand, selects those alternatives that can be made Condorcet winners by *deleting* as few voters as possible.

Furthermore, there is a subtle disparity in the notion of *Condorcet winner* that is employed. Dodgson's rule uses the standard notion and defines a Condorcet winner as an alternative that is preferred to any other alternative by a *strict* majority of voters (see page 97). To avoid confusion, we will use the term *strong Condorcet winner* when referring to such alternatives in the remainder of this chapter. By contrast, Young's rule is defined via the concept of a *weak Condorcet winner*, i.e., an alternative that is preferred to any other alternative by a (not necessarily strict) majority of voters. The example in Figure 37 clarifies the distinction. Clearly, the two notions coincide whenever there are no majority ties.

strong Condorcet winner

weak Condorcet winner

Rothe, Spakowski, and Vogel (2003)—abbreviated RSV hereafter—showed that computing Young winners is Θ_2^p-complete. However, they did *not* use the original version of Young's rule, but the variant that is defined via *strong* Condorcet winners. We show that their proof can be adapted to the original version. Throughout this section, we assume linear preferences.

RSV

R_1	R_2
a	b
b	a
c	c

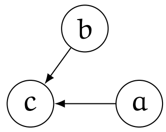


Figure 37: A preference profile R and its corresponding majority graph $G(R)$. R has two weak Condorcet winners (a and b), but no strong Condorcet winner.

8.2.1 *Two Variants of Young's Rule*

To ease comparison, we employ the notation of RSV and let an *election* be defined as a pair (C, V) where C is a finite set of alternatives and V is a preference profile, formally modeled as a multiset of rankings in $\mathscr{L}(C)$.

Definition 8.11. *The* Young score *YoungScore*(C, c, V) *of an alternative* $c \in C$ *in an election* (C, V) *is defined as the size of a largest submultiset* $V' \subseteq V$ *such that* c *is a weak Condorcet winner in* (C, V')*. The SCF* Young *selects those alternatives with the highest Young score.*

Since every candidate is a weak Condorcet winner when V' equals the empty set, the Young score of a candidate is lower-bounded by zero. RSV considered a variant of Young's rule that is defined via strong Condorcet winners. We call the resulting SCF *strongYoung*.

Definition 8.12. *The* strong Young score *of an alternative* $c \in C$ *in an election* (C, V) *is defined as the size of a largest submultiset* $V' \subseteq V$ *such that* c *is a strong Condorcet winner in* (C, V')*, and* 0 *if no such submultiset exists. The SCF* strongYoung *selects those alternatives with the highest strong Young score.*

Figure 38 presents a preference profile for which the choice sets of the two variants differ.

8.2.2 *Complexity of Young Winners*

RSV showed that the winner determination problem for strongYoung is Θ_2^p-complete. In particular, Θ_2^p-hardness was shown via a reduction from the Θ_2^p-complete problem Maximum Set Packing Compare
MSPC (MSPC, for short):

> Given two sets B_1 and B_2, and two collections of sets $\mathscr{S}_1$ and $\mathscr{S}_2$, where each $S \in \mathscr{S}_i$ is a nonempty, finite subset of B_i, is it the case that $\kappa(\mathscr{S}_1) \geqslant \kappa(\mathscr{S}_2)$?

Here, $\kappa(\mathscr{S}_i)$ denotes the maximum number of pairwise disjoint sets in $\mathscr{S}_i$. The proofs of Theorems 2.3 and 2.5 of RSV show how to construct, from a given MSPC instance $I = (B_1, B_2, \mathscr{S}_1, \mathscr{S}_2)$, an election (D, W)

R_1	R_2
a	c
b	b
c	a

Figure 38: A profile R with *Young*$(R) = \{a, b, c\}$ and *strongYoung*$(R) = \{a, c\}$

with two designated alternatives, c and d, such that (a) if $\kappa(\mathscr{S}_1) \geqslant \kappa(\mathscr{S}_2)$, then c and d are strongYoung winners of (D, W), and (b) if $\kappa(\mathscr{S}_2) > \kappa(\mathscr{S}_1)$, then d is the unique strongYoung winner of (D, W).

We now show how this proof can be adapted to work for Young elections. We refer to RSV for definitions and details of the construction, and will only point out the differences here.

Theorem 8.13. *Deciding whether a given alternative is a Young winner is Θ_2^p-complete.*

Proof. Membership in Θ_2^p can be shown by the same argument that was used by RSV for strongYoung.

Given an MSPC instance $I = (B_1, B_2, \mathscr{S}_1, \mathscr{S}_2)$, we construct an election (C', V') such that c and d are designated alternatives in C', and it holds that

$$\text{YoungScore}(C', c, V') = 2\kappa(\mathscr{S}_1) \text{ and } \text{YoungScore}(C', d, V') = 2\kappa(\mathscr{S}_2).$$

Let $C' = C$ and $V' = V \setminus \{v_{(2.4)}, v_{(2.7)}\}$, where $v_{(2.4)}$ is one of the two voters in V referred to as "voters of the form (2.4)" by RSV and $v_{(2.7)}$ is one of the two voters in V referred to as "voters of the form (2.7)" by RSV. One can then define a submultiset $\hat{V}'$ of the voters V' as $\hat{V}' = \hat{V} \setminus v_{(2.4)}$, where $\hat{V}$ is defined on page 381 of RSV. Then, $|\hat{V}'| = 2 \cdot \kappa(\mathscr{S}_1)$ and c is a weak Condorcet winner in $(C', \hat{V}')$, implying that $\text{YoungScore}(C', c, V') \geqslant 2 \cdot \kappa(\mathscr{S}_1)$.

To show that $\text{YoungScore}(C', c, V') \leqslant 2\kappa(\mathscr{S}_1)$, we adapt Lemma 2.4 of RSV in the following way.

Lemma 8.14. *For any λ with $3 < \lambda \leqslant |\mathscr{S}_1| + 1$, let V_λ be a submultiset of V' such that V_λ contains exactly λ voters of the form* (2.4) *or* (2.5) *and* c *is a weak Condorcet winner in* (C', V'). *Then V_λ contains exactly λ voters of the form* (2.3) *and no voters of the form* (2.6)*,* (2.7)*, or* (2.8)*. Moreover, the λ voters of the form* (2.3) *in V_λ represent pairwise disjoint sets from $\mathscr{S}_1$.*

The proof of Lemma 8.14 is similar to the proof of Lemma 2.4 of RSV and we omit it here.

To continue the proof of $\text{YoungScore}(C', c, V') \leqslant 2 \cdot \kappa(\mathscr{S}_1)$, let $k = \text{YoungScore}(C', c, V')$. Let $\hat{V}' \subseteq V'$ be a submultiset of size k such that c is a weak Condorcet winner in $(C, \hat{V}')$ and suppose that there are exactly λ voters of the form (2.4) or (2.5) in $\hat{V}'$. Lemma 8.14 then implies that there are exactly λ voters of the form (2.3) in $\hat{V}'$, those voters represent pairwise disjoint sets from $\mathscr{S}_1$, and $\hat{V}'$ contains no voters of the form (2.6), (2.7), or (2.8). Hence, $k = 2 \cdot \lambda \leqslant 2 \cdot \kappa(\mathscr{S}_1)$.

We thus have $\mathrm{YoungScore}(C', c, V') = 2 \cdot \kappa(\mathscr{S}_1)$. Analogously, one can show that $\mathrm{YoungScore}(C', d, V') = 2 \cdot \kappa(\mathscr{S}_2)$. Therefore,

$$\kappa(\mathscr{S}_1) \geqslant \kappa(\mathscr{S}_2) \Leftrightarrow \mathrm{YoungScore}(C', c, V') \geqslant \mathrm{YoungScore}(C', d, V'),$$

which proves the Θ_2^p-hardness of comparing Young scores. To show that the winner determination problem for Young is Θ_2^p-hard, we alter the profile (C', V') in exactly the same way as (C, V) is altered in Theorem 2.5 of RSV. Let (D', W') be the altered preference profile. One can then show that the Young scores of c and d in (D', W') are the same as in (C', V'), and that all other alternatives have a Young score in (D', W') of at most 2. Thus, comparing the Young scores of c and d in (D', W') is tantamount to deciding whether c is a Young winner.

Altogether, we have that (a) if $\kappa(\mathscr{S}_1) \geqslant \kappa(\mathscr{S}_2)$ then c is a Young winner of (D', W'), and (b) if $\kappa(\mathscr{S}_2) > \kappa(\mathscr{S}_1)$ then d is the unique Young winner of (D', W'). It follows that an MSPC-instance I is in MSPC if and only if c is a Young winner of (D', W'), implying Θ_2^p-hardness of the Young winner determination problem. $\square$

The same proof also shows that deciding whether a given alternative is the *unique* Young winner is Θ_2^p-complete. To see this, observe that I is in MSPC if and only if d is *not* the unique winner of (D', W'). Thus, the *complement* of the unique-winner problem is Θ_2^p-hard. Since Θ_2^p is closed under complement, this proves that the unique-winner problem is Θ_2^p-hard as well.

8.3 DODGSON'S RULE

We now turn to Dodgson's rule. Hemaspaandra, Hemaspaandra, and Rothe (1997)—abbreviated HHR hereafter—have shown that the winner determination problem for Dodgson is Θ_2^p-complete. In this section, we point out an error in their proof and show how it can be corrected. We furthermore show that the winner determination problem for a variant of Dodgson's rule is also Θ_2^p-complete. Throughout this section, we assume linear preferences.

HHR

8.3.1 *Two Variants of Dodgson's Rule*

Similar to Young's rule, Dodgson's rule selects alternatives that are in some sense close to being a Condorcet winner. Closeness in Dodgson's rule is measured in terms of the number of switches that have to be made in the preference profile. A *switch* is an operation that transforms a preference profile R into another preference profile R' such that $R_j = R'_j$ for all $j \in N \setminus \{i\}$ and R'_i is identical to R_i except

switch

R_1	R_2
a	d
b	b
c	c
d	a

Figure 39: A profile R with $\mathit{Dodgson}(R) = \{b\}$ and $\mathit{weakDodgson}(R) = \{a, b, d\}$

that two adjacent alternatives in R_i have swapped their positions in the ranking.[8] We adopt the notation of HHR, which is similar to the notation used in Section 8.2. A *Dodgson triple* (C, c, V) is an election (C, V) with a designated alternative $c \in C$.

Dodgson triple

Definition 8.15. *The* Dodgson score $\mathit{DodgsonScore}(C, c, V)$ *of an alternative* $c \in C$ *in an election* (C, V) *is defined as the minimal number of switches that transform* V *into a preference profile that has* c *as a strong Condorcet winner. The SCF* Dodgson *selects those alternatives with the smallest Dodgson score.*

As in the case of Young's rule, there is a variant of Dodgson's rule that is defined via weak Condorcet winners.

Definition 8.16. *The* weak Dodgson score $\mathit{weakDodgsonScore}(C, c, V)$ *of an alternative* $c \in C$ *in an election* (C, V) *is defined as the minimal number of switches that transform* V *into a preference profile that has* c *as a weak Condorcet winner. The SCF* weakDodgson *selects those alternatives with the smallest weak Dodgson score.*

Whenever the number of voters is odd, both variant of Dodgson's rule coincide. Figure 39 presents an example where the variants differ.

8.3.2 *Complexity of Dodgson Winners*

HHR showed that the winner determination problem for Dodgson is Θ_2^p-complete. In particular, Θ_2^p-hardness was shown by a reduction from the Θ_2^p-hard problem Two Election Ranking (2ER):

2ER

> Given two Dodgson triples (C, c, V) and (D, d, W), where both $|V|$ and $|W|$ are odd and $c \neq d$, is the Dodgson score of c in (C, V) less than or equal to the Dodgson score of d in (D, W)?

The reduction from 2ER to the winner determination problem for Dodgson works by merging the elections $E_1 = (C, V)$ and $E_2 = (D, W)$ into a new election $E_3 = (C', V')$ such that $C \cup D \subseteq C'$ and the following three properties are satisfied:

8 Two elements $a, b \in A$ are called *adjacent* in a ranking $L \in \mathscr{L}(C)$ if there does not exits an alternative $c \in C$ with $a \mathrel{L} c \mathrel{L} b$ or $b \mathrel{L} c \mathrel{L} a$.

(*i*) $\text{DodgsonScore}_{E_3}(c) = \text{DodgsonScore}_{E_1}(c) + 1$,

(*ii*) $\text{DodgsonScore}_{E_3}(d) = \text{DodgsonScore}_{E_2}(d) + 1$, and

(*iii*) $\text{DodgsonScore}_{E_3}(x) > \text{DodgsonScore}_{E_3}(c)$ for all $x \in C' \setminus \{c, d\}$.

Here, $\text{DodgsonScore}_E(x)$ denotes the minimal number of switches required to make alternative x a Condorcet winner in election E. Thus, we immediately have that c is a Dodgson winner in E_3 if and only if $\text{DodgsonScore}_{E_1}(c) \geqslant \text{DodgsonScore}_{E_2}(d)$.

We now define election $E_3 = (C', V')$. For convenience, denote $v = |V|$ and $w = |W|$ and define $m = 2(|C| \cdot v + |D| \cdot w)$. The set C' of alternatives of E_3 is defined as $C' = C \cup D \cup S \cup T$, where C and D are the alternative sets from elections E_1 and E_2, respectively, and $S = \{s_1, \dots s_m\}$ and $T = \{t_1, \dots t_m\}$ contain $2m$ new alternatives (so-called *separating* alternatives).

In order to define the voter set V' of E_3, we introduce the following abbreviations for notational convenience. The symbol $<$ is used to define preferences of voters: $x < y$ if and only if the voter prefers y to x. For an ordered set $A = \{a_1, \dots, a_k\}$, we use $x < A < y$ as a shorthand for $x < a_1 < \dots < a_k < y$ and $x < \overleftarrow{A} < y$ as a shorthand for $x < a_k < \dots < a_1 < y$. Furthermore, let $c_1, \dots, c_{|C|-1}$ and $d_1, \dots, d_{|D|-1}$ be arbitrary enumerations of $C_- = C \setminus \{c\}$ and $D_- = D \setminus \{d\}$, respectively. The voters V' of E_3 consist of the following subgroups:

(a) For every voter in E_1 with preference order C_i over C, there is one voter with preferences

$$d < S < D_- < T < C_i\,.$$

(b) For every voter in E_2 with preference order D_j over D, there is one voter with preferences

$$T < c < S < C_- < D_j\,.$$

(c) There are $\lceil \frac{v}{2} \rceil - \lceil \frac{w}{2} \rceil$ voters with preferences

$$T < c < S < C_- < D_- < d\,.$$

(d) There are $\lceil \frac{v}{2} \rceil$ voters with preferences

$$T < C_- < D_- < \overleftarrow{S} < c < d\,.$$

(e) There are $\lceil \frac{w}{2} \rceil$ voters with preferences

$$T < C_- < D_- < S < d < c\,.$$

The total number of voters is $|V'| = 2v + w + 1$. As both v and w are odd, $|V'|$ is even and a Condorcet winner has to be preferred to every other alternative by at least $\frac{|V'|}{2} + 1 = v + \lceil \frac{w}{2} \rceil + 1$ voters. HHR then claim that properties *(i)*, *(ii)*, and *(iii)* are satisfied.

ERROR We now show that the construction in the proof of HHR does not satisfy property *(iii)*. Consider the following example. Election E_1 is given by $E_1 = (C, V)$, where $C = \{c, c_1, c_2, c_3\}$ and V consists of $v = 3$ voters who all have the same preference order $c < c_1 < c_2 < c_3$. Election E_2 is given by $E_2 = (D, W)$, where $D = \{d, d_1\}$ and W consists of $w = 1$ voter with preference order $d < d_1$.

Thus, Election E_3 has alternatives $C \cup D \cup S \cup T$ with $|S| = |T| = 2(|C| \cdot v + |D| \cdot w) = 2(4 \cdot 3 + 2 \cdot 1) = 28$. The set V' of voters of E_3 consists of the following groups:

(a) There are $v = 3$ voters with preferences

$$d < S < d_1 < T < c < c_1 < c_2 < c_3 \,.$$

(b) There is $w = 1$ voter with preferences

$$T < c < S < c_1 < c_2 < c_3 < d < d_1 \,.$$

(c) There is $\lceil \frac{v}{2} \rceil - \lceil \frac{w}{2} \rceil = 1$ voter with preferences

$$T < c < S < c_1 < c_2 < c_3 < d_1 < d \,.$$

(d) There are $\lceil \frac{v}{2} \rceil = 2$ voters with preferences

$$T < c_1 < c_2 < c_3 < d_1 < \overleftarrow{S} < c < d \,.$$

(e) There is $\lceil \frac{w}{2} \rceil = 1$ voter with preferences

$$T < c_1 < c_2 < c_3 < d_1 < S < d < c \,.$$

Observe that $\text{DodgsonScore}_{E_3}(c) = \text{DodgsonScore}_{E_1}(c) + 1 = 6 + 1 = 7$, as expected. However, $\text{DodgsonScore}_{E_3}(c_3) \leqslant 4$, since two switches in voter (b) and two switches in voter (c) suffice to make c_3 the top choice of a strict majority of the voters (groups (a), (b), and (c)) and therefore a Condorcet winner. Hence, property *(iii)* is violated and the reduction from 2ER does not go through.

The error in the proof argument can be traced back to the end of the proof of Lemma 3.7 (HHR, page 822), where the authors prove that $\text{DodgsonScore}_{E_3}(d_{|D|-1}) > \text{DodgsonScore}_{E_3}(c)$ and claim that "[t]he same argument applies to each element in $(C \cup D) \setminus \{c, d\}$."

CORRECTION Closer inspection of the counter-example reveals the problem in the construction of E_3: in the preference orders of voter groups (b) and (c), alternatives in C_- are *not* separated from alternatives in D by a set of separating alternatives. As a consequence, it is possible to make an alternative in C_- a Condorcet winner by strictly less than $\frac{m}{2}$ switches in those voter groups.[9]

This problem can be avoided by changing the preferences of voter groups (b) and (c) as follows:

(b′) For every voter in E_2 with preference order D_j over D, there is one voter with preferences

$$c < T < C_- < S < D_j .$$

(c′) There are $\lceil \frac{v}{2} \rceil - \lceil \frac{w}{2} \rceil$ voters with preferences

$$c < T < C_- < S < D_- < d .$$

Intuitively, the new construction is more symmetrical than the old one, as the preferences of voter groups (a) and (b′) are defined analogously, with the roles of C and D (and those of S and T) interchanged.

Observe that the alternatives in S are strictly better off in (b′) than in (b) and strictly better off in (c′) than (c). One might worry that this strengthening of alternatives in S reduces their Dodgson scores to such an extent that property *(iii)* is violated. To see that this is not the case, consider some alternative $s \in S$. Observe that the size $m = |S|$ of the alternative set S was chosen in such a way as to ensure that $\text{DodgsonScore}_{E_3}(c) < \frac{m}{2}$. Now suppose that $\text{DodgsonScore}_{E_3}(s) \leqslant \text{DodgsonScore}_{E_3}(c) < \frac{m}{2}$. In order to become a Condorcet winner, s must in particular beat c in a pairwise comparison. As s is preferred to c by only $w + \lceil \frac{v}{2} \rceil - \lceil \frac{w}{2} \rceil = \lceil \frac{v}{2} \rceil + \lceil \frac{w}{2} \rceil - 1$ voters in V', it needs to gain at least $\lceil \frac{v}{2} \rceil + 1$ extra votes over c in voter groups (a), (d), and (e). Gaining one extra vote over c in voter group (a) would require more than m switches, because s is separated from c by T. Thus, switches in (a) are too expensive. Since there are strictly less than $\lceil \frac{v}{2} \rceil + 1$ voters in both (d) and (e), switches need to be performed in both of these voter groups. And since the order of alternatives in S is reversed in (d), the overall number of switches required is greater than $\frac{m}{2}$. Again, this exceeds the Dodgson score of c. This shows that $\text{DodgsonScore}_{E_3}(s) > \text{DodgsonScore}_{E_3}(c)$ for all $s \in S$.

---

9 Whenever $w \geqslant 3$ *or* $v > w$, the voters in groups (a), (b), and (c) constitute a strict majority of voters:

$$v + w + \left(\left\lceil \frac{v}{2} \right\rceil - \left\lceil \frac{w}{2} \right\rceil \right) \geqslant v + \left\lceil \frac{w}{2} \right\rceil + 1 = \frac{|V'|}{2} + 1.$$

8.3.3 *Complexity of weakDodgson Winners*

We now show that the corrected version of the proof of HHR can be adapted to the SCF weakDodgson.

Theorem 8.17. *The winner determination problem for weakDodgson is Θ_2^p-complete.*

Proof. Membership in Θ_2^p is again easy. For hardness, first observe that the problem Weak Two Election Ranking (w2ER), which is defined like 2ER except with "Dodgson score" replaced by "weakDodgson score," inherits Θ_2^p-hardness from 2ER because Dodgson scores and weakDodgson scores coincide for all instances with an odd number of voters. *w2ER*

We present a reduction from w2ER to the winner determination problem for weakDodgson. Given two Dodgson triples (C, c, V) and (D, d, W), denote $E_1 = (C, V)$, $E_2 = (D, W)$, $v = |V|$, and $w = |W|$. Recall that both v and w are odd and assume without loss of generality that $v \geqslant w \geqslant 1$. Define $m = v \cdot |C| + w \cdot |D|$ and observe that m is an upper bound for all weakDodgson scores in E_1 and E_2: even in the worst case (an alternative is least preferred by all voters), $\lceil \frac{v}{2} \rceil \cdot (|C| - 1) < m$ switches suffice to make an alternative a weak Condorcet winner in E_1 (by making it the top choice of $\lceil \frac{v}{2} \rceil$ voters), and an analogous statement holds for E_2.

We now define the new election $E_3 = (C', V')$. The set C' of alternatives of E_3 is defined as $C' = C \cup D \cup S \cup T \cup U$, where C and D are the alternatives from elections E_1 and E_2 and $S = \{s_1, \dots s_m\}$, $T = \{t_1, \dots t_m\}$, and $U = \{u_1, \dots u_m\}$ are $3m$ new alternatives. The voters V' of E_3 consist of the following subgroups:

(a) For every voter in E_1 with preference order C_i over C, there is one voter with preferences

$$U < d < S < D_{-} < T < C_i .$$

(b) For every voter in E_2 with preference order D_j over D, there is one voter with preferences

$$U < c < T < C_{-} < S < D_j .$$

(c) There are $\lceil \frac{v}{2} \rceil - \lceil \frac{w}{2} \rceil$ voters with preferences

$$U < c < T < C_{-} < S < D_{-} < d .$$

(d) There are $\lceil \frac{v}{2} \rceil$ voters with preferences

$$T < C_{-} < D_{-} < S < U < c < d .$$

(e) There are $\lceil \frac{w}{2} \rceil - 1$ voters with preferences

$$T < C_{-} < D_{-} < S < U < d < c.$$

(f) There is one voter with preferences

$$U < d < c < T < C_{-} < D_{-} < S.$$

The total number of voters is $|V'| = 2v + w + 1$. As both v and w are odd, $|V'|$ is even and a weak Condorcet winner has to be preferred to every other alternative by at least $\frac{|V'|}{2} = v + \lceil \frac{w}{2} \rceil$ voters. We now show that the following three properties are satisfied:

(*i*) $\text{weakDodgsonScore}_{E_3}(c) = \text{weakDodgsonScore}_{E_1}(c)$,

(*ii*) $\text{weakDodgsonScore}_{E_3}(d) = \text{weakDodgsonScore}_{E_2}(d)$, and

(*iii*) $\text{weakDodgsonScore}_{E_3}(x) > \text{weakDodgsonScore}_{E_3}(c)$ for all $x \in C' \setminus \{c, d\}$.

For (*i*), observe that c is preferred to every alternative in $C' \setminus C$ by at least $\frac{|V'|}{2}$ of the voters. Thus, in computing the weakDodgson score of c, we only have to take care of alternatives in $C_{-} = \{c_1, \dots, c_{|C|-1}\}$. Let x_i be the number of voters in group (a) that prefer c to c_i. Then, the number of voters in V' that prefer c to c_i is $x_i + \lceil \frac{v}{2} \rceil + \lceil \frac{w}{2} \rceil - 1$. Candidate c is a weak Condorcet winner in E_3 if and only if this number is greater than or equal to $\frac{|V'|}{2} = v + \lceil \frac{w}{2} \rceil$, and this is the case if and only if $x_i \geqslant \lceil \frac{v}{2} \rceil$ for each $i \in \{1, \dots, |C| - 1\}$. But this means that c is a Condorcet winner in E_1. By definition, this can be achieved by k switches, where $k = \text{weakDodgsonScore}_{E_1}(c)$. We have therefore shown the upper bound

$$\text{weakDodgsonScore}_{E_3}(c) \leqslant \text{weakDodgsonScore}_{E_1}(c).$$

Now assume that $\text{weakDodgsonScore}_{E_3}(c) < \text{weakDodgsonScore}_{E_1}(c)$. Due to the construction, all the switches in the optimal sequence occur in voters of group (a), as making c beat any alternative in C_{-} would require more than m switches in all the other relevant voter groups (b), (c), and (f). This means that there is a way to make c a weak Condorcet winner in E_1 with less than $\text{weakDodgsonScore}_{E_1}(c)$ switches, a contradiction. We have thereby shown that

$$\text{weakDodgsonScore}_{E_3}(c) = \text{weakDodgsonScore}_{E_1}(c).$$

An analogous argument proves that property (*ii*) holds.

For (*iii*), recall that m was chosen sufficiently large to be an upper bound on the weakDodgson score of c in E_1, and thus, by (*i*), on the weakDodgson score of c in E_3. We now show that all alternatives in C' other than c and d have a weakDodgson score of at least m in E_3.

Consider an alternative $s \in S$. In order to become a weak Condorcet winner, s must in particular beat (or tie) c in a pairwise comparison. As s is preferred to c by only $w + \lceil \frac{v}{2} \rceil - \lceil \frac{w}{2} \rceil + 1 = \lceil \frac{v}{2} \rceil + \lceil \frac{w}{2} \rceil$ voters in V', it needs to gain at least $\lfloor \frac{v}{2} \rfloor$ extra votes over c in voter groups (a), (d), and (e). But gaining just one extra vote over c would require more than m switches, because s is separated from c by at least m other alternatives in all these voter groups.

A similar argument applies to all other alternatives.

- Candidates in T need $\lfloor \frac{w}{2} \rfloor$ extra votes over d in (b), (c), (d), and (e), but one extra vote requires more than m switches in each of these voters.
- Candidates in U need $\lceil \frac{v}{2} \rceil$ extra votes over $d_i \in D_-$ in (a), (b), (c), and (f), but one extra vote requires more than m switches.
- Candidates in D_- need $\lceil \frac{v}{2} \rceil$ extra votes over c in (a), (d), and (e), but one extra vote requires more than m switches.
- Candidates in C_- need $\lfloor \frac{w}{2} \rfloor$ extra votes over d in (b), (c), (d), and (e), but one extra vote requires more than m switches.

Thus, we have shown that

$$\text{weakDodgsonScore}_{E_3}(x) > m > \text{weakDodgsonScore}_{E_3}(c)$$

for all $x \in C' \setminus \{c, d\}$. It is now easy to see that

(1) if $\text{weakDodgsonScore}_{E_1}(c) \leqslant \text{weakDodgsonScore}_{E_2}(d)$, then c is a weakDodgson winner in E_3, and

(2) if $\text{weakDodgsonScore}_{E_1}(c) > \text{weakDodgsonScore}_{E_2}(d)$, then b is the unique weakDodgson winner in E_3.

Let $I = ((C, c, V), (D, d, W))$ be an instance of w2ER. We have just argued that I is in w2ER if and only if c is a weakDodgson winner in E_3, which immediately implies Θ_2^p-hardness of the weakDodgson winner determination problem. □

8.4 SUMMARY

We have shown that computing winners is intractable for ranked pairs, Young's rule, and Dodgson's rule. In comparison to Young's rule, Dodgson's rule, and many other intractable SCFs, the ranked pairs method is easier in at least two important respects. First, *some* winner can be found efficiently. Second, winner determination is easy in most practical instances. The reason for the latter is that the *expected* number of ties among two or more pairs is rather small. This is particularly true when the number of voters is large compared to the number of alternatives, which is the case in many realistic settings.

It is therefore to be expected that ranked pairs winners are easy to compute *on average* for most reasonable distributions of individual preferences.

Another interesting aspect in the context of the ranked pairs method is the trade-off between tractability and neutrality: while the resolute variant RP_τ is easy to compute, the neutral version RP is intractable.

9

POSSIBLE AND NECESSARY WINNERS OF PARTIAL TOURNAMENTS

In this chapter, we study the problem of computing possible and necessary winners for partially specified weighted and unweighted tournaments. This problem arises naturally in elections with incompletely specified votes, partially completed sports competitions, and more generally in any scenario where the outcome of some pairwise comparisons is not yet fully known. Section 9.1 motivates the setting in more detail and points out differences to related work. After introducing terminology (Section 9.2) and computational problems (Section 9.3) for partial tournaments, we consider unweighted tournament solutions in Section 9.4 and weighted tournament solutions in Section 9.5. We show that for most of the considered solution concepts, possible and necessary winners can be identified in polynomial time. These positive algorithmic results stand in sharp contrast to earlier results concerning possible and necessary winners given partially specified preference profiles.

9.1 MOTIVATION

Tournament solutions play an important role in the mathematical social sciences at large. Application areas include multi-criteria decision analysis (Arrow and Raynaud, 1986; Bouyssou et al., 2006), zero-sum games (Fisher and Ryan, 1995; Laffond et al., 1993b; Duggan and Le Breton, 1996a), coalitional games (Brandt and Harrenstein, 2010), and argumentation theory (Dung, 1995; Dunne, 2007).

When choosing from a tournament, relevant information may only be partly available. This could be because some preferences are yet to be elicited, some matches yet to be played, or certain comparisons yet to be made. In such cases, it is natural to speculate which are the potential and inevitable outcomes on the basis of the information already at hand. Given any tournament solution S, *possible winners* of a partial tournament G are defined as alternatives that are selected by S in *some* completion of G, and *necessary winners* are alternatives that are selected in *all* completions. By a completion we here understand a complete tournament extending G.

In this chapter we address the computational complexity of identifying the possible and necessary winners for a number of weighted and unweighted tournament solutions whose winner determination problem for complete tournaments is tractable. We consider four of the most common tournament solutions—namely, Condorcet winners (*COND*), the Copeland solution (*CO*), the top cycle (*TC*), and the uncovered set (*UC*)—and three common solutions for weighted tournaments—Borda (*BO*), maximin (*MM*) and the resolute variant of ranked pairs (RP_τ). For each of these solution concepts, we characterize the complexity of the following problems: deciding whether a given alternative is a possible winner (*PW*), deciding whether a given alternative is a necessary winner (*NW*), and deciding whether a given subset of alternatives equals the set of winners in some completion (*PWS*). These problems can be challenging, as even unweighted partial tournaments may allow for an exponential number of completions.

Similar problems have been considered before. For Condorcet winners, voting trees and the top cycle, it was already shown that possible and necessary winners are computable in polynomial time (Lang et al., 2012; Pini et al., 2008, 2011). The same holds for the problem of computing possible Copeland winners, which was considered in the context of sports tournaments (Cook et al., 1998).

A more specific setting that is frequently considered within the area of computational social choice differs from our setting in a subtle but important way that is worth being pointed out. There, tournaments are assumed to represent the majority graph for a given preference profile.[1] Since a *partial* preference profile R need not conclusively settle every majority comparison, it may give rise to a *partial* tournament $G(R)$ only. There are two natural ways to define possible and necessary winners for a partial preference profile R and solution concept S. The first is to consider the completions of the incomplete tournament $G(R)$ and the winners under S in these. This approach is covered by the setting in this chapter. The second is to consider the completions of R and the winners under S in the corresponding tournaments.[2] Since every tournament corresponding to a completion of R is also a completion of $G(R)$ but not necessarily the other way round, the second definition gives rise to a *stronger* notion of a possible winner and a *weaker* notion of a necessary winner.[3] Interestingly, and in sharp contrast to our results, determining these stronger possible and weaker necessary winners is intractable for many SCFs (Lang et al., 2012; Xia and Conitzer, 2011).

1 See, e.g., (Baumeister and Rothe, 2010; Betzler and Dorn, 2010; Konczak and Lang, 2005; Walsh, 2007; Xia and Conitzer, 2011) for the basic setting, (Betzler et al., 2009) for parameterized complexity results, (Hazon et al., 2008; Kalech et al., 2011) for probabilistic settings, and (Chevaleyre et al., 2010; Xia et al., 2011) for settings with a variable set of alternatives.

2 Lang et al. (2012) and Pini et al. (2011) compare—both theoretically and experimentally—these two ways of defining possible and necessary winners for three solution concepts: Condorcet winners, voting trees and the top cycle.

3 More precisely, the alternatives that win in a tournament induced by some completion of R form a subset of the possible winners of $G(R)$, and the alternatives that win in all tournaments induced by a completion of R form a superset of the necessary winners of $G(R)$.

In the context of this chapter, we do not assume that tournaments arise from majority comparisons in voting or from any other specific procedure. This approach has a number of advantages. Firstly, it matches the diversity of settings to which tournament solutions are applicable, which goes well beyond social choice and voting. For instance, our results also apply to a question commonly encountered in sports competitions, namely, which teams can still win the cup and which future results this depends on (see, e.g., Cook et al., 1998; Kern and Paulusma, 2004). Secondly, (partial) tournaments provide an informationally sustainable way of representing the relevant aspects of many situations while maintaining a workable level of abstraction and conciseness. For instance, in the social choice setting described above, the partial tournament induced by a partial preference profile is a much more succinct piece of information than the preference profile itself. Finally, specific settings may impose restrictions on the feasible extensions of partial tournaments. The positive algorithmic results in this chapter can be used to efficiently approximate the sets of possible and necessary winners in such settings, where the corresponding problems may be intractable. The voting setting discussed above serves to illustrate this point.

9.2 PARTIAL TOURNAMENTS

Let A denote a finite set of alternatives.[4] A *partial tournament* is an asymmetric directed graph $G = (A, E)$. If E is complete, we have a (complete) tournament. The class of all complete tournaments with vertex set A is denoted by $\mathscr{T}$. Given two partial tournaments $G = (A, E)$ and $G' = (A', E')$, G' is called an *extension* of G, denoted $G \leqslant G'$, if $A = A'$ and $E \subseteq E'$. If E' is complete, G' is called a *completion of* G. We write $[G]$ for the set of completions of G, i.e., $[G] = \{T \in \mathscr{T} : G \leqslant T\}$. *partial tournament* *completion*

For $G = (A, E)$ and $x \in A$, the *dominion* of x is given by $D_E(x) = \{y \in A : (x, y) \in E\}$, and the *dominators* of x are given by $\overline{D}_E(x) = \{y \in A : (y, x) \in E\}$. For $X \subseteq A$, we let $D_E(X) = \bigcup_{x \in X} D_E(x)$ and $\overline{D}_E(X) = \bigcup_{x \in X} \overline{D}_E(x)$. *dominion* *dominators*

For a subset $X \subseteq A$ of alternatives, we further write $E^{X\rightarrow}$ for the set of edges obtained from E by adding all missing edges from alternatives in X to alternatives not in X, i.e.,

$$E^{X\rightarrow} = E \cup \{(x, y) \in X \times (A \setminus X) : (y, x) \notin E\}.$$

We use $E^{X\leftarrow}$ as an abbreviation for $E^{A \setminus X\rightarrow}$, and respectively write $E^{x\rightarrow}$, $E^{x\leftarrow}$, $G^{X\rightarrow}$, and $G^{X\leftarrow}$ for $E^{\{x\}\rightarrow}$, $E^{\{x\}\leftarrow}$, $(A, E^{X\rightarrow})$, and $(A, E^{X\leftarrow})$.

4 In this chapter, typical elements of A will be denoted by x, y instead of a, b, in order not to confuse alternatives with capacity functions of b-matchings (see page 133).

Let n be a positive integer. A *partial* n*-weighted tournament* is a pair $G = (A, w)$ consisting of a finite set of alternatives A and a weight function $w : A \times A \to \{0, \dots, n\}$ such that for each pair $(x, y) \in A \times A$ with $x \neq y$, $w(x, y) + w(y, x) \leqslant n$. We say that $T = (A, w)$ is a *(complete)* n*-weighted tournament* if for all $x, y \in A$ with $x \neq y$, $w(x, y) + w(y, x) = n$. A (partial or complete) *weighted tournament* is a (partial or complete) n-weighted tournament for some $n \in \mathbb{N}$. The class of n-weighted tournaments is denoted by $\mathscr{T}[n]$. Observe that with each partial 1-weighted tournament (A, w) we can associate a partial tournament (A, E) by setting $E = \{(x, y) \in A \times A : w(x, y) = 1\}$. Thus, (partial) n-weighted tournaments can be seen to generalize (partial) tournaments, and we may identify $\mathscr{T}[1]$ with $\mathscr{T}$.

weighted tournament

The notations $G \leqslant G'$ and $[G]$ can be extended naturally to partial n-weighted tournaments $G = (A, w)$ and $G' = (A', w')$ by letting $G \leqslant G'$ if $A = A'$ and $w(x, y) \leqslant w'(x, y)$ for all $x, y \in A$, and $[G] = \{T \in \mathscr{T}[n] : G \leqslant T\}$.

For given $G = (A, w)$ and $X \subseteq A$, we further define $w^{X\to}$ such that for all $x, y \in A$,

$$w^{X\to}(x, y) = \begin{cases} n - w(y, x) & \text{if } x \in X \text{ and } y \notin X, \\ w(x, y) & \text{otherwise,} \end{cases}$$

and set $w^{X\leftarrow} = w^{A\setminus X\to}$. Moreover, $w^{x\to}$, $w^{x\leftarrow}$, $G^{X\to}$, and $G^{X\leftarrow}$ are defined in the obvious way.

9.3 POSSIBLE & NECESSARY WINNERS

We are now ready to formally define the notions of possible and necessary winners, along with the corresponding computational problems. The following paragraphs refer to both the weighted and the unweighted case.

Tournament solutions select alternatives from *complete* tournaments. A partial tournament, on the other hand, can be extended to a number of complete tournaments, and a tournament solution selects a (potentially different) set of alternatives for each of them.

For a given tournament solution S, we can thus define the set of *possible winners* for a partial tournament G as the set of alternatives selected by S from *some* completion of G, i.e., as $PW_S(G) = \bigcup_{T\in[G]} S(T)$. Analogously, the set of *necessary winners* of G is the set of alternatives selected by S from *every* completion of G, i.e., $NW_S(G) = \bigcap_{T\in[G]} S(T)$. We can finally write $PWS_S(G) = \{S(T) : T \in [G]\}$ for the set of *sets* of alternatives that S selects for the different completions of G.

possible winners

necessary winners

Note that $NW_S(G)$ may be empty even if S selects a nonempty set of alternatives for each tournament $T \in [G]$, and that $|PWS_S(G)|$ may be exponential in the number of alternatives of G. It is also easily verified that $G \leqslant G'$ implies $PW_S(G') \subseteq PW_S(G)$ and $NW_S(G) \subseteq$

$NW_S(G')$, and that $PW_S(G) = \bigcup_{G \leqslant G'} NW_S(G')$ and $NW_S(G) = \bigcap_{G \leqslant G'} PW_S(G')$.

Deciding membership in the sets $PW_S(G)$, $NW_S(G)$, and $PWS_S(G)$ for a given solution concept S and a partial tournament G is a natural computational problem. We will refer to these problems as PW_S, NW_S, and PWS_S, respectively.[5]

PW_S, NW_S, PWS_S

For complete tournaments T we have $[T] = \{T\}$ and thus $PW_S(T) = NW_S(T) = S(T)$ and $PWS_S(T) = \{S(T)\}$. As a consequence, for solution concepts S with an NP-hard winner determination problem—like the Banks set and *TEQ*—the problems PW_S, NW_S, and PWS_S are NP-hard as well. We therefore restrict our attention to solution concepts for which winners can be computed in polynomial time.

For irresolute solution concepts, PWS_S may appear a more complex problem than PW_S. We are, however, not aware of a polynomial-time reduction from PW_S to PWS_S. The relationship between these problems may also be of interest for the "classic" possible winner setting with partial preference profiles.

9.4 UNWEIGHTED TOURNAMENTS

In this section, we consider the tournament solutions *COND*, *CO*, *TC*, and *UC*. Weighted tournament solutions will then be considered in Section 9.5.

9.4.1 *Condorcet Winners*

Condorcet winners are a very simple solution concept and will provide a nice warm-up. Since both PW_{COND} and NW_{COND} easily reduce[6] to PWS_{COND}, we focus on the latter problem. Each of the sets in $PWS_{COND}(G)$ is either a singleton or the set A of all alternatives, and determining membership for a singleton is obviously tractable. Checking whether $A \in PWS_{COND}(G)$ is not quite that simple. First observe that $A \in PWS_{COND}(G)$ if and only if there is an extension G' of G in which every alternative is dominated by some other alternative. Given a particular $G = (A, E)$, we can define an extension $G' = (A, E')$ of G by iteratively adding edges from dominated alternatives to undominated ones until this is no longer possible. Formally, let

$$E_0 = E \text{ and } E_{i+1} = E_i \cup \{(x, y) \in X_i \times Y_i : (y, x) \notin E_i\},$$

---

5 Formally, the input for each of the problems consists of an encoding of the partial (n-weighted) tournament G and, for partial n-weighted tournaments, the number n.

6 For a partial tournament G and an alternative x, the following statements hold:

(*i*) $x \in PW_{COND}(G) \Leftrightarrow PWS_{COND}(G) \cap \{A, \{x\}\} \neq \emptyset$, and

(*ii*) $x \in NW_{COND}(G) \Leftrightarrow PWS_{COND}(G) \subseteq \{A, \{x\}\}$.

where X_i and Y_i denote the dominated and undominated alternatives of (A, E_i), respectively. Finally define $E' = \bigcup_{i=0}^{|A|} E_i$, and observe that this set can be computed in polynomial time.

Now, for every undominated alternative x of G' and every dominated alternative y of G', we not only have $(x, y) \in E'$, but also $(x, y) \in E$. This is the case because in the inductive definition of E' only edges from dominated to undominated alternatives are added in every step. It is therefore easily verified that $PWS_{COND}(G)$ contains A if and only if the set of undominated alternatives in G' is either empty or is of size three or more. We have shown the following easy result.

Theorem 9.1. *PW_{COND}, NW_{COND}, and PWS_{COND} can be solved in polynomial time.*

We note that Theorem 9.1 is a corollary of corresponding results for maximin in Section 9.5.2. The reason is that *MM* coincides with *COND* in 1-weighted tournaments.

9.4.2 *Copeland*

Copeland scores coincide with Borda scores in the case of 1-weighted tournaments. As a consequence, the following is a direct corollary of the results in Section 9.5.1.

Theorem 9.2. *NW_{CO}, PW_{CO}, and PWS_{CO} can be solved in polynomial time.*

PW_{CO} can alternatively be solved via a polynomial-time reduction to maximum network flow (see, e.g., Cook et al., 1998, p. 51).

9.4.3 *Top Cycle*

Lang et al. (2012) have shown that possible and necessary winners for *TC* can be computed efficiently by greedy algorithms. For PWS_{TC}, we not only have to check that there exists a completion such that the set in question is dominating, but also that there is no smaller dominating set. It turns out that this can still be done in polynomial time.

Theorem 9.3. *PWS_{TC} can be solved in polynomial time.*

Proof. Consider a partial tournament $G = (A, E)$ and a set $X \subseteq A$ of alternatives. If X is a singleton, the problem reduces to checking whether $X \in PWS_{COND}(G)$. If X is of size two or if one of its elements is dominated by an outside alternative, $X \notin PWS_{TC}(G)$. Therefore, we can without loss of generality assume that $|X| \geqslant 3$ and $(y, x) \notin E$ for all $y \in A \setminus X$ and $x \in X$. The *Smith set* of a partial tournament is defined as the minimal dominant subset of alternatives (Smith, 1973).[7] It can

Smith set

7 For complete tournaments, the Smith set coincides with the top cycle.

be shown that there exists a completion $T \in [G]$ with $TC(T) = X$ if and only if the Smith set of the partial tournament $(X, E|_{X \times X})$ equals the whole set X. Since the Smith set of a partial tournament can be computed efficiently (Brandt et al., 2009a), the theorem follows. $\square$

9.4.4 *Uncovered Set*

We will work with the following alternative characterization of the uncovered set via the *two-step principle*: an alternative is in the uncovered set if and only if it can reach every other alternative in at most two steps.[8] Formally, $x \in UC(T)$ if and only if for all $y \in A \setminus \{x\}$, either $(x, y) \in E$ or there is some $z \in A$ with $(x, z), (z, y) \in E$. We denote the two-step dominion $D_E(D_E(x))$ of an alternative x by $D^2_E(x)$.

two-step principle

We first consider PW_{UC}, for which we check for each alternative whether it can be reinforced to reach every other alternative in at most two steps.

Theorem 9.4. *PW_{UC} can be solved in polynomial time.*

Proof. For a given partial tournament $G = (A, E)$ and an alternative $x \in A$, we check whether x is in $UC(T)$ for some completion $T \in [G]$. Consider the graph $G' = (A, E'')$ where E'' is derived from E as follows. First, we let the dominion of x grow as much as possible by letting $E' = E^{x\rightarrow}$. Then, we do the same for its two-step dominion by defining E'' as $E'^{D_{E'}(x)\rightarrow}$. Now it can be shown that $x \in PW_{UC}(G)$ if and only if $A = \{x\} \cup D_{E''}(x) \cup D^2_{E''}(x)$. $\square$

A similar argument yields the following.

Theorem 9.5. *NW_{UC} can be solved in polynomial time.*

Proof. For a given partial tournament $G = (A, E)$ and an alternative $x \in A$, we check whether x is in $UC(T)$ for all completions $T \in [G]$. Consider the graph $G' = (A, E'')$ with E'' defined as follows. First, let $E' = E^{x\leftarrow}$. Then, expand it to $E'' = E'^{\overline{D}_{E'}(x)\rightarrow}$. Intuitively, this makes it as hard as possible for x to beat alternatives outside of its dominion in two steps. Indeed, it can be shown that $x \in NW_{UC}(G)$ if and only if $A = \{x\} \cup D_{E''}(x) \cup D^2_{E''}(x)$. $\square$

For all solution concepts considered so far, *PW* and *PWS* have the same complexity. One might wonder whether a result like this holds more generally, and whether there could be a polynomial-time reduction from *PWS* to *PW*. The following result shows that this is not the case, unless P=NP.

Theorem 9.6. *PWS_{UC} is NP-complete.*

---

8 In graph theory, vertices satisfying this property are called *kings*.

Proof sketch. Let $G = (A, E)$ be a partial tournament. Given a set $X \subseteq A$ and a completion $T \in [G]$, it can be checked in polynomial time whether $X = UC(T)$. Hence, PWS_{UC} is obviously in NP.

NP-hardness can be shown by a reduction from SAT. For each Boolean formula φ in conjunctive normal-form with a set C of clauses and set P of propositional variables, we construct a partial tournament $G_\varphi = (A_\varphi, E_\varphi)$. Define

$$A_\varphi = C \times \{0,1\} \cup P \times \{0,\dots,5\} \cup \{0,1,2\},$$

i.e., along with three auxiliary alternatives, we introduce for each clause two alternatives and for each propositional variable six. We write c_i, p_i, C_i, and P_i for (c,i), (p,i), $\{c_i : c \in C\}$, and $\{p_i : p \in P\}$, respectively. Let

$$X = C \times \{0\} \cup P \times \{0,1,2\} \cup \{0,1,2\}.$$

Then, E_φ is defined such that it contains no edges between alternatives in $A_\varphi \setminus X$. For alternatives $x \in X$, E_φ is given by the following table, in which each line is of the form $\overline{D}_{E_\varphi}(x) \cap A \setminus X \to x \to D_{E_\varphi}(x) \cap X$ and where it is understood that x dominates all alternatives in $A_\varphi \setminus X$ unless specified otherwise. For improved readability some curly braces have been omitted and a comma indicates set-theoretic union.

$$\begin{aligned}
\{p_3 : p \in c\}, \{p_4 : \bar{p} \in c\}, c_1 &\to c_0 \to 2, P_2, \{p_1 : p \notin c\}, \{p_0 : \bar{p} \notin c\} \\
p_3 &\to p_0 \to 0, p_2, \{c_0 : \bar{p} \in c\} \\
p_4 &\to p_1 \to 0, p_2, \{c_0 : p \in c\} \\
P_3, P_4, p_5 &\to p_2 \to 2, \{q_0, q_1 : q \neq p\} \\
P_3, P_4 &\to \ 0 \ \to 2, C_0, P_2 \\
C_1, P_5 &\to \ 1 \ \to 0, C_0, P_2 \\
\emptyset &\to \ 2 \ \to 1, P_0, P_1
\end{aligned}$$

It now suffices to show that E_φ is specified in such a way that X is the uncovered set of some completion of G_φ if and only if φ is satisfiable.

For every $p \in P$, the edges between p_0, p_1, and 1 are left unspecified. The idea is that p_0 and p_1 are the only candidates to cover p_5, p_0 and 1 are the only candidates to cover p_4, and p_1 and 1 are the only candidates to cover p_3. As $p_0 \in D_{E_\varphi}(p_3)$, $p_1 \in D_{E_\varphi}(p_4)$, and $1 \in D_{E_\varphi}(p_5)$, there are two possibilities of extending G_φ in such a way that p_3, p_4 and p_5 are covered simultaneously and X is the uncovered set. Either all the edges in

(a) $\{(p_0, p_1), (p_1, 1), (1, p_0)\}$, or all those in

(b) $\{(p_1, p_0), (p_0, 1), (1, p_1)\}$

have to be added to E_φ to achieve this (additionally some edges among $A_\varphi \setminus X$ have to be set appropriately as well). Possibility (a) corresponds to setting p to "true." In this case, p_1 also covers c_1 for every clause $c \in C$ that contains p. Possibility (b) corresponds to setting p to "false" and causes p_0 to cover c_1 for every clause $c \in C$ that contains $\bar{p}$. Moreover, for each $c \in C$, the only candidates in X to cover c_1 are p_1 if $p \in c$ and p_0 if $\bar{p} \in c$. Observe that $1 \in D_{E_\varphi}(c_1)$ for all $c \in C$. Thus, if $\bar{p} \in c$, p_1 covering p_3 precludes p_0 covering c_1. Similarly, if $p \in c$, p_0 covering p_4 precludes p_1 covering c_1. Accordingly, if T is a completion of G_φ in which X is the uncovered set, one can read off a valuation satisfying φ from how the edges between p_0, p_1, and 1 are set in T. For the opposite direction, a satisfying valuation for φ is a recipe for extending G_φ to a tournament in which X is the uncovered set. It can be checked that every alternative in X reaches every other alternative in at most two steps, whereas every alternative in $A_\varphi \setminus X$ is covered by some alternative in X. □

9.5 WEIGHTED TOURNAMENTS

We now turn to weighted tournaments and consider the C2 functions *BO*, *MM*, and RP_τ.

9.5.1 *Borda*

Borda scores can be generalized to *partial* weighted tournaments $G = (A, w)$ by letting $s_{BO}(x, G) = \sum_{y \in A \setminus \{x\}} w(x, y)$. Before we proceed further, we define the notion of a b-matching, which will be used in the proofs of two of our results. Let $H = (V_H, E_H)$ be an undirected graph with vertex capacities $b : V_H \to \mathbb{N}_0$. Then, a *b-matching* of H is a function $m : E_H \to \mathbb{N}_0$ such that for all $v \in V_H$, $\sum_{e \in \{e' \in E_H : v \in e'\}} m(e) \leqslant b(v)$. The *size* of b-matching m is defined as $\sum_{e \in E_H} m(e)$. It is easy to see that if $b(v) = 1$ for all $v \in V_H$, then a maximum size b-matching is equivalent to a maximum cardinality matching. In a b-matching problem with upper *and* lower bounds, there further is a function $a : V_H \to \mathbb{N}_0$. A feasible b-matching then is a function $m : E_H \to \mathbb{N}_0$ such that $a(v) \leqslant \sum_{e \in \{e' \in E_H : v \in e'\}} m(e) \leqslant b(v)$.

$s_{BO}(x, G)$

b-*matching*

If H is bipartite, then the problem of computing a maximum size feasible b-matching with upper and lower bounds can be solved in strongly polynomial time (Schrijver, 2003, Chapter 21). We will use this fact to show that PW_{BO} and PWS_{BO} can both be solved in polynomial time. While the following result for PW_{BO} can be shown using Theorem 6.1 of Kern and Paulusma (2004), we give a direct proof that can then be extended to PWS_{BO}.

Theorem 9.7. *PW_{BO} can be solved in polynomial time.*

Proof. Let $G = (A, w)$ be a partial n-weighted tournament and $x \in A$. We give a polynomial-time algorithm for checking whether $x \in PW_{BO}(G)$, via a reduction to the problem of computing a maximum size b-matching of a bipartite graph.

Let $G^{x\rightarrow} = (A, w^{x\rightarrow})$ denote the graph obtained from G by maximally reinforcing x, and $s^* = s_{BO}(x, G^{x\rightarrow})$ the Borda score of x in $G^{x\rightarrow}$. From $G^{x\rightarrow}$, we then construct a bipartite graph $H = (V_H, E_H)$ with vertices $V_H = A \setminus \{x\} \cup E^{<n}$, where $E^{<n} = \{\{i, j\} \subseteq A \setminus \{x\} : w(i, j) + w(j, i) < n\}$,[9] and edges $E_H = \{\{v, e\} : v \in A \setminus \{x\}$ and $v \in e \in E^{<n}\}$. We further define vertex capacities $b : V_H \rightarrow \mathbb{N}_0$ such that $b(\{i, j\}) = n - w(i, j) - w(j, i)$ for $\{i, j\} \in E^{<n}$ and $b(v) = s^* - s_{BO}(v, G^{x\rightarrow})$ for $v \in A \setminus \{x\}$.

Now observe that in any completion $T = (A, w') \in [G^{x\rightarrow}]$, $w'(i, j) + w'(j, i) = n$ for all $i, j \in A$ with $i \neq j$. The sum of the Borda scores in T is therefore $n|A|(|A| - 1)/2$. Some of the weight has already been used up in $G^{x\rightarrow}$; the weight which has not yet been used up is equal to $\alpha = n|A|(|A| - 1)/2 - \sum_{v \in A} s_{BO}(v, G^{x\rightarrow})$. It can be shown that $x \in PW_{BO}(G)$ if and only if H has a b-matching of size at least α. □

This idea can be extended to a polynomial-time algorithm for PWS_{BO} where we use a similar construction for a given candidate set $X \subseteq A$ and a target Borda score s^*. Binary search can be used to efficiently search the interval $I = [\max_{x \in X} s_{BO}(x, G), n(|A| - 1)]$ of possible target scores.

Theorem 9.8. *PWS_{BO} can be solved in polynomial time.*

Proof. Let $G = (A, w)$ be a partial n-weighted tournament, and $X \subseteq A$. We give a polynomial time algorithm for checking whether $X \in PWS_{BO}(G)$, via a bisection method and a reduction to the problem of computing a maximum b-matching of a graph with lower and upper bounds.

Assume that there is a target Borda score s^* and a completion $T \in [G]$ with $X \in PWS_{BO}(T)$ and $s_{BO}(x, T) = s^*$ for all $x \in X$. Then, the maximum Borda score of an alternative not in X is $s^* - 1$.

For a given target Borda score s^*, we construct a bipartite graph $H = (V_H, E_H)$ with vertices $V_H = A \cup E^{<n}$, where $E^{<n} = \{\{i, j\} \subseteq A : i \neq j, w(i, j) + w(j, i) < n\}$, and edges $E_H = \{\{v, e\} : v \in A$ and $v \in e \in E^{<n}\}$. Only the lower bounds $b : V_H \rightarrow \mathbb{N}_0$ and upper bounds $a : V_H \rightarrow \mathbb{N}_0$ depend on s^* and are defined as follows: For vertices $x \in X$, lower and upper bounds coincide and are given by $a(x) = b(x) = s^* - s_{BO}(x, G)$. All other vertices $v \in V_H \setminus X$ have a lower bound of $a(v) = 0$. Upper bounds for these vertices are defined such that $b(v) = s^* - s_{BO}(v, G) - 1$ for $v \in A \setminus X$, and $b(\{i, j\}) = n - w(i, j) - w(j, i)$ for $\{i, j\} \in E^{<n}$. Observe that a feasible b-matching in H corresponds to an extension of G. Such an extension is a completion $T \in [G]$ if and

9 Note that $w(i, j) = w^{x\rightarrow}(i, j)$ for alternatives $i, j \in A \setminus \{x\}$.

only if the b-matching has size $\alpha = n|A|(|A|-1)/2 - \sum_{v\in A} s_{BO}(v, G)$, which equals the weight not yet used up in G. Then, T satisfies $X \in PWS_{BO}(T)$ and $s_{BO}(x, T) = s^*$ for all $x \in X$. If, on the other hand, no s^* gives rise to a graph that has a b-matching of size α, then $X \notin PWS_{BO}(G)$.

In order to obtain a polynomial time algorithm, we need to check whether there exists a target score s^* for which the corresponding graph H admits a b-matching of size α. It is easily verified that any such s^* is contained in the integer interval

$$I = [\max_{x\in X} s_{BO}(x, G), n(|A|-1)].$$

Observe that $|I|$ depends on k and thus is *not* polynomially bounded in the size of G. Checking every integer $s \in I$ is therefore not feasible in polynomial time. However, we now show that we can perform binary search in order to find s^* efficiently. We need the following two observations about the interval I. For $s \in I$, we say that s *admits a feasible* b*-matching* if the corresponding graph H has a feasible b-matching.

First, if an $s' \in I$ admits a feasible b-matching, then every $s'' \in I$ with $s'' \leqslant s'$ also admits a feasible b-matching. This is because removing all weight from edges that exceeds the (reduced) upper bounds gives a feasible b-matching for s''.

Second, with s' as before and α' the size of the corresponding *maximum* feasible b-matching, there cannot be an $s'' \in I$ with $s'' \geqslant s'$ such that the size of a maximum feasible b-matching for s'' is smaller than α'. This is because either *(i)* there is no feasible b-matching since not all lower bounds can be met, or *(ii)* there exists a feasible b-matching but its size is as least α'. To see the latter, note that a decrease in the size of a maximum feasible matching cannot be caused by upper bounds as they were only raised or remained the same. It remains to be shown that the increase in $a(v)$ for $v \in X$ does not result in a smaller maximum b-matching. Since the weight of all edges adjacent to a vertex in X in the b-matching increases, a decrease can only happen in edges $\{v', \{v, v'\}\}$ adjacent to a $v' \in V_H \setminus X$. Lowering the weight on this edge must be caused by a forced increase in weight on $\{v, \{v, v'\}\}$ but since both edges are only coupled by the $b(\{v, v'\})$, the decrease in one cannot exceed the increase in the other and therefore the size of the maximum b-matching does not decrease from s' to s''.

These two observations show that I consists of two intervals where the lower part admits feasible b-matchings of increasing size, whereas the upper part does not admit feasible b-matchings. Therefore, s^* is either at the upper end of the lower part or it does not exist.

Algorithmically, we can check the existence of s^* with the following *binary search* algorithm. Let $[I_{min}, I_{max}]$ be an interval that is initialized to $I = [\max_{x\in X} s_{BO}(x, G), n(|A|-1)]$. Consider the median value s of

binary search

this interval. If the corresponding graph H has no feasible b-matching, continue with the interval $[I_{min}, s-1]$. Otherwise, if the maximum feasible b-matching has size at least α, return "yes". If its size is less than α, continue with $[s+1, I_{max}]$. If $[I_{min}, I_{max}]$ is empty, return "no."

The number of queries of this algorithm is bounded by $\lceil \log_2 |I| \rceil$ and, therefore, polynomial in the size of G. □

We conclude this section by showing that NW_{BO} can be solved in polynomial time as well.

Theorem 9.9. *NW_{BO} can be solved in polynomial time.*

Proof. Let $G = (A, w)$ be a partial weighted tournament and $x \in A$. We give a polynomial-time algorithm for checking whether $x \in NW_{BO}(G)$.

Define $G' = G^{x\leftarrow}$. We want to check whether some other alternative $y \in A \setminus \{x\}$ can achieve a Borda score of more than $s^* = s_{BO}(x, G')$. This can be done separately for each $y \in A \setminus \{x\}$ by reinforcing it as much as possible in G'. If for some y, $s_{BO}(y, G'^{y\rightarrow}) > s^*$, then $x \notin NW_{BO}(G)$. If, on the other hand, $s_{BO}(y, G'^{y\rightarrow}) \leqslant s^*$ for all $y \in A \setminus \{x\}$, then $x \in NW_{BO}(G)$. □

9.5.2 *Maximin*

Like Borda scores, maximin scores can be adapted to partial weighted tournaments in a straightforward way. For $G = (A, w)$, the maximin score of alternative x is given by $s_{MM}(x, G) = \min_{y \in A\setminus\{x\}} w(x, y)$. We first show that PW_{MM} is polynomial-time solvable by reducing it to the problem of finding a maximum cardinality matching of an undirected graph.

$s_{MM}(x, G)$

Theorem 9.10. *PW_{MM} can be solved in polynomial time.*

Proof. We show how to check whether $x \in PW_{MM}(G)$ for a partial n-weighted tournament $G = (A, w)$. Consider $G^{x\rightarrow} = (A, w^{x\rightarrow})$. Then, $s_{MM}(x, G^{x\rightarrow})$ is the best possible maximin score x can get among all completions of G. If $s_{MM}(x, G^{x\rightarrow}) \geqslant \frac{n}{2}$, then we have $s_{MM}(y, T) \leqslant w^{x\rightarrow}(y, x) \leqslant \frac{n}{2}$ for every $y \in A \setminus \{x\}$ and every completion $T \in [G^{x\rightarrow}]$ and therefore $x \in PW_{MM}(G)$. Now consider $s_{MM}(x, G^{x\rightarrow}) < \frac{n}{2}$. We will reduce the problem of checking whether $x \in PW_{MM}(G)$ to that of finding a maximum cardinality matching, which is known to be solvable in polynomial time (Edmonds, 1965). We want to find a completion $T \in [G^{x\rightarrow}]$ such that $s_{MM}(x, T) \geqslant s_{MM}(y, T)$ for all $y \in A \setminus \{x\}$. If there exists a $y \in A \setminus \{x\}$ such that $s_{MM}(x, G^{x\rightarrow}) < s_{MM}(y, G^{x\rightarrow})$, then we already know that $x \notin PW_{MM}(G)$. Otherwise, each $y \in A \setminus \{x\}$ derives its maximin score from at least one particular edge (y, z) where $z \in A \setminus \{x, y\}$ and $w(y, z) \leqslant s_{MM}(x, G^{x\rightarrow})$. Moreover, it is clear that in any completion, y and z cannot both achieve a maximin score of less than $s_{MM}(x, G^{x\rightarrow})$ from edges (y, z) and (z, y) at the same time.

maximum cardinality matching

Construct the following undirected and unweighted graph $H = (V_H, E_H)$ where $V_H = A \setminus \{x\} \cup \{\{i,j\} \subseteq A : i \neq j\}$. Build up E_H such that: $\{i, \{i,j\}\} \in E_H$ if and only if $i \neq j$ and $w^{x\rightarrow}(i,j) \leqslant s_{MM}(x, G^{x\rightarrow})$. In this way, if i is matched to $\{i,j\}$ in H, then i derives a maximin score of less than or equal to $s_{MM}(x, G^{x\rightarrow})$ from his comparison with j. Clearly, H is polynomial in the size of G. Then, it can be shown that $x \in PW_{MM}(G)$ if and only if there exists a matching of cardinality $|A| - 1$ in H. □

For NW_{MM} we apply a similar technique as for NW_{BO}: to see whether $x \in NW_{MM}(G)$, we start with $G^{x\leftarrow}$ and check whether some other alternative can achieve a higher maximin score than x in a completion of $G^{x\leftarrow}$.

Theorem 9.11. *NW_{MM} can be solved in polynomial time.*

Proof. We show how to check whether $x \in NW_{MM}(G)$. Consider the graph $G' = (A, w') = G^{x\leftarrow}$. The maximin score of x in G' is the worst case maximin score of x, among all proper completions of G.

We now check whether there exists a $y \in A \setminus \{x\}$, which can get a higher maximin score than the maximin score of x. For each $y \in A \setminus \{x\}$, construct a graph $G_y = (A, w_y) = G'^{y\rightarrow}$. Now, we have complete information to compute the best possible maximin score of y among completions of G. If the maximin score of y in G_y is more than the maximin score of x in G', then $x \notin NW_{MM}(G)$. Otherwise, repeat the procedure for each $y \in A \setminus \{x\}$. If the maximin score of each y in corresponding G_y is not more than the maximin score of x in G', then $x \in NW_{MM}(G)$. □

We conclude the section by showing that PWS_{MM} can be solved in polynomial time. The proof proceeds by identifying the maximin values that could potentially be achieved simultaneously by all elements of the set in question, and solving the problem for each of these values using similar techniques as in the proof of Theorem 9.10. Only a polynomially bounded number of problems need to be considered.

Theorem 9.12. *PWS_{MM} can be solved in polynomial time.*

Proof. Let $G = (A, w)$ be a partial n-weighted tournament, and $X \subseteq A$. We give a polynomial time algorithm for checking whether $X \in PWS_{MM}(G)$. If $X \in PWS_{MM}(G)$ there must be a completion $T \in [G]$ and $s^* \in \{0, \dots, n\}$ such that $s_{MM}(i, T) = s^*$ for all $i \in X$. We check for $s^* > \frac{n}{2}, s^* = \frac{n}{2}$, and $s^* < \frac{n}{2}$ whether X can be made the set of maximin winners with a maximin score of s^*.

Assume that $s^* > \frac{n}{2}$. Then, $X \in PWS_{MM}$ if and only if X is a singleton $\{x\}$ and $w^{x\rightarrow}(x,j) > \frac{n}{2}$ for all $j \in A \setminus \{x\}$.

For $s^* = \frac{n}{2}$, it is possible that both (i,j) and (j,i) account for the maximin score of i and j in the completion. We create a flow network $N = (V_N, E_N, s, t, c)$ where $V_N = V_H \cup \{s, t\}$. For each $i \in A$, there *flow network* is an edge (s, i) in E_N with capacity 1. For all distinct $i, j \in A$, there are two edges $(i, \{i, j\})$ and $(j, \{i, j\})$ in E_N with capacity 1 if $w(i, j) \leqslant s^* \leqslant n - w(j, i)$; otherwise there are no edges between i, j and $\{i, j\}$ in N. For all $i, j \in X$, there is an edge $(\{i, j\}, t)$ in E_N with capacity 2. For each $i \in A$ and each $j \in A \setminus X$, E_N contains an edge $(\{i, j\}, t)$ with capacity 1. We claim that the maximum value of the flow equals $|A|$ if and only if $X \in PWS_{MM}(G)$. Here, an edge $(i, \{i, j\})$ with nonzero flow in a maximum flow corresponds to $w(i, j) = w(j, i) = s^*$ in the completion if $i, j \in X$ and to $w(i, j) < s^*$ if $i \in A \setminus X$.

Finally, we consider $s^* < \frac{n}{2}$. We will show that there are at most $|A|^2$ possible values for s^* that we need to check. Similarly as in the proof of Theorem 9.10, for a given s^*, we construct an undirected unweighted graph $H = (V_H, E_H)$ with $V_H = A \cup \{\{i, j\} \subseteq A : i \neq j\}$. Build up E_H such that if $i \in X$ then $\{i, \{i, j\}\} \in E_H$ if and only if $w(i, j) \leqslant s^*$ and $s^* \leqslant n - w(j, i)$, and if $i \in A \setminus X$ then $\{i, \{i, j\}\} \in E_H$ if and only if $w(i, j) < s^*$. We claim that there is a matching of cardinality $|A|$ in H if and only if there is a completion T in which for all $i \in X$, $s_{MM}(i, T) = s^*$ and for all $i \in A \setminus X$, $s_{MM}(i, T) < s^*$. Intuitively speaking, an edge $\{i, \{i, j\}\}$ in such a matching corresponds to $w(i, j) = s^*$ in the completion if $i \in X$ and to $w(i, j) < s^*$ if $i \in A \setminus X$. If we increase s^*, the number of edges incident to a vertex i cannot decrease for all $i \in A \setminus X$. For a vertex $i \in X$, increasing s^* can not only add edges $\{i, \{i, j\}\}$ but also remove them from H due to the condition $s^* \leqslant n - w(j, i)$. Fortunately, the latter can only happen $|X| \cdot |A| \leqslant |A|^2$ times as this is the maximum number of distinct values for $w(j, i)$. As these are the only values for s^* where H is about to lose an edge and adding edges cannot decrease the size of the biggest matching, we only need to check for $s^* \in \{n - w(i', i'') : i' \in A, i'' \in X\} \cup \{0, \frac{n}{2} - 1\}$.

matching

Obviously, all cases can be completed in polynomial time. □

9.5.3 *Ranked Pairs*

We have shown in Section 8.1 that computing ranked pairs winners is NP-hard for Tideman's original variant of the method. As mentioned in Section 9.3, this immediately implies that all problems concerning possible or necessary winners are NP-hard as well. Therefore, we focus on the tractable variant RP_τ, which uses a fixed tie-breaking rule τ. For this variant, the possible winner problem turns out to be NP-hard. This also shows that tractability of the winner determination problem, while necessary for tractability of *PW*, is not generally sufficient.

Theorem 9.13. *PW_{RP_τ} is NP-complete.*

Proof. Membership in NP is obvious, as for a given completion and a given tie-breaking rule, the ranked pairs winner can be found efficiently.

NP-hardness can be shown by a reduction from SAT. For a Boolean formula φ in conjunctive normal-form with a set C of clauses and set P of propositional variables, we construct a partial 8-weighted tournament $G_\varphi = (A_\varphi, w_\varphi)$ as follows. For each variable $p \in P$, A_φ contains two *literal alternatives* p and $\bar{p}$ and two *auxiliary alternatives* p' and $\bar{p}'$. For each clause $c \in C$, there is an alternative c. Finally, there is an alternative d for which membership in $PW_{RP_\tau}(G_\varphi)$ is to be decided.

In order to conveniently describe the weight function w_φ, let us introduce the following terminology. For two alternatives $x, y \in A_\varphi$, say that there is a *heavy* edge from x to y if $w_\varphi(x, y) = 8$ (and therefore $w_\varphi(y, x) = 0$). A *medium* edge from x to y means $w_\varphi(x, y) = 6$ and $w_\varphi(y, x) = 2$, and a *light* edge from x to y means $w_\varphi(x, y) = 5$ and $w_\varphi(y, x) = 3$. Finally, a *partial* edge between x and y means $w_\varphi(x, y) = w_\varphi(y, x) = 1$.

We are now ready to define w_φ. For each variable $p \in P$, we have heavy edges from p to $\bar{p}'$ and from $\bar{p}$ to p', and partial edges between p and p' and between $\bar{p}$ and $\bar{p}'$. For each clause $c \in C$, we have a medium edge from c to d and a heavy edge from the literal alternative $\ell_i \in \{p, \bar{p}\}$ to c if the corresponding literal ℓ_i appears in the clause c. Finally, we have heavy edges from d to all auxiliary alternatives and light edges from d to all literal alternatives. For all pairs x, y for which no edge has been specified, we define $w_\varphi(x, y) = w_\varphi(y, x) = 4$.

Observe that the only pairs of alternatives for which w_φ is not fully specified are those pairs that are connected by a partial edge. It can be shown that alternative d is a possible ranked pairs winner in G_φ if and only if φ is satisfiable. Intuitively, choosing a completion w' of w_φ such that $w'(p', p)$ is large and $w'(\bar{p}', \bar{p})$ is small corresponds to setting the variable p to "true." Since the proof is very similar to the one of Theorem 8.7, we omit the details. □

Since RP_τ is resolute, hardness of PWS_{RP_τ} follows immediately.

Corollary 9.14. *PWS_{RP_τ} is NP-complete.*

Computing necessary ranked pairs winners turns out to be coNP-complete. This is again somewhat surprising, as computing necessary winners is often considerably easier than computing possible winners, for both partial tournaments and partial preference profiles (Xia and Conitzer, 2011).

Theorem 9.15. *NW_{RP_τ} is coNP-complete.*

Proof. Membership in coNP is again obvious. For hardness, we give a reduction from UNSAT that is a slight variation of the reduction in the proof of Theorem 9.13. We introduce a new alternative d^*, which has heavy edges to all alternatives in A_φ except d. Furthermore, there is a light edge from d to d^*. It can be shown that d^* is a necessary ranked pairs winner in this partial 8-weighted tournament if and only if φ is unsatisfiable. □

9.6 SUMMARY

In this chapter, we have investigated the problem of computing possible and necessary winners in a setting where partially specified (weighted or unweighted) tournaments instead of preference profiles are given as input. Table 8 summarizes our findings.

A key conclusion is that computational problems for partial tournaments can be significantly easier than their counterparts for partial profiles. For example, possible Borda or maximin winners can be found efficiently for partial tournaments, whereas the corresponding problems for partial profiles are NP-complete (Xia and Conitzer, 2011).

While tractability of the winner determination problem is necessary for tractability of the possible or necessary winners problems, the results for RP_τ show that it is not sufficient. We further considered the problem of deciding whether a given subset of alternatives equals the winner set for some completion of the partial tournament. The results for *UC* imply that this problem cannot be reduced to the computation of possible or necessary winners. Whether a reduction exists in the opposite direction remains an open problem.

Partial tournaments have also been studied in their own right, independent of their possible completions. For instance, Dutta and Laslier (1999) and Peris and Subiza (1999) have generalized several tournament solutions to incomplete tournaments by directly adapting their definitions. In this context, the notion of possible winners suggests a canonical way to generalize a tournament solution to incomplete tournaments, sometimes referred to as the "conservative extension." The positive computational results in this chapter are an indication that this may be a promising approach.

S	PW_S		NW_S		PWS_S	
COND	in P	(Lang et al., 2012)	in P	(Lang et al., 2012)	in P	(Thm. 9.1)
CO	in P	(Thm. 9.2)[a]	in P	(Thm. 9.2)[a]	in P	(Thm. 9.2)
TC	in P	(Lang et al., 2012)[a]	in P	(Lang et al., 2012)	in P	(Thm. 9.3)
UC	in P	(Thm. 9.4)	in P	(Thm. 9.5)	NP-C	(Thm. 9.6)
BO	in P	(Thm. 9.7)[a]	in P	(Thm. 9.9)	in P	(Thm. 9.8)
MM	in P	(Thm. 9.10)[a]	in P	(Thm. 9.11)	in P	(Thm. 9.12)
RP_τ	NP-C	(Thm. 9.13)	coNP-C	(Thm. 9.15)	NP-C	(Cor. 9.14)

[a] This P-time result contrasts with the intractability of the same problem for partial preference profiles (Lang et al., 2012; Xia and Conitzer, 2011).

Table 8: Complexity of computing possible and necessary winners of partial tournaments

10 MINIMAL RETENTIVE SETS IN TOURNAMENTS

We now focus on axiomatic properties of unweighted tournament solutions. For any given tournament solution S, there is another tournament solution $\mathring{S}$ which returns the union of all inclusion-minimal sets that satisfy S-retentiveness, a natural stability criterion with respect to S. Schwartz's tournament equilibrium set (*TEQ*) is defined recursively as $TEQ = \mathring{TEQ}$. After introducing the necessary concepts, we study in Section 10.4 under which circumstances a number of important axiomatic properties are inherited from S to $\mathring{S}$. We obtain sequences of attractive and efficiently computable tournament solutions that "approximate" *TEQ*, which itself is computationally intractable. The asymptotic behavior of these sequences is studied in Section 10.5. In Section 10.6, we finally prove a weaker version of a recently disproved conjecture surrounding *TEQ*, which establishes $\mathring{TC}$—a refinement of the top cycle—as an interesting new tournament solution.

10.1 MOTIVATION

The *tournament equilibrium set* (*TEQ*), introduced by Schwartz (1990), ranks among the most intriguing, but also among the most enigmatic, tournament solutions. Schwartz defined *TEQ* on the basis of the concept of *retentiveness*. For a given tournament solution S, a set B of alternatives is said to be S-*retentive* if S selects from each dominator set of some alternative in B a subset of alternatives that is contained in B. The requirement of retentiveness can be argued for from the perspective of cooperative majority voting, where the voters have to come to an eventual agreement as to which alternative to elect (see Schwartz, 1990, for more details). Additionally, retentiveness strongly resembles the game-theoretic notion of closure under best-response behavior (Basu and Weibull, 1991).

TEQ

retentiveness

Schwartz defines *TEQ* as the union of all inclusion-minimal *TEQ*-retentive sets. This is a proper recursive definition, as the cardinality of the set of dominators of an alternative in a particular set is always smaller than the cardinality of the set itself. Schwartz furthermore conjectured that every tournament contains a *unique* minimal *TEQ*-retentive set. As was shown by Laffond et al. (1993a) and Houy (2009a,b), *TEQ* satisfies any one of a number of important properties such as *monotonicity* if and only if Schwartz's conjecture holds. Brandt et al. (2013) recently disproved Schwartz's conjecture by showing the existence of a counter-example of astronomic proportions. The interest in *TEQ* and retentiveness in general, however, is

Schwartz's conjecture

hardly diminished as concrete counter-examples to Schwartz's conjecture have never been encountered, even when resorting to extensive computer experiments (Brandt et al., 2010). Apparently, *TEQ* satisfies the above mentioned properties for all practical matters. A small number of properties is known to hold independently of Schwartz's conjecture: TEQ is contained in the Banks set (Schwartz, 1990), satisfies composition-consistency (Laffond et al., 1996), and is NP-hard to compute (Brandt et al., 2010).

In this chapter, we intend to shed more light on the fascinating notion of retentiveness by viewing the matter from a more general perspective. For any given tournament solution S, we define another tournament solution $\mathring{S}$ ("S ring") which yields the union of all minimal S-retentive sets. The top cycle, for example, coincides with $\mathring{TRIV}$. By definition, TEQ is the only tournament solution S for which $\mathring{S}$ equals S.

With every tournament solution S we then associate an infinite sequence $(S^{(0)}, S^{(1)}, S^{(2)}, \dots)$ of tournament solutions such that $S^{(0)} = S$ and $S^{(k+1)} = \mathring{S^{(k)}}$ for all $k \geqslant 0$. Our investigation concentrates on three main issues regarding such sequences and the solution concepts therein: *(i)* the inheritance of desirable properties, *(ii)* their asymptotic behavior, and *(iii)* the uniqueness of minimal retentive sets.

First, while *TEQ* itself fails to satisfy the desirable properties mentioned above in very large tournaments, we do know that some less sophisticated tournament solutions such as *TRIV* do. In Section 10.4, we therefore investigate which properties are inherited from S to $\mathring{S}$, and vice versa. We find that the former is the case for most of the properties mentioned above, provided that S always admits a unique minimal S-retentive set, whereas the latter also holds without this assumption. Composition-consistency is a notable exception: we prove that *TEQ* is the *only* composition-consistent tournament solution defined via retentiveness.

Second, we find that for every S the sequence $(S^{(0)}, S^{(1)}, S^{(2)}, \dots)$ converges to *TEQ*. In Section 10.5, we investigate the properties of these sequences in more detail by focusing on the class of tournaments for which Schwartz's conjecture holds. We show that all tournament solutions in the sequence associated with the trivial tournament solution *TRIV* are contained in one another, contain *TEQ*, and, by the inheritance results of Section 10.4, share the desirable properties of *TRIV*. Efficient computability turns out to be inherited from S to $\mathring{S}$ even without any additional assumptions. While this does not imply that *TEQ* itself is efficiently computable, the tournament solutions in the sequence for *TRIV* provide better and better efficiently computable *approximations* of *TEQ*. We also establish tight bounds on the minimal number k such that $S^{(k)}$ is guaranteed to coincide with *TEQ*, relative to the size of the tournament in question.

Third, the sequence associated with each tournament solution gives rise to a corresponding sequence of weaker versions of Schwartz's conjecture. The first such statement regarding the sequence for *TRIV* alleges that every tournament has a unique minimal *TRIV*-retentive set and was proved by Good (1971). In Section 10.6 we prove the second statement: there is a unique minimal *TC*-retentive set in every tournament. We conclude by giving an example of a well-known tournament solution for which the analogous statement does not hold. More precisely, we identify a tournament with disjoint *Copeland-retentive* sets.

10.2 TOURNAMENTS

In this section, we provide the terminology and notation required for our results. For a more extensive treatment of tournament solutions and their properties the reader is referred to Laslier (1997). In contrast to the previous chapter, we define tournament solutions for a *variable agenda*. This allows us to formulate properties that relate choice sets from different subtournaments with each other (see Definition 10.10 on page 148). The definition of a tournament needs to be adapted as well for the purposes of this chapter.

variable agenda

Let U denote a universe of alternatives. A (finite) *tournament* T is a pair $(A, \succ)$, where A is a finite and nonempty subset of U and $\succ$ is an asymmetric and complete binary relation on U, usually referred to as the *dominance relation*. Intuitively, $a \succ b$ signifies that alternative a is preferable to alternative b. In a social choice context, the dominance relation is defined via the strict majority relation P_M, but $\succ$ may have very different interpretations in other contexts. The dominance relation can be extended to sets of alternatives by writing $A \succ B$ when $a \succ b$ for all $a \in A$ and $b \in B$. We further write $\mathscr{T}(U)$ for the set of all tournaments $(A, \succ)$ with $A \subseteq U$.

tournament

dominance relation

The dominion and the dominators of an alternative (see page 127) can be defined with respect to a given subset of alternatives. For a tournament $(A, \succ)$, an alternative $a \in A$, and a subset $B \subseteq A$ of alternatives, we denote by $D_{B,\succ}(a)$ the *dominion of* a *in* B, i.e.,

dominion

$$D_{B,\succ}(a) = \{\, b \in B : a \succ b \,\},$$

and by $D_{B,\succ}(a)$ the *dominators of* a *in* B, i.e.,

dominators

$$D_{B,\succ}(a) = \{\, b \in B : b \succ a \,\}.$$

Whenever the dominance relation $\succ$ is known from the context or B is the set of all alternatives A, the respective subscript will be omitted to improve readability. We further write $T|_B = (B, \{(a, b) \in B \times B : a \succ b\})$ for the restriction of T to B.

The *order* $|T|$ of a tournament $T = (A, \succ)$ refers to the cardinality

order

of A, and $\mathscr{T}_n$ denotes the set of all tournaments with at most n alternatives, i.e.,

$$\mathscr{T}_n = \{T \in \mathscr{T}(U) : |T| \leqslant n\}.$$

We will sometimes write that a statement holds *in* $\mathscr{T}_n$ if the statement holds for all tournaments $T \in \mathscr{T}_n$.

tournament isomorphism

A *tournament isomorphism* of two tournaments $(A, \succ)$ and $(A', \succ')$ is a bijection $\pi\colon A \to A'$ such that for all $a, b \in A$, $a \succ b$ if and only if $\pi(a) \succ' \pi(b)$.

An important structural notion in the context of tournaments is that of a *component*. A component is a subset of alternatives that bear the same relationship to all alternatives not in the set.

Definition 10.1. *Let* $T = (A, \succ)$ *be a tournament. A nonempty subset* B *of* A *is a* component *of* T *if for all* $a \in A \setminus B$*, either* $B \succ \{a\}$ *or* $\{a\} \succ B$.

For a given tournament $\tilde{T}$, a new tournament T can be constructed by replacing each alternative with a component. For notational convenience, we tacitly assume that $\mathbb{N} \subseteq U$.

Definition 10.2. *Let* $B_1, \dots, B_k \subseteq U$ *be pairwise disjoint sets and consider tournaments* $\tilde{T} = (\{1, \dots, k\}, \tilde{\succ})$ *and* $T_1 = (B_1, \succ_1), \dots, T_k = (B_k, \succ_k)$. *The* product *of* $T_1, \dots, T_k$ *with respect to* $\tilde{T}$*, denoted by* $\Pi(\tilde{T}, T_1, \dots, T_k)$*, is the tournament* $(A, \succ)$ *such that* $A = \bigcup_{i=1}^{k} B_i$ *and for all* $b_1 \in B_i, b_2 \in B_j$,

$$b_1 \succ b_2 \quad \textit{if and only if} \quad i = j \textit{ and } b_1 \succ_i b_2\textit{, or } i \neq j \textit{ and } i \mathrel{\tilde{\succ}} j.$$

Let *Cond*(T) denote the set of Condorcet winners of $T = (A, \succ)$, i.e., $\mathit{Cond}(T) = \{a \in A : a \succ b \text{ for all } b \in A \setminus \{a\}\}$.[1] Due to the asymmetry of the dominance relation, every tournament contains at most one Condorcet winner. We are now ready to formally define a tournament solution S with a variable agenda. Call S is *nontrivial* if there exists a tournament $T = (A, \succ)$ such that $S(T)$ is a *strict* subset of A. Following Laslier (1997), we require a tournament solution to be independent of alternatives outside the tournament, invariant under tournament isomorphisms, and to choose the Condorcet winner whenever it exists.

S is nontrivial iff $S \neq$ TRIV

Definition 10.3. *A* tournament solution *is a function* $S : \mathscr{T}(U) \to 2^U \setminus \emptyset$ *such that*

(i) $S(T) = S(T') \subseteq A$ *for all tournaments* $T = (A, \succ)$ *and* $T' = (A, \succ')$ *such that* $T|_A = T'|_A$;

(ii) $S((\pi(A), \succ')) = \pi(S((A, \succ)))$ *for all tournaments* $(A, \succ)$ *and* $(A', \succ')$ *and bijections* $\pi : A \to A'$ *such* π *is a tournament isomorphism of* $(A, \succ)$ *and* $(A', \succ')$*; and*

(iii) $S(T) = \mathit{Cond}(T)$ *whenever* S *is nontrivial and* $\mathit{Cond}(T) \neq \emptyset$.

1 In contrast to the choice set of the C1 function *COND*, *Cond*(T) may be empty.

All the C1 functions introduced in Section 7.2.1 easily generalize to the variable agenda setting and satisfy the conditions of Definition 10.3. To avoid cluttered notation, we write $S(A, \succ)$ instead of $S((A, \succ))$ for a tournament $(A, \succ)$. Furthermore, we frequently write $S(B)$ instead of $S(B, \succ)$ for a subset $B \subseteq A$ of alternatives, if the dominance relation $\succ$ is known from the context.

10.3 RETENTIVE SETS

Motivated by cooperative majority voting, Schwartz (1990) introduced a tournament solution based on a notion he calls *retentiveness*. The intuition underlying retentive sets is that alternative a is only "properly" dominated by alternative b if b is chosen among a's dominators by some underlying tournament solution S. A set of alternatives is then called S-retentive if none of its elements is properly dominated by some alternative outside the set.

Definition 10.4. *Let S be a tournament solution and $T = (A, \succ)$ a tournament. Then, $B \subseteq A$ is* S-retentive *in T if $B \neq \emptyset$ and $S(D(b)) \subseteq B$ for all $b \in B$ such that $D(b) \neq \emptyset$. The set of S-retentive sets for a given tournament $T = (A, \succ)$ will be denoted by $\mathscr{R}_S(T)$, i.e.,*

$$\mathscr{R}_S(T) = \{B \subseteq A : B \text{ is } S\text{-retentive in } T\}.$$

Fix an arbitrary tournament solution S. Since the set of all alternatives is trivially S-retentive, S-retentive sets are guaranteed to exist. If a Condorcet winner exists, it must clearly be contained in any S-retentive set. The union of all (inclusion-)minimal S-retentive sets thus defines a tournament solution.

Definition 10.5. *Let S be a tournament solution. Then, the tournament solution $\mathring{S}$ is given by*

$$\mathring{S}(T) = \bigcup \min_{\subseteq}(\mathscr{R}_S(T)).$$

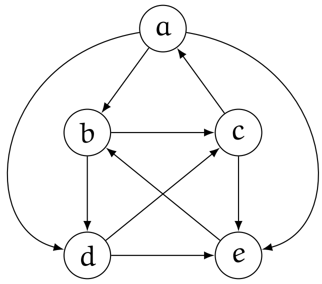
Consider for example the tournament solution *TRIV*, which always selects the set of all alternatives. It is easily verified that there always exists a *unique* minimal *TRIV*-retentive set, and that in fact $\mathring{TRIV} = TC$. See Figure 40 for an example tournament.

For a tournament solution S, we say that $\mathscr{R}_S$ is *pairwise intersecting* if for each tournament T and for all sets $B, C \in \mathscr{R}_S(T)$, $B \cap C \neq \emptyset$. Observe that the nonempty intersection of any two S-retentive sets is itself S-retentive. We thus have the following.

pairwise intersecting

Proposition 10.6. *For every tournament solution S, $\mathscr{R}_S$ admits a unique minimal element if and only if $\mathscr{R}_S$ is pairwise intersecting.*

Schwartz introduced retentiveness in order to recursively define the *tournament equilibrium set* (*TEQ*) as the union of minimal *TEQ*-



x	$D(x)$	$TC(D(x))$
a	$\{c\}$	$\{c\}$
b	$\{a, e\}$	$\{a\}$
c	$\{b, d\}$	$\{b\}$
d	$\{a, b\}$	$\{a\}$
e	$\{a, c, d\}$	$\{a, c, d\}$

Figure 40: Example tournament $T = (\{a, b, c, d, e\}, \succ)$ with $\mathring{TRIV}(T) = TC(T) = \{a, b, c, d, e\}$ and $\mathring{TC}(T) = \{a, b, c\}$. $\mathscr{R}_{TC}(T)$ contains the sets $\{a, b, c\}$, $\{a, b, c, d\}$, and $\{a, b, c, d, e\}$.

retentive sets. This recursion is well-defined because the order of the dominator set of any alternative is strictly smaller than the order of the original tournament.

Definition 10.7 (Schwartz, 1990)**.** *The* tournament equilibrium set *(TEQ) is defined recursively as $TEQ = \mathring{TEQ}$.*

In other words, *TEQ* is the unique fixed point of the ∘-operator. In the tournament of Figure 40, *TEQ* coincides with $\mathring{TC}$. Schwartz conjectured that every tournament admits a *unique* minimal *TEQ*-retentive set. This conjecture was recently disproved by a non-constructive argument using the probabilistic method (Brandt et al., 2013). While this proof showed the *existence* of a counter-example, no concrete counter-example (or even the exact size of one) is known. We let n_{TEQ} denote the largest number n such that $\mathscr{T}_n$ does *not* contain a counter-example.

counter-example

Definition 10.8. *n_{TEQ} denotes the largest integer n such that $\mathscr{R}_{TEQ}$ is pairwise intersecting in $\mathscr{T}_n$.*

Only very rough bounds on n_{TEQ} are known. The proof of Brandt et al. (2013) yields $n_{TEQ} \leqslant 10^{136}$, and an exhaustive computer analysis has shown that $n_{TEQ} \geqslant 12$ (Brandt et al., 2010).

It turns out that the existence of a *unique* minimal S-retentive set is quintessential for showing that $\mathring{S}$ satisfies several important properties to be defined in the next section. Although minimal *TEQ*-retentive sets are not unique in general, it was shown by Laffond et al. (1993a) and Houy (2009a,b) that *TEQ* satisfies these properties *for all tournaments in $\mathscr{T}_{n_{TEQ}}$*.

The ∘-operator can also be applied iteratively. Inductively define

$$S^{(0)} = S \quad \text{and} \quad S^{(k+1)} = \mathring{S^{(k)}},$$

and consider the sequence $(S^{(n)})_{n \in \mathbb{N}_0} = (S^{(0)}, S^{(1)}, S^{(2)}, \dots)$. We say that $(S^{(n)})_n$ *converges* to a tournament solution S' if for each tournament T, there exists $k_T \in \mathbb{N}_0$ such that $S^{(n)}(T) = S'(T)$ for all $n \geqslant k_T$. It turns out that the limit of all these sequences is *TEQ*.

convergence

Theorem 10.9. *Every tournament solution converges to TEQ.*

Proof. Let S be a tournament solution. We show by induction on n that

$$S^{(n-1)}(T) = \mathit{TEQ}(T).$$

for all tournaments $T \in \mathscr{T}_n$. The case $n = 1$ is trivial. For the induction step, let $T = (A, \succ)$ be a tournament of order $|A| = n + 1$. We have to show that $S^{(n)}(T) = \mathit{TEQ}(T)$. Since $S^{(n)}$ is defined as the union of all minimal $S^{(n-1)}$-retentive sets, it suffices to show that a subset $B \subseteq A$ is $S^{(n-1)}$-retentive if and only if it is *TEQ*-retentive. We have the following chain of equivalences:

$$\begin{aligned} B \text{ is } S^{(n-1)}\text{-retentive} &\Leftrightarrow \text{for all } b \in B,\ S^{(n-1)}(D(b)) \subseteq B \\ &\Leftrightarrow \text{for all } b \in B,\ \mathit{TEQ}(D(b)) \subseteq B \\ &\Leftrightarrow B \text{ is } \mathit{TEQ}\text{-retentive.} \end{aligned}$$

In particular, the second equivalence follows from the induction hypothesis, since obviously $|D(a)| \leqslant n$ for all $a \in A$. $\square$

10.4 PROPERTIES OF TOURNAMENT SOLUTIONS BASED ON RETENTIVENESS

In order to compare tournament solutions with one another, a number of desirable properties have been identified. In this section, we review five of the most common properties—*monotonicity*, *independence of unchosen alternatives*, *the weak and strong superset properties*, and $\widehat{\gamma}$—and investigate which of them are inherited from S to $\mathring{S}$ or from $\mathring{S}$ to S. We furthermore show that *composition-consistency* is never inherited.

10.4.1 *Basic Properties*

A tournament solution is monotonic if a chosen alternative remains in the choice set when its dominion is enlarged, while everything else remains unchanged. It is independent of unchosen alternatives if the choice set is invariant under any modification of the dominance relation among the alternatives that are not chosen. A tournament solution satisfies the weak superset property if no new alternatives are chosen when unchosen alternatives are removed, and the strong superset property if in this case the choice set remains unchanged. Finally, $\widehat{\gamma}$ requires that if a the same set of alternatives is selected in two subtournaments $(B_1, \succ)$ and $(B_2, \succ)$ of the same tournament

$(A, \succ)$, then this set is also selected in the tournament $(B_1 \cup B_2, \succ)$.[2] Formally, we have the following definitions.[3]

Definition 10.10. *Let S be a tournament solution.*

(i) S *satisfies* monotonicity (MON) *if for all* $a \in A$, $a \in S(T)$ *implies* $a \in S(T')$ *for all tournaments* $T = (A, \succ)$ *and* $T' = (A, \succ')$ *such that* $T|_{A \setminus \{a\}} = T'|_{A \setminus \{a\}}$ *and* $D_{\succ}(a) \subseteq D_{\succ'}(a)$.

(ii) S *satisfies* independence of unchosen alternatives (IUA) *if* $S(T) = S(T')$ *for all tournaments* $T = (A, \succ)$ *and* $T' = (A, \succ')$ *such that* $T|_{S(T) \cup \{a\}} = T'|_{S(T) \cup \{a\}}$ *for all* $a \in A$.

(iii) S *satisfies the* weak superset property (WSP) *if* $S(B) \subseteq S(A)$ *for all tournaments* $(A, \succ)$ *and* $B \subseteq A$ *such that* $S(A) \subseteq B$.

(iv) S *satisfies the* strong superset property (SSP) *if* $S(B) = S(A)$ *for all tournaments* $(A, \succ)$ *and* $B \subseteq A$ *such that* $S(A) \subseteq B$.

(v) S *satisfies* $\widehat{\gamma}$ *if* $S(B_1) = S(B_2)$ *implies* $S(B_1 \cup B_2) = S(B_1) = S(B_2)$ *for all tournaments* $(A, \succ)$ *and all* $B_1, B_2 \subseteq A$.

basic properties

The five properties just defined—MON, IUA, WSP, SSP, and $\widehat{\gamma}$—will be called *basic properties* of tournament solutions. Observe that SSP implies WSP. Furthermore, the conjunction of MON and SSP implies IUA. To prove that a tournament solution satisfies all basic properties it is therefore sufficient to show that it satisfies MON, SSP, and $\widehat{\gamma}$.

While *TRIV* trivially satisfies all basic properties, more discriminative tournament solutions often fail to satisfy some of them. For example, the Copeland set (*CO*) only satisfies MON and the Banks set (*BA*) and the uncovered set (*UC*) only satisfy MON and WSP. The minimal covering set (*MC*), on the other hand, satisfies all basic properties. The same holds for *TEQ* for all tournaments in $\mathscr{T}_{n_{TEQ}}$ (Laffond et al., 1993a; Houy, 2009a,b). Refer to Section A.3 in the appendix for a complete list of tournament solutions and their properties.

10.4.2 *Inheritance of Basic Properties*

When studying the inheritance of properties from S to $\mathring{S}$ and vice versa, we will make use of the following particular type of decomposable tournament. Let $C_3 = (\{1, 2, 3\}, \succ)$ with $1 \succ 2 \succ 3 \succ 1$, and let I_x be the unique tournament on $\{x\}$. For three tournaments T_1, T_2,

---

2 $\widehat{\gamma}$ is a variant of the better-known expansion property γ, which, together with Sen's α, figures prominently in the characterization of rationalizable choice functions (Brandt and Harrenstein, 2011).

3 Our terminology differs slightly from those of Laslier (1997) and others. *Independence of unchosen alternatives* is also called *independence of the losers* or *independence of non-winners*. The *weak superset property* has been referred to as ϵ^+ or as the *Aïzerman property*.

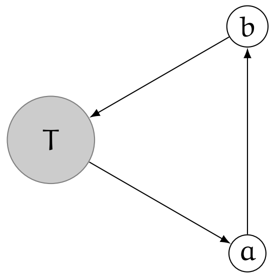


Figure 41: Tournament $C(T, I_a, I_b)$ for a given tournament T. The gray circle represents a component isomorphic to the original tournament T. An edge incident to a component signifies that there is an edge of the same direction incident to each alternative in the component.

and T_3 on disjoint sets of alternatives, let $C(T_1, T_2, T_3)$ be their product with respect to C_3, i.e.,

$$C(T_1, T_2, T_3) = \Pi(C_3; T_1, T_2, T_3).$$

Figure 41 illustrates the structure of $C(T, I_a, I_b)$ for a given tournament T. We have the following lemma.

Lemma 10.11. *Let S be a tournament solution. Then, for each tournament* $T = (A, \succ)$ *and* $a, b \notin A$,

$$\mathring{S}(C(T, I_a, I_b)) = \{a, b\} \cup S(T).$$

Proof. Let $B = \mathring{S}(C(T, I_a, I_b))$ and observe that $B \cap A \neq \emptyset$, because neither $\{a, b\}$ nor any of its subsets is S-retentive. Since a is the Condorcet winner in $D(b) = \{a\}$ and b is the Condorcet winner in $D(c)$ for any $c \in B \cap A$, by S-retentiveness of B we have that $a \in B$ and $b \in B$. Also by retentiveness of B, we have $S(D(a)) = S(T) \subseteq B$. We have thus shown that every S-retentive set must contain $\{a, b\} \cup S(T)$, and that $\{a, b\} \cup S(T)$ is itself S-retentive. □

We are now ready to show that a number of desirable properties are inherited from $\mathring{S}$ to S.

Theorem 10.12. *Let S be a tournament solution. Then each of the five basic properties is satisfied by S if it is satisfied by* $\mathring{S}$.

Proof. We show the following: if S violates one of the five basic properties MON, IUA, WSP, SSP, or $\widehat{\gamma}$, then $\mathring{S}$ violates the same property. Observe that if S violates any of these properties, this is witnessed by a tournament $T = (A, \succ)$ that serves as a counter-example. In the case of SSP (or WSP), there exists a set $B \subset A$ such that $S(A) \subseteq B \subset A$ and $S(B) \neq S(A)$ (or $S(B) \nsubseteq S(A)$, respectively). In the case of MON, there exists $a \in S(T)$ such that $a \notin S(T')$ for a tournament $T' = (A, \succ')$ that satisfies $T|_{A\setminus\{a\}} = T'|_{A\setminus\{a\}}$ and $D_{\succ}(a) \subseteq D_{\succ'}(a)$. In the case of IUA, $S(T) \neq S(T')$ for a tournament $T' = (A, \succ')$ that satisfies

$T|_{S(T)\cup\{a\}} = T'|_{S(T)\cup\{a\}}$ for all $a \in A$. In the case of $\widehat{\gamma}$, there exist subsets $B_1, B_2 \subseteq A$ such that $S(B_1) = S(B_2)$ and $S(B_1 \cup B_2) \neq S(B_1)$.

It thus suffices to show how a counter-example T for S can be transformed into a counter-example T' for $\mathring{S}$. Let $a, b \notin A$ and define $T' = C(T, I_a, I_b)$. Lemma 10.11 implies that $\mathring{S}(T') = \{a, b\} \cup S(T)$. Hence, tournament T' constitutes a counter-example for $\mathring{S}$. □

If $\mathscr{R}_S$ is pairwise intersecting, a similar statement holds for the opposite direction. The conjunction of two properties P and Q is denoted by $P \wedge Q$.

Theorem 10.13. *Let S be a tournament solution such that $\mathscr{R}_S$ is pairwise intersecting, and let P be any of the properties SSP, WSP, IUA, MON $\wedge$ SSP, or $\widehat{\gamma} \wedge$ SSP. Then, P is satisfied by S if and only if it is satisfied by $\mathring{S}$.*

Proof. Assume that $\mathscr{R}_S$ is pairwise intersecting. We need to show that each of the properties SSP, WSP, IUA, MON $\wedge$ SSP, and $\widehat{\gamma} \wedge$ SSP is satisfied by S if and only if it is satisfied by $\mathring{S}$. The direction from right to left follows from Theorem 10.12. We now show that the properties are inherited from S to $\mathring{S}$.

Assume that S satisfies SSP. Let $T = (A, \succ)$ be a tournament, and consider an alternative $x \in A \setminus \mathring{S}(T)$. We need to show that $\mathring{S}(T') = \mathring{S}(T)$, where $T' = (A \setminus \{x\}, \succ)$. Since $\mathscr{R}_S$ is assumed to be pairwise intersecting, it suffices to show that for all $a \in \mathring{S}(T)$, $S(D_A(a)) = S(D_{A\setminus\{x\}}(a))$. To this end, consider an arbitrary $a \in \mathring{S}(T)$. If $x \notin D_A(a)$, then obviously $D_A(a) = D_{A\setminus\{x\}}(a)$ and thus $S(D_A(a)) = S(D_{A\setminus\{x\}}(a))$. Assume on the other hand that $x \in D_A(a)$. Since $a \in \mathring{S}(T)$ and $x \notin \mathring{S}(T)$, it follows that $x \notin S(D_A(a))$, as otherwise $\mathring{S}(T)$ would not be S-retentive. Since S satisfies SSP, we obtain $S(D_A(a)) = S(D_{A\setminus\{x\}}(a))$ as desired.

Assume that S satisfies WSP. Let $T = (A, \succ)$ be a tournament, and consider an alternative $x \in A \setminus \mathring{S}(T)$. We need to show that $\mathring{S}(T') \subseteq \mathring{S}(T)$, where $T' = (A \setminus \{x\}, \succ)$. Since $\mathscr{R}_S$ is assumed to be pairwise intersecting, it suffices to show that $\mathring{S}(T)$ is also S-retentive in T'. To this end, consider an arbitrary $a \in \mathring{S}(T)$. Since S satisfies WSP, we have that $S(D_{A\setminus\{x\}}(a)) \subseteq S(D_A(a))$. Furthermore, by S-retentiveness of $\mathring{S}(T)$, $S(D_A(a)) \subseteq \mathring{S}(T)$ and thus $S(D_{A\setminus\{x\}}(a)) \subseteq \mathring{S}(T)$.

Assume that S satisfies IUA. Let $T = (A, \succ)$ and $T' = (A, \succ')$ be tournaments with $x, y \in A \setminus \mathring{S}(T)$ and $T|_{A\setminus\{x,y\}} = T'|_{A\setminus\{x,y\}}$. We need to show that $\mathring{S}(T) = \mathring{S}(T')$. Since $\mathscr{R}_S$ is assumed to be pairwise intersecting, it suffices to show that for all $a \in \mathring{S}(T)$, $S(D_\succ(a), \succ) = S(D_{\succ'}(a), \succ')$. To this end, consider an arbitrary $a \in \mathring{S}(T)$. By assumption, $a \neq x$ and $a \neq y$. First consider the case when both $x \in D_\succ(a)$ and $y \in D_\succ(a)$. Then, $D_\succ(a) = D_{\succ'}(a)$ and, by S-retentiveness of $\mathring{S}(T)$, $x, y \notin S(D_\succ(a), \succ)$. Since S satisfies IUA, $S(D_\succ(a), \succ) = S(D_{\succ'}(a), \succ')$ as required. Now consider the case when $x \notin D_\succ(a)$ or $y \notin D_\succ(a)$. Then, $T|_{D_\succ(a)} = T'|_{D_{\succ'}(a)}$, and the claim follows immediately.

Assume that S satisfies MON and SSP. We have already seen that SSP is inherited, so it remains to be shown that $\mathring{S}$ satisfies MON. The following argument is adapted from the proof of Proposition 3.6 in Laffond et al. (1993a). Let $T = (A, \succ)$ be a tournament, and consider two alternatives $a, b \in A$ such that $a \in \mathring{S}(T)$ and $b \succ a$. Let $T' = (A, \succ')$ be the tournament with $T|_{A \setminus \{a\}} = T'|_{A \setminus \{a\}}$ and $D_{\succ'}(a) = D_{\succ}(a) \cup \{b\}$. We have to show that $a \in \mathring{S}(T')$. To this end, we claim that for all $c \in A \setminus \{a\}$,

$$a \notin S(D_{\succ'}(c), \succ') \quad \text{implies} \quad S(D_{\succ}(c), \succ) = S(D_{\succ'}(c), \succ'). \tag{8}$$

Consider the case when $c \neq b$ and assume that $a \notin S(D_{\succ'}(c), \succ')$. It follows from monotonicity of S that $a \notin S(D_{\succ}(c), \succ)$. To see this, observe that monotonicity of S implies that $a \in S(D_{\succ'}(c), \succ')$ whenever $a \in S(D_{\succ}(c), \succ)$. Now, since S satisfies SSP,

$$S(D_{\succ'}(c), \succ') = S(D_{\succ'}(c) \setminus \{a\}, \succ') \quad \text{and}$$
$$S(D_{\succ}(c), \succ) = S(D_{\succ}(c) \setminus \{a\}, \succ).$$

It is easily verified that $(D_{\succ'}(c) \setminus \{a\}, \succ') = (D_{\succ}(c) \setminus \{a\}, \succ)$, thus we have $S(D_{\succ'}(c), \succ') = S(D_{\succ}(c), \succ)$.

If $c = b$, then $a \notin S(D_{\succ'}(b), \succ')$ together with SSP of S implies $S(D_{\succ'}(b), \succ') = S(D_{\succ'}(b) \setminus \{a\}, \succ')$. Furthermore, by definition of T and T', $(D_{\succ'}(b) \setminus \{a\}, \succ') = (D_{\succ}(b), \succ)$ and thus $S(D_{\succ'}(b), \succ') = S(D_{\succ}(b), \succ)$. This proves (8).

We proceed to show that $a \in \mathring{S}(T')$. Assume for contradiction that this is not the case. We claim that this implies that

$$\mathring{S}(T') \text{ is S-retentive in } T. \tag{9}$$

To see this, consider $c \in \mathring{S}(T')$. We have to show that $S(D_{\succ}(c), \succ) \subseteq \mathring{S}(T')$. Since, by assumption, $a \notin \mathring{S}(T')$, we have that $a \notin S(D_{\succ'}(c), \succ')$. We can thus apply (8) and get

$$S(D_{\succ}(c), \succ) = S(D_{\succ'}(c), \succ') \text{ for all } c \in \mathring{S}(T'),$$

which, together with the S-retentiveness of $\mathring{S}(T')$ in T', implies (9).

Having assumed that $\mathscr{R}_S$ is pairwise intersecting, it follows from (9) that $\mathring{S}(T) \subseteq \mathring{S}(T')$. Hence, $a \notin \mathring{S}(T)$, a contradiction. This shows that $\mathring{S}$ satisfies MON.

Finally assume that S satisfies $\widehat{\gamma}$ and SSP. We already know from the above that $\mathring{S}$ satisfies SSP, so it remains to be shown that $\mathring{S}$ satisfies $\widehat{\gamma}$. Let $T = (A, \succ)$ be a tournament, and consider two subsets $B_1, B_2 \subseteq A$ such that $\mathring{S}(B_1) = \mathring{S}(B_2) = C$. We have to show that $\mathring{S}(B_1 \cup B_2) = C$. Since $\mathscr{R}_S$ is assumed to be pairwise intersecting, it suffices to show that for all $c \in C$, $S(D_{B_1 \cup B_2}(c)) = S(D_{B_1}(c))$. To this end, consider an arbitrary $c \in C$. As $\mathring{S}(B_1)$ and $\mathring{S}(B_2)$ are S-retentive in B_1 and B_2, respectively, we have $S(D_{B_i}(c)) \subseteq C \subseteq B_1 \cap B_2$ for $i \in \{1, 2\}$. The fact that S satisfies SSP now implies $S(D_{B_1 \cap B_2}(c)) = S(D_{B_1}(c))$ and $S(D_{B_1 \cap B_2}(c)) = S(D_{B_2}(c))$, and thus $S(D_{B_1}(c)) = S(D_{B_2}(c))$. Since S satisfies $\widehat{\gamma}$, we have $S(D_{B_1 \cup B_2}(c)) = S(D_{B_1}(c) \cup D_{B_2}(c)) = S(D_{B_1}(c))$, as desired. $\square$

We proceed by identifying tournament solutions for which Theorem 10.13 can be applied. The following lemma will be useful.

Lemma 10.14. *Let S_1 and S_2 be tournament solutions such that $S_1 \subseteq S_2$ and $\mathscr{R}_{S_1}$ is pairwise intersecting. Then, $\mathscr{R}_{S_2}$ is pairwise intersecting and $\mathring{S}_1 \subseteq \mathring{S}_2$.*

Proof. First observe that $S_1 \subseteq S_2$ implies that every S_2-retentive set is S_1-retentive. Now assume for contradiction that $\mathscr{R}_{S_2}$ is not pairwise intersecting and consider a tournament $(A, \succ)$ with two disjoint S_2-retentive sets $B, C \subseteq A$. Then, by the above observation, B and C are S_1-retentive, which contradicts the assumption that $\mathscr{R}_{S_1}$ is pairwise intersecting.

Furthermore, for every tournament T, $\mathring{S}_2(T)$ is S_1-retentive and thus contains the unique minimal S_1-retentive set, i.e., $\mathring{S}_1(T) \subseteq \mathring{S}_2(T)$. $\square$

Theorem 10.15. *Let S be a tournament solution such that $TEQ \subseteq S$ in $\mathscr{T}_{n_{TEQ}}$. Then, $\mathscr{R}_{S^{(k)}}$ is pairwise intersecting in $\mathscr{T}_{n_{TEQ}}$ for all $k \in \mathbb{N}_0$.*

Proof. We first prove by induction on k that, for all $k \in \mathbb{N}_0$, $TEQ \subseteq S^{(k)}$ in $\mathscr{T}_{n_{TEQ}}$. The case $k = 0$ holds by assumption. Now let T be a tournament in $\mathscr{T}_{n_{TEQ}}$ and suppose that $TEQ(T) \subseteq S^{(k)}(T)$ for some $k \in \mathbb{N}_0$. By definition, $S^{(k+1)}(T)$ is $S^{(k)}$-retentive. We can thus apply the induction hypothesis to obtain that $S^{(k+1)}(T)$ is TEQ-retentive. Since the minimal TEQ-retentive set of T is unique, it is contained in any TEQ-retentive set, and we have that $TEQ(T) \subseteq S^{(k+1)}(T)$. This proves that $TEQ(T) \subseteq S^{(k)}(T)$ for all $T \in \mathscr{T}_{n_{TEQ}}$ and all $k \in \mathbb{N}_0$.

We can now apply Lemma 10.14 with $S_1 = TEQ$ and $S_2 = S^{(k)}$ to show that $\mathscr{R}_{S^{(k)}}$ is pairwise intersecting in $\mathscr{T}_{n_{TEQ}}$ for all $k \in \mathbb{N}_0$. $\square$

Among the tournament solutions that satisfy the conditions of Theorem 10.15 are *TRIV*, *TC*, *MC*, *UC*, and *BA* (see the proof of Theorem 10.19 on 154).

10.4.3 *Composition-Consistency*

We conclude this section by showing that, among all tournament solutions that are defined as the union of all minimal retentive sets with respect to some tournament solution, *TEQ* is the only one that is *composition-consistent*. Composition-consistent tournament solutions choose the "best" alternatives from the "best" components.

Definition 10.16. *A tournament solution* S *is* composition-consistent *if for all tournaments* $T, T_1, \dots, T_k$*, and* $\tilde{T} = (\{1, \dots, k\}, \tilde{\succ})$ *such that* $T = \Pi(\tilde{T}, T_1, \dots, T_k)$*,*

$$S(T) = \bigcup_{i \in S(\tilde{T})} S(T_i).$$

Tournament solutions satisfying this property include *TRIV*, *UC*, *BA*, and *TEQ*. However, $\mathring{S}$ is not composition-consistent unless S equals *TEQ*.

Proposition 10.17. *Let* S *be a tournament solution. Then,* $\mathring{S}$ *is composition-consistent if and only if* $S = TEQ$*.*

Proof. It is well-known that *TEQ* is composition-consistent (Laffond et al., 1996). For the direction from left to right, let S be a tournament solution different from *TEQ*, and assume that $\mathring{S}$ is composition-consistent. Since *TEQ* is the only tournament solution S' such that $S' = \mathring{S}'$, there has to exist a tournament $T = (A, \succ)$ such that $S(T) \neq \mathring{S}(T)$. Let $a, b \notin A$, and define $T^* = C(T, I_a, I_b)$. By Lemma 10.11,

$$\mathring{S}(T^*) = \{a, b\} \cup S(T).$$

On the other hand, by composition-consistency of $\mathring{S}$,

$$\mathring{S}(T^*) = \mathring{S}(T) \cup \mathring{S}(I_a) \cup \mathring{S}(I_b) = \{a, b\} \cup \mathring{S}(T).$$

It follows that $S(T) = \mathring{S}(T)$, a contradiction. □

The *composition-consistent hull* of a tournament solution S, denoted by S^*, is defined as the inclusion-minimal tournament solution that is composition-consistent and contains S (Laffond et al., 1996). It can be shown that $(\mathring{S})^* = S^*$ for all tournament solutions S that satisfy $\mathring{S} \subseteq S$.

10.5 CONVERGENCE TO TEQ

By Theorem 10.9, every tournament solution converges to *TEQ*. Particularly well-behaved types of convergence are those that either yield larger and larger subsets of *TEQ* or smaller and smaller supersets of *TEQ*. The problem with the former type is that no natural refinement of *TEQ* is known and it is doubtful whether any such refinement would be efficiently computable. The latter type, however, turns out to be particularly useful.

Call a sequence $(S^{(n)})_{n \in \mathbb{N}_0}$ *contracting* if for all $k \in \mathbb{N}_0$, $S^{(k+1)} \subseteq S^{(k)}$. Intuitively, the elements of such a sequence constitute better and better "approximations" of *TEQ*. The following lemma identifies a sufficient condition for a sequence to be contracting.

contracting sequences

Lemma 10.18. *Let S be a tournament solution such that $TEQ \subseteq S$ in $\mathscr{T}_{n_{TEQ}}$. If $\mathring{S} \subseteq S$ in $\mathscr{T}_{n_{TEQ}}$, then $S^{(k+1)} \subseteq S^{(k)}$ in $\mathscr{T}_{n_{TEQ}}$ for all $k \in \mathbb{N}_0$.*

Proof. We prove the statement by induction on k for all tournaments in $\mathscr{T}_{n_{TEQ}}$. $\mathring{S} \subseteq S$ holds by assumption. Now suppose that $S^{(k)} \subseteq S^{(k-1)}$ for some $k \in \mathbb{N}_0$. As in the proof of Theorem 10.15, one can show that $TEQ \subseteq S^{(k)}$. Applying Lemma 10.14 with $S_1 = TEQ$ and $S_2 = S^{(k)}$ yields that $\mathscr{R}_{S^{(k)}}$ is pairwise intersecting. Therefore, we can apply Lemma 10.14 again, this time with $S_1 = S^{(k)}$ and $S_2 = S^{(k-1)}$, which gives $S^{(k+1)} \subseteq S^{(k)}$. □

Theorem 10.19. *For all tournaments with at most n_{TEQ} alternatives, the tournament solutions TRIV, TC, MC, UC, and BA give rise to contracting sequences.*

Proof. As *TRIV* obviously satisfies the assumptions of Lemma 10.18, $(TRIV^{(n)})_n$ and $(TC^{(n)})_n$ are contracting. *MC* satisfies the assumptions because $TEQ \subseteq MC$ in $\mathscr{T}_{n_{TEQ}}$ (Laffond et al., 1993a) and $\mathring{MC} \subseteq MC$ in $\mathscr{T}_{n_{TEQ}}$ (Brandt, 2011b). $TEQ \subseteq BA$ was shown by Schwartz (1990), and $TEQ \subseteq UC$ follows from $BA \subseteq UC$. It thus remains to be shown that $\mathring{UC} \subseteq UC$ and $\mathring{BA} \subseteq BA$.

strong retentiveness A tournament solution S satisfies *strong retentiveness* if the choice set of every dominator set is contained in the original choice set, i.e., if $S(D(a)) \subseteq S(A)$ for all $a \in A$ (Brandt, 2011b). It is easy to see that $\mathring{S} \subseteq S$ for every tournament solution S that satisfies strong retentiveness. Indeed, for an arbitrary tournament T, strong retentiveness implies that $S(T)$ is S-retentive and that there do not exist any S-retentive sets disjoint from $S(T)$. Since both *UC* and *BA* satisfy strong retentiveness (Brandt, 2011b), this completes the proof. □

One might wonder if *MC* is contained in the sequence $(TRIV^{(n)})_n$. It is easy to see that this is not the case: while *MC* is known to be composition-consistent (see Laffond et al., 1996), Proposition 10.17 shows that this is not the case for any $TRIV^{(k)}$ with $k \geqslant 1$. Furthermore, there do not exist $i, j \in \mathbb{N}$ with $MC^{(i)} = TRIV^{(j)}$, as otherwise we would have $MC = TRIV^{(j-i)}$.

For a given tournament solution S, one may further want to compare the sequence $(S^{(n)})_n$ with the corresponding sequence $(S^n)_{n\in\mathbb{N}}$ generated by the repeated application of S. Formally,

$$S^1(T) = S(T) \text{ and } S^k(T) = S(S^{k-1}(T)).$$

Since SSP implies that $S^n = S$ for all $n \in \mathbb{N}$, *UC* and *BA* are the only tournament solutions covered by Theorem 10.19 for which such a comparison makes sense. It turns out that for both *UC* and *BA*, the sequences $(S^{(n)})_{n\in\mathbb{N}_0}$ and $(S^n)_{n\in\mathbb{N}}$ are incomparable in the sense that for all $n \in \mathbb{N}$, neither $S^{(n)} \subseteq S^2$ nor $S^n \subseteq \mathring{S}$.

10.5.1 *Iterations to Convergence*

We may ask how many iterated applications of the $\circ$-operator are needed until we arrive at *TEQ*. While we have seen that every tournament solution converges to *TEQ*, it turns out that no solution other than *TEQ* itself does so in a finite number of steps. More precisely, the number of iterations required to reach *TEQ* increases with the order of a tournament and can not be bounded by a constant independent of the order.

For a tournament solution S, let $k_S(n)$ be the smallest $k \in \mathbb{N}_0$ such that $S^{(k)}(T) = \mathit{TEQ}(T)$ for all tournaments $T \in \mathscr{T}_n$.[4]

Proposition 10.20. *Let* $S \neq \mathit{TEQ}$ *be a tournament solution and let* n_0 *be the order of a smallest tournament* T *with* $S(T) \neq \mathit{TEQ}(T)$*. Then, for every* $n \in \mathbb{N}$*,*

$$k_S(n) = \max\left(\left\lfloor \frac{n-n_0}{2} \right\rfloor + 1, 0\right).$$

Proof. Let $f(n) = \max\left(\lfloor \frac{n-n_0}{2} \rfloor + 1, 0\right)$. Our goal is to prove that $f(n)$ is both an upper bound and a lower bound on $k_S(n)$.

For the former, we show that $S^{(f(n))}(T) = \mathit{TEQ}(T)$ for all $T \in \mathscr{T}_n$. Denote by $k_S(T)$ the smallest number k such that $S^{(k)}(T) = \mathit{TEQ}(T)$. Thus, $k_S(n) = \max_{T \in \mathscr{T}_n} k_S(T)$.

A *Condorcet loser* in $(A, \succ)$ is an alternative $a \in A$ such that $D(a) = A \setminus \{a\}$. We claim that the following statements hold for every tournament solution S and every tournament T of order n: *Condorcet loser*

(*i*) If T has a Condorcet loser, then $k_S(T) \leqslant k_S(n-1)$.

(*ii*) If T has no Condorcet loser, then $k_S(T) \leqslant k_S(n-2)+1$.

For (*i*), let a be a Condorcet loser in $T = (A, \succ)$. Then,

$$S^{(k_S(n-1))}(T) = S^{(k_S(n-1))}(A \setminus \{a\}) = \mathit{TEQ}(A \setminus \{a\}) = \mathit{TEQ}(T).$$

The first and the third equality follow from the observations that no minimal retentive set contains a and that a set $B \subseteq A \setminus \{a\}$ is retentive in T if and only if it is retentive in $(A \setminus \{a\}, \succ)$. The second equality is a direct consequence of the definition of k_S. For (*ii*), assume that $T = (A, \succ)$ does not have a Condorcet loser. It follows that $|D(a)| \leqslant n-2$ for all $a \in A$. Similar reasoning as in the proof of Theorem 10.9 implies that a set $B \subseteq A$ is $S^{(k_S(n-2))}$-retentive if and only if B is *TEQ*-retentive. Thus, $S^{(k_S(n-2)+1)}(T) = \mathit{TEQ}(T)$.

We are now ready to show that $k_S(n) \leqslant f(n)$ by induction on n. For $n \leqslant n_0$, $k_S(n) = 0$. Now assume that $k_S(m) \leqslant f(m)$ holds for every $m < n$, and consider a tournament T of order n. If T has a Condorcet loser, (*i*) implies that $k_S(T) \leqslant k_S(n-1) \leqslant f(n-1)$,

4 It can easily be shown that $S^{(\ell)}(T) = \mathit{TEQ}(T)$ for all $T \in \mathscr{T}_n$ and $\ell \geqslant k_S(n)$.

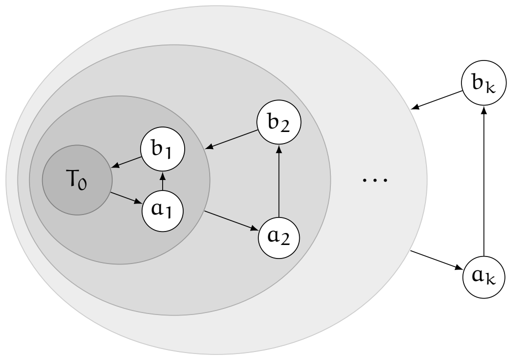


Figure 42: Tournament T_k used in the proof of Proposition 10.20

where the latter inequality follows from the induction hypothesis. If, on the other hand, T does not have a Condorcet loser, *(ii)* implies that $k_S(T) \leqslant k_S(n-2)+1 \leqslant f(n-2)+1$. Thus, $k_S(n) \leqslant \max(f(n-1), f(n-2)+1) = f(n-2)+1$. A simple calculation shows that $f(n-2)+1 = f(n)$ as desired.

In order to show that $k_S(n) \geqslant f(n)$, we inductively define a family of tournaments $T_0, T_1, T_2, \dots$ such that $S^{(f(|T_k|)-1)}(T_k) \neq \mathit{TEQ}(T_k)$. Let $T_0 = (A_0, \succ)$ be a smallest tournament such that $S(T_0) \neq \mathit{TEQ}(T_0)$. By definition, $|A_0| = n_0$. Given $T_{k-1} = (A_{k-1}, \succ)$, let

$$T_k = C(T_{k-1}, I_{a_k}, I_{b_k}),$$

where $a_k, b_k \notin A_{k-1}$ are two new alternatives. Observe that $A_k = A_0 \cup \bigcup_{\ell=1}^{k} \{a_\ell, b_\ell\}$. The structure of T_k is illustrated in Figure 42. Repeated application of Lemma 10.11 yields

$$\begin{aligned} S^{(k)}(T_k) &= \{a_k, b_k\} \cup S^{(k-1)}(T_{k-1}) \\ &= \{a_k, b_k\} \cup \{a_{k-1}, b_{k-1}\} \cup S^{(k-2)}(T_{k-2}) \\ &= \dots \\ &= \bigcup_{\ell=1}^{k} \{a_\ell, b_\ell\} \cup S(T_0). \end{aligned}$$

Since $S(T_0) \neq \mathit{TEQ}(T_0)$, we have that $S^{(k)}(T_k) \neq \mathit{TEQ}^{(k)}(T_k) = \mathit{TEQ}(T_k)$.

We have thus shown that $k_S(n_k) > k$, where $n_k = |A_k|$ is the order of tournament T_k. By definition of T_k, $n_k = n_0 + 2k$, so $k_S(n_k) > k$ implies $k_S(n) > \frac{n-n_0}{2}$ for all $n \geqslant n_0$ such that $n - n_0$ is even. For the case when $n - n_0$ is odd, i.e., when $n = n_0 + 2k + 1$ for some $k \in \mathbb{N}_0$, consider the tournament $T'_k = (A_{k+1} \setminus \{b_{k+1}\}, \succ)$ with $T'_k|_{A_{k+1}\setminus\{b_{k+1}\}} = T_{k+1}|_{A_{k+1}\setminus\{b_{k+1}\}}$. This tournament of order n has a_{k+1} as a Condorcet loser. Thus, $S^{(k)}(T'_k) = S^{(k)}(T_k) \neq \mathit{TEQ}(T_k) = \mathit{TEQ}(T'_k)$. This implies that $k_S(n_0 + 2k + 1) > k$, or, equivalently, $k_S(n) > \lfloor \frac{n-n_0}{2} \rfloor$. $\square$

An easy corollary of Proposition 10.20 is that $k_S(n) \leqslant \lfloor \frac{n}{2} \rfloor$ for all tournament solutions. Since *TRIV* and *TEQ* differ for every tournament with two alternatives, we immediately have $k_{TRIV}(n) = \lfloor \frac{n}{2} \rfloor$. Furthermore, Dutta (1990) constructed a tournament T of order 8 for which $MC(T) \neq TEQ(T)$, and thus $k_{MC}(n) = \max(\lfloor \frac{n}{2} \rfloor - 3, 0)$.

Convergence of the sequence $(S^{(n)})_n$ of tournament solutions should not be confused with convergence of the sequence $(S^{(n)}(T))_n$ of choice sets for a particular tournament T. In particular, $S^{(m)}(T) = S^{(m+1)}(T)$ does *not* imply $S^{(m')}(T) = S^{(m)}(T)$ for all $m' \geqslant m$. For example, the tournaments T_k constructed in the proof of Proposition 10.20 satisfy $TRIV^{(m)}(T_k) = TRIV^{(m')}(T_k) \neq TEQ(T_k)$ for all $m, m' < k_{TRIV}(n_k)$. As a consequence, it might be impossible to *recognize* convergence of $(S^{(n)}(T))_n$ within less than $k_S(|T|)$ iterations.

10.5.2 *Computational Aspects*

The sequences $(TRIV^{(n)})_n$ and $(MC^{(n)})_n$ appear particularly interesting: for all tournaments in $\mathscr{T}_{n_{TEQ}}$, these sequences are contracting, and their members satisfy all basic properties. In addition, *TRIV* and *MC* can be computed efficiently, and we might ask whether this also holds for $TRIV^{(n)}$ and $MC^{(n)}$ when $n \geqslant 1$. This turns out to be the case, as a consequence of the following more general result.

Proposition 10.21. *$\mathring{S}$ is efficiently computable if and only if S is efficiently computable.*

Proof. We show that the computation of S and the computation of $\mathring{S}$ are equivalent under polynomial-time reductions.

To see that $\mathring{S}$ can be reduced to S, consider an arbitrary tournament $T = (A, \succ)$ and define the relation $R = \{(a, b) : a \in S(D(b))\}$. It is easily verified that $\mathring{S}(T)$ is the union of all minimal R-undominated sets[5] or, equivalently, the maximal elements of the asymmetric part of the transitive closure of R. Observing that both R and the minimal R-undominated sets can be computed in polynomial time (see, e.g., Brandt et al., 2009a, for the latter) completes the reduction.

For the reduction from S to $\mathring{S}$, consider a tournament $T = (A, \succ)$ and define $T^* = C(T, I_a, I_b)$ for $a, b \notin A$. By Lemma 10.11, $S(T) = \mathring{S}(T^*) \setminus \{a, b\}$. Clearly, T^* can be computed in polynomial time from T, and S(T) can be computed in polynomial time from $\mathring{S}(T^*)$. □

This result does not imply that *TEQ* can be computed efficiently, despite the fact that both *TRIV* and *MC* converge to *TEQ*. The obvious algorithm for computing $S^{(n)}(T)$ recursively computes $S^{(n-1)}$ for all dominator sets, the number and sizes of which can both be linear in |T|. By Proposition 10.20, the depth of the recursion can be linear in |T| as well, which leads to an exponential number of steps. Brandt

5 A set $B \subseteq A$ is R-undominated if $(a, b) \in R$ for no $b \in B$ and $a \in A \setminus B$.

et al. (2010) have in fact shown that it is NP-hard to decide whether a given alternative is in *TEQ*. Nevertheless, Lemma 10.18 and Proposition 10.21 identify sequences of efficiently computable tournament solutions that provide better and better approximations of TEQ for all tournaments in $\mathscr{T}_{n_{TEQ}}$.

10.6 UNIQUENESS OF MINIMAL RETENTIVE SETS

As shown in Section 10.4, uniqueness of minimal retentive sets plays an important role: if $\mathscr{R}_S$ is pairwise intersecting, then $\mathring{S}$ inherits many desirable properties from S. It is therefore an interesting, and surprisingly difficult, question which tournament solutions are pairwise intersecting. In this section, we answer the question for the top cycle and the Copeland set.

10.6.1 *The Minimal TC-Retentive Set*

We prove that every tournament has a unique minimal *TC*-retentive set, thus establishing $\mathring{TC}$ as an efficiently computable refinement of *TC* that satisfies all basic properties.

Theorem 10.22. *$\mathscr{R}_{TC}$ is pairwise intersecting.*

Proof. Consider an arbitrary tournament $(A, \succ)$, and assume for contradiction that B and C are two disjoint *TC*-retentive subsets of A. Let $b_0 \in B$ and $c_0 \in C$. Without loss of generality we may assume that $c_0 \succ b_0$. Then, $c_0 \in D(b_0)$, and by *TC*-retentiveness of B there has to be some $b_1 \in B$ with $b_1 \in TC(D(b_0))$ and $b_1 \succ c_0$. We claim that for each $m \geqslant 1$ there are $c_1, \ldots, c_m \in C$ such that for all i and j with $0 \leqslant i < j \leqslant m$,

(*i*) $c_{i+1} \in TC(D(c_i))$;

(*ii*) $b_0 \succ c_i$ and $c_i \succ b_1$ if i is odd, and $b_1 \succ c_i$ and $c_i \succ b_0$ otherwise; and

(*iii*) $c_j \succ c_i$ if $j - i$ is odd, and $c_i \succ c_j$ otherwise.

To see that this claim implies the theorem, consider i and j with $0 \leqslant i < j \leqslant m$. Since the dominance relation is irreflexive, and by (*iii*), c_i and c_j must be distinct alternatives. This in turn implies that the size of C is unbounded, contradicting finiteness of A. The situation is illustrated in Figure 43.

The claim itself can be proved by induction on m. First consider the case $m = 1$. Since $b_1 \succ c_0$, and by *TC*-retentiveness of C, there has to be some $c_1 \in C$ with $c_1 \in TC(D(c_0))$ and $c_1 \succ b_1$, showing (*i*). Furthermore, by *TC*-retentiveness of B, $c_1 \notin TC(D(b_0))$ and thus $b_0 \succ c_1$. It follows that (*ii*) and (*iii*) hold as well.

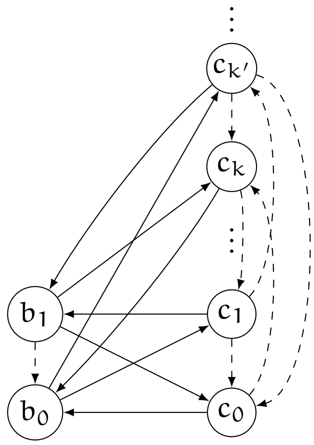


Figure 43: Structure of a tournament with two disjoint TC-retentive sets (k is even and $k' = k+1$ is odd). A dashed edge (a, b) indicates that $a \in TC(D(b))$.

Now assume that the claim holds for all $k \leqslant m$. We show that it also holds for $m+1$.

Consider the case when $m+1$ is even; the case when $m+1$ is odd is analogous. By the induction hypothesis, $b_0 \succ c_m$. Hence, by TC-retentiveness of C, there has to exist some $c_{m+1} \in C$ with $c_{m+1} \in TC(D(c_m))$ and $c_{m+1} \succ b_0$, which together with the induction hypothesis implies *(i)*.

Since $b_1 \in TC(D(b_0))$ and $c_{m+1} \in D(b_0)$, TC-retentiveness of B moreover yields $b_1 \succ c_{m+1}$. Together with the induction hypothesis, this proves *(ii)*.

For *(iii)*, consider an arbitrary $i \in \{1, \dots, m\}$, and first assume that i is odd. We have to prove that $c_{m+1} \succ c_i$. If $i = m$, this immediately follows from *(i)*. If $i < m$, then by the induction hypothesis, $c_i \succ c_m$, $b_0 \succ c_i$, and $b_0 \succ c_m$. Hence, $\{c_{m+1}, c_i, b_0\} \subseteq D(c_m)$. Moreover, as we have already shown, $c_{m+1} \succ b_0$. Assuming for contradiction that $c_i \succ c_{m+1}$, the three alternatives c_{m+1}, c_i, and b_0 would constitute a cycle in $D(c_m)$. Since $c_{m+1} \in TC(D(c_m))$, we would then have that $b_0 \in TC(D(c_m))$, contradicting TC-retentiveness of C. Thus $c_i \not\succ c_{m+1}$. As $c_{m+1} \succ b_0$ and $b_0 \succ c_i$, also $c_{m+1} \neq c_i$. Completeness of $\succ$ implies $c_{m+1} \succ c_i$.

Now assume that i is even. We have to prove that $c_i \succ c_{m+1}$. By the induction hypothesis, $c_m \succ c_i$ and $b_1 \succ c_i$. Assume for contradiction that $c_{m+1} \succ c_i$ and thus $c_{m+1} \in D(c_i)$. Since $i+1$ is odd, we already know that $c_{m+1} \succ c_{i+1}$. Furthermore, $c_{i+1} \in TC(D(c_i))$, and thus $c_{m+1} \in TC(D(c_i))$. However, $b_1 \succ c_{m+1}$ and $b_1 \in D(c_i)$ imply that $b_1 \in TC(D(c_i))$, contradicting TC-retentiveness of C. Therefore $c_{m+1} \not\succ c_i$. Since $c_{m+1} \succ c_m$ and $c_m \succ c_i$, we have $c_{m+1} \neq c_i$ and may conclude that $c_i \succ c_{m+1}$. By virtue of the induction hypothesis we are done. $\square$

Corollary 10.23. *$\mathring{TC}$ is efficiently computable and satisfies all basic properties. Furthermore, $\mathring{TC} \subseteq TC$.*

Proof. Efficient computability follows from Proposition 10.21 and the trivial observation that *TRIV* can be computed efficiently. As $\mathscr{R}_{TC}$ is pairwise intersecting, $\mathring{TC}$ inherits all basic properties from *TC* (Theorem 10.13). Finally, applying Lemma 10.14 with $S_1 = TC$ and $S_2 = TRIV$ yields $\mathring{TC} \subseteq TC$. □

10.6.2 *Copeland-Retentive Sets May Be Disjoint*

For the Copeland set the situation turns out to be quite different: minimal *CO*-retentive sets are not always unique. Our proof makes use of a special class of tournaments called *cyclones*.

Definition 10.24. *Let n be an odd integer and $A = \{a_0, \dots, a_{n-1}\}$ an ordered set of size $|A| = n$. The* cyclone *on A then is the tournament $(A, \succ)$ such that $a_i \succ a_j$ if and only if $j - i \bmod n \in \{1, \dots, \frac{n-1}{2}\}$.*

We are now in a position to prove the following result.

Proposition 10.25. *$\mathscr{R}_{CO}$ is not pairwise intersecting.*

Proof. We construct a tournament T with 70 alternatives that can be partitioned into eight subsets $A, B_0, \dots, B_6$. $A = \{a_0, \dots, a_6\}$ contains seven alternatives, whereas for each $k \in \{0, \dots, 6\}$, $B_k = \{b^k_0, \dots, b^k_8\}$ contains nine. First consider the tournament $\tilde{T} = (\{1, \dots, 14\}, \tilde{\succ})$, where $\tilde{T}|_{\{1,\dots,7\}}$ and $\tilde{T}|_{\{8,\dots,14\}}$ are cyclones on $\{1, \dots, 7\}$ and $\{8, \dots, 14\}$, respectively. For all i and j with $1 \leqslant i \leqslant 7$ and $8 \leqslant j \leqslant 14$, moreover, $j \succ i$ if and only if $j - i \in \{7, 10\}$. Now define T as the product

$$T = \Pi(\tilde{T}, I_{a_0}, \dots, I_{a_6}, T_0, \dots, T_6),$$

where for each $k \in \{0, \dots, 6\}$, T_k is the cyclone on B_k. Thus $B_j \succ \{a_i\}$ if $j \in \{i, i+3 \bmod 7\}$ and $\{a_i\} \succ B_j$ otherwise.

We claim that both $A = \{a_0, \dots, a_6\}$ and $B = B_0 \cup \dots \cup B_6$ are *CO*-retentive in T. For better readability, we will henceforth write a_{x+y} for $a_{x+y \bmod 7}$, B_{x+y} for $B_{x+y \bmod 7}$, and b^k_{x+y} for $b^k_{x+y \bmod 9}$.

For *CO*-retentiveness of A, fix an arbitrary $i \in \{0, \dots, 6\}$ and consider $a_i \in A$. The dominators of a_i are given by

$$D(a_i) = \{a_{i+4}, a_{i+5}, a_{i+6}\} \cup B_i \cup B_{i+3}.$$

Figure 44 illustrates the case where $a_i = a_1$. It is now readily appreciated that in $(D(a_1), \succ)$, a_{i+5} is only dominated by a_{i+4}, whereas all other alternatives are dominated by at least two alternatives. Accordingly, $CO(D(a_i)) = \{a_{i+5}\} \subseteq A$, which implies that A is *CO*-retentive in T.

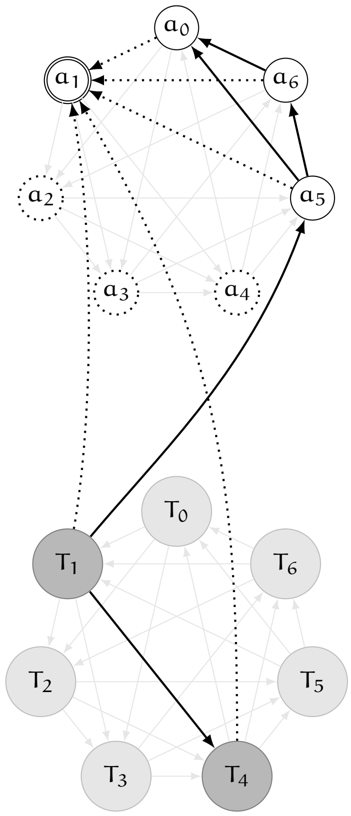


Figure 44: Partial representation of the tournament T used in the proof of Proposition 10.25, illustrating that A is CO-retentive. The case shown is the one where $a_i = a_1$. The dotted edges indicate the dominators of a_1, all missing edges in $(D(a_1), \succ)$ point downward. It is easy to see that a_6 is the Copeland winner in $(D(a_1), \succ)$.

For CO-retentiveness of $B = B_0 \cup \cdots \cup B_6$, fix $k \in \{0, \ldots, 6\}$ and $i \in \{0, \ldots, 8\}$ arbitrarily and consider $b_i^k \in B_k$. The dominators of b_i^k are given by

$$
\begin{aligned}
D(b_i^k) = &\{b_{i+5}^k, b_{i+6}^k, b_{i+7}^k, b_{i+8}^k\} && (10)\\
&\cup \{a_{k+1}, a_{k+2}, a_{k+3}, a_{k+5}, a_{k+6}\} && (11)\\
&\cup B_{k+4} \cup B_{k+5} \cup B_{k+6}. && (12)
\end{aligned}
$$

Figure 45 illustrates the case where $b_i^k = b_2^1$. We now find that $CO(D(b_i^k)) = B_{k+4}$: each alternative $b \in B_{k+4}$ has a Copeland score of $4 + 9 + 9 + 4 + 1 = 27$, whereas each of the alternatives a_{k+1}, a_{k+2}, a_{k+3}, a_{k+5}, a_{k+6} has a score of $2 + 4 + 9 + 9 = 24$ and all other alternatives in $D(b_i^k)$ have a score of at most 19. It follows that B is CO-retentive in T. $\square$

The same construction can also be used to show that $\mathring{CO}$ is not monotonic, which establishes that monotonicity is not inherited in general. To see this, first observe that both A and B are *minimal* re-

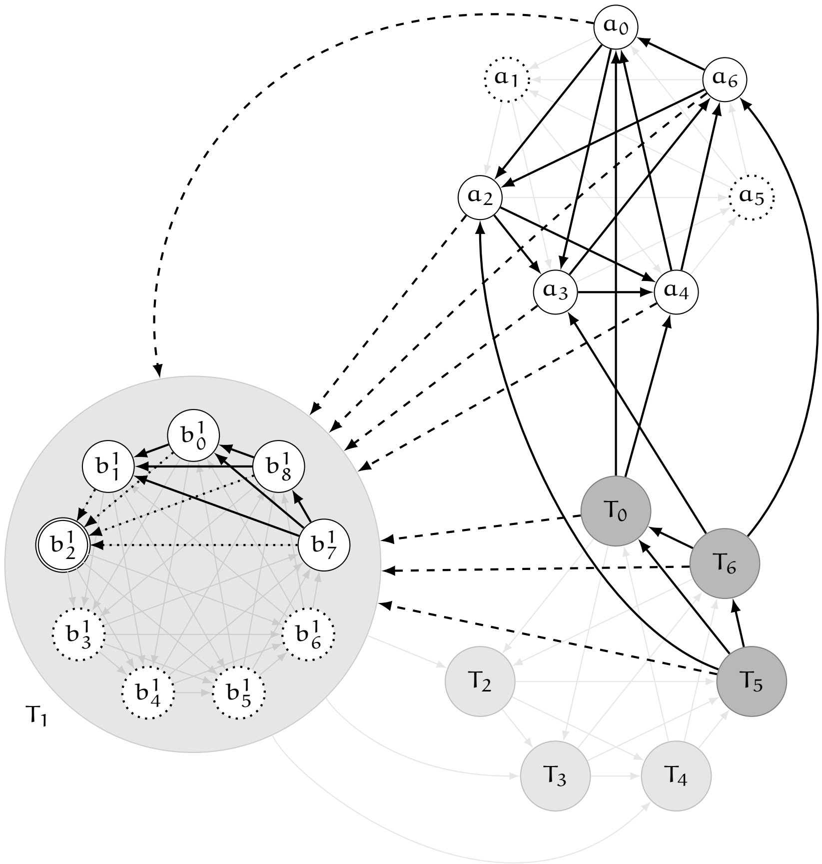


Figure 45: Partial representation of the tournament T used in the proof of Proposition 10.25, illustrating that B is *CO*-retentive. The case shown is the one where $b_i^k = b_2^1$. The dotted and dashed edges indicate the dominators of b_2^1. The dashed edges also represent (part of) the dominance relation inside $D(b_2^1)$. All missing edges in $(D(b_2^1), \succ)$ point downward. It is easy to see that the Copeland winners in $(D(b_2^1), \succ)$ are exactly the alternatives in T_5.

tentive sets in T, i.e., $\mathring{CO}(T) = A \cup B$. Now fix $k \in \{0, \ldots, 6\}$ and $i \in \{0, \ldots, 8\}$ arbitrarily and consider $b_i^k \in B_k$. Let T' be the tournament that is identical to T except that b_i^k is strengthened against all alternatives in B_{k+4}. For example, let $k = 1$. Then $T' = (A \cup B, \succ')$ with $T'|_{A \cup B \setminus \{b_i^1\}} = T|_{A \cup B \setminus \{b_i^1\}}$ and $D_{\succ'}(b_i^1) = D_{\succ}(b_i^1) \setminus B_5$. Since $T'|_{D_{\succ'}(a)} = T|_{D_{\succ}(a)}$ for all $a \in A$, the set A is a minimal *CO*-retentive set in T'. On the other hand, $CO(D_{\succ'}(b_i^1)) = \{a_2\}$, which means that B is not *CO*-retentive in T'. Furthermore, no minimal *CO*-retentive set C can contain b_i^1: every such set would also have to contain $CO(D_{\succ'}(b_i^1)) = \{a_2\}$, and $C' = C \cap A$ would be a strictly smaller *CO*-retentive set. Thus $b_i^1 \notin \mathring{CO}(T')$.

10.7 SUMMARY

Starting with the trivial tournament solution, we have defined an infinite sequence of efficiently computable tournament solutions that, under certain conditions, are strictly contained in one another, strictly contain *TEQ*, and share most of its desirable properties. The implications of these findings are both of theoretical and practical nature.

From a practical point of view, we have outlined an anytime algorithm for computing *TEQ* that returns smaller and smaller supersets of *TEQ*, which furthermore satisfy standard properties suggested in the literature. Previous algorithms for *TEQ* (see, e.g., Brandt et al., 2010) are incapable of providing *any* useful information in general when stopped prematurely.

From a theoretical point of view, the new perspective on *TEQ* as the limit of an infinite sequence of tournament solutions may prove useful to improve our understanding of Schwartz's conjecture. In particular, it yields an infinite sequence of increasingly difficult conjectures, each of them a weaker version of that of Schwartz. We proved the second conjecture in this sequence. Now that Schwartz's conjecture itself has been shown to be false, a natural question is how many statements of this sequence hold. As exemplified in this chapter, both proving and disproving this kind of conjectures turns out to be surprisingly difficult.

11 MANIPULATION OF SOCIAL CHOICE FUNCTIONS

In the final chapter of this thesis, we turn to the issue of strategic manipulation. An SCF is manipulable if one or more voters can misrepresent their preferences in order to obtain a more preferred choice set. While comparing choice sets is trivial for resolute SCFs, this is not the case for irresolute ones. Whether one choice set is preferred to another depends on how the preferences over individual alternatives are to be extended to sets of alternatives. In this chapter, we will be concerned with three of the most well-known preference extensions due to Kelly (1977), Fishburn (1972), and Gärdenfors (1976). After defining these extensions in Section 11.2 and reviewing related work in Section 11.3, we present our results in Sections 11.4 and 11.5. On the one hand, we provide sufficient conditions for strategyproofness and identify social choice functions that satisfy these conditions. For example, we show that the top cycle is strategyproof according to Gärdenfors' set extension, answering a question by Gärdenfors (1976) in the affirmative. On the other hand, we propose necessary conditions for strategyproofness and show that some more discriminatory social choice functions such as the minimal covering set and the bipartisan set, which have recently been shown to be strategyproof according to Kelly's extension, fail to satisfy strategyproofness according to Fishburn's and Gärdenfors' extension.

Throughout this chapter, we assume that preferences are strict, but not necessarily transitive. The reason is that Theorems 11.7 and 11.12–though becoming stronger—are easier to prove for possibly intransitive preferences. Theorems 11.10 and 11.13, on the other hand, become slightly weaker because there exist SCFs that are only manipulable if intransitive preferences are allowed. For all the manipulable SCFs we consider, however, we show that they are manipulable even if transitive preferences are required.

11.1 MOTIVATION

One of the central results in social choice theory states that every nontrivial SCF is susceptible to strategic manipulation (Gibbard, 1973; Satterthwaite, 1975). However, the classic result by Gibbard and Satterthwaite only applies to *resolute* SCFs. For irresolute SCFs, on the other hand, even defining manipulation is nontrivial. The reason is that preference relations over alternatives do not contain enough information to compare sets of alternatives. A number of proposals have been made in the literature, and many of them are surveyed by Taylor (2005) and Barberà (2010).

preferences over sets

How preferences over sets of alternatives relate to or depend on preferences over individual alternatives is a fundamental issue that goes back to at least de Finetti (1937) and Savage (1954). In the context of social choice the alternatives are usually interpreted as mutually exclusive candidates for a unique final choice. For instance, assume a voter prefers a to b, b to c, and a to c. What can we reasonably deduce from this about his preferences over the subsets of $\{a, b, c\}$? It stands to reason to assume that he would strictly prefer $\{a\}$ to $\{b\}$, and $\{b\}$ to $\{c\}$. If a single alternative is eventually chosen from each choice set, it is safe to assume that he also prefers $\{a\}$ to $\{b, c\}$ (Kelly's extension), but whether he prefers $\{a, b\}$ to $\{a, b, c\}$ already depends on (his knowledge about) the final decision process. In the case of a lottery over all pre-selected alternatives according to a known *a priori* probability distribution with full support, he would prefer $\{a, b\}$ to $\{a, b, c\}$ (Fishburn's extension). This assumption is, however, not sufficient to separate $\{a, b\}$ and $\{a, c\}$. Based on a sure-thing principle which prescribes that alternatives present in both choice sets can be ignored, it would be natural to prefer the former to the latter (Gärdenfors' extension). Finally, whether the voter prefers $\{a, c\}$ to $\{b\}$ depends on his attitude towards risk: he might hope for his most-preferred alternative (leximax extension), fear that his worst alternative will be chosen (leximin extension), or maximize his expected utility.

Kelly's extension

Fishburn's extension

Gärdenfors' extension

leximax extension

leximin extension

11.2 PREFERENCE EXTENSIONS

We now formally define the set extensions due to Kelly (1977), Fishburn (1972),[1] and Gärdenfors (1976). Let R_i be a preference relation over A and $X, Y \subseteq A$ two nonempty subsets of A.

- $X \ R_i^K \ Y$ if and only if $x \ R_i \ y$ for all $x \in X$ and all $y \in Y$ (Kelly, 1977).
  One interpretation of this extension is that voters are unaware of the mechanism (e.g., a lottery) that will be used to pick the winning alternative (Gärdenfors, 1979).

- $X \ R_i^F \ Y$ if and only if $x \ R_i \ y$, $x \ R_i \ z$, and $y \ R_i \ z$ for all $x \in X \setminus Y$, $y \in X \cap Y$, and $z \in Y \setminus X$ (Fishburn, 1972)
  One interpretation of this extension is that the winning alternative is picked by a lottery according to some underlying *a priori* distribution and that voters are unaware of this distribution (Ching and Zhou, 2002). Alternatively, one may assume the existence of a tie-breaker with linear, but unknown, preferences.

1 Gärdenfors (1979) attributed this extension to Fishburn because it is the weakest extension that satisfies a certain set of axioms proposed by Fishburn (1972).

- $X\ R_i^G\ Y$ if and only if one of the following conditions is satisfied (Gärdenfors, 1976):
  (*i*) $X \subset Y$ and $x\ R_i\ y$ for all $x \in X$ and $y \in Y \setminus X$,
  (*ii*) $Y \subset X$ and $x\ R_i\ y$ for all $x \in X \setminus Y$ and $y \in Y$, or
  (*iii*) neither $X \subset Y$ nor $Y \subset X$ and $x\ R_i\ y$ for all $x \in X \setminus Y$ and $y \in Y \setminus X$.

  No interpretation in terms of lotteries is known for this set extension.

The definition of Gärdenfors's extension is somewhat "discontinuous," which is not only reflected in the hardly elegant characterization given in Theorem 11.12, but also in the fact that R_i^G might not be transitive, even if R_i is. For an example, let $A = \{a, b, c, d\}$ and $a\ R_i\ b\ R_i\ c\ R_i\ d$. Then $\{a, c\}\ R_i^G\ \{b, c\}$ and $\{b, c\}\ R_i^G\ \{b, d\}$, but not $\{a, c\}\ R_i^G\ \{b, d\}$.[2]

It is easy to see that the set extensions defined above form an inclusion hierarchy.

Fact 11.1. *For all preference relations* R_i *and subsets* $X, Y \subseteq A$,

$$X\ R_i^K\ Y \ \textit{implies}\ \ X\ R_i^F\ Y \ \textit{implies}\ \ X\ R_i^G\ Y.$$

For $\mathscr{E} \in \{K, F, G\}$, let $P_i^{\mathscr{E}}$ denote the strict part of $R_i^{\mathscr{E}}$. As R_i is strict, so is $R_i^{\mathscr{E}}$. Therefore, we have $X\ P_i^{\mathscr{E}}\ Y$ if and only if $X\ R_i^{\mathscr{E}}\ Y$ and $X \neq Y$.

Based on these set extensions, we can now define three different notions of strategyproofness for irresolute SCFs. Note that, in contrast to some related papers, we interpret preference extensions as fully specified (incomplete) preference relations rather than minimal conditions on set preferences.

Definition 11.2. *Let* $\mathscr{E} \in \{K, F, G\}$*. An SCF* f *is* $P^{\mathscr{E}}$-manipulable by a group of voters $C \subseteq N$ *if there exist preference profiles* R *and* R' *with* $R_j = R'_j$ *for all* $j \notin C$ *such that*

$$f(R')\ P_i^{\mathscr{E}}\ f(R) \textit{ for all } i \in C.$$

An SCF is $P^{\mathscr{E}}$-strategyproof *if it is not* $P^{\mathscr{E}}$*-manipulable by single voters. An SCF is* $P^{\mathscr{E}}$-group-strategyproof *if it is not* $P^{\mathscr{E}}$*-manipulable by any group of voters.*

Fact 11.1 implies that P^G-group-strategyproofness is stronger than P^F-group-strategyproofness, which in turn is stronger than P^K-group-strategyproofness.

---

2 It can however be shown that the strict part of Gärdenfors's extension is *acyclic* for transitive preferences.

11.3 RELATED WORK

Barberà (1977a) and Kelly (1977) have shown independently that all nontrivial SCFs that are rationalizable via a quasi-transitive preference relation are P^K-manipulable. However, as witnessed by various other (non-strategic) impossibility results that involve quasi-transitive rationalizability (e.g., Mas-Colell and Sonnenschein, 1972), it appears as if this property itself is unduly restrictive. As a consequence, Kelly (1977, p. 445) concludes his paper by contemplating that "one plausible interpretation of such a theorem is that, rather than demonstrating the impossibility of reasonable strategy-proof social choice functions, it is part of a critique of the regularity [rationalizability] conditions."

Strengthening earlier results by Gärdenfors (1976) and Taylor (2005), Brandt (2011a) showed that no Condorcet extension is P^K-strategyproof. The proof, however, crucially depends on strategic tie-breaking and hence does not work for strict preferences. For this reason, only strict preferences are considered in this chapter.

Brandt (2011a) also provided a sufficient condition for P^K-group-strategyproofness. *Set-monotonicity* can be seen as an irresolute variant of Maskin-monotonicity (Maskin, 1999) and prescribes that the choice set is invariant under the weakening of unchosen alternatives. For a given preference profile R with $b \ R_i \ a$, let $R_{i:(a,b)}$ denote the preference profile

$$R_{i:(a,b)} = (R_1, \dots, R_{i-1}, R_i \setminus \{(b,a)\} \cup \{(a,b)\}, R_{i+1}, \dots, R_n).$$

That is, $R_{i:(a,b)}$ is identical to R except that alternative a is strengthened with respect to b within voter i's preference relation.

Definition 11.3. *An SCF* f *satisfies* set-monotonicity (SET-MON) *if* $f(R_{i:(a,b)}) = f(R)$ *for all preference profiles* R*, voters* i*, and alternatives* a, b *with* $b \notin f(R)$.

Theorem 11.4 (Brandt, 2011a)**.** *Every SCF that satisfies SET-MON is* P^K*-group-strategyproof.*

Set-monotonicity is a demanding condition, but a handful of SCFs are known to be set-monotonic: *PAR*, *OMNI*, *COND*, *TC*, *MC*, and *BP*. With the exception of the Pareto rule and the omninomination rule, all of these SCFs are pairwise Condorcet extensions.

pairwise SCFs

For the class of pairwise SCFs, this condition is also necessary, which shows that many well-known SCFs such as *BO*, *CO*, *UC*, and *BA* are *not* P^K-group-strategyproof.

Theorem 11.5 (Brandt, 2011a)**.** *Every pairwise SCF that is* P^K*-group-strategyproof satisfies SET-MON.*

Strategyproofness according to Kelly's extension thus draws a sharp line within the space of SCFs as almost all established non-pairwise SCFs (such as plurality and all weak Condorcet extensions like Young's rule) are also known to be P^K-manipulable (see, e.g., Taylor, 2005).

The state of affairs for Gärdenfors' and Fishburn's extensions is less clear. Gärdenfors (1976) has shown that *COND* and *OMNI* are P^G-group-strategyproof. In an attempt to extend this result to more discriminatory SCFs, he also claimed that $COND \cap PAR$, which returns the Condorcet winner if it exists and all Pareto-undominated alternatives otherwise, is P^G-strategyproof. However, we show that this is not the case (Proposition 11.17). Gärdenfors (1976, p. 226) concludes that "we have not been able to find any more decisive function which is stable [strategyproof] and satisfies minimal requirements on democratic decision functions." We show that *TC* is such a function (Corollary 11.14).

Apart from a theorem by Ching and Zhou (2002), which uses an unusually strong definition of strategyproofness, we are not aware of any characterization result using Fishburn's extension. Feldman (1979) has shown that the Pareto rule is P^F-strategyproof and Sanver and Zwicker (2012) have shown that the same is true for *TC*.

11.4 NECESSARY AND SUFFICIENT CONDITIONS

We first introduce a new property that requires that modifying preferences between chosen alternatives may only result in smaller choice sets. Set-monotonicity entails a condition called *independence of unchosen alternatives*,[3] which states that the choice set is invariant under modifications of the preferences between unchosen alternatives.[4] Accordingly, the new property will be called *exclusive independence of chosen alternatives*, where "exclusive" refers to the requirement that unchosen alternatives remain unchosen.

Definition 11.6. *An SCF* f *satisfies* exclusive independence of chosen alternatives (EICA) *if* $f(R') \subseteq f(R)$ *for all pairs of preference profiles* R *and* R' *that differ only on alternatives in* $f(R)$, *i.e.,* $R_i|_{\{a,b\}} = R'_i|_{\{a,b\}}$ *for all* $i \in N$ *and all alternatives* a, b *with* $b \notin f(R)$.

It turns out that, together with SET-MON, this new property is sufficient for an SCF to be group-strategyproof according to Fishburn's preference extension.

Theorem 11.7. *Every SCF that satisfies* SET-MON *and* EICA *is* P^F*-group-strategyproof.*

We first need a lemma, which states that EICA together with SET-MON implies the following: if only preferences between chosen alternatives are modified and some alternatives leave the choice set, then

3 See page 148 for a formulation in the context of tournament solutions.
4 See Definition 10.10 on page 148 for a formal definition in the context of tournament solutions.

at least one of them was weakened with respect to an alternative that remains chosen.

Lemma 11.8. *Let* f *be an SCF that satisfies* SET-MON *and* EICA *and consider a pair of profiles* R, R' *that differ only on alternatives in* $f(R)$*. If* $f(R') \subset f(R)$*, then there exist* $i \in N$, $x \in f(R) \setminus f(R')$ *and* $y \in f(R')$ *such that* $x \; R_i \; y$ *and* $y \; R'_i \; x$.

Proof. Assume for contradiction that $R' \setminus R = \bigcup_{i \in N}(R'_i \setminus R_i)$ does not contain a pair (y, x) with $y \in f(R')$ and $x \in f(R) \setminus f(R')$. Then each pair $(y, x) \in R' \setminus R$ belongs to exactly one of the following two classes.

Class 1. $y, x \in f(R')$

Class 2. $y \in f(R) \setminus f(R')$, $x \in A$

We now start with preference profile R' and change the preferences in $R' \setminus R$ one after the other to arrive at profile R. We first change the preferences for all pairs (y, x) from Class 1 and denote the resulting profile by R''. As R' and R'' differ only on alternatives in $f(R')$, EICA implies that $f(R'') \subseteq f(R')$. We then change the preferences for all pairs (y, x) from Class 2. By definition, the resulting profile is R. As $f(R'') \subseteq f(R')$, $y \notin f(R')$ implies $y \notin f(R'')$. Thus, in this second step, only alternatives $y \notin f(R'')$ are weakened and SET-MON implies that $f(R) = f(R'')$. But $f(R) = f(R'') \subseteq f(R')$ contradicts the assumption that $f(R')$ is a strict subset of $f(R)$. □

We are now ready to prove Theorem 11.7.

Proof of Theorem 11.7. Let f be an SCF that satisfies SET-MON and EICA and assume for contradiction that f is not P^F-group-strategyproof. Then, there have to be a group of voters $C \subseteq N$ and two preference profiles R and R' with $R_j = R'_j$ for all $j \notin C$ such that $f(R') \; P^F_i \; f(R)$ for all $i \in C$. We choose R and R' such that $\delta(R, R')$ is minimal, i.e., we look at a "smallest" counter-example in the sense that R and R' coincide as much as possible.[5] Let $f(R) = X$ and $f(R') = Y$. We may assume $\delta(R, R') > 0$ as otherwise $R = R'$ and $X = Y$. Now, consider a pair of alternatives $a, b \in A$ such that, for some $i \in C$, $a \; R_i \; b$ and $b \; R'_i \; a$, i.e., voter i misrepresents his preference relation by strengthening b. The following argument will show that no such a and b exist, which implies that R and R' and consequently X and Y are identical, a contradiction. We need the following two claims.

Claim 1. $b \in Y$

In order to prove this claim, suppose that $b \notin Y$. It follows from SET-MON that $f(R'_{i:(a,b)}) = f(R') = Y$. Thus, R and $R'_{i:(a,b)}$ constitute a smaller counter-example since $\delta(R, R'_{i:(a,b)}) = \delta(R, R') - 1$. This is a contradiction because $\delta(R, R')$ was assumed to be minimal.

5 The distance $\delta(R, R')$ between to preference profiles R and R' was defined on page 98.

Claim 2. $a \in Y$

The following case distinction shows that Claim 2 holds. Suppose $a \notin Y$. If $a \notin X$ either, SET-MON implies that $f(R_{i:(b,a)}) = f(R) = X$. Thus, $R_{i:(b,a)}$ and R' constitute a smaller counter-example since $\delta(R_{i:(b,a)}, R') = \delta(R, R') - 1$. On the other hand, if $a \in X \setminus Y$, $b \in Y$ and $a \; R_i \; b$ contradict the assumption that $Y \; P_i^F \; X$.

We thus have $\{a, b\} \subseteq Y$ for every pair (a, b) such that some voter $i \in C$ misrepresents his preference between a and b. In particular, this means that R and R' differ only on alternatives in $Y = f(R')$. Therefore, Lemma 11.8 implies[6] that either $X = Y$ or $X \subset Y$ and there exist $y \in Y \setminus X$ and $x \in X$ such that $y \; R_i' \; x$ and $x \; R_i \; y$. Both cases contradict the assumption that $Y \; P_i^F \; X$.

Hence, we have shown that no such R and R' exist, which concludes the proof. □

Note that the preferences of voter i in the profile $R_{i:(b,a)}$ might not be transitive. Therefore, one has to be careful when applying the preceding proof to *PAR* and *OMNI*, as those SCFs are only defined for transitive preferences. One can however generalize the definition of both SCFs to intransitive preference profiles in such a way that all arguments in the proof remain valid.[7]

For *pairwise* SCFs, the following weakening of EICA can be shown to be necessary for group-strategyproofness according to Fishburn's extension. It prescribes that modifying preferences among chosen alternatives does not result in a choice set that is a strict superset of the original choice set.

Definition 11.9. *An SCF f satisfies* weak EICA *if $f(R) \not\subset f(R')$ for all pairs of preference profiles R and R' that differ only on alternatives in $f(R)$.*

Theorem 11.10. *Every pairwise SCF that is P^F-group-strategyproof satisfies SET-MON and weak EICA.*

Proof. We need to show that every pairwise SCF that violates either SET-MON or weak EICA is P^F-manipulable.

First, let f be a pairwise SCF that violates SET-MON.[8] Then, there exist a preference profile $R = (R_1, \ldots, R_n)$, a voter i, and two alternatives a, b with $b \; R_i \; a$ and $b \notin f(R) = X$ such that $f(R_{i:(a,b)}) = Y \neq X$.

Let R_{n+1} be a preference relation such that $b \; R_{n+1} \; a$ and $Y \; P_{n+1}^F X$ (such a relation exists because $b \notin X$) and let $R_{n+2} = R_{n+1}^{\leftarrow}$. Let

6 Observe that we apply Lemma 11.8 with $f(R) = Y$ and $f(R') = X$.

7 To see this, define *OMNI*(R) to contain all those alternatives a for which there exists a voter i with $a \; R_i \; b$ for all $b \neq a$. The definition of *PAR* can remain unchanged. Generalized in this way, the choice set of either function may be empty for intransitive preferences. It can however easily be shown that *PAR* and *OMNI* still satisfy SET-MON *in the case of nonempty choice sets*. As the sets X and Y used in the proof of Theorem 11.7 are nonempty, the latter condition is then sufficient for the argument in the proof to go through.

8 Brandt (2011a) has shown that this implies P^K-manipulability in a setting where ties are allowed.

S denote the preference profile $S = (R_1, \dots, R_n, R_{n+1}, R_{n+2})$. It follows from the definition of pairwise SCFs that $f(S) = f(R) = X$ and $f(S_{n+1:(a,b)}) = f(R_{i:(a,b)}) = Y$.

As $Y \; P^F_{n+1} \; X$, we have that f can be manipulated by voter $n+1$ at preference profile S by misstating his preference $b \; R_{n+1} \; a$ as $a \; R_{n+1} \; b$. Hence, f is P^F-manipulable.

Second, let f be a pairwise SCF that violates weak EICA. Then, there exist two preference profiles $R = (R_1, \dots, R_n)$ and $R' = (R'_1, \dots, R'_n)$ that differ only on alternatives in $f(R)$, such that $f(R) = X \subset Y = f(R')$. Let $C \subseteq N$ be the group of voters that have different preferences in R and R', i.e., $C = \{i \in N \mid R_i \neq R'_i\}$. Without loss of generality, we can assume that $C = \{1, \dots, c\}$, where $c = |C|$. For all $i \in \{1, \dots, c\}$, let R_{n+i} be a preference relation such that $Y \; P^F \; X$ and $R_i \setminus R'_i \subseteq R_{n+i}$ (such a preference relation exists because $X \subset Y$ and $R_i \setminus R'_i \subseteq X \times X$) and let $R_{n+c+i} = R^{\leftarrow}_{n+i}$. Furthermore, for all $i \in \{1, \dots, c\}$, let $R'_{n+i} = R_{n+i} \setminus R_i \cup R'_i$. That is, R'_{n+i} differs from R_{n+i} on exactly the same pairs of alternatives as R'_i differs from R_i.

Consider the preference profiles

$$S = (R_1, \dots, R_n, R_{n+1}, \dots, R_{n+c}, R_{n+c+1}, \dots, R_{n+2c}) \text{ and}$$
$$S' = (R_1, \dots, R_n, R'_{n+1}, \dots, R'_{n+c}, R_{n+c+1}, \dots, R_{n+2c}).$$

It follows from the definition of pairwise SCFs that $f(S) = f(R) = X$ and $f(S') = f(R') = Y$. As $Y \; P^F_{n+i} \; X$ for all $i \in \{1, \dots, c\}$, we have that f can be manipulated by the group $\{n+1, \dots, n+c\}$ at preference profile S by misstating their preferences R_{n+i} as R'_{n+i}. Hence, f is P^F-manipulable. $\square$

We now turn to P^G-group-strategyproofness. When comparing two sets, P^G differs from P^F only in the case when neither set is contained in the other. The following definition captures exactly this case.

Definition 11.11. *An SCF f satisfies the* symmetric difference property (SDP) *if either $f(R) \subseteq f(R')$ or $f(R') \subseteq f(R)$ for all pairs of preference profiles R and R' such that $R_i|_{\{a,b\}} = R'_i|_{\{a,b\}}$ for all $i \in N$ and all alternatives a, b with $a \in f(R) \setminus f(R')$ and $b \in f(R') \setminus f(R)$.*

Theorem 11.12. *Every SCF that satisfies SET-MON, EICA, and SDP is P^G-group-strategyproof.*

Proof. Let f be an SCF that satisfies SET-MON, EICA, and SDP, and assume for contradiction that f is not P^G-group-strategyproof. Then, there have to be a group of voters $C \subseteq N$ and two preference profiles R and R' with $R_j = R'_j$ for all $j \notin C$ such that $f(R') \; P^G_i \; f(R)$ for all $i \in C$. Choose R and R' such that $\delta(R, R')$ is minimal and let $X = f(R)$ and $Y = f(R')$.

As P^G coincides with P^F on all pairs where one set is contained in the other set, and, by Theorem 11.7, f is P^F-group-strategyproof, we can conclude that neither $X \subsetneq Y$ nor $Y \subsetneq X$. Thus, SDP implies that there exist pairs (x, y) with $x \in X \setminus Y$ and $y \in Y \setminus X$ such that some voters have modified their preference between x and y, i.e., $(x, y) \in (R_i \setminus R'_i) \cup (R'_i \setminus R_i)$ for some $i \in C$. Each such pair (x, y) thus belongs to at least one of the following two classes:

Class 1. $(x, y) \in R_i \setminus R'_i$ *for some* $i \in C$

Class 2. $(x, y) \in R'_i \setminus R_i$ *for some* $i \in C$

We go on to show that Class 1 contains at least one pair. Assume for contradiction that all pairs belong to Class 2 and let $(x, y) \in R'_i \setminus R_i$ be one of these pairs. As $x \notin Y$, SET-MON implies that $f(R'_{i:(y,x)}) = f(R') = Y$. As $f(R'_{i:(y,x)})\ P^G_i\ f(R)$ for all $i \in C$ and $\delta(R, R'_{i:(y,x)}) = \delta(R, R') - 1$, R and $R'_{i:(y,x)}$ constitute a smaller counter-example, contradicting the minimality of $\delta(R, R')$.

Thus, there is at least one pair (x, y) that belongs to Class 1, i.e., a pair (x, y) with $x \in X \setminus Y$ and $y \in Y \setminus X$ such that $x\ R_i\ y$ for some voter $i \in C$. But this contradicts the assumption that $Y\ P^G_i\ X$ for all $i \in C$, and completes the proof. □

As was the case for Fishburn's extension, a set of necessary conditions for pairwise SCFs can be obtained by replacing EICA with weak EICA.

Theorem 11.13. *Every pairwise SCF that is P^G-group-strategyproof satisfies SET-MON, weak EICA, and SDP.*

Proof. By Theorem 11.10 and the fact that P^G-group-strategyproofness implies P^F-group-strategyproofness, it remains to be shown that every pairwise SCF that violates SDP is P^G-manipulable. Suppose that f is pairwise and violates SDP. Then, there exists two preference profiles $R = (R_1, \dots, R_n)$ and $R' = (R'_1, \dots, R'_n)$ such that $X = f(R)$ and $Y = f(R')$ are not contained in one another, and $R_i|_{\{x,y\}} = R'_i|_{\{x,y\}}$ for all $i \in N$ and all alternatives x, y with $x \in X \setminus Y$ and $y \in Y \setminus X$.

The proof now works analogously as the proof of Theorem 11.10. For each voter i with $R_i \neq R'_i$ we have two new voters R_{n+i} and R_{n+c+i} such that $Y\ P^G_{n+i}\ X$, $R_i \setminus R'_i \subseteq R_{n+i}$, and $R_{n+c+i} = \overleftarrow{R}_{n+i}$. By letting $R'_{n+i} = R_{n+i} \setminus R_i \cup R'_i$ and defining S and S' as in the proof of Theorem 11.10, we can show that f is P^G-manipulable. □

11.5 CONSEQUENCES

We are now ready to study the strategyproofness of the SCFs mentioned in Section 11.3. It can be checked that *COND* and *TC* satisfy SET-MON, EICA, and SDP and thus, by Theorem 11.12, are P^G-group-strategyproof.

Corollary 11.14. *COND and TC are P^G-group-strategyproof.*

Proof. By Theorem 11.12, it is sufficient to show that *COND* and *TC* satisfy SET-MON, EICA, and SDP.

If a preference profile R does not have a Condorcet winner, *COND* trivially satisfies the three properties because all alternatives are chosen. If R has a Condorcet winner, EICA is again trivial and SET-MON and SDP are straightforward.

The (easy) fact that *TC* satisfies SET-MON was shown by Brandt (2011a). To see that *TC* satisfies EICA, consider $b \notin TC(R)$. By definition of *TC*, $b \; R_M^* \; a$ for *no* $a \in TC(R)$. As R and R' differ only on alternatives in $TC(R)$, it follows that $b \; R_M'^* \; a$ for no $a \in TC(R)$, and thus $a \notin TC(R')$.

Finally, to see that *TC* satisfies SDP, observe that $a \; P_M \; b$ for all $a \in TC(R)$ and $b \notin TC(R)$. Thus, if $x \in TC(R) \setminus TC(R')$ and $y \in TC(R') \setminus TC(R)$, we have $x \; P_M \; y$ and $y \; P'_M \; x$. This implies that at least one voter has modified his preference between x and y. □

OMNI, *PAR*, and $COND \cap PAR$ satisfy SET-MON and EICA, but not SDP.

Corollary 11.15. *OMNI, PAR, and $COND \cap PAR$ are P^F-group-strategyproof.*

Proof. By Theorem 11.7, it is sufficient to show that SET-MON and EICA are satisfied. This is easy to see for *OMNI*, so let us focus on $COND \cap PAR$ and *PAR*.

SET-MON holds because a Pareto-dominated alternative remains Pareto-dominated when it is weakened. For EICA, observe that transitivity of Pareto-dominance implies that each Pareto-dominated alternative is dominated by a Pareto-undominated one. Therefore, if $a \notin PAR(R)$ and R and R' differ only on alternatives in $PAR(R)$, then $a \notin PAR(R')$ because a is still Pareto-dominated by the same alternative. □

As *OMNI*, *PAR*, and $COND \cap PAR$ are not pairwise, the fact that they violate SDP does *not* imply that they are P^G-manipulable. In fact, it turns out that *OMNI* is strategyproof according to Gärdenfors' extension, while *PAR* and $COND \cap PAR$ are not.

Proposition 11.16. *OMNI is P^G-group-strategyproof.*

Proof. Assume for contradiction that *OMNI* is P^G-manipulable. Then, there have to be a group of voters $C \subseteq N$ and two preference profiles R and R' with $R_j = R'_j$ for all $j \notin C$ such that $OMNI(R') \; P_i^G \; OMNI(R)$ for all $i \in C$. Denote $X = OMNI(R)$ and $Y = OMNI(R')$.

As P^G coincides with P^F on all pairs where one set is contained in the other set, and, by Theorem 11.7, *OMNI* is P^F-group-strategyproof, we can conclude that neither $X \subseteq Y$ nor $Y \subseteq X$. Choose $x \in X \setminus Y$ and $y \in Y \setminus X$ arbitrarily. On the one hand, $x \in X \setminus Y$ implies the existence of a voter $i \in C$ with $x \; R_i \; a$ for all $a \neq x$. On the other hand, $Y \; P_i^G \; X$ implies $y \; R_i \; x$, a contradiction. □

Proposition 11.17. *PAR and COND ∩ PAR are P^G-manipulable.*

Proof. Consider the following profile $R = (R_1, R_2, R_3, R_4)$.

R_1	R_2	R_3	R_4
c	c	a	a
d	d	b	b
b	a	c	c
a	b	d	d

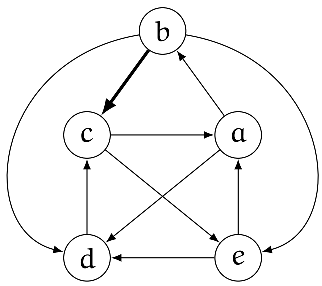
It is easily verified that $PAR(R) = \{a, b, c\}$. Now let $R' = (R'_1, R_2, R_3, R_4)$ where $d \; R'_1 \; c \; R'_1 \; a \; R'_1 \; b$. Obviously, $PAR(R') = \{a, c, d\}$ and $\{a, c, d\} \; P^G_1 \; \{a, b, c\}$ because $d \; R_1 \; b$. I.e., the first voter can obtain a preferable choice set by misrepresenting his preferences. As neither R nor R' has a Condorcet winner, the same holds for *COND* ∩ *PAR*. □

Finally, we show that *MC* and *BP* violate weak EICA, which implies that both rules are manipulable according to Fishburn's extension.

Corollary 11.18. *MC and BP are P^F-manipulable.*

Proof. By Theorem 11.10 and the fact that both *MC* and *BP* are pairwise, it suffices to show that *MC* and *BP* violate weak EICA. To this end, consider the following profile $R = (R_1, R_2, R_3, R_4, R_5)$ and the corresponding majority graph representing P_M.

R_1	R_2	R_3	R_4	R_5
d	c	b	e	d
e	b	c	a	c
a	a	e	b	a
b	e	a	d	b
c	d	d	c	e



It can be checked that $MC(R) = BP(R) = \{a, b, c\}$. Define $R' = R_{1:(c,b)}$, i.e., the first voter strengthens c with respect to b. Observe that P_M and P'_M disagree on the pair $\{b, c\}$, and that $MC(R') = BP(R') = \{a, b, c, d, e\}$. Thus, both *MC* and *BP* violate weak EICA and the first voter can manipulate because $\{a, b, c, d, e\} \; P^F_1 \; \{a, b, c\}$. □

11.6 SUMMARY

We investigated the effect of various preference extensions on the manipulability of irresolute SCFs. We proposed necessary and sufficient conditions for strategyproofness according to Fishburn's and Gärdenfors' set extensions and used these conditions to illuminate the strategyproofness of a number of well-known SCFs. Our results are summarized in Table 9.[9] As mentioned in Section 11.3, some of these results were already known or—in the case of P^F-strategyproofness of the top cycle—have been discovered independently by other authors. In contrast to the papers by Gärdenfors (1976), Feldman (1979), and Sanver and Zwicker (2012), which more or less focus on particular SCFs, our axiomatic approach yields unified proofs of most of the statements in the table.

Many interesting open problems remain. For example, it is not known whether there exists a Pareto-optimal pairwise SCF that is strategyproof according to Gärdenfors' extension. Recently, the study of the manipulation of irresolute SCFs by other means than untruthfully representing one's preferences—e.g., by abstaining the election (Pérez, 2001; Jimeno et al., 2009)—has been initiated. It would be desirable to characterize SCFs that cannot be manipulated by abstention and, more generally, to improve our understanding of the interplay between both types of manipulation. For instance, it is not difficult to show that the negative results in Corollary 11.18 also extend to manipulation by abstention.

Another interesting related question concerns the epistemic foundations of the above extensions. Most of the literature in social choice theory focusses on well-studied economic models where agents have full knowledge of a random selection process, which is often assumed to be a lottery with uniform probabilities. The study of more intricate distributed protocols or computational selection devices that justify certain set extensions appears to be very promising. For instance, Kelly's set extension could be justified by a distributed protocol for "unpredictable" random selections that do not permit a meaningful prior distribution.

	P^K-strategyproof	P^F-strategyproof	P^G-strategyproof
OMNI	✓	✓	✓[a]
COND	✓	✓	✓[a]
TC	✓	✓[b]	✓
PAR	✓	✓[c]	–
COND ∩ *PAR*	✓	✓	–
MC	✓	–	–
BP	✓	–	–

[a] Gärdenfors (1976)

[b] Sanver and Zwicker (2012)

[c] Feldman (1979)

Table 9: Strategyproofness of irresolute SCFs

9 The results concerning P^K-strategyproofness are due to Brandt (2011a) and are included for the sake of completeness.

APPENDIX

A
SUMMARY TABLES

A.1 PROPERTIES OF DOMINANCE STRUCTURES

	MON	EFF	MAX	SING	OI [a]
strict dominance (S)	✓	✓	✓	✓	✓
weak dominance (W)	–	✓	✓	✓	–
very weak dominance (V)	✓	✓	–	✓	–
Börgers dominance (B)	✓	✓	✓	–	✓
mixed strict dominance (S^*)	✓	✓	✓	–	✓
mixed weak dominance (W^*)	–	✓	✓	–	–
mixed very weak dominance (V^*)	✓	✓	–	–	–
covering (C_M)	–	✓	✓	✓	✓
deep covering (C_D)	–	✓	✓	✓	✓

[a] OI = order-independence

A.2 PROPERTIES OF SYMMETRIC DOMINANCE STRUCTURES

	weak MON	TRA	SUB-COM	UNI
covering (C_M)	✓	✓	✓	✓
deep covering (C_D)	✓	✓	✓	✓

A.3 PROPERTIES OF TOURNAMENT SOLUTIONS

	MON	IUA	WSP	SSP	$\widehat{\gamma}$
trivial rule (*TRIV*)	✓	✓	✓	✓	✓
Copeland set (*CO*)	✓	–	–	–	–
Condorcet rule (*COND*)	✓	✓	✓	✓	✓
top cycle (*TC*)	✓	✓	✓	✓	✓
uncovered set (*UC*)	✓	–	✓	–	–
minimal covering set (*MC*)	✓	✓	✓	✓	✓
bipartisan set (*BP*)	✓	✓	✓	✓	✓
Banks set (*BA*)	✓	–	✓	–	–
tournament equilibrium set (*TEQ*)	–	–	–	–	–
TEQ in $\mathscr{T}_{n_{TEQ}}$	✓	✓	✓	✓	✓

BIBLIOGRAPHY

K. R. Apt. Uniform proofs of order independence for various strategy elimination procedures. *Contributions to Theoretical Economics*, 4(1), 2004.

K. R. Apt. Direct proofs of order independence. *Economics Bulletin*, 31 (1):106–115, 2011.

K. J. Arrow. *Social Choice and Individual Values*. New Haven: Cowles Foundation, 1st edition, 1951. 2nd edition 1963.

K. J. Arrow and H. Raynaud. *Social Choice and Multicriterion Decision-Making*. MIT Press, 1986.

R. J. Aumann and A. Brandenburger. Epistemic conditions for Nash equilibrium. *Econometrica*, 63(5):1161–1180, 1995.

D. Austen-Smith and J. S. Banks. *Positive Political Theory I: Collective Preference*. University of Michigan Press, 2000.

J. S. Banks. Sophisticated voting outcomes and agenda control. *Social Choice and Welfare*, 3:295–306, 1985.

S. Barberà. Manipulation of social decision functions. *Journal of Economic Theory*, 15(2):266–278, 1977a.

S. Barberà. The manipulation of social choice mechanisms that do not leave "too much" to chance. *Econometrica*, 45(7):1573–1588, 1977b.

S. Barberà. Strategy-proof social choice. In K. J. Arrow, A. K. Sen, and K. Suzumura, editors, *Handbook of Social Choice and Welfare*, volume 2, chapter 25, pages 731–832. Elsevier, 2010.

K. Basu and J. Weibull. Strategy subsets closed under rational behavior. *Economics Letters*, 36:141–146, 1991.

D. Baumeister and J. Rothe. Taking the final step to a full dichotomy of the possible winner problem in pure scoring rules. In *Proceedings of the 19th European Conference on Artificial Intelligence (ECAI)*, pages 1019–1020, 2010.

M. Benisch, G. B. David, and T. Sandholm. Algorithms for closed under rational behavior (CURB) sets. *Journal of Artificial Intelligence Research*, 38:513–534, 2010.

B. D. Bernheim. Rationalizable strategic behavior. *Econometrica*, 52(4): 1007–1028, 1984.

N. Betzler and B. Dorn. Towards a dichotomy for the possible winner problem in elections based on scoring rules. *Journal of Computer and System Sciences*, 76(8):812–836, 2010.

N. Betzler, S. Hemmann, and R. Niedermeier. A multivariate complexity analysis of determining possible winners given incomplete votes. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI)*, pages 53–58, 2009.

K. Binmore. Interpersonal comparison of utility. In H. Kincaid and D. Ross, editors, *The Oxford Handbook of Philosophy of Economics*, pages 540–559. Oxford University Press, 2009.

E. Borel. La théorie du jeu et les équations intégrales à noyau symétrique. *Comptes Rendus de l'Académie des Sciences*, 173:1304–1308, 1921.

T. Börgers. Pure strategy dominance. *Econometrica*, 61(2):423–430, 1993.

D. Bouyssou, T. Marchant, M. Pirlot, A. Tsoukiàs, and P. Vincke. *Evaluation and Decision Models: Stepping Stones for the Analyst*. Springer-Verlag, 2006.

F. Brandt. Group-strategyproof irresolute social choice functions. In T. Walsh, editor, *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI)*, pages 79–84. AAAI Press, 2011a.

F. Brandt. Minimal stable sets in tournaments. *Journal of Economic Theory*, 146(4):1481–1499, 2011b.

F. Brandt and F. Fischer. On the hardness and existence of quasi-strict equilibria. In B. Monien and U.-P. Schroeder, editors, *Proceedings of the 1st International Symposium on Algorithmic Game Theory (SAGT)*, volume 4997 of *Lecture Notes in Computer Science (LNCS)*, pages 291–302. Springer-Verlag, 2008a.

F. Brandt and F. Fischer. Computing the minimal covering set. *Mathematical Social Sciences*, 56(2):254–268, 2008b.

F. Brandt and P. Harrenstein. Characterization of dominance relations in finite coalitional games. *Theory and Decision*, 69(2):233–256, 2010.

F. Brandt and P. Harrenstein. Set-rationalizable choice and self-stability. *Journal of Economic Theory*, 146(4):1721–1731, 2011.

F. Brandt, F. Fischer, and P. Harrenstein. The computational complexity of choice sets. *Mathematical Logic Quarterly*, 55(4):444–459, 2009a.

F. Brandt, F. Fischer, P. Harrenstein, and Y. Shoham. Ranking games. *Artificial Intelligence*, 173(2):221–239, 2009b.

F. Brandt, F. Fischer, P. Harrenstein, and M. Mair. A computational analysis of the tournament equilibrium set. *Social Choice and Welfare*, 34(4):597–609, 2010.

F. Brandt, M. Chudnovsky, I. Kim, G. Liu, S. Norin, A. Scott, P. Seymour, and S. Thomassé. A counterexample to a conjecture of Schwartz. *Social Choice and Welfare*, 40:739–743, 2013.

Y. Chevaleyre, J. Lang, N. Maudet, and J. Monnot. Possible winners when new candidates are added: The case of scoring rules. In *Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI)*, pages 762–767. AAAI Press, 2010.

S. Ching and L. Zhou. Multi-valued strategy-proof social choice rules. *Social Choice and Welfare*, 19:569–580, 2002.

V. Conitzer and T. Sandholm. Complexity of (iterated) dominance. In *Proceedings of the 6th ACM Conference on Electronic Commerce (ACM-EC)*, pages 88–97. ACM Press, 2005.

V. Conitzer, M. Rognlie, and L. Xia. Preference functions that score rankings and maximum likelihood estimation. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI)*, pages 109–115, 2009.

W. J. Cook, W. H. Cunningham, W. R. Pulleyblank, and A. Schrijver. *Combinatorial Optimization*. Wiley and Sons, 1998.

B. de Finetti. La prévison: ses lois logiques, ses sources subjectives. *Annales de l'institut Henri Poincaré*, 7(1):1–68, 1937.

B. Debord. Caractérisation des matrices des préférences nettes et méthodes d'agrégation associées. *Mathématiques et sciences humaines*, 97:5–17, 1987.

M. Dufwenberg and M. Stegeman. Existence and uniqueness of maximal reductions under iterated strict dominance. *Econometrica*, 70 (5):2007–2023, 2002.

M. Dufwenberg, H. Norde, H. Reijnierse, and S. Tijs. The consistency principle for set-valued solutions and a new direction for normative game theory. *Mathematical Methods of Operations Research*, 54(1):119–131, 2001.

J. Duggan. Uncovered sets. *Social Choice and Welfare*, 2012. Forthcoming.

J. Duggan and M. Le Breton. Dutta's minimal covering set and Shapley's saddles. *Journal of Economic Theory*, 70:257–265, 1996a.

J. Duggan and M. Le Breton. Dominance-based solutions for strategic form games. Mimeo, 1996b.

J. Duggan and M. Le Breton. Mixed refinements of Shapley's saddles and weak tournaments. *Social Choice and Welfare*, 18(1):65–78, 2001.

P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77:321–357, 1995.

P. E. Dunne. Computational properties of argumentation systems satisfying graph-theoretic constraints. *Artificial Intelligence*, 171(10-15):701–729, 2007.

B. Dutta. Covering sets and a new Condorcet choice correspondence. *Journal of Economic Theory*, 44:63–80, 1988.

B. Dutta. On the tournament equilibrium set. *Social Choice and Welfare*, 7(4):381–383, 1990.

B. Dutta and J.-F. Laslier. Comparison functions and choice correspondences. *Social Choice and Welfare*, 16(4):513–532, 1999.

J. Edmonds. Paths, trees and flowers. *Canadian Journal of Mathematics*, 17:449–467, 1965.

J. Elster and J. Roemer. *Interpersonal Comparisons of Well-Being*. Cambridge University Press, 1991.

A. Feldman. Manipulation and the Pareto rule. *Journal of Economic Theory*, 21:473–482, 1979.

P. C. Fishburn. Even-chance lotteries in social choice theory. *Theory and Decision*, 3:18–40, 1972.

P. C. Fishburn. *The Theory of Social Choice*. Princeton University Press, 1973.

P. C. Fishburn. Condorcet social choice functions. *SIAM Journal on Applied Mathematics*, 33(3):469–489, 1977.

P. C. Fishburn. Non-cooperative stochastic dominance games. *International Journal of Game Theory*, 7(1):51–61, 1978.

P. C. Fishburn. Probabilistic social choice based on simple voting comparisons. *Review of Economic Studies*, 51(167):683–692, 1984.

D. C. Fisher and J. Ryan. Tournament games and positive tournaments. *Journal of Graph Theory*, 19(2):217–236, 1995.

W. Gaertner. *A Primer in Social Choice Theory*. LSE Perspectives in Economic Analysis. Oxford University Press, 2006.

P. Gärdenfors. Manipulation of social choice functions. *Journal of Economic Theory*, 13(2):217–228, 1976.

P. Gärdenfors. On definitions of manipulation of social choice functions. In J. J. Laffont, editor, *Aggregation and Revelation of Preferences*. North-Holland, 1979.

A. Gibbard. Manipulation of voting schemes. *Econometrica*, 41:587–602, 1973.

I. Gilboa, E. Kalai, and E. Zemel. On the order of eliminating dominated strategies. *Operations Research Letters*, 9:85–89, 1990.

I. Gilboa, E. Kalai, and E. Zemel. The complexity of eliminating dominated strategies. *Mathematics of Operations Research*, 18(3):553–565, 1993.

I. J. Good. A note on Condorcet sets. *Public Choice*, 10:97–101, 1971.

K. A. Hansen, P. B. Miltersen, and T. B. Sørensen. The computational complexity of trembling hand perfection and other equilibrium refinements. In S. Kontogiannis, E. Koutsoupias, and P. Spirakis, editors, *Proceedings of the 3rd International Symposium on Algorithmic Game Theory (SAGT)*, volume 6386 of *Lecture Notes in Computer Science (LNCS)*, pages 198–209. Springer-Verlag, 2010.

J. C. Harsanyi. Oddness of the number of equilibrium points: A new proof. *International Journal of Game Theory*, 2:235–250, 1973.

J. C. Harsanyi and R. Selten. *A General Theory of Equilibrium Selection in Games*. MIT Press, 1988.

N. Hazon, Y. Aumann, S. Kraus, and M. Wooldridge. Evaluation of election outcomes under uncertainty. In *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 959–966, 2008.

E. Hemaspaandra, L. Hemaspaandra, and J. Rothe. Exact analysis of Dodgson elections: Lewis Carroll's 1876 voting system is complete for parallel access to NP. *Journal of the ACM*, 44(6):806–825, 1997.

N. Houy. Still more on the tournament equilibrium set. *Social Choice and Welfare*, 32:93–99, 2009a.

N. Houy. A few new results on TEQ. Mimeo, 2009b.

O. Hudry. A note on "Banks winners in tournaments are difficult to recognize" by G. J. Woeginger. *Social Choice and Welfare*, 23:113–114, 2004.

S. Hurkens. *Games, rules and solutions*. PhD thesis, Tilburg University, 1995.

J. L. Jimeno, J. Pérez, and E. García. An extension of the Moulin No Show Paradox for voting correspondences. *Social Choice and Welfare*, 33(3):343–459, 2009.

T. Jungbauer and K. Ritzberger. Strategic games beyond expected utility. *Economic Theory*, 48(2–3):377–398, 2011.

M. Kalech, S. Kraus, G. A. Kaminka, and C. V. Goldman. Practical voting rules with partial information. In *Proceedings of the 10th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 151–182, 2011.

J. S. Kelly. Strategy-proofness and social choice functions without single-valuedness. *Econometrica*, 45(2):439–446, 1977.

W. Kern and D. Paulusma. The computational complexity of the elimination problem in generalized sports competitions. *Discrete Optimization*, 1(2):205–214, 2004.

K. Konczak and J. Lang. Voting procedures with incomplete preferences. In *Proceedings of the Multidisciplinary Workshop on Advances in Preference Handling*, pages 124–129. 2005.

G. Laffond, J.-F. Laslier, and M. Le Breton. More on the tournament equilibrium set. *Mathématiques et sciences humaines*, 31(123):37–44, 1993a.

G. Laffond, J.-F. Laslier, and M. Le Breton. The bipartisan set of a tournament game. *Games and Economic Behavior*, 5:182–201, 1993b.

G. Laffond, J. Lainé, and J.-F. Laslier. Composition-consistent tournament solutions and social choice functions. *Social Choice and Welfare*, 13:75–93, 1996.

J. Lang, M. S. Pini, F. Rossi, D. Salvagnin, K. B. Venable, and T. Walsh. Winner determination in voting trees with incomplete preferences and weighted votes. *Journal of Autonomous Agents and Multi-Agent Systems*, 25(1):130–157, 2012.

J.-F. Laslier. *Tournament Solutions and Majority Voting*. Springer-Verlag, 1997.

M. Le Breton. On the uniqueness of equilibrium in symmetric two-player zero-sum games with integer payoffs. *Économie publique*, 17 (2):187–195, 2005.

R. D. Luce and H. Raiffa. *Games and Decisions: Introduction and Critical Survey*. John Wiley & Sons Inc., 1957.

L. M. Marx and J. M. Swinkels. Order independence for iterated weak dominance. *Games and Economic Behavior*, 18:219–245, 1997.

A. Mas-Colell and H. Sonnenschein. General possibility theorems for group decisions. *Review of Economic Studies*, 39(2):185–192, 1972.

A. Mas-Colell, M. D. Whinston, and J. R. Green. *Microeconomic Theory*. Oxford University Press, 1995.

E. Maskin. Nash equilibrium and welfare optimality. *Review of Economic Studies*, 66(26):23–38, 1999.

R. D. McKelvey. Covering, dominance, and institution-free properties of social choice. *American Journal of Political Science*, 30(2):283, 1986.

R. D. McKelvey and P. C. Ordeshook. Symmetric spatial games without majority rule equilibria. *The American Political Science Review*, 70(4):1172–1184, 1976.

H. Moulin. *Axioms of Cooperative Decision Making*. Cambridge University Press, 1988a.

H. Moulin. Condorcet's principle implies the no show paradox. *Journal of Economic Theory*, 45:53–64, 1988b.

H. Moulin. *Fair Division and Collective Welfare*. The MIT Press, 2003.

R. B. Myerson. *Game Theory: Analysis of Conflict*. Harvard University Press, 1991.

J. F. Nash. Non-cooperative games. *Annals of Mathematics*, 54(2):286–295, 1951.

H. Norde. Bimatrix games have quasi-strict equilibria. *Mathematical Programming*, 85:35–49, 1999.

S. Obraztsova and E. Elkind. On the complexity of voting manipulation under randomized tie-breaking. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI)*, pages 319–324. AAAI Press, 2011.

M. J. Osborne and A. Rubinstein. *A Course in Game Theory*. MIT Press, 1994.

G. Owen. *Game Theory*. Academic Press, 2nd edition, 1982.

C. H. Papadimitriou. *Computational Complexity*. Addison-Wesley, 1994.

D. Parkes and L. Xia. A complexity-of-strategic-behavior comparison between Schulze's rule and ranked pairs. In *Proceedings of the 26th AAAI Conference on Artificial Intelligence (AAAI)*, pages 1429–1435. AAAI Press, 2012.

D.G. Pearce. Rationalizable strategic behavior and the problem of perfection. *Econometrica*, 52(4):1029–1050, 1984.

A. Perea. A one-person doxastic characterization of Nash strategies. *Synthese*, 158:251–271, 2007.

J. Pérez. The strong no show paradoxes are a common flaw in Condorcet voting correspondences. *Social Choice and Welfare*, 18(3):601–616, 2001.

J. E. Peris and B. Subiza. Condorcet choice correspondences for weak tournaments. *Social Choice and Welfare*, 16(2):217–231, 1999.

M. S. Pini, F. Rossi, K. B. Venable, and T. Walsh. Dealing with incomplete agents' preferences and an uncertain agenda in group decision making via sequential majority voting. In *Proceedings of the 11th International Conference on Principles of Knowledge Representation and Reasoning (KR)*, pages 571–578. AAAI Press, 2008.

M. S. Pini, F. Rossi, K. B. Venable, and T. Walsh. Possible and necessary winners in voting trees: Majority graphs vs. profiles. In *Proceedings of the 10th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 311–318, 2011.

C. R. Plott. Axiomatic social choice theory: An overview and interpretation. *American Journal of Political Science*, 20(3):511–596, 1976.

J. Richelson. Conditions on social choice functions. *Public Choice*, 31 (1):79–110, 1977.

W. H. Riker. *The art of political manipulation*. Yale University Press, 1986.

K. W. S. Roberts. Interpersonal comparability and social choice theory. *Review of Economic Studies*, 47(2):421–439, 1980.

J. Rothe, H. Spakowski, and J. Vogel. Exact complexity of the winner problem for Young elections. *Theory of Computing Systems*, 36(4): 375–386, 2003.

L. Samuelson. Dominated strategies and common knowledge. *Games and Economic Behavior*, 4:284–313, 1992.

M. R. Sanver and W. S. Zwicker. Monotonicity properties and their adaption to irresolute social choice rules. *Social Choice and Welfare*, 39(2–3):371–398, 2012.

M. A. Satterthwaite. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10:187–217, 1975.

L. J. Savage. *The Foundations of Statistics*. Wiley Publications in Statistics. Wiley and Sons, 1954.

A. Schrijver. *Combinatorial Optimization - Polyhedra and Efficiency*. Springer, 2003.

T. Schwartz. On the possibility of rational policy evaluation. *Theory and Decision*, 1(1):89–106, 1970.

T. Schwartz. *The Logic of Collective Choice*. Columbia University Press, 1986.

T. Schwartz. Cyclic tournaments and cooperative majority voting: A solution. *Social Choice and Welfare*, 7:19–29, 1990.

L. Shapley. Order matrices. I. Technical Report RM-1142, The RAND Corporation, 1953a.

L. Shapley. Order matrices. II. Technical Report RM-1145, The RAND Corporation, 1953b.

L. Shapley. Some topics in two-person games. In M. Dresher, L. S. Shapley, and A. W. Tucker, editors, *Advances in Game Theory*, volume 52 of *Annals of Mathematics Studies*, pages 1–29. Princeton University Press, 1964.

J. H. Smith. Aggregation of preferences with variable electorate. *Econometrica*, 41(6):1027–1041, 1973.

R. Tarjan. Depth-first search and linear graph algorithms. *SIAM Journal on Computing*, 1(2):146–160, 1972.

A. D. Taylor. *Social Choice and the Mathematics of Manipulation*. Cambridge University Press, 2005.

T. N. Tideman. Independence of clones as a criterion for voting rules. *Social Choice and Welfare*, 4(3):185–206, 1987.

E. van Damme. *Refinements of the Nash Equilibrium Concept*. Springer-Verlag, 1983.

J. von Neumann. Zur Theorie der Gesellschaftspiele. *Mathematische Annalen*, 100:295–320, 1928.

J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 1944.

K. Wagner. More complicated questions about maxima and minima, and some closures of NP. *Theoretical Computer Science*, 51:53–80, 1987.

T. Walsh. Uncertainty in preference elicitation and aggregation. In *Proceedings of the 22nd AAAI Conference on Artificial Intelligence (AAAI)*, pages 3–8. AAAI Press, 2007.

R. B. Wilson. The finer structure of revealed preference. *Journal of Economic Theory*, 2(4):348–353, 1970.

G. J. Woeginger. Banks winners in tournaments are difficult to recognize. *Social Choice and Welfare*, 20:523–528, 2003.

L. Xia and V. Conitzer. Determining possible and necessary winners under common voting rules given partial orders. *Journal of Artificial Intelligence Research*, 41:25–67, 2011.

L. Xia, M. Zuckerman, A. D. Procaccia, V. Conitzer, and J. S. Rosenschein. Complexity of unweighted coalitional manipulation under some common voting rules. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI)*, pages 348–353. AAAI Press, 2009.

L. Xia, J. Lang, and J. Monnot. Possible winners when new alternatives join: New results coming up! In *Proceedings of the 10th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 829–836, 2011.

H. P. Young. An axiomatization of Borda's rule. *Journal of Economic Theory*, 9:43–52, 1974.

T. M. Zavist and T. N. Tideman. Complete independence of clones in the ranked pairs rule. *Social Choice and Welfare*, 6(2):167–173, 1989.

R. Zeckhauser. Voting systems, honest preferences, and Pareto optimality. *American Political Science Review*, 67(3):934–946, 1973.

W. S. Zwicker. The voter's paradox, spin, and the Borda count. *Mathematical Social Sciences*, 22:187–227, 1991.