An Axiomatic and Computational Analysis of Altruism, Fairness, and Stability in Coalition Formation Games

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Anna Maria Kerkmann
aus Düsseldorf

Düsseldorf, Mai 2022

aus dem Institut für Informatik
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Berichterstatter:

1. Prof. Dr. Jörg Rothe

2. Jun.-Prof. Dr. Robert Bredereck

Tag der mündlichen Prüfung:

Acknowledgements

First of all, I would like to thank my supervisor Jörg Rothe who has introduced me to the world of scientific research and guided me through my whole doctorate time. Without his guidance and constant encouragement, the publication of so many papers and the writing of this thesis would not have been possible. Also, Jörg's effort and expertise have enabled me to become a better researcher. Next, I would like to express my gratitude to Dorothea Baumeister. I'm thankful for her ongoing support, her willingness to listen to all my problems, and all her constructive and detailed feedback. Moreover, I am much obliged to my co-authors who have spared no effort to improve our joint research.

Shortly after the start of my doctoral time I was given the great opportunity to join the "NRW-Forschungskolleg Online-Partizipation," a PhD-program which is supported by the North Rhine-Westphalian funding scheme "Forschungskollegs." I very much appreciate my time as a member of the program, during which I was able to profit from its interdisciplinary scope. I thank the coordinators for their guidance and all other PhD students from the program for giving me interesting insights into their research.

Many thanks go to my colleagues who enriched my time as a doctoral student. I very much enjoyed the supportive atmosphere of our group. And I feel lucky that I was able to share an office with some of my colleagues for quite some time before the COVID-19 pandemic.

My family and friends deserve much thanks for their ongoing love and support. Particularly, I would like to thank my father for sparking my interest in computer science, my mother for always being there for me and helping me to reach all my goals, and my siblings, David and Kathi, for playing dozens of card games with me. Above all, I owe my deepest gratitude to Marcel for making my whole life so enjoyable. Thank you for all your love and affection, your faith in my abilities, and for going on so many fun adventures with me. Last but not least, I would like to thank my friends for proofreading parts of my thesis, inviting me to many exciting climbing trips, providing me with amusing riddles, and spending numerous great days together.

Zusammenfassung

Diese Arbeit beschäftigt sich mit Koalitionsbildungsspielen, welche zum Forschungsbereich der kooperativen Spieltheorie gehören. Bei diesen Spielen geht es darum, wie sich Spieler auf Grundlage ihrer individuellen Präferenzen in Gruppen, auch Koalitionen genannt, aufteilen. Bei unseren Untersuchungen konzentrieren wir uns größtenteils auf hedonische Koalitionsbildungsspiele, kurz hedonische Spiele, bei welchen vorausgesetzt wird, dass die Präferenzen der Spieler nur von ihren eigenen Koalitionen abhängen. Ein zentrales Thema in Bezug auf diese Spiele ist die Suche nach sinnvollen Formaten zur Abgabe der Präferenzen. Diese Formate sollten einfach zu erheben, möglichst ausdrucksstark und zugleich kompakt darstellbar sein. In der einschlägigen Literatur wurden bereits einige solcher Formate vorgestellt, die auch wir in dieser Arbeit behandeln werden. Ein zweiter wichtiger Punkt bei der Erforschung von hedonischen Spielen ist die Untersuchung von Stabilität, Fairness und Optimalität. Klassische Stabilitätskonzepte behandeln beispielsweise die Frage, ob einzelne Spieler oder Gruppen von Spielern einen Anreiz haben, von ihren Koalitionen abzuweichen. Zu den bekanntesten solcher Konzepte gehören Nash-Stabilität und Kernstabilität.

Auf Grundlage des aktuellen Stands der Literatur führen wir in dieser Arbeit neue Modelle für (hedonische) Koalitionsbildungsspiele ein und untersuchen diese im Hinblick auf axiomatische Eigenschaften, Stabilität, Fairness und Optimalität. Dabei spielen insbesondere Untersuchungen der Berechnungskomplexität eine wichtige Rolle.

Zuerst stellen wir verschiedene Modelle für Altruismus in Koalitionsbildungsspielen vor. Wir konzentrieren uns dabei zunächst auf den Kontext von hedonischen Spielen und erweitern die Modelle anschließend auf allgemeinere Koalitionsbildungsspiele, bei denen eine weitreichendere Form des Altruismus' möglich ist. Wir untersuchen unsere Modelle axiomatisch und vergleichen diese dabei untereinander und mit anderen Modellen. Zudem analysieren wir die Entscheidungsprobleme, die sich bei der Betrachtung klassischer Stabilitätskonzepte im Kontext von altruistischen Spielen ergeben, in Hinblick auf ihre Berechnungskomplexität.

Anschließend definieren wir drei schwellwertbasierte Fairnessbegriffe für hedonische Spiele. Diese werden in den Kontext einschlägiger Stabilitäts- und Fairnesskonzepte eingeordnet und im Hinblick auf ihre Berechnungskomplexität erforscht. Außerdem untersuchen wir den Einfluss, den unsere Fairnesskonzepte auf die soziale Wohlfahrt haben.

Schließlich führen wir ein weiteres Präferenzformat ein, bei dem die Spieler zwischen Freunden, neutralen Spielern und Feinden unterscheiden. Sie geben dementsprechend eine dreigeteilte schwache Ordnung ab. Da die Präferenzen, welche sich aus diesen Ordnungen ableiten lassen, nicht vollständig sein müssen, unterscheiden wir in den entstehenden Spielen zwischen möglicher und notwendiger Stabilität. Auch hier führen wir eine Komplexitätsanalyse der Probleme durch, die sich bezüglich bekannter Stabilitätskonzepte ergeben.

Abstract

This thesis deals with coalition formation games, which belong to the research area of cooperative game theory. In these games, players divide into groups, also called coalitions, based on their individual preferences. In our research, we mainly focus on hedonic coalition formation games, hedonic games for short, in which players' preferences are assumed to depend only on the coalitions containing themselves. A central problem in hedonic games research is finding reasonable formats for the elicitation of preferences. These preference representations should be easy to elicit, reasonably expressive, and succinct. Many such formats have already been presented in related literature, some of which we will also discuss in this thesis. A second central point in research concerning hedonic games is the investigation of stability, fairness, and optimality. For instance, common stability concepts deal with the question of whether individual players or groups of players might have an incentive to deviate from their current coalitions. Among those notions are, for example, Nash and core stability.

Based on the current state of research, we introduce new models for (hedonic) coalition formation games and investigate them with respect to axiomatic properties, stability, fairness, and optimality. In particular, investigations of the computational complexity of the associated decision problems play an important role.

We start with introducing several models for altruism in coalition formation games. First, we focus on the context of hedonic games and then extend the models to more general coalition formation games, where a broader form of altruism is possible. We conduct an axiomatic analysis of our models and compare them to related models and to each other. In addition, we study the problems, that arise when considering classical stability concepts in the context of altruistic coalition formation games, with respect to their computational complexity.

Subsequently, we define three threshold-based fairness notions for hedonic games. These notions are considered local fairness notions in the sense that the agents only have to inspect their own coalitions to decide whether a coalition structure is fair to them. We study the relations of these notions to other common stability and fairness concepts and examine them with respect to their computational complexity. Furthermore, we investigate the price of local fairness, i.e., the impact that our fairness concepts have on the social welfare.

Finally, we introduce another preference format in which players distinguish between friends, neutral players, and enemies. Accordingly, they cast their preferences by submitting a weak rankings that is separated by two thresholds. Since the preferences that can be derived from these rankings are not necessarily complete, we distinguish between possible and necessary stability in the resulting games. Again, we perform a computational complexity analysis of the problems that arise with respect to common stability concepts.

Contents

Chapter 1

Introduction

Nowadays, there is an enormous range of research that is concerned with topics of artificial intelligence (AI). In fact, AI research is not only about mimicking human intelligence but also about a variety of solution concepts that apply knowledge from different fields of science including not only natural sciences such as biology and physics but also social sciences such as sociology and economics. Two fields that have gained major interest on AI conferences are *multiagent systems* and *game theory*. While the research concerning these areas is very broad, there is not always a clear distinction between them.

Research concerning multiagent systems mainly deals with distributed problem solving, i.e, the cooperation of agents whose aim is to collectively solve some problem. Those systems are often applied to problems that might be more difficult or not at all solvable for a single agent. Inspiring examples of very successful multiagent systems can be found in nature: Ant colonies use their communication abilities and division of labor to master complex problems that would never be feasible for a single ant.

Game theory deals with the interaction among individual agents which are mostly assumed to be selfishly pursuing their own goals. Research in this area roughly started with the works due to Borel [22], Neumann [103], and Neumann and Morgenstern [104] and is commonly divided into *noncooperative* and *cooperative* game theory. While noncooperative game theory focuses on the preferences and actions of individual agents, cooperative game theory also sees individual preferences but rather focuses on the formation of groups and allows them to take joint actions. Examples of noncooperative game theory include the famous *prisoners' dilemma* [118], the *Monty Hall problem* (see, e.g., Selvin [131, 132] or the German book by Randow [120]), but also classic combinatorial games such as tic-tac-toe, nim, chess, go, or sudoku. The focus of noncooperative game theory is mainly on studying equilibria, i.e., stable states where no agent has a reason to deviate from her current strategy. In cooperative games, agents may form coalitions and take joint actions. For more background on multiagent systems and game theory see, e.g., the textbooks by Shoham and Leyton-Brown [133] and by Rothe [126].

The focus of this thesis is on *coalition formation games* which are a key topic in cooperative game theory. Their applications range from technical, engineering, and economic problems to social and even political problems. Drèze and Greenberg [53] initiated the study of coalition formation games with *hedonic* preferences. These games were later formalized by Banerjee et al. [15] and Bogomolnaia and Jackson [21]. The key idea of such games is that agents have to form partitions while only caring about the coalitions that they are part of. In the general framework of hedonic games, the agents have arbitrary preferences over all coalitions containing themselves. Yet, it is not reasonable to elicit rankings over all such coalitions in practice. Rather, reasonable preference representations are needed. Ideally, such formats should be succinct, expressive, and easy to elicit. The determination of reasonable preference representations has been a fundamental part of hedonic games research. Well-established representations include cardinal formats such as the *additive encoding* due to Bogomolnaia and Jackson [21] and the *fractional encoding* by Aziz et al. [11]. Other formats are based on the partitioning of the agents into friends and enemies [50, 111, 17] or on the usage of propositional formulas [56, 9].

Another crucial branch of hedonic games research addresses problems related to notions of *stability*, *fairness*, and *optimality*. The determination of such notions constitutes a major part of past research. In particular, several common notions of stability deal with single player deviations. For instance, a partition of the agents in a hedonic game is said to be *Nash stable* (or in a Nash equilibrium) if no agent wants to deviate to another coalition of the partition [21]. Other stability notions concern the deviation of groups. *Core stability* is probably the most important notion of group stability in hedonic games (see, e.g., the early paper by Banerjee et al. [15] and the survey by Woeginger [147]). Informally, a group of players blocks a given partition of the players with respect to the notion of core stability if all players in this group prefer it to the groups assigned by the partition. A partition is core stable (or *in the core*) if there is no blocking coalition [15]. Relevant notions of optimality include *Pareto-optimality*, *popularity*, and the maximization of *utilitarian* or *egalitarian social welfare*. Interesting notions of fairness, for instance *envy-freeness*, have been adopted from the research of fair division and resource allocation (see Foley [61] and the book chapters by Bouveret et al. [24] and by Lang and Rothe [95] for background on these topics).

Given such notions of stability, optimality, or fairness, we are interested in the identification of sufficient conditions for such notions, i.e., we ask which properties guarantee the stability, fairness, or optimality of outcomes. Also, stable, fair, or optimal outcomes might not even exist for certain preference profiles. A decent amount of research has been focusing on identifying properties that guarantee the existence of such outcomes. For instance, Bogomolnaia and Jackson [21] showed that Nash stable coalition structure are guaranteed to exist in symmetric additively separable hedonic games. Yet, it was later shown that deciding whether a Nash stable coalition structure exists in an (asymmetric) additively separable hedonic game is NP-complete [136]. Determining the complexity of such existence problems has generally been an important research aspect. For core stability and strict core stability, the existence problem has been proved to be Σ_2^p-complete for additively separable hedonic games [146, 116, 111]. Yet, there again exist conditions that simplify the existence problem. Burani and Zwicker [35] have shown that there always exist core stable outcomes for symmetric additively separable hedonic games with *purely cardinal* preferences. Dimitrov et al. [50] proved that the existence problem is trivial for friend-oriented and enemy-oriented hedonic games.

In this thesis, we build on the current state of research and introduce further succinct preference representations. Tackling the problem of finding an expressive, compact, and easy to elicit preference format, we introduce *weak rankings with double thresholds*. These rankings are more expressive than purely ordinal rankings (the *individually rational encoding* [14]), the *friends-and-enemies encoding* [50], or the *singleton encoding* [40]. Yet, our format is cognitively plausible and easy to elicit from the agents — probably easier than, for example, propositional formulas (as in the case of *hedonic coalition nets* [56] or the *boolean hedonic encoding* [9]). Furthermore, we do not make strong assumptions on the nature of the preferences (as the *anonymous encoding* [15] which only takes coalition sizes into account or the boolean hedonic encoding which was designed for dichotomous preferences) and our format is succinct. In conclusion, our format provides a satisfactory balance between the three requirements.

A second important aspect that we tackle in this thesis leads to a new branch of preference modeling. Since the beginnings of game theory, agents were commonly considered as completely rational and self-interested individuals (see Neumann and Morgenstern [104]). We challenge this assumption and aim for a more realistic representation of real-world coalition formation scenarios: We introduce *altruism* into coalition formation games. In our models, agents are not narrowly selfish but take the opinions of their friends into account when comparing different coalition structures. We present a variety of altruistic models and compare them with regard to their axiomatic properties. After concentrating on hedonic models, we also introduce models of altruism that drop the hedonic restriction. The changes we make for these nonhedonic models bring some axiomatic improvements and, in our opinion, an even more realistic model of altruism.

A third part of this thesis is concerned with fairness in hedonic games. Previous literature considers *envy-freeness* as a notion of fairness [21, 10, 148, 114]. Yet, to verify this notion, agents have to inspect the coalitions of other agents. To some extend, this is in conflict with the hedonic assumption which states that agents only care about their own coalitions. Furthermore, we want to avoid the need to compare large numbers of coalitions. Hence, we introduce three notions of *local fairness* that can be decided solely based on the agents' own coalitions and their preferences.

Besides these conceptual contributions, this thesis also contains several technical contributions. We investigate the *FEN-hedonic games* that result from lifting weak rankings with double thresholds to preferences over coalitions. We characterize stability in these games and study the problems of verifying stable coalition structures and of checking their existence. Furthermore, we not only axiomatically study altruistic coalition formation games but also provide elaborate computational analyses of the associated stability verification and existence problems. Our results cover many common stability notions such as Nash stability, core stability, Pareto optimality, and popularity. Concerning our notions of local fairness, we determine the complexity of computing local fairness thresholds and deciding whether locally fair coalitions structures exist for additively separable hedonic games. Moreover, we study the price of our local fairness notions.

1.1 Outline

In Chapter 2, we provide the required background for this thesis and explain all concepts that are needed to comprehend the following chapters. The provided background includes an introduction to computational complexity in Section 2.1, a brief overview of graph theory in Section 2.2, and a survey of the relevant aspects of coalition formation games in Section 2.3. This survey contains not only basic definitions and observations but also references to related work. The major part of our research starts in Chapter 3 where we study different aspects of altruism in coalition formation games. More precisely, Chapter 3 divides into three sections. First, we explore altruistic hedonic games in Section 3.1. After introducing such games, we conduct an axiomatic analysis of our altruistic models and investigate the problems of verifying stable outcomes and of deciding whether stable outcomes exist in such games. In Section 3.2, we further analyze altruistic hedonic games while concentrating on the notions of popularity and strict popularity. Subsequently, we study altruism in a more general scope of coalition formation games. In particular, Section 3.3 expounds an altruistic coalition formation model which is not restricted to hedonic preferences but allows for a more far-reaching altruistic behavior. We identify some advantages that this extended model offers compared to the altruistic hedonic model and study stability in these games. In Chapter 4, we continue with research concerning notions of local fairness in hedonic games. After proposing three such notions, we relate them to other popular notions of stability, determine the computational complexity of the associated decision problems, and study the price of our local fairness notions. In Chapter 5, we introduce and study FEN-hedonic games where agents divide the other agents into friends, enemies, and neutral players while additionally ranking their friends and enemies respectively. We then investigate problems concerning the verification and existence of possibly or necessarily stable coalition structures. We conclude with Chapter 6 where we recap this thesis, highlight some important contributions, and identify some possible directions for future research.

Chapter 2

Background

In this chapter, we provide background information for all subjects studied in the following chapters. We illustrate the essential models and notions that are important for understanding this thesis. We start with an introduction to computational complexity theory in Section 2.1. In Section 2.2, we provide a brief introduction to graph theory. Furthermore, we give an insight into coalition formation in Section 2.3. For literature on the more general topic of cooperative game theory, see the textbooks by Chalkiadakis et al. [41], Shoham and Leyton-Brown [133], Peleg and Sudhölter [113], or the book chapters by Elkind and Rothe [55] and Elkind et al. [57].

2.1 Computational Complexity

A main part of this thesis will be the study of different decision problems and the determination of their computational complexity. But what is a decision problem, how do we measure its complexity, and what does it mean that a problem is 'hard' or 'easy'? We will answer these and other questions in the following section and give a short introduction to computational complexity theory. For more background on this topic we refer to the textbooks by Rothe [125, 128], Papadimitriou [112], and Arora and Barak [4].

2.1.1 Computational Problems

The objective of computational complexity theory is to classify *computational problems* based on their difficulty. In general, a computational problem can be any kind of problem that could be solved by a computer. There are different types of computational problems such as *decision problems*, *optimization problems*, and *search problems*. In this section, we will concentrate on decision problems which are basically questions that can be answered either by yes or no. We will represent any decision problem by specifying its name, an input format, and a question concerning the input. One of the most important decision problems in computational complexity theory is the boolean satisfiability problem (SAT) [66]:

SATISFIABILITY (SAT)	
Given:	A boolean formula φ in conjunctive normal form.
Question:	Is there a truth assignment for the variables in φ that satisfies φ?

Now, given any decision problem, any concrete input that satisfies the specified input requirements is called an *instance* of the problem. An instance is a *yes-instance* if and only if the answer to the specified question is 'yes' for this instance. Otherwise, the instance is called *no-instance.* Decision problems can also be represented by the set of their yes-instances. For instance, SAT can be written as

$$\text{SAT} = \{\varphi \mid \varphi \text{ is a satisfiable boolean formula in conjunctive normal form }\}.$$

2.1.2 Algorithms, Runtimes, and Complexity Classes

In computer science, we use algorithms to solve problems.[1] Informally, a *deterministic algorithm* for problem A is a finite sequence of explicit instructions that, when executed for a given input I, outputs the answer to problem A for input I. Formally, algorithms can be modeled via *Turing machines* which were invented by Turing [141, 142] in 1936. We will not give a formal definition of Turing machines here but give some intuitive explanations instead. We refer to the textbooks by Rothe [125, 128] and Papadimitriou [112] for more background on Turing machines.

A Turing machine M that solves a problem A can be started with any instance I of the problem. Starting with an initial configuration that is based on the input I, the Turing machine M then does some computations which lead to further configurations. After a finite number of computation steps, M might reach a final configuration where it *accepts* the input I. The set of all inputs that M accepts is called the *language of M* and is denoted by $L(M)$. We say that *M accepts* the problem A if it accepts all its yes-instances and none of its no-instances, i.e., if $L(M) = A$.

We further distinguish between deterministic and nondeterministic Turing machines. *Deterministic Turing machines* (DTMs) represent deterministic algorithms and the computation of a DTM is a deterministic sequence of configurations. That means that its computation can be represented by a single unique path of configurations and it accepts the input exactly if it accepts the input on this one path. For a DTM M that accept language $L(M)$, we also say that it *decides* the problem $L(M)$. In contrast to that, *nondeterministic Turing machines* (NTMs) represent nondeterministic algorithms and can have more than one computation path. In the computation of a NTM, there can occur configurations for which the next computation step is not unique; rather, there might be multiple possible successor configurations. In this case,

[1] There are problems that are not solvable by algorithms, e.g., the halting problem, but we will only concentrate on solvable problems in this thesis.

the computation can be represented by a tree where every fork of the tree represents a nondeterministic situation. For a NTM M, we say that M accepts the input I if I is accepted on at least one path of the computation tree.

When developing algorithms[2] for a given problem, there will certainly be more than one possible solution. So the following questions arise: What is the best algorithm for the given problem? And how do we even compare two algorithms? In computational complexity theory, we compare algorithms based on their *computation times* (or *runtimes*) and *space requirements*. The runtime is measured by the number of elementary computational steps that are needed when executing the algorithm. The space requirements are measured by the size of the memory that is used while executing the algorithm. In this thesis, we will concentrate only on the runtime of algorithms.

Now, the goal of algorithmics is to find algorithms that have low runtimes. But of course, the runtime of an algorithm may vary based on the concrete input instance. For example, algorithm M might solve a given problem faster than algorithm N for a given instance while N solves the same problem faster than M for another instance. So, how do we compare the runtimes of these two algorithms? Common answers to this question are to compare either the *best-case*, *average-case*, or *worst-case runtimes*. In this thesis, we will concentrate on the latter. This means that we are looking for upper bounds on the runtimes of algorithms. Further, we always measure runtimes based on the size of the input (which is usually encoded in binary). We can then group problems into *complexity classes* which specify upper bounds on their *worst-case time complexity*, i.e., given a problem A, we ask for the maximal number of computation steps that the fastest algorithm solving A might need for any instance of A.

For any computable total function $f : \mathbb{N} \to \mathbb{N}$, we define

- $\mathrm{DTIME}(f)$ as the set of all problems A for which there exists a DTM M with $L(M) = A$ that decides any instance of A of size n in time at most $f(n)$; and
- $\mathrm{NTIME}(f)$ as the set of all problems A for which there exists an NTM M with $L(M) = A$ that accepts any yes-instance of A of size n in time at most $f(n)$.

As we are only interested in the asymptotic behavior of the worst-case runtime, we group complexity classes with similar functions together. In particular, we denote the set of all polynomial functions[3] by $\mathbb{P}\mathrm{ol}$ and define all problems that can be decided by a DTM (respectively NTM) in polynomial time as follows:

$$\mathrm{P} = \bigcup_{f \in \mathbb{P}\mathrm{ol}} \mathrm{DTIME}(f) \qquad \text{and} \qquad \mathrm{NP} = \bigcup_{f \in \mathbb{P}\mathrm{ol}} \mathrm{NTIME}(f).$$

As each DTM is also a NTM (that just does not make use of the nondeterminism), it directly follows from the definitions of P and NP that $\mathrm{P} \subseteq \mathrm{NP}$. However, it is still unknown

[2] Keep in mind that we use Turing machines as models of algorithms.

[3] A polynomial function f has the form $f(n) = \sum_{i=0}^{m} c_i \cdot n^i = c_m \cdot n^m + \ldots + c_2 \cdot n^2 + c_1 \cdot n + c_0$ where $n, m, c_m, \ldots, c_0 \in \mathbb{N}$.

whether $P \subset NP$ or $P = NP$ holds. This open question is known as the "P versus NP problem" [45] and is one of the seven Millennium Prize Problems [36] that were declared by the Clay Mathematics Institute in 2000. In this thesis, we will follow the common belief and assume that $P \subset NP$. Actually, if $P = NP$ would hold, many results from the literature would become meaningless. For examples of such results and further discussions on the P versus NP problem see, e.g., the book chapters by Rothe [127] and Arora and Barak [5].

For all problems in P we also say that they are *solvable in (deterministic) polynomial time* and that they can be solved *efficiently*. We sometimes also say that these problems are *easy*. For all problems that are not in P, i.e., not solvable in deterministic polynomial time, we say that they are *computationally intractable*.[4]

Additionally to the two well-known complexity classes P and NP, we will also consider the complexity class coNP which is the class of the complements of all problems in NP. It is defined as

$$\mathrm{coNP} = \{\bar{A} \mid A \in \mathrm{NP}\}$$

where, for any decision problem A, the *complement* of A is defined by

$$\bar{A} = \{I \mid I \text{ is a no-instance of } A\}.$$

Note that NP can analogously be defined as the set of all decision problems for which a yes-instance can be verified in deterministic polynomial time while coNP can be defined as the set of all decision problems for which a no-instance can be verified in deterministic polynomial time.

2.1.3 Polynomial-Time Many-One Reducibility and Hardness

We have just seen that complexity classes specify upper bounds on the complexity of the contained problems. This subsection will be on how to specify lower bounds on the complexity of problems. We use *reductions* to show that one problem is at least as complex (or *hard*) as another one.

We say that problem A is *polynomial-time many-one reducible* to problem B (denoted by $A \leq_m^p B$) if and only if there exists a polynomial-time computable total function f that maps instances of A to instances of B such that, for every instance I of A,

$$I \in A \Longleftrightarrow f(I) \in B,$$

i.e., I is a yes-instance of A if and only if $f(I)$ is a yes-instance of B. Note that the relation $\leq_m^p$ is reflexive (i.e., $A \leq_m^p A$ for any problem A) and transitive (i.e., $A \leq_m^p C$ for any problems A, B, and C with $A \leq_m^p B$ and $B \leq_m^p C$). Furthermore, a problem is said to be *hard* for a complexity

[4] Cobham [43] and Edmonds [54] were the first to identify the set of *tractable problems* with the class P (see the *Cobham-Edmonds thesis*).

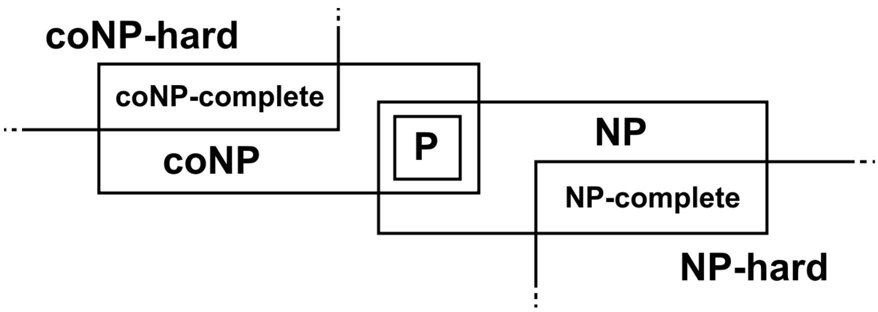


Figure 2.1: Assumed relations among the complexity classes P, NP, and coNP

class if it is at least as hard as every other problem in the class. Problems that are hard for a class and are also contained in it are called *complete* for the class. Formally, problem A is $\leq_m^p$-hard for a complexity class $\mathscr{C}$ if $B \leq_m^p A$ for every problem B in $\mathscr{C}$. We then also say that A is $\mathscr{C}$-hard. Moreover, problem A is $\leq_m^p$-complete for the complexity class $\mathscr{C}$ if $A \in \mathscr{C}$ and A is $\mathscr{C}$-hard. We then also say that A is $\mathscr{C}$-complete. The complexity classes P, NP, and coNP are closed under $\leq_m^p$-reducibility which means that for $\mathscr{C} \in \{\mathrm{P}, \mathrm{NP}, \mathrm{coNP}\}$ and any two problems A and B, $A \leq_m^p B$ and $B \in \mathscr{C}$ implies $A \in \mathscr{C}$.

Note that SAT was the first problem that was shown to be NP-complete. This was shown independently by Cook [44] in 1971 and by Levin [97] in 1973. While they had to use quite sophisticated constructions to show the hardness of SAT,[5] we will use the following helpful implications to proof the hardness of problems. They follow directly from the transitivity of $\leq_m^p$ and because P is closed under $\leq_m^p$:

- For any complexity class $\mathscr{C}$, if A is $\mathscr{C}$-hard and $A \leq_m^p B$ then B is $\mathscr{C}$-hard.
- If problem A is NP-complete or coNP-complete, then $A \in \mathrm{P}$ if and only if $\mathrm{P} = \mathrm{NP}$.

By the second implication and under the assumption that $\mathrm{P} \neq \mathrm{NP}$, there is no polynomial-time algorithm for any NP-complete or coNP-complete problem. Assuming that $\mathrm{P} \neq \mathrm{NP}$, $\mathrm{NP} \neq \mathrm{coNP}$, and $\mathrm{P} \neq \mathrm{NP} \cap \mathrm{coNP}$, the relations between the three complexity classes and the sets of NP-hard and coNP-hard problems can be illustrated as shown in Figure 2.1.

2.1.4 Some Hard Problems

The list of decision problems that have been shown to be NP-complete grows steadily since the Cook-Levin theorem [44, 97] was published. For instance, Karp [82] showed the NP-completeness of many problems and a collection of many hard problems can be found in the book by Garey and Johnson [66]. We will now present some decision problems that will be used in this thesis.

[5] You can find proofs of the *Cook-Levin theorem*, e.g., in Garey and Johnson [66, Section 2.6] or Rothe [125, Section 3.5.3].

One of the decision problems that Karp [82] showed to be NP-complete is EXACT COVER BY 3-SETS (X3C). It is defined as follows.

EXACT COVER BY 3-SETS (X3C)	
Given:	Integers $k \geq 2$ and $m \geq 2$, a set $B = \{b_1, \ldots, b_{3k}\}$, and a collection $\mathscr{S} = \{S_1, \ldots, S_m\}$ of 3-element subsets of B.
Question:	Is there an exact cover of B in $\mathscr{S}$, i.e., a subset $\mathscr{S}' \subseteq \mathscr{S}$ of size k such that every element of B occurs in exactly one set in $\mathscr{S}'$?

We will often make use of a restricted version of X3C. Gonzalez [68] showed that the problem remains NP-complete even when every element of the set occurs exactly three times in the 3-element subset collection.

RESTRICTED EXACT COVER BY 3-SETS (RX3C)	
Given:	An integer $k \geq 2$, a set $B = \{b_1, \ldots, b_{3k}\}$, and a collection $\mathscr{S} = \{S_1, \ldots, S_{3k}\}$ of 3-element subsets of B, where each element of B occurs in exactly three sets in $\mathscr{S}$.
Question:	Does there exist an exact cover of B in $\mathscr{S}$, i.e., a subset $\mathscr{S}' \subseteq \mathscr{S}$ of size k such that every element of B occurs in exactly one set in $\mathscr{S}'$?

We illustrate RX3C with the following example.

Example 2.1. Let $k = 3$, $B = \{1, \ldots, 9\}$, and $\mathscr{S} = \{S_1, \ldots, S_9\}$ with

$$\begin{aligned} S_1 &= \{1,2,3\}, & S_2 &= \{1,5,6\}, & S_3 &= \{1,5,9\}, \\ S_4 &= \{2,4,6\}, & S_5 &= \{2,7,8\}, & S_6 &= \{3,4,5\}, \\ S_7 &= \{3,7,8\}, & S_8 &= \{4,6,9\}, & S_9 &= \{7,8,9\}. \end{aligned}$$

Then, the question is whether there is a subset $\mathscr{S}'$ of $\mathscr{S}$ of size 3 such that each element of B occurs exactly one time in $\mathscr{S}'$. In fact, there exists such a subset, namely $\mathscr{S}' = \{S_3, S_4, S_7\} = \{\{1,5,9\}, \{2,4,6\}, \{3,7,8\}\}$. Hence, the given instance is a yes-instance of RX3C.

We now turn to the following graph problem that was shown to be NP-complete by Karp [82]. Also note that we will give some more background on graph theory in Section 2.2.

CLIQUE	
Given:	An integer $k \geq 1$ and an undirected graph $G = (V, E)$.
Question:	Is there a clique of size k in G, i.e., a subset $V' \subseteq V$ of the vertices such that all vertices in V' are pairwise connected?

2.1.5 Beyond P and NP

P, NP, and coNP are not the only complexity classes out there. In fact, there are various hierarchies of complexity classes beyond NP. For instance, Meyer and Stockmeyer [99] and Stockmeyer [134] introduced the *polynomial hierarchy* which makes use of *oracle Turing machines*. For two complexity classes $\mathscr{C}$ and $\mathscr{D}$, the class $\mathscr{C}^{\mathscr{D}}$ contains all problems that can be solved by an algorithm according to class $\mathscr{C}$ that additionally has access to an *oracle* which verifies instances of a set $D \in \mathscr{D}$ in a single computation step. The polynomial hierarchy is defined inductively by $\Delta_0^p = \Sigma_0^p = \Pi_0^p = \mathrm{P}$ and, for $i \geq 0$,

$$\begin{aligned}
\Delta_{i+1}^p &= \mathrm{P}^{\Sigma_i^p}, \\
\Sigma_{i+1}^p &= \mathrm{NP}^{\Sigma_i^p}, \text{ and} \\
\Pi_{i+1}^p &= \mathrm{co}\Sigma_{\mathrm{i+1}}^{\mathrm{p}}.
\end{aligned}$$

Additionally, $\mathrm{PH} = \bigcup_{i \geq 0} \Sigma_i^p$. For the first layer of the polynomial hierarchy, the definitions imply that $\Delta_1^p = \mathrm{P}$, $\Sigma_1^p = \mathrm{NP}$, and $\Pi_1^p = \mathrm{coNP}$. The second layer is given by $\Delta_2^p = \mathrm{P}^{\mathrm{NP}}$, $\Sigma_2^p = \mathrm{NP}^{\mathrm{NP}}$, and $\Pi_2^p = \mathrm{coNP}^{\mathrm{NP}}$. For more details on the polynomial hierarchy, the reader is referred to, e.g., the textbooks by Rothe [125, 128].

There are many further interesting aspects of complexity theory such as *parameterized complexity* (see, e.g., the books by Downey and Fellows [52, 51] and Flum and Grohe [60]) and *probabilistic complexity* (see Gill [67] and, e.g., Balcázar et al. [13]).

2.2 Graph Theory

We now give some basics of graph theory. For an extensive introduction to graph theory, see, e.g., the textbooks by West [144] and Gurski et al. [70].

Formally, a *graph* is a pair $G = (V, E)$ where V is a set of vertices (or *nodes*) and E is a set of edges. In the case of an *undirected graph*, the edges are undirected and we have $E \subseteq \{\{u, v\} \mid u, v \in V, u \neq v\}$. In the case of a *directed graph*, the edges are directed and we have $E \subseteq V \times V$ where any $(u, v) \in E$ is a directed edge from u to v. By removing the directions of the edges in any directed graph (V, E), we obtain its *underlying undirected graph* $(V, \{\{u, v\} \mid (u, v) \in E\})$.

We will now illustrate some important notions of graph theory. While most of the following terms can also be defined similarly for directed graphs, we will concentrate on the case of undirected graphs. We define the following notions for any undirected graph $G = (V, E)$.

G isomorphic to another undirected graph $G' = (V', E')$ if there is a bijection $f : V \to V'$ with $\{u, v\} \in E \iff \{f(u), f(v)\} \in E'$. We say that two vertices u and v are *neighbors* if $\{u, v\} \in E$. The set of all neighbors of v is denoted by $N(v) = \{u \in V | \{u, v\} \in E\}$.

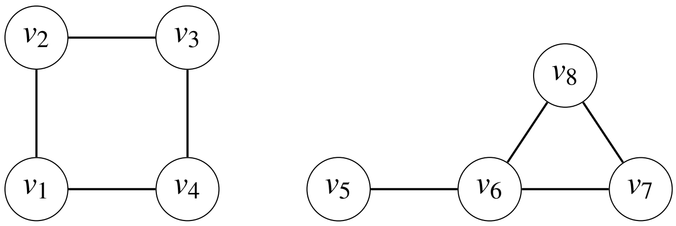


Figure 2.2: An undirected graph with two connected components from Example 2.2

A *path* from vertex v_1 to vertex v_k is a sequence $p = (v_1, \ldots, v_k)$ of vertices with $k \geq 1$ and $\{v_i, v_{i+1}\} \in E$ for all $i \in \{1, \ldots, k-1\}$. The *length* of a path $p = (v_1, \ldots, v_k)$ is the number of contained edges, i.e., $k-1$. A *cycle* is a path $p = (v_1, \ldots, v_k)$ where $\{v_k, v_1\} \in E$. The *distance* between vertices u and v is the length of a shortest path between u and v and is denoted by $d(u,v)$. If there is no path between u and v, then $d(u,v) = \infty$. The *diameter* of G is the maximal distance between two vertices in G, i.e., $\max_{u,v \in V} d(u,v)$.

A graph $G' = (V', E')$ is a *subgraph* of G if $V' \subseteq V$ and $E' \subseteq E$. Any subset $V' \subseteq V$ of the vertices *induces* a subgraph which is defined by $G|_{V'} = (V', E \cap \{\{u,v\} \mid u,v \in V'\})$. Hence, $G|_{V'}$ consists of all vertices in V' and all edges from G between the vertices in V'. G is *connected* if there exists a path from u to v for each two vertices $u,v \in V$ with $u \neq v$. Furthermore, G is a *tree* if it is connected and contains no cycles. An induced subgraph $G|_{V'}$ of G (with $V' \subseteq V$) is a *connected component* of G if $G|_{V'}$ is connected and there is no superset V'' with $V' \subset V'' \subseteq V$ for which $G|_{V''}$ is connected. Note that G can be partitioned into connected components in linear time. This can, for instance, be done via *depth-first search*. Finally, a set $V' \subseteq V$ is a *clique* in G if there is an edge $\{u,v\} \in E$ between any two vertices $u,v \in V', u \neq v$.

We complete this section with the following example which illustrates the above definitions.

Example 2.2. Let $G = (V,E)$ be an undirected graph with vertices $V = \{v_1, \ldots, v_8\}$ and edges $E = \{\{v_1,v_2\},\{v_2,v_3\},\{v_3,v_4\},\{v_4,v_1\},\{v_5,v_6\},\{v_6,v_7\},\{v_7,v_8\},\{v_6,v_8\}\}$. This graph is depicted in Figure 2.2. First, it can be observed that G has two connected components: the subgraphs induced by the vertex sets $V_1 = \{v_1,v_2,v_3,v_4\}$ and $V'' = \{v_5,v_6,v_7,v_8\}$. Furthermore, $p = (v_1,v_2,v_3,v_4)$ is a path of length 3 in G. p is also a cycle because $\{v_4,v_1\} \in E$. The vertex sets $\{v_6,v_7,v_8\}$ and $\{v_5,v_6\}$ are examples of cliques in G. The distance between vertices v_4 and v_5 is ∞ since there is no path connecting these two vertices. But the distance between vertices v_5 and v_8 is 2 since this is the length of a shortest path between them. The diameter of G is ∞ since G is not connected. However, the induced subgraphs $G|_{V'}$ and $G|_{V''}$ both have a diameter of 2. The two induced subgraphs $G|_{\{v_1,v_3,v_4\}}$ and $G|_{\{v_5,v_6,v_7\}}$ are isomorphic to each other while this is not the case for the two subgraphs $G|_{V'}$ and $G|_{V''}$. Last, the induced subgraph $G|_{\{v_1,v_3,v_5,v_7,v_8\}}$ is not connected and has four connected components.

2.3 Coalition Formation Games

In this thesis, we consider *coalition formation games* as a subclass of *non-transferable utility (NTU) games*. In these games, agents form groups based on their individual preferences where, in general, any partition of the agents is a possible outcome of the game. The agents evaluate the possible outcomes based on individual preferences.

We will now give an introduction to coalition formation games. After introducing some basic concepts, we will provide some background on hedonic games. They form an important subclass of coalition formation games where agents only care about the coalitions that they belong to. Afterwards, we describe some common preference representations, including cardinal formats, representations based on the categorization into friends and enemies, and many more. We then define some stability, optimality, and fairness notions that are of interest when studying hedonic games and explain some interesting decision problems which are associated with these notions. We complete the chapter by surveying the literature on this topic and summarizing some interesting results.

For more overviews of coalition formation games, see the survey by Hajduková [72] or the textbook by Chalkiadakis et al. [41]. For more background on NTU games see, for example, Section 5.1 in the textbook by Chalkiadakis et al. [41]. For literature on hedonic games, we refer to the book chapter by Aziz and Savani [8] and the survey by Woeginger [147].

2.3.1 Basic Definitions

Let $N = \{1, \dots, n\}$ be the set of *agents* (which we also call *players*). Subsets $C \subseteq N$ of the agents are called *coalitions* and, for any player $i \in N$, we denote the set of all coalitions containing i by $\mathscr{N}^i = \{C \subseteq N \mid i \in C\}$. It holds that $|\mathscr{N}^i| = 2^{n-1}$, which means that the number of coalitions containing i is exponential in the number of agents. Coalitions that contain only one player are also called *singleton coalitions* or *singletons*, for short. The coalition N that consists of all players is also called the *grand coalition*. A *coalition structure* is a partition $\Gamma = \{C_1, \dots, C_k\}$ of the set N of agents. As for every partition, it holds that $\bigcup_{i=1}^{k} C_i = N$ and $C_i \cap C_j = \emptyset$ for all $i, j \in \{1, \dots, k\}$ with $i \neq j$. There is no general restriction on the number k of coalitions in a coalition structure which means that k can range anywhere between 1 and n. The unique coalition in Γ that contains agent i is denoted by $\Gamma(i)$. Moreover, the set of all coalition structures for a set of agents N is denoted by $\mathscr{C}_N$. Note that the size of $\mathscr{C}_N$ grows exponentially with the number n of agents and equals the nth Bell number [18, 123]. For example, the first six Bell numbers are $B_1 = 1$, $B_2 = 2$, $B_3 = 5$, $B_4 = 15$, $B_5 = 52$, and $B_6 = 203$, which means that there are 203 possible partitions of a set of six agents.

Based on these notions, a *coalition formation game* is a pair $(N, \succeq)$, where $N = \{1, \dots, n\}$ is a set of agents and $\succeq = (\succeq_1, \dots, \succeq_n)$ is the profile of preferences of the agents. For each agent $i \in N$, $\succeq_i$ denotes her preference relation which is a complete weak order over all coalition structures, i.e., $\succeq_i \subseteq \mathscr{C}_N \times \mathscr{C}_N$. For two coalition structures $\Gamma, \Delta \in \mathscr{C}_N$, we say that agent i

weakly prefers Γ *to* Δ if $\Gamma \succeq_i \Delta$, that i *prefers* Γ *to* Δ (denoted by $\Gamma \succ_i \Delta$) if $\Gamma \succeq_i \Delta$ but not $\Delta \succeq_i \Gamma$, and that i is *indifferent between* Γ *and* Δ (denoted by $\Gamma \sim_i \Delta$) if $\Gamma \succeq_i \Delta$ and $\Delta \succeq_i \Gamma$.

2.3.2 Hedonic Games

The focus of this thesis will mainly be on coalition formation games with hedonic preferences, *hedonic games* for short. They were introduced independently by Banerjee et al. [15] and Bogomolnaia and Jackson [21]. The key idea of hedonic games (going back to Drèze and Greenberg [53]) is that agents only care about the coalitions that they are part of and not about the rest of a coalition structure. More formally, let any coalition formation game $(N, \succeq)$ be given. Then, the preference $\succeq_i$ of player i is hedonic if it only depends on the coalitions that i is part of, i.e., if for any two coalition structures $\Gamma, \Delta \in \mathscr{C}_N$, it holds that $\Gamma(i) = \Delta(i)$ implies $\Gamma \sim_i \Delta$. If the preferences of all agents $i \in N$ are hedonic, $(N, \succeq)$ is also called hedonic. For such a hedonic (coalition formation) game $(N, \succeq)$, the preferences are usually represented by complete weak orders over the set of coalitions containing an agent, i.e., $\succeq_i \subseteq \mathscr{N}^i \times \mathscr{N}^i$ for all $i \in N$. For two coalitions $A, B \in \mathscr{N}^i$, we then say that i *weakly prefers* A to B if $A \succeq_i B$, that i *prefers* A to B if $A \succ_i B$, and that i is *indifferent* between A and B if $A \sim_i B$. It follows from the definition of hedonic games that $\Gamma \succeq_i \Delta$ if and only if $\Gamma(i) \succeq_i \Delta(i)$.

Note that there are subclasses of hedonic games where only coalitions of certain sizes are allowed. For example, *marriage* and *roommate games* [64, 124] are hedonic games where all coalitions must have a size of at most two. These games and many other matching models are studied in *matching theory*. For more background on this topic we refer to the book chapter by Klaus et al. [94] and the textbooks by Roth and Sotomayor [124], Manlove [98], and Gusfield and Irving [71]. For other subclasses of hedonic games, the agents are assumed to divide into two types. In *hedonic diversity games* [31], an agent's preference depends on the fractions of agents of each type in a coalition. In this thesis however, we will only concentrate on general hedonic games where agents have no types and arbitrary coalition sizes are allowed.

We now give a simple example of a hedonic game.

Example 2.3. Let the set of players be given by $N = \{1,2,3\}$. Then, there are four different coalitions containing agent 1, namely $C_1 = \{1\}$, $C_2 = \{1,2\}$, $C_3 = \{1,3\}$, and $C_4 = \{1,2,3\}$. Here, C_1 is a singleton coalition and C_4 is the grand coalition. The set of all possible coalition structures $\mathscr{C}_N$ contains exactly five coalition structures:

$$\Gamma_1 = \{\{1\},\{2\},\{3\}\}, \ \Gamma_2 = \{\{1\},\{2,3\}\}, \ \Gamma_3 = \{\{1,2\},\{3\}\},$$
$$\Gamma_4 = \{\{1,3\},\{2\}\}, \text{ and } \Gamma_5 = \{\{1,2,3\}\}.$$

In particular, we have $\mathscr{C}_N = \{\Gamma_1,\Gamma_2,\Gamma_3,\Gamma_4,\Gamma_5\}$.

Further consider the following preference profile $\succeq = (\succeq_1, \succeq_2, \succeq_3)$ that, together with the set of agents N, defines a hedonic game $\mathscr{G} = (N, \succeq)$:

$$\{1,2,3\} \succ_1 \{1,2\} \succ_1 \{1,3\} \succ_1 \{1\},$$
$$\{1,2\} \succ_2 \{1,2,3\} \succ_2 \{2\} \succ_2 \{2,3\},$$
$$\{3\} \succ_3 \{1,3\} \sim_3 \{2,3\} \succ_3 \{1,2,3\}.$$

For this hedonic game, agent 1 prefers coalition $\Gamma_4(1) = \{1,3\}$ to coalition $\Gamma_2(1) = \{1\}$. Therefore, 1 also prefers Γ_4 to Γ_2. In contrast, agent 3 is indifferent between Γ_4 and Γ_2 because she is indifferent between $\Gamma_4(3) = \{1,3\}$ and $\Gamma_2(3) = \{2,3\}$.

2.3.3 Preference Representations

Even when considering the restricted case of hedonic coalition formation games, it is not reasonable to elicit full preferences in practice. Collecting a full preference over $\mathscr{N}^i$ for every agent $i \in N$ would not only lead to a preference profile of exponential size (in the number of agents) but would also present an extreme cognitive burden for the agents. Hence, we are looking for succinct representations of the preferences that are still reasonably expressive and easy to elicit.

Cardinal Preference Representations

There is a broad literature that concerns the problem of finding compact representations for hedonic preferences. Commonly used representations include the *additive encoding* due to Bogomolnaia and Jackson [21], the *fractional encoding* due to Aziz et al. [11], the *modified fractional encoding* due to Olsen [110], and the *friends-and-enemies encoding* due to Dimitrov et al. [50]. All these four representations have in common that they can be specified via cardinal valuation functions, i.e., they belong to the class of *cardinal hedonic games*. In these games, each agent i assigns a cardinal value to every other agent j that indicates how much i likes j. The agents' preferences can then be inferred from their valuation functions. The four representations differ in the range of valuations and in how the preferences are inferred.

Additively Separable Hedonic Games A hedonic game $(N, \succeq)$ is *additively separable* if, for every player $i \in N$, there exists a valuation function $v_i : N \to \mathbb{Q}$ such that for any two coalitions $A, B \in \mathscr{N}^i$ it holds that

$$A \succeq_i B \iff \sum_{j \in A} v_i(j) \geq \sum_{j \in B} v_i(j).$$

Hence, an additively separable hedonic game can also be represented by a tuple (N, v) consisting of a set of agents and a collection of valuation functions. It is commonly assumed that $v_i(i) = 0$ for every $i \in N$.[6] In additively separable hedonic games, agent i's valuation of a coalition $A \in \mathscr{N}^i$ is defined as $v_i^{\text{add}}(A) = \sum_{j\in A} v_i(j)$. Additively separable hedonic games [21] were studied, e.g., by Sung and Dimitrov [137, 136], Aziz et al. [10], and Woeginger [146].

Example 2.4. Again, consider the hedonic game $\mathscr{G} = (N, \succeq)$ from Example 2.3. $\mathscr{G}$ is additively separable as it can be represented via the following valuation functions:

i	$v_i(1)$	$v_i(2)$	$v_i(3)$
1	0	2	1
2	2	0	-1
3	-1	-1	0

We validate that these valuation functions indeed lead to the preferences from Example 2.3 using agent 2 as an example. We compute agent 2's valuations for the four coalitions:

$$\begin{aligned} &v_2^{\text{add}}(\{1,2\}) = 2+0 = 2, && v_2^{\text{add}}(\{1,2,3\}) = 2+0-1 = 1, \\ &v_2^{\text{add}}(\{2\}) = 0, \text{ and} && v_2^{\text{add}}(\{2,3\}) = 0-1 = -1. \end{aligned}$$

Since $v_2^{\text{add}}(\{1,2\}) > v_2^{\text{add}}(\{1,2,3\}) > v_2^{\text{add}}(\{2\}) > v_2^{\text{add}}(\{2,3\})$, agent 2's valuation function v_2 indeed corresponds to the preference $\{1,2\} \succ_2 \{1,2,3\} \succ_2 \{2\} \succ_2 \{2,3\}$.

Fractional Hedonic Games In fractional hedonic games, the value of a coalition is the average value of the members of the coalition. Hence, given a valuation function v_i of agent i, i's fractional value for a coalition $A \in \mathscr{N}^i$ is $v_i^{\text{frac}}(A) = \frac{1}{|A|}\sum_{j\in A} v_i(j)$ and a hedonic game $(N, \succeq)$ is *fractional* if for every player $i \in N$ there exists a valuation function $v_i : N \to \mathbb{Q}$ such that for any two coalitions $A, B \in \mathscr{N}^i$ it holds that

$$A \succeq_i B \iff v_i^{\text{frac}}(A) \geq v_i^{\text{frac}}(B).$$

Again, giving a fractional hedonic game by a tuple (N, v) of agents and valuation functions, it is commonly assumed that $v_i(i) = 0$ for all agents $i \in N$. Fractional hedonic games [11] have been studied, e.g., by Bilò et al. [19], Brandl et al. [25], Kaklamanis et al. [80], and Carosi et al. [37].

Modified Fractional Hedonic Games Modified fractional hedonic games are defined analogously to fractional hedonic games besides that the valuation of a player $i \in N$ for coalition $A \in \mathscr{N}^i$ is defined by

$$v_i^{\text{mfrac}}(A) = \begin{cases} \frac{1}{(|A|-1)}\sum_{j\in A} v_i(j) & \text{if } A \neq \{i\}, \\ 0 & \text{if } A = \{i\}. \end{cases}$$

[6] This is a normalization assumption. For each additively separable preference $\succeq_i$, there exists a valuation function v_i with $v_i(i) = 0$.

Modified fractional hedonic games [110] were studied, e.g., by Elkind et al. [58], Kaklamanis et al. [80], Monaco et al. [101, 102], Bullinger [33], and Bullinger and Kober [34].

The Friends-and-Enemies-Encoding In the friends-and-enemies encoding due to Dimitrov et al. [50], each player $i \in N$ partitions the other players into a set of friends $F_i \subseteq N \setminus \{i\}$ and a set of enemies $E_i = N \setminus (F_i \cup \{i\})$. Based on this representation, Dimitrov et al. [50] distinguish between the *friend-oriented* and the *enemy-oriented* preference extension. Under the friend-oriented model, agents prefer coalitions with more friends to coalitions with fewer friends, and in the case that two coalitions contain the same number of friends, they prefer the coalition with fewer enemies. Formally, a hedonic game $(N, \succeq)$ is friend-oriented if, for any agent $i \in N$, there exist a set of friends $F_i \subseteq N \setminus \{i\}$ and a set of enemies $E_i = N \setminus (F_i \cup \{i\})$ such that for any two coalitions $A, B \in \mathscr{N}^i$ it holds that

$$A \succeq_i B \iff |A \cap F_i| > |B \cap F_i| \text{ or } \big(|A \cap F_i| = |B \cap F_i| \text{ and } |A \cap E_i| \leq |B \cap E_i|\big). \tag{2.1}$$

Analogously, a hedonic game $(N, \succeq)$ is enemy-oriented if, for any agent $i \in N$, there exist a set of friends $F_i \subseteq N \setminus \{i\}$ and a set of enemies $E_i = N \setminus (F_i \cup \{i\})$ such that for any two coalitions $A, B \in \mathscr{N}^i$ it holds that

$$A \succeq_i B \iff |A \cap E_i| < |B \cap E_i| \text{ or } \big(|A \cap E_i| = |B \cap E_i| \text{ and } |A \cap F_i| \geq |B \cap F_i|\big). \tag{2.2}$$

Friend-oriented and enemy-oriented hedonic games can be seen as the subclasses of additively separable hedonic games where the valuation functions of the agents map only to $\{-1, n\}$ and $\{-n, 1\}$, respectively. In particular, in friend-oriented hedonic games, agents assign value n to their friends and value -1 to their enemies. In enemy-oriented hedonic games, agents assign value 1 to their friends and value $-n$ to their enemies. These cardinal values assure that the resulting additively separable hedonic preferences in fact satisfy the conditions from Equations 2.1 and 2.2. Agent i's friend-oriented respectively enemy-oriented value for coalition $A \in \mathscr{N}^i$ is then given by

$$\begin{aligned} v_i^{\text{fo}}(A) &= \sum_{j \in A} v_i(j) = n|A \cap F_i| - |A \cap E_i| \text{ and} \\ v_i^{\text{eo}}(A) &= \sum_{j \in A} v_i(j) = |A \cap F_i| - n|A \cap E_i|. \end{aligned}$$

Note that friend- and enemy-oriented hedonic games are also referred to as hedonic games with *appreciation of friends* and *aversion to enemies*. Friend- and enemy-oriented hedonic games [50] were studied, e.g, by Sung and Dimitrov [137, 136], Aziz and Brandl [7], Rey et al. [122], and Igarashi et al. [79].

Visual Presentation All these classes of *cardinal hedonic games* can be represented by complete weighted directed graphs with the agents as vertices where the weight of an edge (i, j) from agent i to agent j is i's value for j. Sometimes some edges with equal weights, e.g., all edges with weight zero, are omitted in the graph representation. In the case

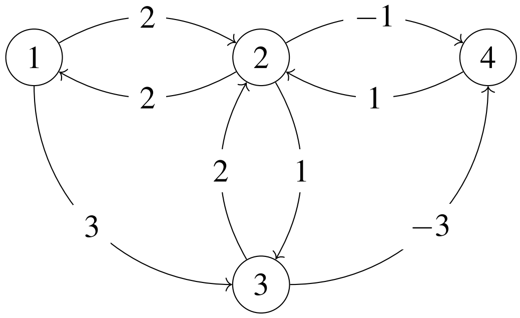


Figure 2.3: Graph representing the modified fractional hedonic game in Example 2.5. All omitted edges have weight zero.

of the friends-and-enemies encoding, all weights can be omitted. Instead, the game can be visualized by a directed graph where an edge from agent i to agent j indicates that j is i's friend. This graph is also called *network of friends*.

We call a cardinal hedonic game (N, v) *symmetric* if $v_i(j) = v_j(i)$ for all $i, j \in N$ and *simple* if $v_i(j) \in \{0, 1\}$ for all $i, j \in N$. For symmetric friend-oriented and symmetric enemy-oriented hedonic games, we also say that the friendship relations are *mutual*. In this case, the network of friends is an undirected graph where an edge $\{i, j\}$ represents the mutual friendship between agents i and j.

We now give examples of a modified fractional hedonic game and a friend-oriented hedonic game, respectively.

Example 2.5. We consider a modified fractional hedonic game (N, v) with four agents $N = \{1, 2, 3, 4\}$. The valuation functions of the agents are given by the graph in Figure 2.3 where all omitted edges represent valuations of zero. According to this graph, the valuation functions of the agents are:

i	$v_i(1)$	$v_i(2)$	$v_i(3)$	$v_i(4)$
1	0	2	3	0
2	2	0	1	−1
3	0	2	0	−3
4	0	1	0	0

We now compute the modified fractional preference of agent 2. First note that the set $\mathcal{N}^2$ of coalitions containing agent 2 has size $2^{n-1} = 2^3 = 8$. Agent 2's modified fractional valuations for these eight coalitions are given in the following table:

C	$\{2\}$	$\{1,2\}$	$\{2,3\}$	$\{2,4\}$	$\{1,2,3\}$	$\{1,2,4\}$	$\{2,3,4\}$	$\{1,2,3,4\}$
$v_2^{\mathrm{mfrac}}(C)$	0	$2/1$	$1/1$	$-1/1$	$3/2$	$1/2$	$0/2$	$2/3$

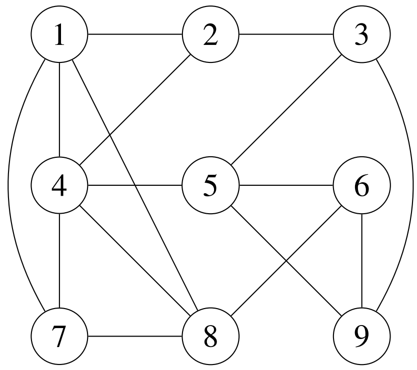


Figure 2.4: Graph representing the friend-oriented hedonic game with mutual friendship relations in Example 2.6

Sorting these valuations leads to the following modified fractional preferences of agent 2:

$$\{1,2\} \succ_2 \{1,2,3\} \succ_2 \{2,3\} \succ_2 \{1,2,3,4\} \succ_2 \{1,2,4\} \succ_2 \{2\} \sim_2 \{2,3,4\} \succ_2 \{2,4\}.$$

Note that agent 2's additively separable preferences for the graph in Figure 2.3 differ from the above modified fractional preferences. For example, agent 2 prefers $\{1,2,3\}$ to $\{1,2\}$ under additively separable preferences.

Next, we give a short example of a friend-oriented hedonic game.

Example 2.6. We consider a friend-oriented hedonic game $(N,\succeq)$ with nine agents, i.e., $N = \{1,\dots,9\}$. The mutual friendship relations among the agents are given by the network of friends in Figure 2.4. Furthermore, we consider the two coalitions $A = \{1,2,3,4,5,6,9\}$ and $B = \{1,2,4,7,8\}$.

For agent 1, it holds that she has two friends and four enemies in A while she has four friends and no enemies in B. Therefore, 1 prefers B to A under friend-oriented preferences. Actually, B is agent 1's most preferred coalition as it contains all of her friends and none of her enemies.

Considering agent 2, we can observe that 2 has three friends and three enemies in A while she has two friends and two enemies in B. Although the proportions of friends and enemies are the same for both coalitions, agent 2 prefers coalition A to B under friend-oriented preferences. This is because she compares the absolute numbers of friends in the two coalitions, which is greater for A than for B. Using the cardinal representation of the preferences with value $n = 9$ for friends and value -1 for enemies, agent 2's friend-oriented valuations for A and B are

$$v_2^{\text{fo}}(A) = n|A \cap F_2| - |A \cap E_2| = 9 \cdot |\{1,3,4\}| - |\{5,6,9\}| = 9 \cdot 3 - 3 = 24 \text{ and}$$
$$v_2^{\text{fo}}(B) = n|B \cap F_2| - |B \cap E_2| = 9 \cdot |\{1,4\}| - |\{7,8\}| = 9 \cdot 2 - 2 = 16.$$

Preference Representations Based on Friends and Enemies

Apart from the friends-and-enemies encoding due to Dimitrov et al. [50], there has been quite some research concerning preference representations that are based on the partitioning of agents into different groups.

For instance, Ota et al. [111] study hedonic games where agents specify their preferences by partitioning the other agents into friends, enemies, and *neutral agents*. In their model, an agent's preference is independent of all agents that she is neutral to. They then distinguish between the *friend appreciation* and *enemy aversion* due to Dimitrov et al. [50] and consider the problems of verifying (strict) core stability and checking the existence of (strictly) core stable coalition structures. They show that the neutral agents have an impact of on the computational complexity of these problems.

Similarly, Barrot et al. [17] study hedonic games where the agents partition each other into friends, enemies, and *unknown agents*. In contrast to Ota et al. [111], Barrot et al. [17] do not assume that agents are neutral to agents that they do not know. Instead, they distinguish between *extraverted* and *introverted agents* who either appreciate the presence of unknown agents or prefer coalitions with fewer unknown agents. They then investigate the impact of unknown agents on core stability and individual stability.

Another preference representation that is based on the partitioning of agents into friends, enemies, and neutral players is described in Chapter 5. In particular, we introduce *FEN-hedonic games* where agents represent their preferences via *weak rankings with double threshold*. That means that each agent partitions the other agents into friends, enemies, and neutral players and additionally specifies weak rankings on her friends and on her enemies, respectively. For more details on FEN-hedonic games, see Chapter 5. Weak rankings with double threshold are also studied by Rey and Rey [121] who obtain preferences over coalitions by measuring the distance between any given coalition and the specified ranking.

Further Preference Representations

There are several other preference representations for hedonic games. We will now give a brief overview of some prominent of these representations.

Under the *singleton encoding* by Cechlárová and Romero-Medina [40], the agents specify rankings over single agents. Cechlárová and Romero-Medina [40] define two preference extensions that lead to so-called $\mathscr{B}$-preferences and $\mathscr{W}$-preferences, respectively. Agents with $\mathscr{B}$-preferences rank coalitions only based on the most preferred player in each coalition. Agents with $\mathscr{W}$-preferences only care about the least preferred member of their coalitions. These two preference extensions are also studied by Cechlárová and Hajduková [38, 39].

In the *individually rational encoding* due to Ballester [14], agents only rank the coalitions that they prefer to being alone. This leads to a succinct representation whenever the number of those coalitions is small.

Under the *anonymous encoding* defined by Banerjee et al. [15], the agents' preferences only depend on their coalition sizes. This means that, under anonymous preferences, the agents are indifferent among any two coalitions of the same size and do not care about the identity of the agents. Anonymous hedonic games have also been studied by Darmann et al. [47].

Elkind and Wooldridge [56] proposed a very expressive representation: *hedonic coalition nets* where the agents specify their preferences by giving a set of propositional formulas. With these formulas, the agents can specify which combination of agents they would like to have in their coalitions. For instance, agent i might specify the formula $j \wedge k \mapsto_i 8$ which means that i obtains utility 8 if she is in a coalition with agents j and k. These propositional formulas can also be more complex and contain the Boolean operators $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$, and $\neg$. An agent's total utility for a given coalition is the sum of all formulas that are satisfied by the coalition. Elkind and Wooldridge [56] show that hedonic coalition nets generalize several other preference representations such as hedonic games with $\mathscr{B}$- or $\mathscr{W}$-preferences [40], the individually rational encoding [14], additively separable hedonic games [21], and anonymous hedonic games [15].

Aziz et al. [9] consider hedonic games with *dichotomous preferences*. Formally, player i's preference is dichotomous if she can partition the set $\mathscr{N}^i$ of coalitions containing herself into two groups, satisfactory coalitions and unsatisfactory coalitions, such that she strictly prefers any satisfactory coalition to any unsatisfactory coalition and is indifferent between any two coalitions of the same group. Aziz et al. [9] introduce the *boolean hedonic encoding*, a succinct representation for hedonic games with dichotomous preferences. In this encoding, each agent's preference is given by a single propositional formula that characterizes this agent's satisfactory set of coalitions. Hedonic games with dichotomous preferences are also studied by Peters [114]. He studies the computational complexity of finding stable and optimal coalition structures in such hedonic games. While doing so, he distinguishes between several representations of such games, including the boolean encoding.

2.3.4 Stability and Optimality in Hedonic Games

Central questions in coalition formation are which coalition structures are likely to form and which coalition structures are desirable outcomes. There is a broad literature that studies such desirable properties in coalition formation. The solution concepts are concerned with *optimality*, *stability*, and *fairness*. In this section, we will consider several notions of stability and optimality.

There are various stability notions that have been proposed in the literature. Those notions mainly concern the question whether there are agents that would like to deviate from a given coalition structure. We distinguish different categories of stability notions. First, there are concepts based on *single player deviations* such as *Nash stability*, *individual stability*, or *individual rationality* that capture whether there are agents that would like to perform a deviation to another coalition on their own. Second, there exist notions of *group stability* such as *core stability* that capture whether groups of agents would want to deviate together. And third, there are notions that are based on the comparison of coalition structures such as *Pareto optimality* or *popularity*. These notions can also be seen as *optimality concepts*. Further optimality criteria are concerned with the maximization of *social welfare* or other measurements of the agents' satisfaction.

We now define some common stability notions and start with some classic notions. For any given hedonic game $(N, \succeq)$, coalition structure $\Gamma \in \mathscr{C}_N$ is said to be

- *perfect* (PF)[7] if every agent is in her most preferred coalition, i.e., every agent $i \in N$ weakly prefers $\Gamma(i)$ to every other coalition $C \in \mathscr{N}^i$.
- *individually rational* (IR) if every agent weakly prefers her current coalition to being alone, i.e., every agent $i \in N$ weakly prefers $\Gamma(i)$ to $\{i\}$.

Note that perfectness (formulated by Aziz et al. [12]) and individually rationality are two of the most extreme stability notions that we consider here. While perfectness is stronger than almost all other stability notions (except for strict popularity), individual rationality imposes only a minimal requirement and is implied by many other notions.

We continue with some further classic notions that are concerned with single player deviations and were formulated by Bogomolnaia and Jackson [21]. Coalition structure $\Gamma \in \mathscr{C}_N$ is

- *Nash stable* (NS) if no agent wants to deviate to another coalition in $\Gamma \cup \{\emptyset\}$, i.e., every agent $i \in N$ weakly prefers $\Gamma(i)$ to every coalition $C \cup \{i\}$ with $C \in \Gamma \cup \{\emptyset\}$.
- *individually stable* (IS) if no agent wants to deviate to another coalition C in $\Gamma \cup \{\emptyset\}$ and can do so without making any agent in C worse off. Formally, Γ is IS if for all agents $i \in N$ and all coalitions $C \in \Gamma \cup \{\emptyset\}$, it holds that i weakly prefers $\Gamma(i)$ to $C \cup \{i\}$ or there is a player $j \in C$ who prefers C to $C \cup \{i\}$.
- *contractually individually stable* (CIS) if no agent i wants to deviate to another coalition C in $\Gamma \cup \{\emptyset\}$ and can do so without making any agent in C or $\Gamma(i)$ worse off. Formally, Γ is CIS if for all agents $i \in N$ and all coalitions $C \in \Gamma \cup \{\emptyset\}$, it holds that i weakly prefers $\Gamma(i)$ to $C \cup \{i\}$ or there is a player $j \in C$ who prefers C to $C \cup \{i\}$ or there is a player $k \in \Gamma(i) \setminus \{i\}$ who prefers $\Gamma(i)$ to $\Gamma(i) \setminus \{i\}$.

Additionally, Sung and Dimitrov [138] introduced contractual Nash stability and some other related notions. We say that coalition structure $\Gamma \in \mathscr{C}_N$ is

- *contractually Nash stable* (CNS) if no agent i wants to deviate to another coalition in $\Gamma \cup \{\emptyset\}$ and can do so without making any agent in $\Gamma(i)$ worse off. Formally, Γ is CNS if for all agents $i \in N$ and all coalitions $C \in \Gamma \cup \{\emptyset\}$, it holds that i weakly prefers $\Gamma(i)$ to $C \cup \{i\}$ or there is a player $k \in \Gamma(i) \setminus \{i\}$ who prefers $\Gamma(i)$ to $\Gamma(i) \setminus \{i\}$.

[7] In the context of the friends-and-enemies encoding [50], perfectness is sometimes also called "wonderful stability", e.g., by Woeginger [147], Elkind and Rothe [55], and Rey et al. [122].

We now turn to core stability which is a classic notion of group stability that was already studied by Banerjee et al. [15]. Later, core stability and strict core stability have also been extensively studied by Dimitrov et al. [50]. For any coalitions structure $\Gamma \in \mathscr{C}_N$ and any nonempty coalition $C \subseteq N$, C is said to block Γ if every agent $i \in C$ prefers C to $\Gamma(i)$. C is said to weakly block Γ if all agents $i \in C$ weakly prefer C to $\Gamma(i)$ and at least one agent $j \in C$ prefers C to $\Gamma(j)$. Coalition structure $\Gamma \in \mathscr{C}_N$ is

- *core stable* (CS) if no nonempty coalition blocks Γ.
- *strictly core stable* (SCS) if no nonempty coalition weakly blocks Γ.

Karakaya [81] and Aziz and Brandl [7] formulated some more related notions. For a coalition $C \subseteq N$, we say that coalition structure $\Delta \in \mathscr{C}_N$ is reachable from coalition structure $\Gamma \in \mathscr{C}_N$, $\Gamma \neq \Delta$, by coalition C if, for all $i, j \in N \setminus C$, it holds that $\Gamma(i) = \Gamma(j) \iff \Delta(i) = \Delta(j)$. In other words, if Δ is reachable from Γ by C, then all agents in C might deviate to other coalitions while all agents in $N \setminus C$ have to stay together as before. Then, a coalition $C \subseteq N$, $C \neq \emptyset$,

- *strong Nash blocks* coalition structure Γ if there exists a coalition structure Δ that is reachable from Γ by C such that all agents $i \in C$ prefer $\Delta(i)$ to $\Gamma(i)$.
- *weakly Nash blocks* Γ if there exists a coalition structure Δ that is reachable from Γ by C such that all agents $i \in C$ weakly prefer $\Delta(i)$ to $\Gamma(i)$ and there is an agent $j \in C$ who prefers $\Delta(j)$ to $\Gamma(j)$.
- *strong individually blocks* Γ if there exists a coalition structure Δ that is reachable from Γ by C such that all agents $i \in C$ prefer $\Delta(i)$ to $\Gamma(i)$ and there is an agent $j \in C$ such that all $k \in \Delta(j)$ weakly prefer $\Delta(k)$ to $\Gamma(k)$.

Based on these notions it holds that coalition structure $\Gamma \in \mathscr{C}_N$ is

- *strong Nash stable* (SNS) [81] if there is no coalition $C \subseteq N$ that strong Nash blocks Γ.
- *strictly strong Nash stable* (SSNS) [7] if there is no coalition $C \subseteq N$ that weakly Nash blocks Γ.
- *strong individually stable* (SIS) [7] if there is no coalition $C \subseteq N$ that strong individually blocks Γ.

We now turn to some concepts that are based on the comparison of coalition structures. For two coalition structures $\Gamma, \Delta \in \mathscr{C}_N$, we say that Δ *Pareto-dominates* Γ if every agent $i \in N$ weakly prefers $\Delta(i)$ to $\Gamma(i)$ and there is an agent $j \in N$ who prefers $\Delta(j)$ to $\Gamma(j)$. Coalition structure $\Gamma \in \mathscr{C}_N$ is

- *Pareto optimal* (PO) if there is no coalition structure that Pareto-dominates Γ.

Popularity is another notion that is based on the comparison of coalition structures. The notion was first proposed in the context of marriage games by Gärdenfors [65]. In the context of hedonic games, popularity and strict popularity were formulated by Aziz et al. [10] and Lang et al. [96]. A coalition structure $\Gamma \in \mathscr{C}_N$ is

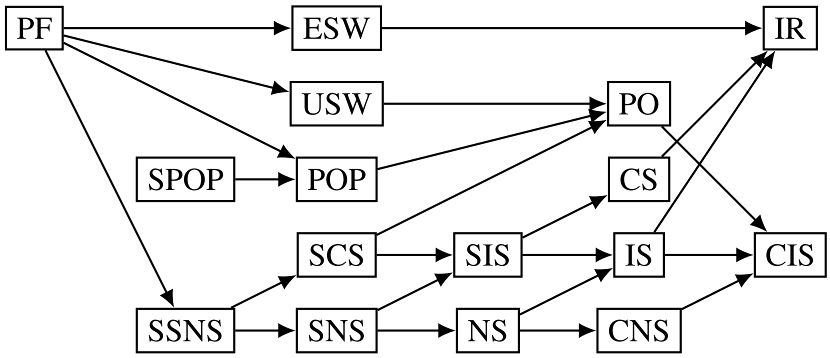


Figure 2.5: Relations among the stability and optimality notions from Section 2.3.4 where a notion A implies a notion B exactly if there is a directed path from A to B

- *popular* (POP) if for every coalition structure $\Delta \in \mathscr{C}_N$, at least as many agents prefer Γ to Δ as the other way around; formally, this means for all $\Delta \in \mathscr{C}_N$ with $\Delta \neq \Gamma$ that

$$|\{i \in N \mid \Gamma(i) \succ_i \Delta(i)\}| \geq |\{i \in N \mid \Delta(i) \succ_i \Gamma(i)\}|.$$

- *strictly popular* (SPOP) if for every coalition structure $\Delta \in \mathscr{C}_N$, more agents prefer Γ to Δ than the other way around; formally, this means for all $\Delta \in \mathscr{C}_N$ with $\Delta \neq \Gamma$ that

$$|\{i \in N \mid \Gamma(i) \succ_i \Delta(i)\}| > |\{i \in N \mid \Delta(i) \succ_i \Gamma(i)\}|.$$

We consider two further concepts that were formulated by Aziz et al. [10] and are concerned with *social welfare maximization*. For any cardinal hedonic game (N, v), we say that $\Gamma \in \mathscr{C}_N$ *maximizes*

- *utilitarian social welfare* (USW) if $\sum_{i\in N} v_i(\Gamma(i)) \geq \sum_{i\in N} v_i(\Delta(i))$ for all $\Delta \in \mathscr{C}_N$.

- *egalitarian social welfare* (ESW) if $\min_{i\in N} v_i(\Gamma(i)) \geq \min_{i\in N} v_i(\Delta(i))$ for all $\Delta \in \mathscr{C}_N$.

We also say that a coalition structure Γ is USW or ESW by which we mean that Γ maximizes USW or ESW. We further use all abbreviations from this section as nouns and adjectives; for example, we say that a coalition structure is CS (core stable) or that it satisfies CS (core stability).

There are a lot of relations among these stability and optimality notions. Some of them follow directly from the definitions, e.g., NS trivially implies IS which in turn implies CIS. The relations among all notions from this section are visualized in Figure 2.5. For more background on these relations, see, e.g., Bogomolnaia and Jackson [21] (for relations among PO, NS, IS, CIS, and CS), Sung and Dimitrov [138] (for relations among SCS, CS, NS, IS, CNS, and CIS), Aziz and Brandl [7] (for relations among SSNS, SNS, SIS, and previous notions), or Kerkmann [83] (for relations among SPOP, POP, PO, and other notions).

2.3.5 Fairness in Hedonic Games

Besides the stability concepts from the previous section, other important notions in hedonic games are concerned with *fairness*. Some of these notions are inspired from the field of *fair division* where three classic fairness criteria are *equitability*, *proportionality*, and *envy-freeness*. Further fairness notions in fair division are *jealousy-freeness* due to Gourvès et al. [69], *envy-freeness up to one good* due to Budish [32], the *max-min fair share criterion* by Budish [32], and the *min-max fair share criterion* by Bouveret and Lemaître [23]. For background on fair division theory see, e.g., the book chapters by Lang and Rothe [95] and Bouveret et al. [24].

For hedonic games, it was proposed to use envy-freeness as a notion of fairness [21, 10, 114, 115]. We say that a coalition structure $\Gamma \in \mathscr{C}_N$ is *envy-free by replacement* (EFR) if no agent envies another agent for her coalition, i.e., if for all agents $i, j \in N$ with $\Gamma(i) \neq \Gamma(j)$, agent i weakly prefers $\Gamma(i)$ to $(\Gamma(j) \setminus \{j\}) \cup \{i\}$. Perfectness is the only notion from Section 2.3.4 that implies EFR. Also, EFR does not imply any of the notion from Section 2.3.4. The following two examples illustrate EFR and some of the notions from Section 2.3.4 while showing that EFR is independent from all notions besides perfectness.

Example 2.7. Consider the hedonic game $\mathscr{G} = (N, \succeq)$ with $N = \{1,2,3\}$ and the following preference profile $\succeq = (\succeq_1, \succeq_2, \succeq_3)$:

$$\{1,2\} \succ_1 \{1,2,3\} \succ_1 \{1\} \succ_1 \{1,3\},$$
$$\{1,2\} \succ_2 \{1,2,3\} \succ_2 \{2\} \succ_2 \{2,3\},$$
$$\{1,3\} \succ_3 \{3\} \succ_3 \{1,2,3\} \succ_3 \{2,3\}.$$

Further consider the coalition structure $\Gamma = \{\{1,2\},\{3\}\}$. We first observe that this coalition structure is not EFR because agent 3 envies agent 2 for her coalition. In particular, we have

$$(\Gamma(2) \setminus \{2\}) \cup \{3\} = \{1,3\} \succ_3 \{3\} = \Gamma(3).$$

Yet, agents 1 and 2 prefer Γ to every other coalition structure which implies that Γ is SPOP. Moreover, Γ is SSNS since there is no coalition that weakly Nash blocks Γ: First, observe that agents 1 and 2 can not be part of any weakly Nash blocking coalition because Γ is their unique most preferred coalition structure. Hence, $\{3\}$ is the only remaining coalition that could weakly Nash block Γ. However, deviating from $\{3\}$ to $\{1,2\} \cup \{3\}$ does not present an improvement to agent 3. Thus, there is no weakly Nash blocking coalition.

Note that $\mathscr{G}$ is additively separable and can be represented via the following valuations:

i	$v_i(1)$	$v_i(2)$	$v_i(3)$
1	0	2	-1
2	2	0	-1
3	1	-2	0

For these valuation functions, Γ maximizes USW and ESW. In particular, it holds that the USW for Γ is $2+2+0=4$ while the ESW for Γ is $\min\{2,2,0\}=0$.

Summing up, we have shown that Γ is SPOP, SSNS, USW, and ESW but not EFR. This shows that none of SPOP, SSNS, USW, and ESW implies EFR. Since all other notions from Section 2.3.4 except for perfectness are implied by SPOP, SSNS, USW, or ESW, none of these notions implies EFR either.

The next example shows that no notions from Section 2.3.4 is implied by EFR.

Example 2.8. Consider a very simple hedonic game $\mathscr{G}=(N,\succeq)$ with two players $N=\{1,2\}$ and the preferences $\{1\}\succ_1\{1,2\}$ and $\{2\}\succ_2\{1,2\}$. While coalition structure $\{\{1,2\}\}$ is EFR, it is neither IR nor CIS. Hence, EFR does not imply IR or CIS. Since all notions from Section 2.3.4 imply IR or CIS, none of these notions is implied by EFR.

In order to decide whether EFR is satisfied, agents have to inspect not only their own but also the coalitions of other agents. In Chapter 4, we introduce three further notions of *local fairness* that can be decided while all agents only inspect their own coalitions. The three local fairness notions, namely *min-max fairness*, *grand-coalition fairness*, and *max-min fairness*, are defined via individual threshold coalitions. In Chapter 4, we study the relations among these three local fairness notions and also relate them to other notions of stability. We further study the computational complexity of the related existence problems and of computing the threshold coalitions.

Further works studying envy-freeness in coalition formation games are due to Wright and Vorobeychik [148], Ueda [143], and Barrot and Yokoo [16]. For instance, Ueda [143] introduces and studies *justified envy-freeness*, a weakening of EFR, that is implied by CS.

2.3.6 Decision Problems for Hedonic Games

There are some natural questions that arise when studying the above stability, optimality, and fairness notions. For instance, we are interested in whether a given notion can be guaranteed for any hedonic game or whether there are hedonic games that do not allow any coalition structures that satisfy this notion. For any notion α, we are further interested in the computational complexity of the *verification problem* and the *existence problem*, which are defined as follows:

α-VERIFICATION	
Given:	A hedonic game $(N,\succeq)$ and a coalition structure $\Gamma\in\mathscr{C}_N$.
Question:	Does Γ satisfy α in $(N,\succeq)$?

α-EXISTENCE	
Given:	A hedonic game $(N,\succeq)$.
Question:	Is there a coalition structure $\Gamma\in\mathscr{C}_N$ that satisfies α in $(N,\succeq)$?

Note that there is a link between the complexities of these two problems: If α-VERIFICATION is in P for a concept α, then α-EXISTENCE is in NP as instances can be guessed nondeterministically and verified in polynomial time.

For any notion α, the following search problem is of interest as well:

α-SEARCH	
Input:	A hedonic game $(N, \succeq)$.
Output:	A coalition structure $\Gamma \in \mathscr{C}_N$ that satisfies α in $(N, \succeq)$ or "no" if there does not exist such a coalition structure.

Obviously, any upper bound on the computational complexity of α-SEARCH carries over to α-EXISTENCE, e.g., α-SEARCH $\in$ P implies α-EXISTENCE $\in$ P. Similarly, lower bounds on the computational complexity of α-EXISTENCE carry over to α-SEARCH, e.g., α-EXISTENCE being NP-hard implies α-SEARCH being NP-hard. Also, if α-VERIFICATION is in P, then α-SEARCH is in NP.

Stability Results

We will now summarize some results concerning the above problems for the stability, optimality, and fairness notions from Sections 2.3.4 and 2.3.5. Some of these results can be deduced directly from their definitions; some results are known from the literature.

Easy Verification First observe that α-VERIFICATION with $\alpha \in \{\mathrm{IR}, \mathrm{NS}, \mathrm{IS}, \mathrm{CIS}, \mathrm{CNS}, \mathrm{EFR}\}$ is easy for any hedonic game for which the preferences can be accessed in polynomial time. For all these notions, we can find the answer to α-VERIFICATION by iterating over all agents and checking a polynomial number (in the number of agents) of preference relations. This leads to a polynomial time algorithm if single preference relations can be checked in polynomial time. Also, whenever we can determine the agents' most preferred coalitions in polynomial time, PF-VERIFICATION is easy. For all other notions from Section 2.3.4, α-VERIFICATION is not easy in general. Indeed, it was shown that α-VERIFICATION is coNP-complete for $\alpha \in \{\mathrm{CS}, \mathrm{SCS}, \mathrm{PO}, \mathrm{POP}, \mathrm{SPOP}, \mathrm{USW}, \mathrm{ESW}\}$ even if the preferences are additively separable (see Table 2.1).

Guaranteed Existence The three stability notions IR, CIS, and PO impose rather mild restrictions on coalition structures and can be fulfilled for any hedonic game. For example, for any hedonic game $(N, \succeq)$ with $N = \{1, \dots, n\}$, the coalition structure $\{\{1\}, \dots, \{n\}\}$ consisting only of singleton coalitions is IR. This follows directly from the definition of IR. Turning to PO, it can be easily seen that a PO coalition structure is guaranteed to exist by the following observations: Whenever a coalition structure Γ_2 Pareto-dominates coalition structure Γ_1, every agent weakly prefers Γ_2 to Γ_1 and at least one agent prefers Γ_2 to Γ_1. This means that the overall satisfaction grows when switching from Γ_1 to the Pareto-dominating coalition structure Γ_2. Now, assuming that there is no PO coalition structure would mean that there is an infinite sequence of coalition structures $(\Gamma_1, \Gamma_2, \dots)$. such that Γ_{i+1} Pareto-dominates Γ_i for every $i \geq 1$. Since there is only a finite number of coalition structures and since no coalition structure can occur twice in the sequence (due to the growth of satisfaction), such a sequence can not exist and there has to be a PO coalition structure. Since every PO coalition structure is CIS, this also implies the existence of a CIS coalition structure. The corresponding result for CIS was also shown by Ballester [14]. Finally, due to the guaranteed existence of these three notions, we can deduce that α-EXISTENCE is trivially in P for any hedonic game and $\alpha \in \{\mathrm{IR}, \mathrm{CIS}, \mathrm{PO}\}$. In addition, Bogomolnaia and Jackson [21] show that, for any hedonic game with strict preferences,[8] there exists a coalition structure that is PO, IR, and CIS at the same time.

EFR coalition structures are guaranteed to exist for any hedonic game as well. In fact, the coalition structures $\{\{N\}\}$ consisting of the grand coalition and $\{\{1\}, \dots, \{n\}\}$ consisting only of singleton coalitions are always EFR by definition. Yet, Ueda [143] shows that there exist hedonic games where no coalition structure besides these two trivial ones is EFR.

For any cardinal hedonic game, coalition structures maximizing USW and ESW are guaranteed to exist as well. Again, this follows directly from the definitions.

Properties that Guarantee Existence For all other notions from Section 2.3.4, coalition structures that satisfy these notions are not guaranteed to exist in general hedonic games. However, some work has been done, studying properties that guarantee the existence of stable coalition structures. For example, Bogomolnaia and Jackson [21] study properties that guarantee the existence of PO, CS, NS, IS, or CIS coalition structures. They show that, for any symmetric ASHG (from here on, "additively separable hedonic game" is also abbreviated with "ASHG"), USW implies NS [21, proof of Proposition 2]. Since USW coalition structures are guaranteed to exist in ASHGs, this means that any symmetric ASHG admits a NS coalition structure. The same holds for IS and CNS coalition structures because NS implies IS and CNS. Suksompong [135] generalizes the result by Bogomolnaia and Jackson [21] and shows that NS coalition structures are even guaranteed to exist for *subset-neutral* hedonic games, a generalization of symmetric ASHGs. Moreover, Bogomolnaia and Jackson [21] show that there exists a coalition structure that simultaneously satisfies PO, IR, and IS for any ASHG with strict preferences. Banerjee et al. [15] study the existence of CS coalition structures under different restrictions of hedonic games. Motivated by the fact that there even may not be a CS coalition structure for hedonic games that satisfy rather strong restrictions, e.g., for anonymous ASHGs, they introduce the *weak top-coalition property* which guarantees the existence of a CS coalition structure. Burani and Zwicker [35] show that all symmetric ASHGs that have *purely cardinal* preference profiles admit a coalition structure that is both NS and CS. The existence of CS is also studied by Dimitrov et al. [50]. They show that CS and SCS coalition structures exist for any friend-oriented and enemy-oriented hedonic

[8] That means that no player is indifferent between any two different coalitions.

game. Furthermore, Alcalde and Revilla [1] introduce a property called *top responsiveness* that guarantees the existence of CS coalition structures. Dimitrov and Sung [48] strengthen the result of Alcalde and Revilla [1] by showing that top responsiveness even guarantees the existence of SCS coalition structures. Dimitrov and Sung [49] additionally prove that top responsiveness together with *mutuality* ensures the existence of NS coalition structures. As a counterpart to top responsiveness, Suzuki and Sung [139] introduce *bottom refuseness* (which is later called *bottom responsiveness* by Aziz and Brandl [7]). They show that, similar to top responsiveness, bottom refuseness guarantees the existence of CS coalition structures. Since friend-oriented hedonic games fulfill top responsiveness while enemy-oriented hedonic games fulfill bottom refuseness, the existence results by Alcalde and Revilla [1] for top responsiveness and by Suzuki and Sung [139] for bottom refuseness generalize the existence results by Dimitrov et al. [50] for friend-oriented and enemy-oriented hedonic games. Sung and Dimitrov [138] study the existence of CNS coalition structures and show that any hedonic game that satisfies *separability* (a generalization of additive separability) and *weak mutuality* admits a CNS coalition structure. Karakaya [81] establishes two properties that guarantee the existence of a SNS coalition structure: the *weak top-choice property* and the *descending separability* of preferences. Aziz and Brandl [7] show that the existence of a SSNS coalition structure is guaranteed in hedonic games that satisfy top responsiveness and *mutuality*. Yet, these two properties do not guarantee the existence of PF coalition structures. They also show that SIS coalition structures are guaranteed in hedonic games that satisfy *bottom responsiveness* while the existence of SNS coalition structures is guaranteed in hedonic games that satisfy *strong bottom responsiveness* and *mutuality*. Furthermore, Aziz and Brandl [7] study the existence of stable coalition structures in friend-oriented and enemy-oriented hedonic games. They show that each symmetric friend-oriented hedonic game admits a SSNS coalition structure. Moreover, each enemy-oriented hedonic game admits a SIS coalition structure and even a SNS coalition structure if the game is symmetric. They further show that SCS coalition structures are guaranteed to exist in hedonic games with strict $\mathscr{B}$-preferences [40]. Finally, Brandl et al. [25] show that CS, NS, and IS coalition structures are not guaranteed to exist in fractional hedonic games.

Complexity Results for ASHGs Without applying suiting restrictions, many classes of hedonic games do not admit stable coalition structures in general. In these cases, the related existence problems are not trivial. And even if coalition structures satisfying a given notion are guaranteed to exist, the problem of finding such coalition structures might still be intractable. Hence, there has been some research on the computational complexity of the existence and search problems for several classes of hedonic games and various stability notions. For example, Brandl et al. [25] not only show that CS, NS, and IS coalition structures are not guaranteed to exist in fractional hedonic games, but they also show that α-EXISTENCE with $\alpha \in \{\mathrm{CS}, \mathrm{NS}, \mathrm{IS}\}$ is NP-hard for fractional hedonic games.

We will now illustrate some results from the literature. While doing so, we concentrate on the popular class of ASHGs. The results are summarized in Table 2.1. Sung and Dimitrov [136] show that, for any ASHG, α-EXISTENCE with $\alpha \in \{\mathrm{NS}, \mathrm{IS}\}$ is NP-complete and that α-

α	α-VERIFICATION	α-EXISTENCE	α-SEARCH
PF	in P [10]	in P [10]	in P [10]
IR	in P [10]	trivial [10]	trivial [10]
NS	in P [10]	NP-complete [136], trivial if sym [21, 135]	NP-hard [136], PLS-complete [62]
IS	in P [10]	NP-complete [136], trivial if sym [21]	NP-hard [136], PLS-complete [63]
CIS	in P [10]	trivial [14]	in P [10]
CNS	in P	in NP, trivial if sym [21, 138]	in NP, in P if sym [138, 63]
CS	coNP-complete [137]	Σ_2^p-complete [146, 116, 111], trivial if fn. 1 holds	Σ_2^p-complete [146, 116, 111]
SCS	coNP-complete [10]	Σ_2^p-complete [116, 111]	Σ_2^p-complete [116, 111]
PO	coNP-complete [10, 33]	trivial	in P if fn. 2 holds, in P if fn. 3 holds
POP	coNP-complete [10, 26]	NP-hard [10, 26], coNP-hard [26]	NP-hard [10, 26], coNP-hard [26]
SPOP	coNP-complete [26]	coNP-hard [26]	coNP-hard [26]
USW	coNP-complete [10]	trivial [10]	NP-hard [10]
ESW	coNP-complete [10]	trivial [10]	NP-hard [10]
EFR	in P [10]	trivial [10, 143]	trivial [10, 143]

[1] If the game is symmetric and preferences are *purely cardinal* [35], if the game is friend-oriented [50], or if the game is enemy-oriented [50].
[2] If all preferences are strict [10].
[3] If the game is *mutually indifferent* [33]. Note that mutual indifference implies symmetry.

Table 2.1: Computational complexity of the problems from Section 2.3.6 for additively separable hedonic games and the stability, optimality, and fairness notions from Sections 2.3.4 and 2.3.5. Some additional results are given for subclasses of additively separable hedonic games, e.g., "if sym" indicates that a result holds for symmetric additively separable hedonic games.

EXISTENCE with $\alpha \in \{\mathrm{CS}, \mathrm{SCS}\}$ is NP-hard. Recall that all these hardness results carry over to the corresponding search problems. Gairing and Savani [62] strengthen the above result for NS by showing that NS-SEARCH is PLS-complete for symmetric ASHGs. In a follow-up work, Gairing and Savani [63] define some new stability concepts for symmetric ASHGs. In particular, they define *vote-in stability* which is equivalent to IS and *vote-out stability* which is equivalent to CNS. The combination of vote-in and vote-out stability is equivalent to CIS. They show that, for symmetric ASHGs, IS-SEARCH is PLS-complete while CNS-SEARCH is in P. Sung and Dimitrov [137] show that CS-VERIFICATION is coNP-complete for enemy-oriented hedonic games. Since enemy-oriented hedonic games are a subclass of ASHGs, the hardness extends to the general case of ASHGs. Aziz et al. [10] extend the result by showing that SCS-VERIFICATION is coNP-complete for enemy-oriented hedonic games as well. Furthermore, Aziz et al. [10] show many more results concerning the complexity of verification, existence, and search problems in ASHGs. For instance, they present an algorithm that finds a CIS coalition structure for any ASHG, i.e., CIS-SEARCH is in P. Woeginger [146] shows that C-EXISTENCE is Σ_2^p-complete for ASHGs. Afterwards, Woeginger [147] surveys the results and open problems concerning CS and SCS. Peters [116] extends the hardness result by Woeginger [146] and shows that SC-EXISTENCE is Σ_2^p-complete for ASHGs. Peters and Elkind [117] establish metatheorems that show the NP-hardness of α-EXISTENCE for several stability notions α. They apply these theorems to several classes of hedonic games such as ASHGs and fractional hedonic games. For ASHGs, their metatheorems reveal that α-EXISTENCE is NP-hard for $\alpha \in \{\mathrm{NS}, \mathrm{IS}, \mathrm{CS}, \mathrm{SCS}, \mathrm{SSNS}, \mathrm{SNS}, \mathrm{SIS}\}$. Brandt and Bullinger [26] study POP and SPOP in ASHGs. They show that POP-EXISTENCE is NP-hard and coNP-hard for symmetric ASHGs. Thus, they conclude that it is likely that this problem is even Σ_2^p-complete. They also show that, for symmetric ASHGs, SPOP-EXISTENCE is coNP-hard and α-VERIFICATION with $\alpha \in \{\mathrm{POP}, \mathrm{SPOP}\}$ is coNP-complete. Furthermore, Aziz et al. [10] and Bullinger [33] study the combination of PO with other stability notions in ASHGs. They, e.g., show that it is hard to find coalition structures that are PO and EF or PO and IR. Last, Peters [115] and Hanaka and Lampis [73] study stability in ASHGs (and other classes of hedonic games) from the viewpoint of parameterized complexity.

Outlook Needless to say, there exists more interesting related literature. For instance, some research deals with the prices of stability, optimality, and fairness. These prices measure the losses of social welfare that come with certain stability, optimality, or fairness notions. For example, the price of NS is the worst-case ratio between the maximum social welfare and the social welfare of any NS coalition structure. Bilò et al. [19] study the price of NS in fractional hedonic games. Elkind et al. [58] consider the price of PO in additively separable, fractional, and modified fractional hedonic games. Brânzei and Larson [29] investigate the price of CS in *coalitional affinity games* which are equivalent to ASHGs. In Chapter 4, we study the price of *local fairness* in ASHGs.

Another recent branch of research studies the dynamics of deviations in hedonic games. Bilò et al. [19] study best-response Nash dynamics in fractional hedonic games. Hoefer et al. [77] analyze the impact of structural constraints (locality and externality) on the convergence in hedonic games. Carosi et al. [37] introduce *local core stability* and study the convergence of local core dynamics in simple symmetric fractional hedonic games. They also study the price of local core stability. Brandt et al. [27] investigate how deviations according to the notion of IS converge in various classes of hedonic games including anonymous and fractional hedonic games. Brandt et al. [28] study dynamics based on single-player deviations in ASHGs.

Further interesting research concerns the robustness of stability against the deletion of agents (*agent failure*) [79], strategyproof mechanisms that prevent strategical agent behavior [59], or hedonic games where the communication of the agents is restricted by an underlying graph such that agents can only form a coalition if they are connected in the graph [78].

Chapter 3

Altruism in Coalition Formation Games

In game theory, it is usually assumed that the agents are completely self-interested and act perfectly rational to accomplish their individual goals. Hence, the agents are assumed to always take those actions that lead them to their own optimal outcomes. This idea is related to the notion of the *homo economicus*. However, there has been some recent research from evolutionary biologists that shows that this approach is obsolete. In 2020, Hare and Woods [74] rephrased the Darwinian evolutionary thesis "survival of the fittest" with the thesis "survival of the friendliest". They studied the social behavior of several animal species, including dogs but also chimpanzees and bonobos. They observe that species with highly developed social skills and friendly behavior towards other individuals of their own and other species have an evolutionary advantage. They even argue that friendliness was essential for the success of the human species.

Along the same lines, there has been some research that attempts to integrate social aspects into models of cooperative game theory. Some authors introduce social aspects as altruism via a social network among the agents [6, 20, 76, 2]. Others directly integrate an agent's degree of selfishness or altruism into her utility function [75, 42, 3, 119]. Rothe [129] surveyed the approaches to altruism in game theory.

In the following three sections, we will study altruism in the scope of coalition formation games. In Section 3.1, we introduce *altruistic hedonic games* where agents are not narrowly selfish but also take the opinions of their friends into account when comparing two coalitions. We distinguish between several models of altruism and investigate them with respect to their axiomatic properties and the computational complexity of the associated decision problems. We continue our study in Sections 3.2, concentrating on the notions of popularity and strict popularity. In Section 3.3, we extend the models of altruism to the more general scope of coalition formation games and show that this extension brings some axiomatic advantages.

Related work Since the first introduction of altruistic hedonic games (see the preceding conference version [107] of the paper that we present in the next section [91]), there has appeared some follow-up research concerning aspects of altruism in hedonic games. For example, Schlueter and Goldsmith [130] introduce so-called *super altruistic hedonic games*. In their model, agents also behave altruistically towards agents that are further away in a social network but weight their altruistic consideration with their distances to them. This approach is related to the social distance games by Brânzei and Larson [30]. Bullinger and Kober [34] also generalize the preceding models of altruistic hedonic games. They introduce what they call *loyalty* in hedonic games. For any cardinal hedonic game, they consider agents to be loyal to any other agent that yields a positive utility when being with her in a coalition of size two. Miles [100] provides a useful online tool that can be used to simulate altruistic, friend-oriented, fractional, or additively separable hedonic games.

3.1 Altruistic Hedonic Games

This section is about the following journal article that introduces and studies *altruistic hedonic games*.

Publication (Kerkmann et al. [91])

A. Kerkmann, N. Nguyen, A. Rey, L. Rey, J. Rothe, L. Schend, and A. Wiechers. "Altruistic Hedonic Games". Submitted to the *Journal of Artificial Intelligence Research*. 2022

3.1.1 Summary

While previous literature on hedonic games focuses mainly on selfish players, this work introduces and studies *altruism in hedonic games*. The main idea while introducing our concepts of altruism is that players do not only care about their own valuations of coalitions but also about the valuations of others. We assume that the players have mutual friendship relations which are represented by a *network of friends*. We then assume that agents care about all their friends, i.e., their neighbors in the network of friends. When introducing the altruistic behavior, we incorporate the opinions of an agent's friends into her utility. While doing so, we make sure that the game is still hedonic: Agents only care about their own coalitions; hence, they only consider those friends that are in the same coalition. We focus on friend-oriented valuations of coalitions [50] and distinguish three degrees of altruism. First, we define a *selfish-first* degree where an agent first looks at her own friend-oriented valuation of a coalition and, only in the case that she values two coalitions the same, she looks at the her friends' valuations. Second, in the case of *equal-treatment* preferences, an agent treats herself and her friends in her coalition the same and aggregates all valuations with equal weights. Last, we introduce *altruistic-treatment* preferences where an agent first asks her friends for their valuations and, only in the case that her friends value two coalitions the same, she decides based on her own friend-oriented valuations. When aggregating the friends' valuations, we further distinguish between two aggregation methods. For *average-based* hedonic preferences, we aggregate the valuations by taking the average and, for *minimum-based* hedonic preferences, we aggregate by taking the minimum. This change of the aggregation function might seem minor but in fact makes a major difference in the altruistic behavior.

After introducing the different models of altruism in hedonic games, we differentiate them from the literature and study some axiomatic properties. In particular, our models can express preferences that can not be expressed by other models known from the related literature. Furthermore, they satisfy some desirable properties such as reflexivity, transitivity, polynomial-time computability of single preferences, and anonymity. After finishing our axiomatic study, we further consider the problems of verifying stable coalition structures in altruistic hedonic games and of deciding whether a stable outcome exists for a given altruistic hedonic game.

We study both problems for several common stability notions, such as Nash stability, core stability, and perfectness. While studying these problems, we not only concentrate on altruistic hedonic games where all agents act according to the same average-based or minimum-based degree of altruism but also consider the case of general altruistic hedonic games where each agent might individually behave according to a different degree of altruism. For selfish-first altruistic hedonic games, we provide a complete picture of the complexity of all considered problems. In particular, we show that there exist individually rational, Nash stable, individually stable, and contractually individually stable coalition structures for any altruistic hedonic game. For selfish-first altruistic hedonic games, even core stable and strictly core stable coalition structures are guaranteed to exist and the existence of perfect coalition structures can be decided in polynomial time. Concerning the verification problem, we proof that, for general altruistic hedonic games, individual rationality, Nash stability, individual stability, and contractual individual stability can be verified in polynomial time while core stability and strict core stability verification are coNP-complete. For selfish-first altruistic hedonic games, we further show that perfectness verification is in P.

3.1.2 Personal Contribution and Preceding Versions

This journal paper largely extends and improves multiple preceding conference papers that were published by Nhan-Tam Nguyen, Anja Rey, Lisa Rey, Jörg Rothe, and Lena Schend at AAMAS'16 [107], by Alessandra Wiechers and Jörg Rothe at STAIRS'20 [145], and by Jörg Rothe at AAAI'21 [129]. Parts of the AAMAS'16 paper were also presented at CoopMAS'16 [108] and COMSOC'16 [109]. Furthermore, Jörg Rothe and I presented some axiomatic properties of altruistic hedonic games at COMSOC'21 [86].

The modeling of the average-based altruistic hedonic games is due to the authors of the AAMAS'16 paper [107] and the modeling of the minimum-based variation is due to the authors of the STAIRS'20 paper [145].

My contributions are the merging and reorganization of the individual conference papers, additional related work in Section 2, the revision of various proofs from the AAMAS'16 paper [107] (the proofs of Propositions 4.2, 5.2, 6.14, and 6.17, Theorems 5.3, 5.5, 5.6, 6.5, and 6.6, and Lemma 6.1), additional visualizations (Figure 2 and Tables 2 and 4), the extension of various results to the more general case where agents might act according to different degrees of altruism, and additional results. In particular, I contributed all results concerning the properties of min-based altruistic preferences (see Section 5.3), the detailed results regarding the property of friend-dependence in Theorem 5.4, the results for type-I-monotonicity under average-based EQ and AL preferences in Theorem 5.6, Lemmas 6.3 and 6.13, Example 6.7, Theorem 6.9, Corollaries 6.10, 6.12, and 6.15, and the proofs of Propositions 6.11 and 6.16.

The writing of this journal paper was done jointly with all co-authors. The finalization and polishing were done by Jörg Rothe and me.

3.1.3 Publication

The full article [91] is appended here.

Journal of Artificial Intelligence Research 1 (2022) XX–XX Submitted X/22; published X/XX

Altruistic Hedonic Games

Anna Maria Kerkmann ANNA.KERKMANN@HHU.DE
Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany

Nhan-Tam Nguyen NHAN-TAM.NGUYEN@HHU.DE
Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany

Anja Rey ANJA.REY@HHU.DE
Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany

Lisa Rey LISA.REY@HHU.DE
Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany

Jörg Rothe ROTHE@HHU.DE
Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany

Lena Schend LENA.SCHEND@GOOGLEMAIL.COM
Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany

Alessandra Wiechers ALESSANDRA.WIECHERS@HHU.DE
Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany

Abstract

Hedonic games are coalition formation games in which players have preferences over the coalitions they can join. For a long time, all models of representing hedonic games were based upon selfish players only. Among the known ways of representing hedonic games compactly, we focus on friend-oriented hedonic games and propose a novel model for them that takes into account not only the players' own preferences but also their friends' preferences. Depending on the order in which players look at their own or their friends' preferences, we distinguish three degrees of altruism: selfish-first, equal-treatment, and altruistic-treatment preferences. We study both the axiomatic properties of these games and the computational complexity of problems related to various common stability concepts.

1. Introduction

The breathtakingly rapid development of artificial intelligence (AI) is largely based on mimicking—by means of tools, methods, and insights from computer science, mathematics, and other fields of science—human intelligence and human properties, attributes, and behavior as individuals and in society. Interaction among agents in a multiagent system—a key topic in AI—is typically modeled via game-theoretic means. From the early beginnings of (noncooperative) game theory due to von Neumann and Morgenstern (1944), a player (or agent—we will use the terms *player* and *agent* synonymously) in a game has been viewed as a *homo economicus*: Such players are perfectly rational, narrowly selfish, and interested only in maximizing their own gains, no matter what the costs to the other players are. In spirit, this assumption is somewhat related to Darwin's thesis of *"survival of the fittest,"* where "survival" essentially is measured by the ability of reproduction. However, even in terms of biology and evolution, there are reasonable doubts if selfishness alone (in the sense that more aggressive behavior yields more offspring) is really the key to success. Recently, 1 — 2 — 3   4

Figure 1: Example of a network of friends

Hare and Woods (2020) countered Darwin's thesis with their *"survival of the friendliest."* Specifically, one of their many arguments is that of the two species making up the genus Pan among the great apes, bonobos and chimpanzees, the bonobos benefit from their much friendlier behavior: The most successful male bonobo has more progenies than the most successful male chimpanzee, i.e., has a higher reproduction rate. Hare and Woods (2020) also argue that the evolutionary supremacy of the human species is mainly due to their *friendly* behavior, which made it possible for them to form larger social groups and even societies.

Now, if we agree that AI is best off when mimicking natural life and simulating real-world human behavior, the *homo economicus* from the early days of game theory is obsolete and better models are needed. Indeed, relentlessly aiming at one's own advantage and maximizing one's own utility regardless of the consequences for others may in fact not only diminish an agent's individual gains, but it may also harm the society of agents in a multiagent system as a whole. With this in mind, there have been some approaches of taking ethics, psychology, emotions, and behavioral dynamics into consideration in collective decision-making (Regenwetter, Grofman, Marley, & Tsetlin, 2006; Popova, Regenwetter, & Mattei, 2013; Rothe, 2019). This paper integrates *altruism* into the model of hedonic games.

Hedonic games, originally proposed by Drèze and Greenberg (1980) and later formally modeled by Bogomolnaia and Jackson (2002) and Banerjee, Konishi, and Sönmez (2001), are coalition formation games in which players have preferences over coalitions (subsets of players) they can be part of. In the context of decentralized coalition formation, several stability concepts and representations have been studied from an axiomatic and a computational complexity point of view; see the survey by Woeginger (2013a) on this topic and the book chapters by Aziz and Savani (2016) and Elkind and Rothe (2015) for an overview.

Dimitrov, Borm, Hendrickx, and Sung (2006) proposed a model that allows for compact representation of hedonic games, namely, the *friend-and-enemy encoding* of the players' preferences, where each player divides the other players into friends and enemies. Based on this encoding, they suggest two models of preference extensions: appreciation of friends and aversion to enemies. In *friend-oriented hedonic games*, a coalition A is preferred to another coalition B if A contains either more friends than B or the same number of friends as B but fewer enemies than B. This setting corresponds to a *network of friends* represented as a graph. We focus on the natural restriction of symmetric friendship relations and assume that the graph is undirected. For example, suppose there are four players, 1, 2, 3, and 4, and let 1 be friends with 2 but neither with 3 nor with 4, while 2 and 3 are friends with each other but not with 4. The corresponding network is displayed in Figure 1. Now, in the friend-oriented extension, player 2 prefers teaming up with 1 and 3 to forming a coalition with 1 and 4. Player 1, on the other hand, is indifferent between coalitions $\{1, 2, 3\}$ and $\{1, 2, 4\}$ because they both contain the one friend of 1's (namely, 2) and one of 1's enemies (either 3 or 4). However, following the paradigm of the *"survival of the friendliest,"* 1 can be expected to care about her friend 2's interests and thus might prefer a coalition in which 2 is satisfied ($\{1, 2, 3\}$) to one in which 2 is less satisfied ($\{1, 2, 4\}$). Indeed, player 1 would have a direct advantage of respecting 2's interests, since 2 and 3—being friends—can be expected to cooperate better than 2 and 4. In order to model such preferences, starting from friend-oriented hedonic games, we will introduce three degrees of altruism.

1.1 Our Contribution

Focusing on the friend-oriented extension of preferences due to Dimitrov et al. (2006) and considering the idea of players caring about their friends' preferences, we propose hedonic games with three degrees of altruistic influences: from being selfish first and considering one's friends' preferences to be of lower priority, over aggregating one's own opinions and those of one's friends equally, to truly altruistically letting one's friends decide first. The latter is the most altruistic case we consider, as we assume that from a player's perspective only friends can be consulted, while players further away (such as a friend's friend that is one's own enemy) cannot be communicated with or cannot be trusted or do not provoke the need to help. In a social network, for example, the whole set of players other than one's own friends might not even be known.[1]

Since we consider friends to be equally important, we first focus on their average valuation when comparing two coalitions. To distinguish the above-mentioned three degrees of altruism, we assign a sufficiently large weight either to a player's own valuation (the selfish-first case), or to the average valuation of this player's friends (the truly altruistic case), or to none of them (the equal-treatment case).

As an alternative, we also propose a minimum-based variant where we replace the average by the minimum in our previous definitions. As innocent as this small change appears to be, it is in fact as fundamental as considering egalitarian social welfare instead of utilitarian social welfare in multiagent resource allocation.[2] Minimum-based altruism may be more suitable than average-based altruism when the well-being of the entire group of agents crucially suffers from their unhappiest member.

All of the proposed games are compactly representable but not fully expressive. However, our representations of altruistic hedonic games can express *other* hedonic games than those expressible by different compact representations common in the literature. We provide a two-part study of these newly introduced games: First, we analyze the defined preferences

1. It may be debatable whether *"altruism"* is really the best term to capture our model. After all, even though the players' utilities for a coalition don't depend on their own preferences alone, they do not depend on *all* the players' preferences either but merely on *their own and their friends' preferences*—so one might be tempted to call this *"empathy among friends"* rather than *"altruism."* However, we have argued in the previous paragraph why it does make sense to consider only one's friends in the network. And even if our agents may not be completely selfless, they do behave altruistically toward their friends. Another important reason to not change the term *"altruistic hedonic game"* is that, meanwhile, quite a number of papers (listed in Sections 1.2 and 2.2.2) have adopted this term, so renaming it now would only cause confusion in the literature.
2. As noted by Nguyen, Nguyen, Roos, and Rothe (2014, p. 257), "*utilitarian social welfare* sums up the agents' individual utilities in a given allocation, thus providing a useful measure of the overall—and also of the average—benefit for society. For instance, in a combinatorial auction the auctioneer's aim is to maximize the auction's revenue (i.e., the sum of the prizes paid for the items auctioned), no matter which agent can realize which utility.

   In contrast, *egalitarian social welfare* gives the utility of the agent who is worst off in a given allocation, which provides a useful measure of fairness in cases where the minimum needs of all agents are to be satisfied."

with respect to axiomatic properties such as anonymity, monotonicity, and friend dependence; second, we consider various common stability concepts and show that our games always permit (contractually) individually stable and Nash stable solutions, and that testing whether a given solution is stable with respect to these concepts is tractable. We furthermore characterize when perfect solutions exist, and we analyze the computational complexity of the verification and the existence problem of core stable solutions.

1.2 Preliminary Conference and Workshop Versions

This paper merges and extends preliminary versions that appeared in the proceedings of several conferences: At the *15th International Conference on Autonomous Agents and Multiagent Systems* (AAMAS'16), Nguyen, Rey, Rey, Rothe, and Schend (2016) introduced altruistic hedonic games and established the first related results, which they also presented at the *7th International Workshop on Cooperative Games in Multiagent Systems* (CoopMAS'16) in Singapore and at the *6th International Workshop on Computational Social Choice* (COMSOC'16) in Toulouse, France (the latter two without archival proceedings). At the *9th European Starting AI Researcher Symposium* (STAIRS'20), Wiechers and Rothe (2020) introduced minimum-based altruistic hedonic games, and Rothe (2021) surveyed altruism in game theory for the senior-member track of the *35th AAAI Conference on Artificial Intelligence* (AAAI'21). Kerkmann and Rothe (2021) presented their study of the axiomatic properties of (minimum-based) altruistic hedonic games at the *8th International Workshop on Computational Social Choice* (COMSOC'21, without archival proceedings) in Haifa, Israel. We have extended these preliminary versions by merging them and adding many more examples, discussion, omitted proofs, and further results (including Example 6.7, Lemma 6.3, Theorems 5.4, 5.5, 5.6, and 6.9, and Corollaries 6.10 and 6.15).

1.3 Organization

In Section 2, we present related work, focusing on notions of altruism in noncooperative and cooperative games and on related literature on hedonic games. In Section 3, we give the basic definitions of hedonic games and some of the most common stability concepts. We formally introduce altruistic hedonic games in Section 4, discuss them in comparison with related but different notions from the literature, and study their axiomatic properties in Section 5. In Section 6, we deal with stability concepts and study the related existence and verification problems in terms of their complexity. Section 7 concludes the paper and raises some interesting open questions.

2. Related Work

In this section, we present related work, in particular regarding various ways of introducing notions of altruism into existing game-theoretic models, both in noncooperative and cooperative game theory (Sections 2.1 and 2.2). Since the literature about altruism in noncooperative games is older and richer than in cooperative games, we start with the former.

2.1 Altruism in Noncooperative Games

Game theory more or less started with the early papers by Borel (1921) and von Neumann (1928) and the book by von Neumann and Morgenstern (1944) who explored *noncooperative* games in which all players are on their own, competing with each other to win the game and to maximize their own profit. For more background on noncooperative game theory and its algorithmic aspects, the reader is referred, e.g., to the book edited by Nisan, Roughgarden, Tardos, and Vazirani (2007) and the book chapter by Faliszewski, Rothe, and Rothe (2015). Altruism in games has been considered mainly for noncooperative games to date. We give a short overview, starting with models of social components through a network of players.

2.1.1 Games with Social Networks

Ashlagi, Krysta, and Tennenholtz (2008) introduced *social context games* by embedding a strategic game into a *social context* that consists of a graph of neighborhood among the players and an aggregation function. The resulting social context game has the same players and strategies as the underlying strategic game. However, the players' payoffs in the resulting game do not only depend on their original payoffs but also on the neighborhood graph and the aggregation functions that express the social context. Ashlagi et al. (2008) focus on *resource selection games* (a famous subclass of *congestion games*[3]) as the underlying strategic games and on the following four social contexts: They obtain the payoffs of the social context game by either taking the minimum, maximum, or average of the players' and their neighbors' original payoffs (so-called *best-member*, *min-max*, or *surplus collaborations*) or they aggregate by so-called *competitive rankings*. Bilò, Celi, Flammini, and Gallotti (2013) apply the model of social context games by Ashlagi et al. (2008) to *linear congestion games* and *Shapley cost-sharing games* with the aggregation functions min, max, and sum (or average). They characterize the graph topologies modeling these social contexts such that the existence of pure Nash equilibria (as defined in Footnote 3) is guaranteed.

Hoefer, Penn, Polukarova, Skopalik, and Vöcking (2011) also consider players being embedded in a social network and assume that certain constraints specify which sets of coalitions may jointly deviate from their actual strategies in the game. When doing so, however, they assume that the players are *considerate* not to hurt others: Players ignore (i.e., choose to not carry out) potentially profitable group deviations whenever those would cause their neighbors' utilities to decrease. Exploring the properties of so-called *considerate equilibria* in resource selection games, Hoefer et al. (2011) show that there exists a state that is stable with respect to selfish and considerate behavior at the same time.

Anagnostopoulos, Becchetti, de Keijzer, and Schäfer (2013) study *altruism* and *spite* in strategic games. They consider directed weighted social networks where player i assigning a positive (negative) weight to player j means that i is altruistic (spiteful) towards j. They consider three classes of strategic games, namely, *congestion games*, *minsum scheduling games*, and *generalized second price auctions*, and study the *price of anarchy* (relating the worst-case cost of an equilibrium to the cost of an optimal outcome; see Koutsoupias and Papadimitriou (1999)) for these games.

3. A fundamental property of congestion games is that they always have a *Nash equilibrium in pure strategies*, i.e., there always exists a profile of pure strategies such that no player has an incentive to deviate from her strategy in the profile, provided the other players also stay with their strategies in the profile.

2.1.2 Games with an Altruistic Factor

We now turn to work that models altruism in strategic games by means of an altruistic factor that is integrated into the agents' cost or payoff functions.

Hoefer and Skopalik (2013) consider *atomic congestion games*. There are myopic selfish players and a set of resources, each with a nondecreasing delay function. Every player chooses a strategy by selecting (or allocating) a subset of resources, and experiences a delay corresponding to the total delay on all selected resources, which depends on the number of players that have allocated any of these resources. The goal of each player is to minimize the experienced delay. Now, altruism is introduced by Hoefer and Skopalik (2013) into such games as follows. They assume that the players are partly selfish and partly altruistic, which is formalized by an *altruism level* $\beta_i \in [0, 1]$ for each player i, where $\beta_i = 0$ means i is purely selfish and $\beta_i = 1$ means i is purely altruistic. These players' incentive is to optimize a linear combination of personal cost (their individually experienced delay) and social cost (the total delay of all players). Hoefer and Skopalik (2013) study under which conditions there exist pure Nash equilibria in various types of such games. They also show that *optimal stability thresholds* (the minimum number of altruists such that there exists an optimal Nash equilibrium) and *optimal anarchy thresholds* (the minimum number of altruists such that every Nash equilibrium is optimal) can be computed in polynomial time. Chen, de Keijzer, Kempe, and Schäfer (2014) study a similar model for *nonatomic congestion games*.

Apt and Schäfer (2014) introduce so-called *selfishness levels* for strategic games, which are based on the so-called "altruistic games" due to Ledyard (1995) (and, more recently, De Marco & Morgan, 2007). Selfishness levels measure the discrepancy between the social welfare in a Nash equilibrium and in a social optimum. After showing that their model is equivalent to some previous models of altruism due to Chen et al. (2014), Elias, Martignon, Avrachenkov, and Neglia (2010), and Caragiannis, Kaklamanis, Kanellopoulos, Kyropoulou, and Papaioannou (2010), Apt and Schäfer (2014) determine the selfishness levels of several well-studied strategic games, such as fair cost-sharing games, linear congestion games, the *n-player prisoner's dilemma*, the *n-player public goods game*, and the *traveler's dilemma game*. While these games have finite selfishness levels, Apt and Schäfer (2014) also show that other specific games like *Cournot competition*, *tragedy of the commons*, and *Bertrand competition* have an infinite selfishness level.

Rahn and Schäfer (2013) introduce yet another class of games, which they call *social contribution games*. They are motivated by the fact that altruistic behavior may actually render equilibria more inefficient (e.g., in congestion games) and may thus harm society as a whole (Anagnostopoulos et al., 2013). This is not the case for so-called valid utility games, though (Chen et al., 2014). Therefore, a question naturally arises: What is it that causes or influences the inefficiency of equilibria in games with altruistic players? In social contribution games, players' individual costs are set to the cost they cause for society just because of their presence, thus providing a useful abstraction of games with altruistic players when the robust price of anarchy is to be analyzed. Rahn and Schäfer (2013) in particular show that social contribution games are what they call *altruism-independently smooth*, which means that the robust price of anarchy in these games remains unaltered under arbitrary altruistic extensions.

2.2 Altruism in Cooperative Games

In a cooperative game, players may work together by forming groups, so-called *coalitions*, and may take joint actions so as to realize their goals better than if they were on their own. If a *coalition structure* (i.e., a partition of the players into coalitions) has formed, the question arises how stable it is, i.e., whether some players may have an incentive to leave their coalition and to join another one. There are plenty of special types of cooperative games and of stability notions some of which we will encounter below. For more background on cooperative game theory, the reader is referred, e.g., to the books by Peleg and Sudhölter (2003) and Chalkiadakis, Elkind, and Wooldridge (2011) and to the book chapter by Elkind and Rothe (2015).

2.2.1 Hedonic Games

Hedonic games are cooperative games with nontransferable utility. After Drèze and Greenberg (1980) introduced the concept, Banerjee et al. (2001) and Bogomolnaia and Jackson (2002) formally modeled them. In such coalition formation games, players have preferences over the coalitions they can be a member of.

Since every player in a hedonic game needs to rank (by a weak order) exponentially many (in the number of players) coalitions, it is crucial to find compact representations for these games. One such representation is the *friends-and-enemies encoding* by Dimitrov et al. (2006) where players partition the other players into two groups: *friends* and *enemies*. Based on this representation, they suggest two preference extensions. In the *friend-oriented preference extension*, which will be formally defined in Section 3, players prefer coalitions with more friends, and only if two coalitions have the same number of friends, players prefer to be with fewer enemies. In the *enemy-oriented preference extension*, on the other hand, players prefer coalitions with fewer enemies, and only if two coalitions have the same number of enemies, players prefer to be with more friends.

In addition to representation issues, much work has been done regarding the properties of hedonic games such as various notions of stability (see, e.g., Cechlárová & Hajduková, 2003, 2004; Aziz, Brandt, & Harrenstein, 2013a; Dimitrov et al., 2006) and studying the related problems in terms of their computational complexity (see, e.g., Ballester, 2004; Sung & Dimitrov, 2007, 2010; Woeginger, 2013b; Peters & Elkind, 2015; Peters, 2016). Needless to say that most of the papers just listed contribute to more than one of these goals; for example, Dimitrov et al. (2006) both introduce new ways of representing hedonic games and study their stability.

The friends-and-enemies encoding of Dimitrov et al. (2006) has inspired a lot of follow-up work. For example, Ota, Barrot, Ismaili, Sakurai, and Yokoo (2017) allow for neutral agents in addition to friends and enemies and study their impact on (strict) core stability. Also considering friends, enemies, and neutral agents, Kerkmann, Lang, Rey, Rothe, Schadrack, and Schend (2020) propose a bipolar extension of the responsive extension principle and use it to derive *partial* preferences over coalitions, characterize coalition structures that necessarily or possibly satisfy certain stability concepts, and study the related problems in terms of their complexity. Barrot, Ota, Sakurai, and Yokoo (2019) study the impact of additional *unknown* agents, and Rey, Rothe, Schadrack, and Schend (2016) study wonderful stability (a.k.a. perfectness) and strict core stability in enemy-oriented hedonic games. Peters and

Elkind (2015) establish metatheorems that help proving NP-hardness results for the problem of checking whether given hedonic games admit stable coalition structures.

For more background on hedonic games, the reader is referred to the book chapters by Aziz and Savani (2016) and Elkind and Rothe (2015) and to the excellent survey by Woeginger (2013a).

2.2.2 Altruism in Hedonic Games

Since we first introduced *altruistic hedonic games* in 2016 (Nguyen et al., 2016),[4] there has been some follow-up literature on this topic.

Schlueter and Goldsmith (2020) introduced *super altruistic hedonic games* and studied them with respect to various stability notions. In their model, players in the same coalition have a different impact on a player based on their distances in the underlying network of friends. Their model is also related to the *social distance games* by Brânzei and Larson (2011) where the utility of a player is defined by measuring her distance to the other members of her coalition.

Kerkmann and Rothe (2020) extend the altruistic hedonic games to coalition formation games in general. While our model is *hedonic* in the sense that players' preferences only depend on the coalitions that they are part of, in their model players behave altruistically also towards their friends in other coalitions, which makes their games nonhedonic.

Recently, Bullinger and Kober (2021) introduced *loyalty in cardinal hedonic games.* In their model, each player has a *loyalty set* that contains all players for which she has positive utility when being together with this player in a coalition of size two. The *loyal variant* of a cardinal hedonic game is then defined to have the same players as the original game but with redefined utilities. More specifically, the utility of a player for a coalition structure is derived by taking the minimum of her own original utility and the original utilities of all players in her loyalty set who are also in her coalition. Bullinger and Kober (2021) also define a *k-fold loyal variant* where the loyal variant is applied k times. Note that our min-based altruistic hedonic games with equal-treatment preferences are equivalent to their loyal variant of symmetric friend-oriented hedonic games.

Based on the social context games (Ashlagi et al., 2008) described in Section 2.1.1, Monaco, Moscardelli, and Velaj (2018) introduced *social context hedonic games* (which, in fact, are *nonhedonic* games). Such games are based on additively separable utilities, an altruism factor, and a social network among the players. A player's utility for a coalition structure is then defined to be the sum of her own additively separable utility for her coalition and the utilities of her neighbors in the network, the latter weighted by the altruism factor.

Remotely related to our min-based altruistic hedonic games is the work of Monaco, Moscardelli, and Velaj (2019) who study the *modified fractional hedonic games* introduced by Olsen (2012). These games behave qualitatively different than the fractional hedonic games due to Aziz, Brandl, Brandt, Harrenstein, Olsen, and Peters (2019). In particular, Monaco et al. (2019) study the performance of Nash (and, to some extent, core) stable outcomes in modified fractional hedonic games with egalitarian social welfare.

4. (Nguyen et al., 2016) is one of the preliminary conference versions of the present paper (recall Section 1.2) and has thus been incorporated into it.

3. Preliminaries

A *hedonic game* is given by a pair $(N, \succeq)$, where $N = \{1, \dots, n\}$ is a set of players and $\succeq = (\succeq_1, \dots, \succeq_n)$ is a list of the players' preferences. For $i \in N$, let $\mathcal{N}^i = \{C \subseteq N \mid i \in C\}$ denote the set of coalitions containing i. Player i's preference relation $\succeq_i \subseteq \mathcal{N}^i \times \mathcal{N}^i$ induces a complete, weak preference order over $\mathcal{N}^i$. For $A, B \in \mathcal{N}^i$, we say that player i *weakly prefers A to B* if $A \succeq_i B$, that i *prefers A to B* ($A \succ_i B$) if $A \succeq_i B$ but not $B \succeq_i A$, and that i is *indifferent between A and B* ($A \sim_i B$) if $A \succeq_i B$ and $B \succeq_i A$. We call $C \in \mathcal{N}^i$ *acceptable for player i* if $C \succeq_i \{i\}$. A *coalition structure* is a partition $\Gamma = \{C_1, \dots, C_k\}$ of the players into k coalitions $C_1, \dots, C_k \subseteq N$ (i.e., $\bigcup_{r=1}^{k} C_r = N$ and $C_r \cap C_s = \emptyset$ for all distinct $r, s \in \{1, \dots, k\}$). The unique coalition in Γ containing player $i \in N$ is denoted by $\Gamma(i)$. The set of all coalition structures for a set of agents N is denoted by $\mathcal{C}_N$. For two coalition structures $\Gamma, \Delta \in \mathcal{C}_N$, we say that agent i prefers Γ to Δ if i prefers $\Gamma(i)$ to $\Delta(i)$ (and analogously so for weak preference and indifference).

3.1 Friend-Oriented Preference Extension

In order to avoid preference orders that are exponentially long in the number of players, a common way to represent players' preferences is to consider a *network of friends* (Dimitrov et al., 2006): Every player $i \in N$ has a set of friends $F_i \subseteq N \setminus \{i\}$ and a set of enemies $E_i = N \setminus (F_i \cup \{i\})$. Visually, we represent the players in N by the vertices in a graph $G = (N, H)$, and let a directed edge $(i, j) \in H$ denote that j is i's friend, that is, the open neighborhood of i represents the set of i's friends $F_i = \{j \mid (i, j) \in H\}$. Since in the context of stability it is reasonable to consider symmetric friendship relations only (as noted, e.g., by Woeginger, 2013a), we will focus on undirected graphs representing networks of friends.

In the *friend-oriented* preference extension (Dimitrov et al., 2006), players prefer coalitions with more friends, and only if two coalitions have the same number of friends, players prefer to be with fewer enemies. Formally, define

$$A \succeq_i^F B \iff |A \cap F_i| > |B \cap F_i| \text{ or } (|A \cap F_i| = |B \cap F_i| \text{ and } |A \cap E_i| \leq |B \cap E_i|). \tag{1}$$

Note that friend-oriented preferences can be represented additively, by assigning a value of $n = |N|$ to each friend and a value of -1 to each enemy (Dimitrov et al., 2006): For any player $i \in N$ and any coalition $A \in \mathcal{N}^i$, define the *value of a coalition* by

$$v_i(A) = n|A \cap F_i| - |A \cap E_i|. \tag{2}$$

It then holds that $-(n-1) \leq v_i(A) \leq n(n-1)$, and $v_i(A) > 0$ if and only if $|A \cap F_i| > 0$. For $A, B \in \mathcal{N}^i$, we have

$$A \succeq_i^F B \iff v_i(A) \geq v_i(B). \tag{3}$$

3.2 Stability Concepts

The following stability concepts are commonly studied in hedonic games (Aziz & Savani, 2016). The relations between these concepts are illustrated in Figure 2.

Definition 3.1. *Let $(N, \succeq)$ be a hedonic game and Γ a coalition structure. A coalition $C \subseteq N$* blocks Γ *if for each $i \in C$ it holds that $C \succ_i \Gamma(i)$. If there is at least one $i \in C$*

$$\begin{tikzcd}
 & \text{NS} \arrow[r] & \text{IS} \arrow[r] \arrow[dr] & \text{CIS} \\
\text{perfect} \arrow[ur] \arrow[r] & \text{SCS} \arrow[r] \arrow[ur] & \text{CS} \arrow[r] & \text{IR}
\end{tikzcd}$$

Figure 2: Relations between the stability notions defined in Section 3.2. In this figure, there is a directed path from notion A to notion B if and only if A implies B.

with $C \succ_i \Gamma(i)$ while $C \succeq_j \Gamma(j)$ holds for the other players $j \neq i$ in C, we call C weakly blocking. *A coalition structure Γ is said to be*

1. individually rational *(IR) if for all $i \in N$, $\Gamma(i)$ is acceptable;*
2. Nash stable *(NS) if for all $i \in N$ and for each $C \in \Gamma \cup \{\emptyset\}$ with $\Gamma(i) \neq C$, it holds that $\Gamma(i) \succeq_i C \cup \{i\}$;*
3. individually stable *(IS) if for all $i \in N$ and for each $C \in \Gamma \cup \{\emptyset\}$, it either holds that $\Gamma(i) \succeq_i C \cup \{i\}$ or there is a player $j \in C$ with $C \succ_j C \cup \{i\}$;*
4. contractually individually stable *(CIS) if for all $i \in N$ and for each $C \in \Gamma \cup \{\emptyset\}$, it either holds that $\Gamma(i) \succeq_i C \cup \{i\}$, or there is a player $j \in C$ with $C \succ_j C \cup \{i\}$, or there is a player $k \in \Gamma(i)$ with $i \neq k$ and $\Gamma(i) \succ_k \Gamma(i) \setminus \{i\}$;*
5. core stable *(CS) if there is no nonempty coalition that blocks Γ;*
6. strictly core stable *(SCS) if there is no coalition that weakly blocks Γ; and*
7. perfect *if for all $i \in N$ and for all $C \in \mathcal{N}^i$, it holds that $\Gamma(i) \succeq_i C$.*

4. Altruistic Hedonic Games

In this section, we introduce our new model that refines friend-oriented hedonic games by taking altruistic influences into account. In this model, players still want to be with as many friends (and, secondarily, with as few enemies) as possible, but in addition they want their friends to be as satisfied as possible.

4.1 Failure of a Naïve Approach

A first attempt to formalize this idea (that will turn out to fail) is the following. Consider the scenario where $i \in N$ has a friend-oriented preference extension (according to Equivalence (1)) except that, whenever the number of friends in A and B is the same and so is the number of enemies in A and B (i.e., $A \sim_i^F B$), i now prefers A to B if more of i's friends that are contained in A and B prefer A to B than B to A (again according to Equivalence (1)). Formally:

$$\begin{aligned}
A \succeq_i^{NA} B \iff & |A \cap F_i| > |B \cap F_i| \text{ or} \\
& (|A \cap F_i| = |B \cap F_i| \text{ and } |A \cap E_i| < |B \cap E_i|) \text{ or} \\
& (|A \cap F_i| = |B \cap F_i| \text{ and } |A \cap E_i| = |B \cap E_i| \text{ and} \\
& \ |\{j \in A \cap B \cap F_i \mid A \succ_j^F B\}| \geq |\{j \in A \cap B \cap F_i \mid B \succ_j^F A\}|).
\end{aligned} \tag{4}$$

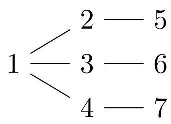


Figure 3: Network of friends representing the hedonic game in Example 4.1

Intuitively, according to (4), a player is selfish first, but as soon as she is indifferent between two coalitions in the sense of (1), she cares about her friends' preferences. A major disadvantage of this definition, however, is that *irrational* preference orders can arise, i.e., preference orders that are not transitive in general, as the following example shows.

Example 4.1. *Consider the hedonic game* $(N, \succeq^{NA})$ *with* $N = \{1, 2, 3, 4, 5, 6, 7\}$ *and the network of friends from Figure 3. For coalitions* $A = \{1, 2, 3, 5\}$*,* $B = \{1, 2, 4, 7\}$*, and* $C = \{1, 3, 4, 6\}$*, it holds that* $A \succ_1^{NA} B$ *and* $B \succ_1^{NA} C$*, yet* $C \succ_1^{NA} A$*, violating transitivity.*

In order to ensure transitivity, we have to add an extra condition to Equivalence (4). One idea would be to demand indifference between all coalitions that are involved in a $\succ_i^{NA}$-cycle by (4). This, however, can lead to a comparison of all coalitions containing a player, so determining a relation between two coalitions might comprise an exponential number of steps in the number of players. Then it would have been easier to give an arbitrary preference order as an input in the first place. Another idea would be to include the preferences of all friends, not only of those contained in the considered coalitions, but this would still lead to preference orders that are not transitive and would also contradict the concept of hedonic games. In the following, we take a different approach that does not violate transitivity.

4.2 Modeling Altruism Based on the Friend-Oriented Preference Extension

Given the failure of extending friend-oriented preferences by breaking ties with "majority voting," we consider the following model instead: When comparing two coalitions A and B, player i considers two aspects. First, i takes her own friend-oriented value for the coalitions into account and, second, she also incorporates the opinions of her friends in A and B. As i incorporates her friends' opinions, she aggregates their friend-oriented values by either taking the average or the minimum. While the first variant gives equal weights to all her friends, the second variant is better motivated in situations when i wishes to improve the satisfaction of the friend that is worst off because she would always suffer with her unhappiest friend.

Recall that player i's value for coalition $A \in \mathcal{N}^i$ in the friend-oriented encoding is given by $v_i(A) = n|A \cap F_i| - |A \cap E_i|$. We denote the average value of i's friends in A and the average value of i and her friends in A by

$$\mathrm{avg}_i^F(A) = \sum_{a \in A \cap F_i} \frac{v_a(A)}{|A \cap F_i|} \qquad \text{and} \qquad \mathrm{avg}_i^{F+}(A) = \sum_{a \in (A \cap F_i) \cup \{i\}} \frac{v_a(A)}{|(A \cap F_i) \cup \{i\}|}. \tag{5}$$

Note that normalization by the number of i's friends in a coalition prevents a "tyranny of the many" (otherwise, large coalitions might be preferred merely because they contain more friends). Similarly, we denote the corresponding minimum values by

$$\min_i^F(A) = \min_{a \in A \cap F_i} \{v_a(A)\} \qquad \text{and} \qquad \min_i^{F+}(A) = \min_{a \in (A \cap F_i) \cup \{i\}} \{v_a(A)\} \tag{6}$$

where the minimum of the empty set is defined as zero.

Assigning a weight to player i's own contribution in comparison to her friends' influence on her preferences, we will distinguish between three *degrees of altruism*:

(a) Selfish First (SF): A player is selfish first and asks her friends only in case of indifference, i.e., initially she decides which of two coalitions she prefers friend-orientedly, and if and only if she is indifferent between them, she asks her friends for a vote. For a constant $M \geq n^2$, we use the utility functions

$$u_i^{avgSF}(A) = M \cdot v_i(A) + \mathrm{avg}_i^F(A) \text{ and} \tag{7}$$

$$u_i^{minSF}(A) = M \cdot v_i(A) + \min_i^F(A) \tag{8}$$

to define agent i's

- *avg-based SF altruistic preferences* by $A \succeq_i^{avgSF} B \iff u_i^{avgSF}(A) \geq u_i^{avgSF}(B)$;
- *min-based SF altruistic preferences* by $A \succeq_i^{minSF} B \iff u_i^{minSF}(A) \geq u_i^{minSF}(B)$.

(b) Equal Treatment (EQ): A player's and her friends' friend-oriented opinions are treated equally for the decision. For the utility functions

$$u_i^{avgEQ}(A) = \mathrm{avg}_i^{F+}(A) \text{ and} \tag{9}$$

$$u_i^{minEQ}(A) = \min_i^{F+}(A), \tag{10}$$

define agent i's

- *avg-based EQ altruistic preferences* by $A \succeq_i^{avgEQ} B \iff u_i^{avgEQ}(A) \geq u_i^{avgEQ}(B)$;
- *min-based EQ altruistic preferences* by $A \succeq_i^{minEQ} B \iff u_i^{minEQ}(A) \geq u_i^{minEQ}(B)$.

(c) Altruistic Treatment (AL): A player first asks her friends for their opinions on a coalition they are contained in and adopts their average or minimum value; if and only if the consensus is indifference, the player decides for herself. For $M \geq n^4$, we use the utility functions

$$u_i^{avgAL}(A) = v_i(A) + M \cdot \mathrm{avg}_i^F(A) \text{ and} \tag{11}$$

$$u_i^{minAL}(A) = v_i(A) + M \cdot \min_i^F(A) \tag{12}$$

to define agent i's

- *avg-based AL altruistic preferences* by $A \succeq_i^{avgAL} B \iff u_i^{avgAL}(A) \geq u_i^{avgAL}(B)$;
- *min-based AL altruistic preferences* by $A \succeq_i^{minAL} B \iff u_i^{minAL}(A) \geq u_i^{minAL}(B)$.

The next proposition shows that the definitions of the SF and AL preferences indeed capture the intuitive ideas behind them: In the case of SF preferences, the own value is the first decisive factor, and in the case of AL preferences, the friends' values are the first decisive factor.

Proposition 4.2. *For $M \geq n^4$, the following statements hold for each $i \in N$ and for any two coalitions $A, B \in \mathcal{N}^i$:*

1. $v_i(A) > v_i(B)$ *implies* $A \succ_i^{avgSF} B$,
2. $v_i(A) > v_i(B)$ *implies* $A \succ_i^{minSF} B$,
3. $\mathrm{avg}_i^F(A) > \mathrm{avg}_i^F(B)$ *implies* $A \succ_i^{avgAL} B$, *and*
4. $\min_i^F(A) > \min_i^F(B)$ *implies* $A \succ_i^{minAL} B$.

Proof. We just state the proofs for statements (1) and (3). The proofs for (2) and (4) are quite similar.

We start with statement (1). The claim clearly holds for $\mathrm{avg}_i^F(A) \geq \mathrm{avg}_i^F(B)$. For $\mathrm{avg}_i^F(A) < \mathrm{avg}_i^F(B)$, it holds if and only if $M > \frac{\mathrm{avg}_i^F(B) - \mathrm{avg}_i^F(A)}{v_i(A) - v_i(B)}$. The numerator is upper-bounded by $|B \cap F_i| \cdot \frac{n(n-1)}{|B \cap F_i|} - |A \cap F_i| \cdot \frac{-(n-1)}{|A \cap F_i|} = n^2 - 1$. For the denominator, we have $v_i(A) - v_i(B) > 0$. Since $v_i(A)$ and $v_i(B)$ are integral, $v_i(A) - v_i(B) \geq 1$. Thus $M > n^2 - 1$ suffices.

We now turn to statement (3). The claim clearly holds for $v_i(A) \geq v_i(B)$. For $v_i(A) < v_i(B)$, the claim holds if and only if $M > \frac{v_i(B) - v_i(A)}{\mathrm{avg}_i^F(A) - \mathrm{avg}_i^F(B)}$. The numerator is upper-bounded by $n(n-1) + (n-1) = n^2 - 1 < n^2$.

We further show that the denominator is lower-bounded by $\frac{1}{n^2}$: First, for the sake of readability, let $\alpha = \sum_{a \in A \cap F_i} v_a(A)$ and $\beta = \sum_{b \in B \cap F_i} v_b(B)$. Then α and β are integral by the integrality of v_a and v_b. Note that the premise $\mathrm{avg}_i^F(A) > \mathrm{avg}_i^F(B)$ is equivalent to $\frac{\alpha}{|A \cap F_i|} > \frac{\beta}{|B \cap F_i|}$. This implies $\alpha|B \cap F_i| - \beta|A \cap F_i| \geq 1$, since α, β, $|A \cap F_i|$, and $|B \cap F_i|$ are each integral. Thus

$$\mathrm{avg}_i^F(A) - \mathrm{avg}_i^F(B) = \frac{\alpha}{|A \cap F_i|} - \frac{\beta}{|B \cap F_i|} = \frac{\alpha|B \cap F_i| - \beta|A \cap F_i|}{|A \cap F_i||B \cap F_i|} \geq \frac{1}{|A \cap F_i||B \cap F_i|} \geq \frac{1}{n^2}.$$

Overall, $M \geq n^4$ suffices for statement (3). ❑

Now, an altruistic hedonic game (AHG) is a hedonic game where the preference profile consists of any mixture of avg-based and min-based SF, EQ, and AL altruistic preferences. The subclasses of AHGs where all agents have the same type of altruistic preferences are, e.g, called avg-based SF AHGs (with avg-based SF preferences), min-based AHGs (with min-based SF, EQ, or AL preferences), or EQ AHGs (with avg- or min-based EQ preferences), etc. We will sometimes abuse notation and just write u_i for player i's utility (or $\succeq_i$ for i's preference) when the considered altruistic model is clear from the context or when we talk about multiple models.

The following examples illustrate the different approaches to altruism in hedonic games. We start with explaining the three avg-based preferences in Example 4.3.

Example 4.3. *Consider a game with five agents where the network of friends forms a path as shown in Figure 4a. Table 4b gives an overview of the relevant values and average values needed to determine player* 1*'s utilities for a number of acceptable coalitions depending on the degree of altruism. A dash indicates that a value does not exist.*

It can be seen that the friend-oriented preference and all three avg-based altruistic preferences are different. Under the friend-oriented preference extension (1)*, player* 1*'s weak*

4 — 3 — 1 — 2 — 5

(a) Network of friends

C	$\{1,2,3\}$	$\{1,2,3,4\}$	$\{1,2,3,5\}$	N	$\{1,2\}$	$\{1,3\}$	$\{1,2,4\}$	$\{1,2,5\}$	$\{1,3,4\}$
$v_1(C)$	10	9	9	8	5	5	4	4	4
$v_2(C)$	4	3	9	8	5	–	4	10	–
$v_3(C)$	4	9	3	8	–	5	–	–	10
$\mathrm{avg}_1^F(C)$	4	6	6	8	5	5	4	10	10
$\mathrm{avg}_1^{F+}(C)$	6	7	7	8	5	5	4	7	7

(b) Player 1's average values for certain coalitions

Figure 4: Different approaches to altruism in the avg-based AHG from Example 4.3

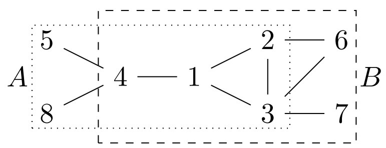


Figure 5: Network of friends with coalitions A and B in Example 4.4

preference order $\succeq_1^F$ *is given in the first line according to the values of* v_1. *For* avg-based SF preferences*, the order remains the same; however, indifferences are resolved based on the average of the friends' values* $\mathrm{avg}_1^F(C)$*, as is the case here with* $\{1,2,5\} \succ_1^{avgSF} \{1,2,4\}$*. Under* avg-based EQ preferences, $\mathrm{avg}_1^{F+}(C)$ *is considered and the grand coalition is the most preferred one; intuitively, because all friends have a large number of friends at the same time. Finally, under* avg-based AL preferences*, the average of player* 1*'s friends* $\mathrm{avg}_1^F(C)$ *is considered first. Player* 1*'s friends consider* $\{1,2,5\}$ *and* $\{1,3,4\}$ *to be the best coalitions. As player* 1 *values these two coalitions the same, she adopts this opinion and is indifferent between them. Player* 1*'s friends are also indifferent between* $\{1,2,3\}$ *and* $\{1,2,4\}$*. Since player* 1 *assigns a higher value to* $\{1,2,3\}$*, she resolves this tie with* $\{1,2,3\} \succ_1^{avgAL} \{1,2,4\}$*.*

The next example shows that the min-based preferences are different from the avg-based preferences.

Example 4.4. *Let* $N = \{1,\dots,8\}$ *be the set of players with the network of friends displayed in Figure 5. Calculating the values of players* 1, 2, 3, *and* 4 *for the coalitions* $A = \{1,2,3,4,5,8\}$ *and* $B = \{1,2,3,4,6,7\}$ *reveals that* $v_1(A) = v_1(B) = v_2(B) = v_4(A) = 3\cdot 8 - 2 = 22$, $v_2(A) = v_3(A) = 2\cdot 8 - 3 = 13$, $v_3(B) = 4\cdot 8 - 1 = 31$*, and* $v_4(B) = 1\cdot 8 - 4 = 4$*. The resulting avg-based and min-based utilities of player* 1 *for these two coalitions are shown in Table 1. They reveal that, for all three degrees of altruism, player* 1 *prefers* A *to* B *when taking the minimum, yet prefers* B *to* A *when taking the average.*

	u_1^{avgSF}	u_1^{avgEQ}	u_1^{avgAL}	u_1^{minSF}	u_1^{minEQ}	u_1^{minAL}
A	$22M + 16$	17.5	$22 + 16M$	$22M + 13$	13	$22 + 13M$
B	$22M + 19$	19.75	$22 + 19M$	$22M + \ \ 4$	4	$22 + \ \ 4M$

Table 1: Player 1's utilities for coalitions A and B in Example 4.4

5. Properties of Altruistic Hedonic Games

In this section, we study which desirable properties are satisfied by altruistic preferences. First, however, we start with a short discussion of expressiveness and explain how our models differ from other representations known from the literature.

5.1 Expressiveness and a Short Discussion of Our and Other Models

First, our models are not fully expressive because players are indifferent between friends and enemies, respectively. Further, for coalitions that only consist of enemies, all our altruistic preference extensions correlate with the original definition of friend-oriented preferences.

Second, we focus on avg-based EQ preferences and show that their expressiveness is incomparable to additively separable hedonic games, fractional hedonic games, hedonic games with $\mathcal{B}$- or $\mathcal{W}$-preferences, and FEN-hedonic games (for the definitions of these representations, see Aziz, Brandt, & Seedig, 2013b; Aziz et al., 2019; Cechlárová & Hajduková, 2003, 2004; Kerkmann et al., 2020). Note that in all of the above models two players' preference orders are independent from each other, but in our model they might depend on each other. Players are free in making friends; however, the induced preferences crucially depend on their friends' relations to other players—indeed, this is the key point of introducing our model of altruism. In other words, players' preferences are constrained by their friends' preferences. More concretely, the avg-based EQ extension can express preferences that are not additively separable:

Example 5.1. *Consider a game with three players ($N = \{1, 2, 3\}$, so $n = 3$) where the network of friends is a path: 1 — 2 — 3. Recall that for any $i \in N$ and $A \in \mathcal{N}^i$, we have*

$$u_i^{avgEQ}(A) = \text{avg}_i^{F+}(A) = \sum_{a \in A \cap (F_i \cup \{i\})} \frac{n|A \cap F_a| - |A \cap E_a|}{|A \cap (F_i \cup \{i\})|}.$$

Then $u_1^{avgEQ}(\{1, 2\}) = \frac{3+3}{2} = 3$ and $u_1^{avgEQ}(\{1, 2, 3\}) = \frac{(3-1)+3\cdot 2}{2} = 4$. Consequently, $\{1, 2, 3\} \succ_1^{avgEQ} \{1, 2\}$, but $\{1\} \succ_1^{avgEQ} \{1, 3\}$ because agent 3 is an enemy of 1's, and $\{1, 2\} \succ_1^{avgEQ} \{1\}$ because agent 2 is a friend of 1's.

In particular, the preferences considered in Example 5.1 cannot be expressed in additively separable hedonic games. However, additively separable preferences can express strict preferences over coalitions of size two that only contain the considered agent and a single additional agent. This is not possible for EQ preferences because of indifference between friends and enemies, respectively. (Comparing three coalitions of size two under EQ preferences, there will always be an indifference.)

Similarly, fractional, FEN-hedonic, $\mathcal{B}$- and $\mathcal{W}$-preferences, can express strict preferences over three size-two coalitions but cannot express the avg-based EQ preference of agent 1 from Example 5.1.

Overall, neither are avg-based EQ preferences more expressive than any of the other considered models nor the other way around. Similar examples also exist for the SF and AL extensions and the min-based model.

5.2 Properties of Preference Extensions

Next, we give a selection of properties of preference extensions that are inspired by properties from various related fields such as social choice theory and resource allocation, which also are concerned with preferences, and we adapt them appropriately to our setting.

Let $N = \{1, \dots, n\}$ be a set of players and F_i and E_i the sets of player i's friends and enemies, respectively. Let $G = (N, H)$ be the corresponding network of friends. Let $G' = (N', H')$ be some network of friends that is isomorphic to G by the isomorphism $\varphi : N \to N'$. Consider player i's preference relation $\succeq_i$ on $\mathcal{N}^i$ and $\varphi(i)$'s preference relation $\succeq'_{\varphi(i)}$ on $\mathcal{N}'^{\varphi(i)}$ that were deduced from G and G' by the same preference extension.

We say $\succeq_i$ is *reflexive* if $A \succeq_i A$ for each coalition $A \in \mathcal{N}^i$; $\succeq_i$ is *transitive* if for any three coalitions $A, B, C \in \mathcal{N}^i$, $A \succeq_i B$ and $B \succeq_i C$ implies $A \succeq_i C$; $\succeq_i$ is *polynomial-time computable* if for two given coalitions $A, B \in \mathcal{N}^i$, it can be decided in polynomial time whether or not $A \succeq_i B$; and $\succeq_i$ is *anonymous* if renaming the players in N does not change the structure of i's preference, i.e., if for any two coalitions $A, B \in \mathcal{N}^i$, it holds that $A \succeq_i B$ if and only if $\{\varphi(a) \mid a \in A\} \succeq'_{\varphi(i)} \{\varphi(b) \mid b \in B\}$.

Clearly, the first three properties are necessary to have efficiently computable and rational preferences, and anonymity means that only the structure of the friendship network is important. We further define the following properties.

Weak Friend-Orientedness: If coalition A is acceptable for i, then $A \cup \{f\}$ is also acceptable for i, where $f \in F_i \setminus A$.

Favoring Friends: If $x \in F_i$ and $y \in E_i$ then $\{x, i\} \succ_i \{y, i\}$.

Indifference between Friends: If $x, y \in F_i$ then $\{x, i\} \sim_i \{y, i\}$.

Indifference between Enemies: If $x, y \in E_i$ then $\{x, i\} \sim_i \{y, i\}$.

Note that these four properties hold for friend-oriented preferences, see the work of Alcantud and Arlegi (2012).[5] The next property is inspired by the property *"citizens' sovereignty"* from social choice theory, which says that only the voters shall decide on who has won an election, so for a voting rule to satisfy this property it is required that every candidate can be made a winner for suitably chosen voter preferences (see, e.g., Baumeister & Rothe, 2015).[6] Similarly, we require that only the players shall decide on which coalitions turn out to be their most preferred ones, under a suitably chosen network of friends.

5. Alcantud and Arlegi (2012) define so-called weighted GNB rankings (where objects are classified into three categories: good, neutral, and bad), which are a generalization of friend-oriented preferences in hedonic games.
6. This property is also known as *non-imposition*.

Sovereignty of Players: For a fixed player i and each $C \in \mathcal{N}^i$, there exists a network of friends such that C ends up as i's most preferred coalition.

We now introduce two types of monotonicity. *Type-I-monotonicity* ensures that if i (weakly) prefers A over B, this should still be true after an enemy j of i's, who is contained in both coalitions and weakly prefers A to B friend-orientedly, turns into i's friend. *Type-II-monotonicity* is similarly defined but requires that j is only in A (hence has no opinion on B or its relation to A), but still i's preference of A over B should not be altered by j turning from an enemy of i's into i's friend.

Monotonicity: Let $j \neq i$ be a player with $j \in E_i$ and let $A, B \in \mathcal{N}^i$. Let further $\succeq_i'$ be the preference relation resulting from $\succeq_i$ when j turns from being i's enemy into being i's friend (all else remaining equal). We call $\succeq_i$

- *type-I-monotonic* if it holds that (1) if $A \succ_i B$, $j \in A \cap B$, and $A \succeq_j^F B$, then $A \succ_i' B$, and (2) if $A \sim_i B$, $j \in A \cap B$, and $A \succeq_j^F B$, then $A \succeq_i' B$.
- *type-II-monotonic* if it holds that (1) if $A \succ_i B$ and $j \in A \setminus B$, then $A \succ_i' B$, and (2) if $A \sim_i B$ and $j \in A \setminus B$, then $A \succeq_i' B$.

The next property is *local friend dependence*. It says that an agent's preferences over some coalitions can change if the sets of this agent's friends' friends change. These friends also have to be members of the coalition that is under consideration. Thus local friend dependence is a crucial property that characterizes the essence of the proposed altruistic preferences and distinguishes them from previous models, e.g., from *additively separable* (Aziz et al., 2013b) or *friend-oriented* preferences (Dimitrov et al., 2006).

Local Friend Dependence: The preference order $\succeq_i$ can depend on the sets of friends $F_1, \dots, F_n$ of some agents. Let $A, B \in \mathcal{N}^i$. We say that comparison (A, B) is

- *friend-dependent in* $\succeq_i$ if $A \succeq_i B$ is true (or false) and can be made false (or true) by changing the set of friends of some players in $N \setminus \{i\}$ (while not changing any relation to i);
- *locally friend-dependent in* $\succeq_i$ if $A \succeq_i B$ is true (or false), can be made false (or true) by changing the set of friends of some players in $(A \cup B) \cap F_i$ (while not changing any relation to i), and changing the set of friends of any of the other players in $N \setminus (\{i\} \cup (F_i \cap (A \cup B)))$ (while not changing any relation to any player in $\{i\} \cup (F_i \cap (A \cup B)))$ does not affect the status of the comparison.

We say $\succeq_i$ is *friend-dependent* if there are $A, B \in \mathcal{N}^i$ such that (A, B) is friend-dependent in $\succeq_i$.

We say $\succeq_i$ is *locally friend-dependent* if $\succeq_i$ is friend-dependent and every (A, B) that is friend-dependent in $\succeq_i$ is locally friend-dependent in $\succeq_i$.

Finally, we turn to *local unanimity*: If two coalitions A and B contain the same friends of a player i, and if i and all these friends value A higher than B, then we want i to prefer A over B. This is a desirable property as it means that an unanimous opinion of agent i and her friends will always be reflected in i's preference.

	Avg-based preferences			Min-based preferences		
	$\succeq^{avgSF}$	$\succeq^{avgEQ}$	$\succeq^{avgAL}$	$\succeq^{minSF}$	$\succeq^{minEQ}$	$\succeq^{minAL}$
Reflexivity	✓	✓	✓	✓	✓	✓
Transitivity	✓	✓	✓	✓	✓	✓
Polynomial-time computability	✓	✓	✓	✓	✓	✓
Anonymity	✓	✓	✓	✓	✓	✓
Weak friend-orientedness	✓	✓	✓	✓	✓	✓
Favoring friends	✓	✓	✓	✓	✓	✓
Indifference between friends	✓	✓	✓	✓	✓	✓
Indifference between enemies	✓	✓	✓	✓	✓	✓
Sovereignty of players	✓	✓	✓	✓	✓	✓
Type-I-monotonicity	✓	✗	✗	✗	✗	✗
Type-II-monotonicity	✓	✗	✗	✓	✗	✗
Local friend dependence	✓[1]	✓[1]	✓[1]	✓[1]	✓[2]	✓[1]
Local unanimity	✓	✓	✓	✓	✓	✓

[1] If there are at least four agents and the considered agent has at least one friend.
[2] If the considered agent has at least two friends.

Table 2: Properties satisfied (✓) or not (✗) by the altruistic preferences from Section 4.2

Local Unanimity: Let $A, B \in \mathcal{N}^i$ and $A \cap F_i = B \cap F_i$. We say that $\succeq_i$ is *locally unanimous* if $v_a(A) > v_a(B)$ for each $a \in (F_i \cup \{i\}) \cap A$ implies that $A \succ_i B$.

The above definition covers all cases where the same subset of friends is consulted who all have an unanimous opinion in terms of friend-oriented preferences; in particular, it covers the case where all friends are consulted: $F_i \subseteq A \cap B$.

5.3 Study of Desirable Properties

We now consider the desirable properties from Section 5.2. Table 2 summarizes our results for all three degrees of avg-based and min-based altruistic preferences.

Proposition 5.2. *Under all three degrees of avg-based and min-based altruistic preferences, the following properties are satisfied: reflexivity, transitivity, polynomial-time computability, as well as anonymity.*

Proof. Reflexivity follows immediately from the definition.

Transitivity follows from the fact that the relation $\geq$ is transitive for rational numbers.

Furthermore, each value (as defined in (2)) that an agent assigns to a coalition can obviously be computed in polynomial time. Hence, each summand in the friends' average value (defined in (5)) and each element in the friends' minimum value (defined in (6)) can be computed in polynomial time. The number of summands (and elements in the minimum) is bounded by the number n of players, which implies that both sums and minima can be computed in polynomial time, which in turn allows to determine the utilities for any two coalitions in polynomial time.

Finally, by renaming the players, the numbers of friends and enemies of the players do not change. Therefore, the calculations of the utilities do not change either, leading to no change in the relation between two coalitions. ❑

Theorem 5.3. *Under all three degrees of avg-based and min-based altruistic preferences, weak friend-orientedness, favoring friends, indifference between friends, indifference between enemies, sovereignty of players, and local unanimity are satisfied.*

Proof. We show these properties for avg-based EQ preferences only. The proofs for the other five models of altruism work analogously and are therefore omitted.

Weak Friend-Orientedness: Suppose that A is acceptable for $i \in A$, that is, we have $A \succeq_i^{avgEQ} \{i\}$, which is equivalent to the inequality

$$\sum_{a \in A \cap (F_i \cup \{i\})} \frac{v_a(A)}{|A \cap (F_i \cup \{i\})|} \geq v_i(\{i\}) = 0. \tag{13}$$

For any $f \in F_i \setminus A$, we show that the coalition $A \cup \{f\}$ is acceptable for i as well. By definition, $A \cup \{f\} \succeq_i^{avgEQ} \{i\}$ if and only if $\sum_{a \in (A \cup \{f\}) \cap (F_i \cup \{i\})} v_a(A \cup \{f\}) \geq v_i(\{i\}) = 0$. The left-hand side of this inequality equals

$$\sum_{a \in A \cap (F_i \cup \{i\})} v_a(A) + \sum_{a \in A \cap (F_i \cup \{i\})} v_a(\{f\}) + v_f(A \cup \{f\}). \tag{14}$$

The first sum is nonnegative by the premise (13). The second sum and $v_f(A \cup \{f\})$ are positive because i is f's friend (and vice versa) and there are at most $n-1$ enemies for f. Thus, in total, we have that (14) is nonnegative and, therefore, $A \cup \{f\}$ is acceptable for i.

Favoring Friends: Let $i \in N$, $x \in F_i$, and $y \in E_i$. This property holds because

$$\begin{aligned} u_i^{avgEQ}(\{x,i\}) &= \sum_{a \in \{x,i\} \cap (F_i \cup \{i\})} \frac{v_a(\{x,i\})}{|\{x,i\} \cap (F_i \cup \{i\})|} \\ &= \frac{v_x(\{x,i\}) + v_i(\{x,i\})}{|\{x,i\} \cap (F_i \cup \{i\})|} = \frac{n+n}{2} = n > -1 \\ &= v_i(\{y,i\}) = \sum_{b \in \{y,i\} \cap (F_i \cup \{i\})} \frac{v_b(\{y,i\})}{|\{y,i\} \cap (F_i \cup \{i\})|} = u_i^{avgEQ}(\{y,i\}). \end{aligned}$$

Indifference between Friends: Let $x, y \in F_i$. As i's utility for both coalitions, $\{x,i\}$ and $\{y,i\}$, is n, we have $\{x,i\} \sim_i^{avgEQ} \{y,i\}$.

Indifference between Enemies: For $x, y \in E_i$, i's utility for $\{x,i\}$ and $\{y,i\}$ is -1, which also implies indifference.

Sovereignty of Players: Let $i \in N$ and $C \in \mathcal{N}^i$. We construct the network of friends G such that for all pairs of players $x, y \in C$, $x \neq y$, there is an edge $\{x,y\}$ in G, while there are no other edges in G. Then C is i's most preferred coalition.

Local Unanimity: Let $A, B \in \mathcal{N}^i$ with $A \cap F_i = B \cap F_i$ and let $v_j(A) > v_j(B)$ for each $j \in A \cap (F_i \cup \{i\})$. Then $A \cap (F_i \cup \{i\}) = B \cap (F_i \cup \{i\})$. Hence, it is obvious that

$$u_i^{avgEQ}(A) = \sum_{j \in A \cap (F_i \cup \{i\})} \frac{v_j(A)}{|A \cap (F_i \cup \{i\})|} > \sum_{j \in B \cap (F_i \cup \{i\})} \frac{v_j(B)}{|B \cap (F_i \cup \{i\})|} = u_i^{avgEQ}(B).$$

Thus $A \succ_i^{avgEQ} B$, showing local unanimity.

This completes the proof for avg-based EQ altruistic preferences; the proofs for the other altruistic preferences, as mentioned above, are very similar. ❑

Turning to local friend dependence, we can show that this property holds for all three degrees of avg-based and min-based altruistic preferences except for some edge cases where there are not enough agents or no friends at all. In particular, if i has no friends, her altruistic preferences coincide with her friend-oriented preferences (3). Additionally, $\succeq_i^{minEQ}$ coincides with the friend-oriented preference extension if i has only one friend. In these cases, the preferences are not friend-dependent.[7]

Theorem 5.4. *The preference $\succeq_i^\star$ of agent $i \in N$ is (locally) friend-dependent exactly if*

1. *in case $\star = avgSF$ or $\star = minSF$, i has at least one friend and $n \geq 4$;*
2. *in case $\star = avgEQ$, i has at least one friend and $n \geq 4$ or i has exactly one friend and $n = 3$;*
3. *in case $\star = minEQ$, i has at least two friends (and thus $n \geq 3$);*
4. *in case $\star = avgAL$ or $\star = minAL$, i has at least one friend and $n \geq 3$.*

Proof. First, for any altruistic preference $\succeq_i^\star$ obtained under any of the three degrees of avg-based or min-based altruism, it is easy to see that every pair (A, B) of coalitions that is friend-dependent under $\succeq_i^\star$ is also locally friend-dependent under $\succeq_i^\star$. This holds because i's utilities for A and B only depend on the set of i's friends and the sets of friends of i's friends in A and B, respectively. In other words, i's utilities for A and B can only be changed by changing some F_a with $a \in \{i\} \cup (F_i \cap (A \cup B))$. Hence, $\succeq_i^\star$ is locally friend-dependent if and only if $\succeq_i^\star$ is friend-dependent.

Second, we show that $\succeq_i^\star$ is friend-dependent, i.e., there exists a pair (A, B) of coalitions that is friend-dependent under $\succeq_i^\star$ if and only if i has at least one (or two) friends and N is sufficiently large. We omit the cases of one, two, and three players as these few small examples can easily be verified and the results are as stated in the theorem.

Only if: If i has no friends, then there are no friends whose sets of friends could be changed. So, there is obviously no pair of coalitions that is friend-dependent under any degree of altruism and, thus, $\succeq_i^\star$ is not friend-dependent. Moreover, we consider $\succeq_i^{minEQ}$ for the case that i has exactly one friend. Then $v_i(C)$ is the minimum valuation in $u_i^{minEQ}(C)$

7. Note that friend dependence is a crucial property that distinguishes our altruistic preferences from previous models, e.g., from *additively separable* (Aziz et al., 2013b) or *friend-oriented* preferences (Dimitrov et al., 2006).

for any coalition C because i has at most one friend in C while this friend might have more friends in C. Hence, $\succeq_i^{minEQ}$ coincides with $\succeq_i^F$ and is not friend-dependent.

If: First, for $\star \in \{avgSF, avgEQ, avgAL\}$, we show that there is a pair $(A, B) \in \mathcal{N}^i \times \mathcal{N}^i$ that is friend-dependent under $\succeq_i^\star$ if $n \geq 4$ and $|F_i| > 0$.

Case 1: There are at least two agents $e_1, e_2 \in N \setminus \{i\}$ that are i's enemies (and at least one friend $f \in F_i$ due to $|F_i| > 0$). It holds that $v_i(\{i, f, e_1\}) = v_i(\{i, f, e_2\})$. Hence, i's utility depends on $\text{avg}_i^F(\{i, f, e_1\})$ and $\text{avg}_i^F(\{i, f, e_2\})$. If $\text{avg}_i^F(\{i, f, e_1\}) = \text{avg}_i^F(\{i, f, e_2\})$, we change F_f such that $\text{avg}_i^F(\{i, f, e_1\}) \neq \text{avg}_i^F(\{i, f, e_2\})$, and vice versa. (This is possible by adding e_1 to F_f or deleting e_1 from F_f.) This changes i's preference over $\{i, f, e_1\}$ and $\{i, f, e_2\}$ under all three degrees of altruism. Hence, $(\{i, f, e_1\}, \{i, f, e_2\})$ is friend-dependent.

Case 2: i is friends with all but one agent $e_1 \in N \setminus \{i\}$ and thus has at least two friends $f_1, f_2 \in F_i$ (due to $n \geq 4$). Then $(\{i, f_1, e\}, \{i, f_2, e\})$ is friend-dependent (by adding e to F_{f_1} or deleting e from F_{f_1}).

Case 3: If i is friends with all (at least $n - 1 \geq 3$) agents $f_1, \dots, f_{n-1} \in N \setminus \{i\}$, then $(\{i, f_1, f_2\}, \{i, f_1, f_3\})$ is friend-dependent (by adding f_2 to F_{f_1} or deleting f_2 from F_{f_1}).

For $\star \in \{minSF, minAL\}$, the proof that $\succeq_i^\star$ is friend-dependent if $n \geq 4$ and $|F_i| > 0$ is very similar to the above argumentation and is therefore omitted.

Finally, we show that $\succeq_i^{minEQ}$ is friend-dependent if $|F_i| \geq 2$. Assuming $|F_i| \geq 2$, there are $f_1, f_2 \in F_i$. If f_1 and f_2 are friends of each other, then $\{i, f_1, f_2\} \succ_i^{minEQ} \{i, f_1\}$. Otherwise, $\{i, f_1\} \succ_i^{minEQ} \{i, f_1, f_2\}$. Hence, we can change $\succeq_i^{minEQ}$ by changing the friendship relation between f_1 and f_2, which means that $\succeq_i^{minEQ}$ is friend-dependent. ❑

We now turn to our two types of monotonicity. Interestingly, both types of monotonicity hold for avg-based SF preferences and type-II-monotonicity also holds for min-based SF preferences, but both types of monotonicity are violated for all other altruistic preferences.

Theorem 5.5. *Avg-based SF preferences are type-I-monotonic and type-II-monotonic. Min-based SF preferences are type-II-monotonic.*

Proof. We start with avg-based SF preferences. Let $G = (N, H)$ be a network of friends and let $i \in N$, $A, B \in \mathcal{N}^i$, and $j \in E_i$. We denote with $G' = (N, H \cup \{\{i, j\}\})$ the network of friends resulting from G when j turns from being i's enemy into being i's friend (all else being equal). Then, for any player $a \in N$ and coalition $C \in \mathcal{N}^a$, we denote a's value for C in G' with $v'_a(C)$, her SF preference in G' with $\succeq_a^{avgSF\prime}$ and her new friend and enemy sets with F'_a and E'_a. Hence, we have $F'_i = F_i \cup \{j\}$, $E'_i = E_i \setminus \{j\}$, $F'_j = F_j \cup \{i\}$, and $E'_j = E_j \setminus \{i\}$. Further, v'_i, v'_j, and $\succeq_i^{avgSF\prime}$ differ from v_i, v_j, and $\succeq_i^{avgSF}$. The friends, enemies, and values of all other players stay the same, i.e., $F'_a = F_a$, $E'_a = E_a$, and $v'_a = v_a$ for all $a \in N \setminus \{i, j\}$.

Type-I-Monotonicity: Let $j \in A \cap B$ and $A \succeq_j^F B$, i.e., $v_j(A) \geq v_j(B)$. It then holds that $v'_i(A) = n|A \cap F'_i| - |A \cap E'_i| = n|A \cap F_i| + n - |A \cap E_i| + 1 = v_i(A) + n + 1$. Equivalently, $v'_i(B) = v_i(B) + n + 1$, $v'_j(A) = v_j(A) + n + 1$, and $v'_j(B) = v_j(B) + n + 1$.

Furthermore,

$$\begin{aligned}\operatorname{avg}_i^{F\prime}(A) &= \sum_{a\in A\cap F_i'} \frac{v_a'(A)}{|A\cap F_i'|} = \sum_{a\in(A\cap F_i)\cup\{j\}} \frac{v_a'(A)}{|(A\cap F_i)\cup\{j\}|}\\ &= \sum_{a\in A\cap F_i} \frac{v_a(A)}{|A\cap F_i|+1} + \frac{v_j'(A)}{|A\cap F_i|+1}\\ &= \frac{|A\cap F_i|}{|A\cap F_i|+1}\cdot \operatorname{avg}_i^F(A) + \frac{v_j(A)+n+1}{|A\cap F_i|+1} \quad \text{and} \qquad (15)\\ \operatorname{avg}_i^{F\prime}(B) &= \frac{|B\cap F_i|}{|B\cap F_i|+1}\cdot \operatorname{avg}_i^F(B) + \frac{v_j(B)+n+1}{|B\cap F_i|+1}. \qquad (16)\end{aligned}$$

If $A \succ_i^{avgSF} B$ then either (i) $v_i(A) = v_i(B)$ and $\operatorname{avg}_i^F(A) > \operatorname{avg}_i^F(B)$, or (ii) $v_i(A) > v_i(B)$. In case (i), $v_i(A) = v_i(B)$ implies $v_i'(A) = v_i'(B)$ and $|A \cap F_i| = |B \cap F_i|$. Applying $|A \cap F_i| = |B \cap F_i|$, $\operatorname{avg}_i^F(A) > \operatorname{avg}_i^F(B)$, and $v_j(A) \geq v_j(B)$ to (15) and (16), we get $\operatorname{avg}_i^{F\prime}(A) > \operatorname{avg}_i^{F\prime}(B)$. Then, $v_i'(A) = v_i'(B)$ and $\operatorname{avg}_i^{F\prime}(A) > \operatorname{avg}_i^{F\prime}(B)$ implies $A \succ_i^{avgSF\prime} B$. In case (ii), $v_i(A) > v_i(B)$ implies $v_i'(A) > v_i'(B)$. Hence, $A \succ_i^{avgSF\prime} B$.

If $A \sim_i^{avgSF} B$ then $v_i(A) = v_i(B)$ and $\operatorname{avg}_i^F(A) = \operatorname{avg}_i^F(B)$. $v_i(A) = v_i(B)$ implies $v_i'(A) = v_i'(B)$ and $|A \cap F_i| = |B \cap F_i|$. Applying $|A \cap F_i| = |B \cap F_i|$, $\operatorname{avg}_i^F(A) = \operatorname{avg}_i^F(B)$, and $v_j(A) \geq v_j(B)$ to (15) and (16), we get $\operatorname{avg}_i^{F\prime}(A) \geq \operatorname{avg}_i^{F\prime}(B)$. Hence, $v_i'(A) = v_i'(B)$ and $\operatorname{avg}_i^{F\prime}(A) \geq \operatorname{avg}_i^{F\prime}(B)$ implies $A \succeq_i^{avgSF\prime} B$.

Type-II-Monotonicity: Let $j \in A \setminus B$. It follows that $v_i'(A) = v_i(A) + n + 1$ and $v_i'(B) = v_i(B)$.

If $A \succ_i^{avgSF} B$ then $v_i(A) \geq v_i(B)$. Hence, $v_i'(A) = v_i(A) + n + 1 \geq v_i(B) + n + 1 > v_i(B) = v_i'(B)$. This implies $A \succ_i^{avgSF\prime} B$.

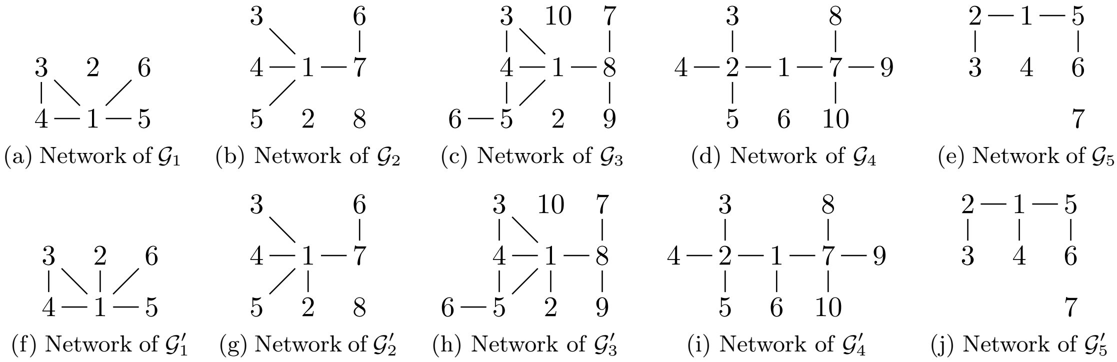
If $A \sim_i^{avgSF} B$ then $v_i(A) = v_i(B)$. Again, $v_i'(A) = v_i(A) + n + 1 = v_i(B) + n + 1 > v_i(B) = v_i'(B)$. Hence, $A \succ_i^{avgSF\prime} B$.

This completes the proof for avg-based SF preferences. The proof that min-based SF preferences are type-II-monotonic is identical to the proof that avg-based SF preferences are type-II-monotonic. ❑

Theorem 5.6. *Min-based SF preferences are not type-I-monotonic. Avg-based and min-based EQ and AL preferences are neither type-I-monotonic nor type-II-monotonic.*

Proof. As a counterexample for all min-based altruistic preferences, consider the game $\mathcal{G}_1$ with the network of friends in Figure 6a. To see that none of the three degrees of min-based altruistic preferences is type-I-monotonic, consider players $i = 1$ and $j = 2 \notin F_1$ and coalitions $A = \{1, 2, 3, 4\}$ and $B = \{1, 2, 5, 6\}$. Then $v_1(A) = v_1(B) = 11$, $v_2(A) = v_2(B) = -34$, $v_3(A) = v_4(A) = 11$, and $v_5(B) = v_6(B) = 4$. Hence, $\min_1^F(A) = 11$ and $\min_1^F(B) = 4$. It follows that $A \succ_1^{minSF} B$, $A \succ_1^{minEQ} B$, and $A \succ_1^{minAL} B$.

Making $j = 2$ a friend of $i = 1$, we get the game $\mathcal{G}_1'$ with the network of friends shown in Figure 6f. For this network, we have $v_1(A) = v_1(B) = 18$, $v_2(A) = v_2(B) = 4$, $v_3(A) = v_4(A) = 11$, and $v_5(B) = v_6(B) = 4$. Then $\min_1^F(A) = 4 = \min_1^F(B)$. Hence,



(a) Network of $\mathcal{G}_1$ (b) Network of $\mathcal{G}_2$ (c) Network of $\mathcal{G}_3$ (d) Network of $\mathcal{G}_4$ (e) Network of $\mathcal{G}_5$

(f) Network of $\mathcal{G}'_1$ (g) Network of $\mathcal{G}'_2$ (h) Network of $\mathcal{G}'_3$ (i) Network of $\mathcal{G}'_4$ (j) Network of $\mathcal{G}'_5$

Figure 6: Networks of friends in the proof of Theorem 5.6

	Type	Games	i	j	Violation
$\succeq^{avgAL}$	I	$\mathcal{G}_3$ (Fig. 6c), $\mathcal{G}'_3$ (Fig. 6h)	1	2	For $A = \{1,2,7,8,9,10\}$ and $B = \{1,\dots,6\}$: $A \succ_1^{avgAL} B$, $2 \in A \cap B$, $v_2(A) \geq v_2(B)$ in $\mathcal{G}_3$ but $B \succ_1^{avgAL} A$ in $\mathcal{G}'_3$
$\succeq^{avgEQ}$	II	$\mathcal{G}_4$ (Fig. 6d), $\mathcal{G}'_4$ (Fig. 6i)	1	6	For $A = \{1,2,3,4,5,6\}$ and $B = \{1,5,7,8,9,10\}$: $A \sim_1^{avgEQ} B$, $6 \in A \setminus B$ in $\mathcal{G}_4$ but $B \succ_1^{avgEQ} A$ in $\mathcal{G}'_4$
$\succeq^{avgAL}$	II	$\mathcal{G}_5$ (Fig. 6e), $\mathcal{G}'_5$ (Fig. 6j)	1	4	For $A = \{1,2,3,4\}$ and $B = \{1,5,6,7\}$: $A \sim_1^{avgAL} B$, $4 \in A \setminus B$ in $\mathcal{G}_5$ but $B \succ_1^{avgAL} A$ in $\mathcal{G}'_5$

Table 3: Violation of type-I- and type-II-monotonicity in the proof of Theorem 5.6

$A \sim_1^{minSF} B$, $A \sim_1^{minEQ} B$, and $A \sim_1^{minAL} B$, which contradicts type-I-monotonicity for the three degrees of min-based altruistic preferences.

To see that $\succeq^{minEQ}$ and $\succeq^{minAL}$ violate type-II-monotonicity, consider the same game $\mathcal{G}_1$ from Figure 6a and again players $i = 1$ and $j = 2 \notin F_1$, but now coalitions $A = \{1,2,3,4\}$ and $B = \{1,5,6\}$. Then $A \succ_1^{minEQ} B$ and $A \succ_1^{minAL} B$. However, considering $\mathcal{G}'_1$, we get $B \succ_1^{minEQ} A$ and $B \succ_1^{minAL} A$, violating type-II-monotonicity for min-based EQ and AL preferences.

We now turn to avg-based EQ preferences and type-I-monotonicity. Let $\mathcal{G}_2$ be a game with the network of friends shown in Figure 6b. We consider players $i = 1$, $j = 2 \notin F_1$, and coalitions $A = \{1,2,3,4,5\}$ and $B = \{1,2,6,7,8\}$. Then $A \sim_1^{avgEQ} B$, $2 \in A \cap B$, and $v_2(A) \geq v_2(B)$. Making 2 a friend of 1 leads to the new game $\mathcal{G}'_2$ with the network of friends shown in Figure 6g. However, in $\mathcal{G}'_2$ we have $B \succ_1^{avgEQ} A$, violating type-I-monotonicity for avg-based EQ preferences.

With analogous arguments, avg-based EQ preferences are not type-II-monotonic and avg-based AL preferences are neither type-I- nor type-II-monotonic, as shown in Table 3 that lists all counterexamples showing the violations of the respective properties. ❑

Note that the above results of Theorems 5.5 and 5.6 are a desirable outcome since this behavior exactly captures the intuition behind the definitions of the different altruistic preferences. In particular, if a player i gets an additional friend who is really unsatisfied in i's coalition then this should diminish player i's utility under EQ and AL preferences. Also, under min-based SF preferences, type-I-monotonicity does not have to be satisfied as an additional friend that is added to two coalitions might impose the same upper bound on the friends' minimum valuation for both coalitions. In the case of adding an additional friend to only one of the two coalitions (as in the case of type-II-monotonicity), however, both models of SF preferences ensure that the additional friend will increase i's utility.

In addition to the axiomatic properties from Section 5.2, one could consider notions of independence (for a characterization of friend-oriented preferences using an independence axiom, see, e.g., Alcantud & Arlegi, 2012). Classic independence axioms say that a relation between two coalitions, A and B, continues to hold even if a new (and the same) player is introduced to both coalitions. However, independence axioms of this type are not desirable in our model because the new player can be valued very differently in both coalitions. This would be the case, for example, if the new player were an enemy to most of i's friends in coalition A but were a friend to most of i's friends in coalition B. Similarly, $\mathcal{B}$- and $\mathcal{W}$-preferences (Cechlárová & Romero-Medina, 2001) are natural extensions from singleton encodings that are not independent.

6. Stability in Altruistic Hedonic Games

In this section, we study stability in AHGs. We mostly concentrate on the cases of general AHGs and SF AHGs but also provide some results for EQ and AL AHGs. We study the common stability concepts that were defined in Section 3.2. Questions of interest are how hard it is to verify whether a given coalition structure satisfies a certain concept in a given AHG and whether stable coalition structures for certain concepts always exist. If, for some concept, such coalition structures are not guaranteed to exist, we are also interested in the computational complexity of deciding whether or not such coalition structures exist in a given hedonic game. Formally, in the *verification problem* for a stability notion σ, we are given an AHG $(N, \succeq)$ and a coalition structure $\Gamma \in \mathcal{C}_N$, and we ask whether Γ satisfies σ. In the *existence problem* for σ, we are given an AHG $(N, \succeq)$, and we ask whether there exists a coalition structure $\Gamma \in \mathcal{C}_N$ that satisfies σ. Our complexity results for these problems are summarized in Table 4.

We start with two lemmas that will be very useful for the subsequent analysis.

Lemma 6.1. *For all three degrees of avg-based and min-based altruistic preferences, the following two statements hold:*

1. *A player i has a positive utility for a coalition $C \in \mathcal{N}^i$ if and only if i has at least one friend in C.*
2. *If a player i has at least one friend, i's most preferred coalition contains at least one friend of i's.*

Proof. For the first statement, if i has no friends in C then i's utility for C is at most zero under all altruistic preferences because $v_i(C) \leq 0$, $\text{avg}_i^F(C) = 0$, and $\min_i^F(C) = 0$. If i

	Verification		Existence	
	general	SF	general	SF
ind. rationality	in P	in P	YES	YES
contr. ind. stability	in P	in P	YES	YES
ind. stability	in P	in P	YES	YES
Nash stability	in P	in P	YES	YES
core stability	coNP-complete	coNP-complete[1]	in Σ_2^p	YES
str. core stability	coNP-complete	coNP-complete[1]	in Σ_2^p	YES
perfectness	in coNP	in P	in Σ_2^p[2]	in P

[1] Even for avg-based SF and min-based SF AHGs, i.e., if all agents have avg-based SF preferences or all agents have min-based SF preferences.

[2] In coNP for avg-based EQ and avg-based AL AHGs.

Table 4: Overview of complexity results of stability verification and existence problems in altruistic hedonic games. The columns "general" give the results for general AHGs (with any mixture of avg-based and min-based SF, EQ, and AL preferences). The columns "SF" give the results for SF AHGs (with any mixture of avg-based and min-based SF preferences). A "YES" entry for one of the existence problems indicates that there always exists a coalition structure that fulfills the considered stability concept in the considered class of AHGs. A gray entry indicates that only trivial upper bounds are known.

has at least one friend in C then these friends also have at least one friend in C (namely, i). Hence, i and all friends of i's in C assign a positive value to C, i.e., $v_i(C) > 0$, $\text{avg}_i^F(C) > 0$, and $\min_i^F(C) > 0$, which implies a positive utility.

The second statement follows directly from the first one. ❑

The first statement of the above lemma provides the following characterization of individual rationality.

Corollary 6.2. *For any AHG $(N, \succeq)$, the following two statements hold:*

1. *A coalition $C \in \mathcal{N}^i$ is acceptable for player $i \in N$ if she has at least one friend in C or is the only player in C.*
2. *A coalition structure Γ is individually rational if and only if for each player $i \in N$, it holds that i has at least one friend in $\Gamma(i)$ or i is the only player in $\Gamma(i)$.*

By Corollary 6.2, individual rationality verification is in P and existence is trivial because $\Gamma = \{\{1\}, \dots, \{n\}\}$ is always individually rational.

The next useful lemma only considers SF preferences. We omit its straightforward proof.

Lemma 6.3. *Given avg-based or min-based SF preferences and coalitions $C, D \in \mathcal{N}^i$, where D is a clique of size k, it holds that*

- *if i prefers C to D then i has at least k friends in C and*
- *if i is indifferent between C and D then C is a clique of size k.*

We continue with Nash stability and (contractually) individual stability and get general results for AHGs with any mixture of altruistic preferences.

Proposition 6.4. *For any AHG $(N, \succeq)$, it can be tested in polynomial time whether a given coalition structure is Nash stable, individually stable, or contractually individually stable.*

Proof. Let Γ be a coalition structure. For Nash stability, we need to check if for each player $i \in N$ and for each coalition $C \in \Gamma \cup \{\emptyset\}$, i prefers $\Gamma(i)$ to being added to C. For n players, there are at most $n+1$ such coalitions, and the preference relation can be verified in polynomial time by Proposition 5.2. Similar arguments apply to individual and contractually individual stability. ❑

The existence problem is trivial for Nash stability and (contractually) individual stability because there always exist stable coalition structures.

Theorem 6.5. *For any AHG $(N, \succeq)$, there exists a Nash stable, individually stable, and contractually individually stable coalition structure.*

Proof. Let $C = \{i \in N \mid F_i = \emptyset\}$ be the set of players without friends and rename its members by $C = \{1, \ldots, k\}$. The coalition structure $\{\{1\}, \ldots, \{k\}, N \setminus C\}$ is Nash stable: Each $i \in C$ has a utility of 0 when being alone. As i has no friends, by Lemma 6.1.1 this is the highest utility i can get. Hence, i does not want to move to another coalition. Every $j \in N \setminus C$ has at least one friend in $N \setminus C$, which implies $u_j(N \setminus C) > 0$ by Lemma 6.1.1. Hence, j would rather like to stay in $N \setminus C$ than to move to any of her enemies $1, \ldots, k$ or to the empty coalition.

Nash stability implies individual stability, which in turn implies contractually individual stability (see, e.g., Figure 2 on page 10 or Aziz & Savani, 2016). ❑

We now turn to core stability. Theorem 6.6 is inspired by a result of Dimitrov et al. (2006).

Theorem 6.6. *For any SF AHG $(N, \succeq)$, there always exists a strictly core stable (and thus core stable) coalition structure.*

Proof. We show that the coalition structure Γ consisting of the connected components of the underlying network of friends is strictly core stable (and thus core stable). We know that the players from different coalitions in Γ are not friends: Each $i \in N$ has all of her friends in $\Gamma(i)$.

For the sake of contradiction, assume that Γ is not strictly core stable, i.e., that there is a coalition $C \neq \emptyset$ that weakly blocks Γ. We then have $C \succeq_i^{SF} \Gamma(i)$ for all $i \in C$ and $C \succ_j^{SF} \Gamma(j)$ for some $j \in C$. Consider any player $i \in C$. Since i weakly prefers C to $\Gamma(i)$, there have to be at least as many friends of i's in C as in $\Gamma(i)$. Since $\Gamma(i)$ contains all of i's friends, C also has to contain all friends of i's. Then all these friends of i's also have all their friends in C for the same reason (and so on). Consequently, C contains all players

(a) Network of friends in Example 6.7 (b) Network of friends in Example 6.8

Figure 7: Two networks of friends for Examples 6.7 and 6.8

from the connected component $\Gamma(i)$, i.e., $\Gamma(i) \subseteq C$. Since C weakly blocks Γ, C cannot be equal to $\Gamma(i)$ and so contains some player $k \notin \Gamma(i)$. However, this is a contradiction because k is an enemy of i's and i would prefer $\Gamma(i)$ to C if C contained the same number of friends but more enemies than $\Gamma(i)$. ❑

Note that the proof of Theorem 6.6—showing that the coalition structure consisting of the connected components in the underlying network of friends is (strictly) core stable for SF preferences—does not carry over to general AHGs. Indeed, the following example shows that this coalition structure may not be core stable in this case.

Example 6.7. *Consider an AHG $(N, \succeq)$ with six agents and the network of friends in Figure 7a. The players have different altruistic preferences: Players* 2 *and* 5 *have SF preferences while players* 1, 3, 4, *and* 6 *have AL preferences. (In this example, it does not matter whether the preferences are avg-based or min-based.)*

First, consider the coalition structure $\Gamma = \{N\}$ consisting of the grand coalition, i.e., of the one connected component in this network of friends. It holds that $C = \{1,2,3\}$ blocks Γ: 1 *prefers C to N because* $\mathrm{avg}_1^F(C) = \min_1^F(C) = v_2(C) = 2n > 2n-3 = v_2(N) = \mathrm{avg}_1^F(N) = \min_1^F(N)$; 2 *prefers C to N because* $v_2(C) = 2n > 2n-3 = v_2(N)$; *and* 3 *prefers C to N because* $\mathrm{avg}_3^F(C) = \min_3^F(C) = v_2(C) = 2n > 2n-3 = \mathrm{avg}_3^F(N) = \min_3^F(N)$.

Further, consider coalition structure $\Delta = \{\{1,2,3\},\{4,5,6\}\}$ that results from the core deviation of C. It then holds that every agent is in one of her most preferred coalitions. Thus Δ is perfect and therefore also strictly core stable (again, see Figure 2 on page 10 or Aziz & Savani, 2016).

Example 6.7 shows that the coalition structure consisting of the connected components is not always core stable while there may exist another coalition structure that is even strictly core stable. However, the next example shows that a strictly core stable coalition structure does not need to exist in general. Example 6.8 shows this for the case of min-based EQ and AL preferences.

Example 6.8. *Consider the network of friends in Figure 7b. We show that there does not exist a strictly core stable coalition structure under min-based EQ preferences. A direct calculation gives the following: Under min-based EQ preferences,*

1. *the unique most preferred coalition of players* 1 *and* 2 *is* $A = \{1,2,3\}$.

2. *the unique most preferred coalition of players* 4 *and* 5 *is* $B = \{3,4,5\}$.

3. *player* 3 *has exactly two most preferred coalitions:* A *and* B.

With these observations, we can directly conclude the following: If a given coalition structure Γ does not contain A, the coalition A weakly blocks Γ. So any strictly core stable coalition structure has to contain A. However, the same holds for B, and since A and B are not disjoint, they cannot both be contained in the same coalition structure. Thus there does not exist a strictly core stable coalition structure under min-based EQ preferences. Similar arguments work for the same example and min-based AL preferences.

Turning to the verification problem of (strict) core stability, we have the following results.

Theorem 6.9. *For general AHGs, (strict) core stability verification is in* coNP*. For avg-based SF AHGs and min-based SF AHGs, (strict) core stability verification is even* coNP-*complete.*

Proof. We start with showing the membership of (strict) core stability verification in coNP for general AHGs. Let G be a network of friends on agents N and Γ a coalition structure. Γ is not (strictly) core stable if there is a coalition $C \subseteq N$ that (weakly) blocks Γ. Hence, we nondeterministically guess a coalition $C \subseteq N$ and check whether C blocks Γ. This can be done in polynomial time since we only need to check a linear number of preference relations, which in turn can be done in polynomial time (in the number of agents) for all models (see Proposition 5.2).

To show coNP-hardness of (strict) core stability verification under avg-based and min-based SF AHGs, we make use of a restricted variant of the NP-complete problem Exact Cover by 3-Sets (Garey, Johnson, & Stockmeyer, 1976); it was shown by Gonzalez (1985) that this problem remains NP-complete even when each element of the set occurs in exactly three of the 3-element subsets:

Restricted Exact Cover by 3-Sets (RX3C)	
Given:	An integer $k \geq 2$, a set $B = \{1, \ldots, 3k\}$ and a collection $\mathscr{S} = \{S_1, \ldots, S_{3k}\}$ of 3-element subsets of B ($S_i \subseteq B$ and $\lvert S_i\rvert = 3$ for $1 \leq i \leq 3k$), where each element of B occurs in exactly three sets in $\mathscr{S}$.
Question:	Does there exist an exact cover of B in $\mathscr{S}$, i.e., a subset $\mathscr{S}' \subseteq \mathscr{S}$ of size k such that every element of B occurs in exactly one set in $\mathscr{S}'$?

We provide a polynomial-time many-one reduction from RX3C to the complements of our problems.

Let $(B, \mathscr{S})$ be an instance of RX3C, consisting of a set $B = \{1, \ldots, 3k\}$ and a collection $\mathscr{S} = \{S_1, \ldots, S_{3k}\}$ of 3-element subsets of B. We may assume, without loss of generality, that $k > 5$. From $(B, \mathscr{S})$ we construct the following AHG. The set of players is given by

$$N = \{\beta_b \mid b \in B\} \cup \{\zeta_S \mid S \in \mathscr{S}\} \cup \{\alpha_{S,1}, \alpha_{S,2}, \alpha_{S,3} \mid S \in \mathscr{S}\} \cup \{\delta_{S,1}, \ldots, \delta_{S,4k-3} \mid S \in \mathscr{S}\}.$$

Define the sets

$$\begin{aligned} \mathit{Beta} &= \{\beta_b \mid b \in B\}, \\ \mathit{Zeta} &= \{\zeta_S \mid S \in \mathscr{S}\}, \text{ and} \\ Q_S &= \{\alpha_{S,1}, \alpha_{S,2}, \alpha_{S,3}, \delta_{S,1}, \ldots, \delta_{S,4k-3}\} \text{ for each } S \in \mathscr{S}. \end{aligned}$$

The network of friends is given in Figure 8, where a dashed circle around a group of players means that all these players are friends of each other:

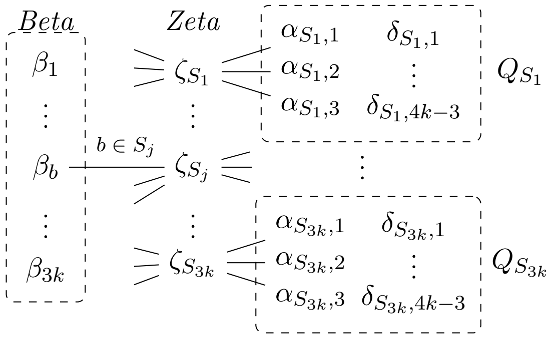


Figure 8: Network of friends in the proof of Theorem 6.9

- All players in *Beta* are friends of each other.
- For every $S \in \mathscr{S}$, all players in Q_S are friends of each other.
- For every $S \in \mathscr{S}$, ζ_S is friends with $\alpha_{S,1}$, $\alpha_{S,2}$, $\alpha_{S,3}$, and the three β_b with $b \in S$.

Consider the coalition structure $\Gamma = \{Beta\} \cup \{\{\zeta_S\} \cup Q_S \mid S \in \mathscr{S}\}$.

We claim that $\mathscr{S}$ contains an exact cover for B if and only if Γ is not (strictly) core stable under the avg-based SF model if and only if Γ is not (strictly) core stable under the min-based SF model.

We start by showing the equivalence for the avg-based SF model (and for both core stability and strict core stability).

Only if: Assume that there is an exact cover $\mathscr{S}' \subseteq \mathscr{S}$, $|\mathscr{S}'| = k$, for B. Consider $C = Beta \cup \{\zeta_S \mid S \in \mathscr{S}'\}$. Then C blocks Γ (i.e., $C \succ_i^{avgSF} \Gamma(i)$ for all $i \in C$) because

- every $\beta_b \in Beta$ has $3k$ friends in C and only $3k-1$ friends in $\Gamma(\beta_b) = Beta$ and
- every ζ_S, $S \in \mathscr{S}'$, has 3 friends and $4k-4$ enemies in C and 3 friends and $4k-3$ enemies in $\Gamma(\zeta_S) = \{\zeta_S\} \cup Q_S$.

Hence, Γ is not core stable (and thus not strictly core stable) under the avg-based SF model.

If: Assume that Γ is not strictly core stable under the avg-based SF model. Then there is a coalition $C \subseteq N$ that weakly blocks Γ, i.e., $C \succeq_i^{avgSF} \Gamma(i)$ for all $i \in C$ and $C \succ_j^{avgSF} \Gamma(j)$ for some $j \in C$. First observe that in Γ every $\alpha_{S,j}$ for $S \in \mathscr{S}$ and $1 \leq j \leq 3$ is together with all her friends and none of her enemies. Hence, $\Gamma(\alpha_{S,j})$ is already $\alpha_{S,j}$'s unique most preferred coalition, so there is no other coalition that $\alpha_{S,j}$ would like to deviate to. Thus, for all $S \in \mathscr{S}$ and $1 \leq j \leq 3$, we have $\alpha_{S,j} \notin C$. However, this implies that also no $\delta_{S,l}$ with $S \in \mathscr{S}$ and $1 \leq l \leq 4k-3$ is in C because $\delta_{S,l}$ cannot weakly prefer C to $\Gamma(\delta_{S,l})$ if C contains no α-player. Hence, we have shown that $C \subseteq Beta \cup Zeta$.

Define $\#_\beta = |Beta \cap C|$ and $\#_\zeta = |Zeta \cap C|$ as the numbers of, respectively, β-players and ζ-players in C. We will show that $\#_\beta = 3k$ and $\#_\zeta = k$.

It is easy to see that there has to be at least one β-player in C. Consider some $\beta_b \in C$. Since β_b weakly prefers C to $\Gamma(\beta_b) = Beta$, which is a clique of size $3k$, and since β_b is not contained in any other clique of size $3k$, by Lemma 6.3 β_b has at least $3k$ friends in C. Since

β_b has three ζ-friends in total, at least $3k-3$ of β_b's friends in C are β-players. Taking β_b herself into account, we have $\#_\beta \geq 3k-2$. Since these $3k-2$ β-players have at least $3k$ friends in C, they all need to have at least one ζ-player as a friend in C. Therefore, $\#_\zeta \geq k$.

Consider some $\zeta_S \in C$. Since ζ_S has three friends and $4k-3$ enemies in $\Gamma(\zeta_S)$, at most three friends in C, and she weakly prefers C to $\Gamma(\zeta_S)$, ζ_S has exactly three friends (three β-players) and at most $4k-3$ enemies in C. Hence, C contains at most $4k-3+3+1 = 4k+1$ players and $|C| = \#_\beta + \#_\zeta \leq 4k+1$.

For a contradiction, assume that $\#_\beta = 3k-2$. Then each of these β-players has only $3k-3$ β-friends in C and additionally needs at least 3 ζ-friends in C. Since each β-player has exactly three ζ-friends and vice versa, we then have at least $(3k-2)\cdot 3 = 9k-6$ edges between the β- and ζ-players in C. Then there are at least $3k-2$ ζ-players in C. Thus $\#_\beta + \#_\zeta \geq (3k-2)+(3k-2) = 6k-4$, which is a contradiction (for $k > 2$) to $\#_\beta + \#_\zeta \leq 4k+1$. Analogously, we get a contradiction when assuming that $\#_\beta = 3k-1$. Consequently, $\#_\beta = 3k$.

So far, we have $\#_\beta = 3k$, $\#_\zeta \geq k$, and $\#_\beta + \#_\zeta \leq 4k+1$. Hence, $\#_\zeta = k$ or $\#_\zeta = k+1$. For the sake of contradiction, assume that $\#_\zeta = k+1$. First, recall that each of these ζ-players has three β-friends in C. Then there are exactly $(k+1)3 = 3k+3$ edges between the β-players and the ζ-players in C. Since every β-player has at least on ζ-friend in C, every β-player has at least one edge to a ζ-player in C. Hence, there are at least $3k-3$ β-players who have exactly one edge to a ζ-player in C and at most three β-players who have more than one edge to a ζ-player in C. Consider a $\zeta_S \in C$ who is friends with three β-players who have only one ζ-friend in C. (There has to be such a ζ-player because otherwise there would be $k+1$ ζ-players with a β-friend who has two ζ-friends. However, at most six ζ-players can be friends with one of these β-players. And $6 < k+1$ for $k > 5$.) Since ζ_S weakly prefers C to $\Gamma(\zeta_S)$ and has exactly 3 friends and $4k-3$ enemies in C, it holds that $\text{avg}^F_{\zeta_S}(C) \geq \text{avg}^F_{\zeta_S}(\Gamma(\zeta_S))$. In C, ζ_S has three β-friends who all have the same valuation: $\text{avg}^F_{\zeta_S}(C) = n\cdot(3k)-k$. In $\Gamma(\zeta_S)$, ζ_S has three α-friends who all have the same valuation: $\text{avg}^F_{\zeta_S}(\Gamma(\zeta_S)) = n\cdot(4k)$. Hence, $\text{avg}^F_{\zeta_S}(C) < \text{avg}^F_{\zeta_S}(\Gamma(\zeta_S))$, which is a contradiction. Hence, $\#_\zeta = k$.

Finally, since every of the $3k$ β-players in C has one of the k ζ-players in C as a friend, it holds that $\{S \mid \zeta_S \in C\}$ is an exact cover for B.

This completes the proof for coNP-hardness of (strict) core stability verification under the avg-based SF model.

To show coNP-hardness of (strict) core stability verification under min-based SF AHGs, we can use the exact same reduction and arguments as before. Only in the very end of the proof, we once consider the minimum over the friends' valuations instead of the average. However, since in both coalitions the valuations are the same for all friends, the minimum and the average lead to the same result in this case. ❑

From Theorem 6.9 we immediately get the following corollary.

Corollary 6.10. *For general AHGs, (strict) core stability verification is* coNP*-complete.*

We now turn to perfectness and establish the following characterization under SF preferences.

Proposition 6.11. *For any SF AHG, a coalition structure is perfect if and only if it consists of the connected components of the underlying network of friends and all these components are cliques.*

Proof. From left to right, assume that the coalition structure $\Gamma \in \mathcal{C}_N$ is perfect. It then holds for every agent $i \in N$ that she weakly prefers $\Gamma(i)$ to every coalition $C \in \mathcal{N}^i$. It follows that every agent $i \in N$ is in her most preferred coalition, where she is together with all her friends and none of her enemies. This implies that each coalition in Γ is a connected component and a clique. The implication from right to left is obvious. ❑

From Proposition 6.11 we get the following characterization of perfect coalition structures.

Corollary 6.12. *For any SF AHG, there exists a perfect coalition structure if and only if all connected components of the underlying network of friends are cliques.*

By Proposition 6.11 and Corollary 6.12, perfectness verification and existence are in P for SF AHGs.

We now turn to avg-based EQ and AL preferences. The next lemma will be used in the proof of Proposition 6.14 and says that if player j has a friend k in coalition C and k has another friend $\ell \notin C$ who is not j's friend, then j prefers $C \cup \{\ell\}$ to C under avg-based EQ and AL preferences. Hence, a coalition structure that contains C is not perfect.

Lemma 6.13. *Let $C \subseteq N$, $k, j \in C$, and $\ell \in N \setminus C$ with $j \in F_k$, $k \in F_\ell$, and $j \notin F_\ell$. Then $C \cup \{\ell\} \succ_j^{avgEQ} C$ and $C \cup \{\ell\} \succ_j^{avgAL} C$.*

Proof. Let $C \subseteq N$, $k, j \in C$, and $\ell \in N \setminus C$ with $j \in F_k$, $k \in F_\ell$, and $j \notin F_\ell$. When ℓ joins the coalition C, k's valuation increases by n while j's valuation and the valuation of at most $n-3$ of j's friends ($|(C \cap F_j) \setminus \{k\}| \leq n-3$) decreases by one. Since the number of j's friends is the same in $C \cup \{\ell\}$ and C, this means that $u_j^{avgEQ}(C \cup \{\ell\}) > u_j^{avgEQ}(C)$ and $u_j^{avgAL}(C \cup \{\ell\}) > u_j^{avgAL}(C)$. Thus $C \cup \{\ell\} \succ_j^{avgEQ} C$ and $C \cup \{\ell\} \succ_j^{avgAL} C$. ❑

The next proposition shows that whenever a perfect coalition structure exists in an avg-based EQ AHG or an avg-based AL AHG, it is unique and consists of the connected components of the underlying network of friends. This in particular means that every agent is together with all her friends.

Proposition 6.14. *Whenever a perfect coalition structure exists under avg-based EQ or AL preferences, it is unique and consists of the connected components of the underlying network of friends.*

Proof. We will concentrate on the proof for EQ. The proof for AL is very similar.

Let C be a coalition in a perfect coalition structure. Then C is the most preferred coalition of every player in C.

First, for a contradiction, suppose C were not connected. Then, for any player $i \in C$, there is a player $k \in C$ that is an enemy of i's and of all of i's friends in C. Hence, removing k from C increases the valuations of all players in $(C \cap F_i) \cup \{i\}$ and thus increases i's utility, which is a contradiction to C being i's most preferred coalition.

Second, observe that C contains an entire connected component: Suppose C is a proper subset of a connected component. Then there exist two players $k \in C$ and $\ell \notin C$ that are friends of each other. By Lemma 6.1(2), there exists another friend $j \in F_k \cap C$. We distinguish the following cases which all lead to contradictions:

Case 1: There exists a friend $j \in F_k \cap C$ with $\ell \notin F_j$. Then, by Lemma 6.13, this is a contradiction to C being j's most preferred coalition because $C \cup \{\ell\} \succ_j^{avgEQ} C$.

Case 2: For each $j \in F_k \cap C$, it holds that $\ell \in F_j$ (and $j \in F_\ell$ by symmetry).

Case 2.1: There exists some $x \in C \setminus F_k$ with $\ell \notin F_x$.

Case 2.1.1: $x \in E_j$ for all $j \in F_k \cap C$. Then $C \setminus \{x\} \succ_k^{avgEQ} C$, which is a contradiction to C being k's most preferred coalition.

Case 2.1.2: There is a $j \in F_k \cap C$ with $x \in F_j$. Then $C \cup \{\ell\} \succ_x^{avgEQ} C$ by Lemma 6.13, which again is a contradiction.

Case 2.2: For each $x \in C \setminus F_k$, $\ell \in F_x$. This implies that all players in C are ℓ's friends and $v_\ell(C \cup \{\ell\}) = n \cdot |C|$. Thus, comparing coalitions $C \cup \{\ell\}$ and C from k's point of view and letting λ denote $|F_k \cap C|$, we obtain:

$$
\begin{aligned}
& u_k^{avgEQ}(C \cup \{\ell\}) - u_k^{avgEQ}(C) \\
= \; & \frac{v_k(C \cup \{\ell\}) + \sum_{j \in F_k \cap C} v_j(C \cup \{\ell\}) + v_\ell(C \cup \{\ell\})}{1 + \lambda + 1} - \frac{v_k(C) + \sum_{j \in F_k \cap C} v_j(C)}{1 + \lambda} \\
= \; & \frac{1}{(2+\lambda)(1+\lambda)} \Bigg((1+\lambda) v_k(C \cup \{\ell\}) - (2+\lambda) v_k(C) + (1+\lambda) \sum_{j \in F_k \cap C} v_j(C \cup \{\ell\}) \\
& \qquad - (2+\lambda) \sum_{j \in F_k \cap C} v_j(C) + (1+\lambda) v_\ell(C \cup \{\ell\}) \Bigg).
\end{aligned}
$$

With

$$
\begin{aligned}
(1+\lambda) v_k(C \cup \{\ell\}) - (2+\lambda) v_k(C) \; &= \; n + |E_k \cap C| \quad \text{and} \\
(1+\lambda) \sum_{j \in F_k \cap C} v_j(C \cup \{\ell\}) - (2+\lambda) \sum_{j \in F_k \cap C} v_j(C) \; &\geq \; \lambda(1+\lambda) n - \lambda \cdot |C| \cdot n,
\end{aligned}
$$

we get

$$
\begin{aligned}
u_k^{avgEQ}(C \cup \{\ell\}) - u_k^{avgEQ}(C) \geq \; & \frac{n + |E_k \cap C| + \lambda(1+\lambda)n - \lambda \cdot |C| \cdot n + (1+\lambda)(n \cdot |C|)}{(2+\lambda)(1+\lambda)} \\
= \; & \frac{n + |E_k \cap C| + \lambda(1+\lambda)n + n \cdot |C|}{(2+\lambda)(1+\lambda)} > 0.
\end{aligned}
$$

Therefore, $C \cup \{\ell\} \succ_k^{avgEQ} C$, which again is a contradiction.

Since all cases lead to a contradiction, C has to be an entire connected component. ❑

Note that by Proposition 6.14, perfectness existence is in coNP under avg-based EQ and AL preferences: There exists a perfect coalition structure for a given game if and only if the coalition structure Γ that consists of the connected components of the network of friends is perfect. Hence, to show that there is no perfect coalition structure, we nondeterministically guess a player $i \in N$ and a coalition $C \in \mathcal{N}^i$ and check whether $C \succ_i^{avgEQ} \Gamma(i)$ (or $C \succ_i^{avgAL} \Gamma(i)$) in polynomial time.

Corollary 6.15. *In avg-based EQ AHGs and avg-based AL AHGs, perfectness existence is in* coNP.

Under avg-based EQ and AL preferences, we further have the following restriction on perfect coalition structures.

Proposition 6.16. *If there exists a perfect coalition structure for an avg-based EQ or avg-based AL AHG, all connected components have a diameter of at most two.*

Proof. Assume that there is a coalition C in Γ that has a diameter greater than 2. Then there are agents $i, j \in C$ with a distance greater than 2, i.e., j is an enemy of i's and all of i's friends. Hence, i and of all her friends have a higher valuation for $C \setminus \{j\}$ than for C. It follows that i prefers $C \setminus \{j\}$ to C under avg-based EQ and AL preferences. Consequently, Γ is not perfect. ❑

Interestingly, however, there also exist networks with a diameter of at most two that do not allow a perfect coalition structure, e.g., stars (i.e., one central vertex connected to a number of leaves).

Proposition 6.17. *Under avg-based EQ preferences, trees with at least four vertices do not allow a perfect coalition structure. Under avg-based AL preferences, trees with at least three vertices do not allow a perfect coalition structure.*

Proof. Trees with a diameter of more than two do not allow a perfect coalition structure by Proposition 6.16.

Trees with a diameter of two are stars. Let i be the central player and j a leaf. It holds that $N \setminus \{j\} \succ_i^{avgEQ} N$ (and $N \setminus \{j\} \succ_i^{avgAL} N$) such that $\{N\}$ is not perfect, which in turn implies that there cannot be a perfect coalition structure by Proposition 6.14. ❑

Finally, turning to general AHGs, it is interesting to see that Proposition 6.14 does not extend to the general case. As we have seen in Example 6.7, general AHGs allow for perfect coalition structures that do not consist of the connected components of the underlying network of friends.

7. Conclusions and Future Work

We have introduced and studied altruism in hedonic games where the agents' utility functions may depend on their friends' preferences. We have distinguished between three degrees of altruism, depending on the order in which an agent looks at her own and at her friends' preferences, and between an average- and minimum-based aggregation of some agents' preferences.

Axiomatically, we have defined desirable properties and have shown which of these are satisfied by which of our models and which are not. In particular, we have shown that all our altruistic preferences fulfill basic properties, such as reflexivity, transitivity, polynomial-time computability of utilities, and anonymity. Moreover, we have studied properties such as local unanimity, local friend dependence, and monotonicity. Specifically, we have considered two types of monotonicity, which combined with our six models of AHGs give 12 cases to study. Interestingly, monotonicity holds in only three of these cases while it fails to hold in the other nine cases. This contrasts with the results in *altruistic coalition formation games* (Kerkmann & Rothe, 2020) where monotonicity *fails* in three out of the corresponding 12 cases and is satisfied in the other nine cases (Kerkmann & Rothe, 2021).

Comparing altruistic hedonic games to other hedonic games from the literature, we have seen that they can express different preferences than the commonly studied representations. In terms of stability, altruistic hedonic games always admit, e.g., Nash stable coalition structures. In the case of selfish-first preferences, also core stable and strictly core stable outcomes are guaranteed to exist; both the verification and the existence problem for perfectness is polynomial-time solvable; yet the verification problems for core stability and strict core stability are computationally intractable, i.e., coNP-complete. We have also established characterizations for two of the stability notions, namely individual rationality and perfectness.

We consider it important future work to complete the characterization of all stability notions (e.g., to characterize when the grand coalition is perfect under equal-treatment and altruistic-treatment preferences). Also, while the complexity results in Table 4 are complete for selfish-first altruistic hedonic games, they are not yet complete for the general case. It would therefore be interesting to see if we can find matching lower and upper bounds in those cases where there is still a complexity gap.

Further, it might be useful to extend the model and normalize by the size of the coalition to consider only relative valuations. This can be seen as an altruistic version of a friend-oriented fractional hedonic game (Dimitrov et al., 2006; Aziz et al., 2019). For example, one could define

$$A \succeq_i^{avgEQ_f} B \iff \sum_{a \in (A \cap F_i) \cup \{i\}} \frac{v_a(A)}{|A| \cdot |(A \cap F_i) \cup \{i\}|} \geq \sum_{b \in (B \cap F_i) \cup \{i\}} \frac{v_b(B)}{|B| \cdot |(B \cap F_i) \cup \{i\}|}$$

and study the common stability notions, etc. for these preferences.

In addition, we propose to consider restrictions of the input such as constraining networks to special graph classes (e.g., interval graphs, where the width of an interval represents an agent's "tolerance") and studying problems of strategic influence (e.g., misreporting preferences to friends, pretending to be a friend while one in fact is an enemy, asserting control over the game as a whole).

One could also consider different weights for equal-treatment preferences because, as the number of friends increases, the weight of one's own opinion becomes diluted. This can be handled by weighting one's own opinion by $1/2$ and the aggregated opinion of one's friends by $1/2$. It could also be interesting to study the changes concerning stability if directed (non-mutual) networks of friends are considered. In a similar vein, the model can be extended to edge-weighted graphs, where the intensity of influence of a friend is given by the edge weight.

Acknowledgments

We are grateful to the anonymous AAMAS'16, STAIRS'20, and AAAI'21 reviewers for their helpful comments and suggestions. This work was supported in part by DFG grants RO-1202/14-2 and RO-1202/21-1 and the project "Online Participation" funded by the NRW Ministry of Culture and Science.

References

Alcantud, J., & Arlegi, R. (2012). An axiomatic analysis of ranking sets under simple categorization. *SERIEs – Journal of the Spanish Economic Association*, *3*(1–2), 227–245.

Anagnostopoulos, A., Becchetti, L., de Keijzer, B., & Schäfer, G. (2013). Inefficiency of games with social context. In *Proceedings of the 6th International Symposium on Algorithmic Game Theory*, pp. 219–230. Springer-Verlag *Lecture Notes in Computer Science #8146*.

Apt, K., & Schäfer, G. (2014). Selfishness level of strategic games. *Journal of Artificial Intelligence Research*, *49*, 207–240.

Ashlagi, I., Krysta, P., & Tennenholtz, M. (2008). Social context games. In *Proceedings of the 4th International Workshop on Internet & Network Economics*, pp. 675–683. Springer-Verlag *Lecture Notes in Computer Science #5385*.

Aziz, H., Brandl, F., Brandt, F., Harrenstein, P., Olsen, M., & Peters, D. (2019). Fractional hedonic games. *ACM Transactions on Economics and Computation*, *7*(2), 6:1–6:29.

Aziz, H., Brandt, F., & Harrenstein, P. (2013a). Pareto optimality in coalition formation. *Games and Economic Behavior*, *82*, 562–581.

Aziz, H., Brandt, F., & Seedig, H. (2013b). Computing desirable partitions in additively separable hedonic games. *Artificial Intelligence*, *195*, 316–334.

Aziz, H., & Savani, R. (2016). Hedonic games. In Brandt, F., Conitzer, V., Endriss, U., Lang, J., & Procaccia, A. (Eds.), *Handbook of Computational Social Choice*, chap. 15, pp. 356–376. Cambridge University Press.

Ballester, C. (2004). NP-completeness in hedonic games. *Games and Economic Behavior*, *49*(1), 1–30.

Banerjee, S., Konishi, H., & Sönmez, T. (2001). Core in a simple coalition formation game. *Social Choice and Welfare*, *18*(1), 135–153.

Barrot, N., Ota, K., Sakurai, Y., & Yokoo, M. (2019). Unknown agents in friends oriented hedonic games: Stability and complexity. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence*, pp. 1756–1763. AAAI Press.

Baumeister, D., & Rothe, J. (2015). Preference aggregation by voting. In Rothe, J. (Ed.), *Economics and Computation. An Introduction to Algorithmic Game Theory, Computational Social Choice, and Fair Division*, Springer Texts in Business and Economics, chap. 4, pp. 197–325. Springer-Verlag.

Bilò, V., Celi, A., Flammini, M., & Gallotti, V. (2013). Social context congestion games. *Theoretical Computer Science*, *514*, 21–35.

Bogomolnaia, A., & Jackson, M. (2002). The stability of hedonic coalition structures. *Games and Economic Behavior*, *38*(2), 201–230.

Borel, É. (1921). La théorie du jeu et les équations intégrales à noyau symétrique gauche. In *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, Vol. 173. 1304–1308.

Brandt, F., Fischer, F., Harrenstein, P., & Shoham, Y. (2009). Ranking games. *Artificial Intelligence*, *173*(2), 221–239.

Brânzei, S., & Larson, K. (2011). Social distance games. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence*, pp. 91–96. AAAI Press/IJCAI.

Bullinger, M., & Kober, S. (2021). Loyalty in cardinal hedonic games. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence*, pp. 66–72. ijcai.org.

Caragiannis, I., Kaklamanis, C., Kanellopoulos, P., Kyropoulou, M., & Papaioannou, E. (2010). The impact of altruism on the efficiency of atomic congestion games. In *Proceedings of the 5th International Symposium on Trustworthy Global Computing*, pp. 172–188. Springer-Verlag *Lecture Notes in Computer Science #6084*.

Cechlárová, K., & Hajduková, J. (2003). Computational complexity of stable partitions with B-preferences. *International Journal of Game Theory*, *31*(3), 353–364.

Cechlárová, K., & Hajduková, J. (2004). Stable partitions with $\mathcal{W}$-preferences. *Discrete Applied Mathematics*, *138*(3), 333–347.

Cechlárová, K., & Romero-Medina, A. (2001). Stability in coalition formation games. *International Journal of Game Theory*, *29*(4), 487–494.

Chalkiadakis, G., Elkind, E., & Wooldridge, M. (2011). *Computational Aspects of Cooperative Game Theory.* Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan and Claypool Publishers.

Chen, P., de Keijzer, B., Kempe, D., & Schäfer, G. (2014). Altruism and its impact on the price of anarchy. *ACM Transactions on Economics and Computation*, *2*(4), Article 17, 17:1–17:45.

De Marco, G., & Morgan, J. (2007). Slightly altruistic equilibria in normal form games. Tech. rep. 185, Centre for Studies in Economics and Finance, University of Salerno, Fisciano, Italy.

Dimitrov, D., Borm, P., Hendrickx, R., & Sung, S. (2006). Simple priorities and core stability in hedonic games. *Social Choice and Welfare*, *26*(2), 421–433.

Drèze, J., & Greenberg, J. (1980). Hedonic coalitions: Optimality and stability. *Econometrica*, *48*(4), 987–1003.

Elias, J., Martignon, F., Avrachenkov, K., & Neglia, G. (2010). Socially-aware network design games. In *Proceedings of the 29th Conference on Information Communications*, pp. 41–45. IEEE.

Elkind, E., & Rothe, J. (2015). Cooperative game theory. In Rothe, J. (Ed.), *Economics and Computation. An Introduction to Algorithmic Game Theory, Computational Social Choice, and Fair Division*, Springer Texts in Business and Economics, chap. 3, pp. 135–193. Springer-Verlag.

Faliszewski, P., Rothe, I., & Rothe, J. (2015). Noncooperative game theory. In Rothe, J. (Ed.), *Economics and Computation. An Introduction to Algorithmic Game Theory, Computational Social Choice, and Fair Division*, Springer Texts in Business and Economics, chap. 2, pp. 41–134. Springer-Verlag.

Garey, M., Johnson, D., & Stockmeyer, L. (1976). Some simplified NP-complete graph problems. *Theoretical Computer Science*, *1*, 237–267.

Gonzalez, T. (1985). Clustering to minimize the maximum intercluster distance. *Theoretical Computer Science*, *38*, 293–306.

Hare, B., & Woods, V. (2020). *Survival of the Friendliest: Understanding Our Origins and Rediscovering Our Common Humanity*. Penguin Random House, New York.

Hayrapetyan, A., Tardos, É., & Wexler, T. (2006). The effect of collusion in congestion games. In *Proceedings of the 38th ACM Symposium on Theory of Computing*, pp. 89–98. ACM Press.

Hoefer, M., Penn, M., Polukarova, M., Skopalik, A., & Vöcking, B. (2011). Considerate equilibrium. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence*, pp. 234–239. AAAI Press/IJCAI.

Hoefer, M., & Skopalik, A. (2013). Altruism in atomic congestion games. *ACM Transactions on Economics and Computation*, *1*(4), 21:1–21:21.

Kerkmann, A., Lang, J., Rey, A., Rothe, J., Schadrack, H., & Schend, L. (2020). Hedonic games with ordinal preferences and thresholds. *Journal of Artificial Intelligence Research*, *67*, 705–756.

Kerkmann, A., & Rothe, J. (2020). Altruism in coalition formation games. In *Proceedings of the 29th International Joint Conference on Artificial Intelligence*, pp. 347–353. ijcai.org.

Kerkmann, A., & Rothe, J. (2021). Four faces of altruistic hedonic games. In *Proceedings of the 8th International Workshop on Computational Social Choice*. Nonarchival proceedings.

Koutsoupias, E., & Papadimitriou, C. (1999). Worst-case equilibria. In *Proceedings of the 16th Annual Symposium on Theoretical Aspects of Computer Science*, pp. 404–413. Springer.

Kuniavsky, S., & Smorodinsky, R. (2011). Equilibrium and potential in coalitional congestion games. Tech. rep. arXiv:1111.3933v1 [cs.GT], ACM Computing Research Repository (CoRR).

Ledyard, J. (1995). Public goods: A survey of experimental research. In Kagel, J., & Roth, A. (Eds.), *The Handbook of Experimental Economics*, chap. 2, pp. 111–194. Princeton University Press.

Monaco, G., Moscardelli, L., & Velaj, Y. (2018). Hedonic games with social context. In *Proceedings of the 19th Italian Conference on Theoretical Computer Science*, Vol. 2234, pp. 24–35. CEUR-WS.org.

Monaco, G., Moscardelli, L., & Velaj, Y. (2019). On the performance of stable outcomes in modified fractional hedonic games with egalitarian social welfare. In *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems*, pp. 873–881. IFAAMAS.

Nguyen, N., Nguyen, T., Roos, M., & Rothe, J. (2014). Computational complexity and approximability of social welfare optimization in multiagent resource allocation. *Journal of Autonomous Agents and Multi-Agent Systems*, *28*(2), 256–289.

Nguyen, N., Rey, A., Rey, L., Rothe, J., & Schend, L. (2016). Altruistic hedonic games. In *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems*, pp. 251–259. IFAAMAS.

Nisan, N., Roughgarden, T., Tardos, É., & Vazirani, V. (Eds.). (2007). *Algorithmic Game Theory*. Cambridge University Press.

Olsen, M. (2012). On defining and computing communities. In *Proceedings of the 18th Computing: The Australasian Theory Symposium*, Vol. 128, pp. 97–102.

Ota, K., Barrot, N., Ismaili, A., Sakurai, Y., & Yokoo, M. (2017). Core stability in hedonic games among friends and enemies: Impact of neutrals. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, p. 359–365. AAAI Press.

Peleg, B., & Sudhölter, P. (2003). *Introduction to the Theory of Cooperative Games*. Kluwer Academic Publishers.

Peters, D. (2016). Complexity of hedonic games with dichotomous preferences. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence*, pp. 579–585. AAAI Press.

Peters, D., & Elkind, E. (2015). Simple causes of complexity in hedonic games. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence*, pp. 617–623. AAAI Press/IJCAI.

Popova, A., Regenwetter, M., & Mattei, N. (2013). A behavioral perspective on social choice. *Annals of Mathematics and Artificial Intelligence*, *68*(1–3), 5–30. Special Issue: Algorithms, Approximation, and Empirical Studies in Behavioral and Computational Social Choice, edited by J. Goldsmith and J. Rothe.

Rahn, M., & Schäfer, G. (2013). Bounding the inefficiency of altruism through social contribution games. In *Proceedings of the 9th International Workshop on Internet & Network Economics*, pp. 391–404. Springer-Verlag *Lecture Notes in Computer Science #8289*.

Regenwetter, M., Grofman, B., Marley, A., & Tsetlin, I. (2006). *Behavioral Social Choice: Probabilistic Models, Statistical Inference, and Applications*. Cambridge University Press.

Rey, A., Rothe, J., Schadrack, H., & Schend, L. (2016). Toward the complexity of the existence of wonderfully stable partitions and strictly core stable coalition structures in enemy-oriented hedonic games. *Annals of Mathematics and Artificial Intelligence*, *77*(3), 317–333.

Rothe, J. (2019). How can we model emotional and behavioral dynamics in collective decision making? In Laslier, J., Moulin, H., Sanver, R., & Zwicker, W. (Eds.), *The Future of Economic Design*, Studies in Economic Design, pp. 245–251. Springer.

Rothe, J. (2021). Thou shalt love thy neighbor as thyself when thou playest: Altruism in game theory. In *Proceedings of the 35th AAAI Conference on Artificial Intelligence*, pp. 15070–15077. AAAI Press.

Schlueter, J., & Goldsmith, J. (2020). Super altruistic hedonic games. In *Proceedings of the 33rd International Florida Artificial Intelligence Research Society Conference*, pp. 160–165. AAAI Press.

Sung, S., & Dimitrov, D. (2007). On core membership testing for hedonic coalition formation games. *Operations Research Letters*, *35*(2), 155–158.

Sung, S., & Dimitrov, D. (2010). Computational complexity in additive hedonic games. *European Journal of Operational Research*, *203*(3), 635–639.

von Neumann, J. (1928). Zur Theorie der Gesellschaftsspiele. *Mathematische Annalen*, *100*(1), 295–320.

von Neumann, J., & Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton University Press.

Wiechers, A., & Rothe, J. (2020). Stability in minimization-based altruistic hedonic games. In *Proceedings of the 9th European Starting AI Researchers' Symposium*, Vol. 2655. CEUR-WS.org.

Woeginger, G. (2013a). Core stability in hedonic coalition formation. In *Proceedings of the 39th International Conference on Current Trends in Theory and Practice of Computer Science*, pp. 33–50. Springer-Verlag *Lecture Notes in Computer Science #7741*.

Woeginger, G. (2013b). A hardness result for core stability in additive hedonic games. *Mathematical Social Sciences*, *65*(2), 101–104.

3.2 Popularity and Strict Popularity in Average-Based and Minimum-Based Altruistic Hedonic Games

The next article studies the problem of verifying popular and strictly popular coalition structures in average-based and minimum-based altruistic hedonic games.

Publication (Kerkmann and Rothe [88])

A. Kerkmann and J. Rothe. "Popularity and Strict Popularity in Average-Based and Minimum-Based Altruistic Hedonic Games". Submitted to the *47th International Symposium on Mathematical Foundations of Computer Science*. 2022

3.2.1 Summary

Considering hedonic games, the question of what accounts for a 'good' coalition structure naturally arises. There are several notions of stability in hedonic games that indicate whether an agent or a group of agents have an incentive to deviate from a given coalition structure. These concepts include, e.g., Nash stability, individual stability, and core stability. By contrast, this work studies popularity and strict popularity in hedonic games. These two concepts measure whether a given coalition structure is preferred to every other possible coalition structure by a (strict) majority of the agents.

We study popularity and strict popularity in (minimum-based) altruistic hedonic games [107, 145, 91] and determine the complexities of two decision problems. First, we consider the problem of verifying whether a given coalition structure in a given altruistic hedonic game is (strictly) popular. Second, we consider the existence problem which asks whether there exists a (strictly) popular coalition structure for a given altruistic hedonic game. While the complexity of these problems has been partly determined for strict popularity by Nguyen et al. [107] and Wiechers and Rothe [145], the problems have not been considered before for the notion of popularity. We solve all cases of strict popularity verification in (minimum-based) altruistic hedonic games that were left open by Nguyen et al. [107] and Wiechers and Rothe [145]. Furthermore, we completely determine the complexity of popularity verification in (minimum-based) altruistic hedonic games for all degrees of altruism. Our results reveal that all considered verification problems are coNP-complete. Additionally, we obtain some coNP-hardness results for strict popularity existence in equal-treatment and altruistic-treatment altruistic hedonic games. Besides, we infer that popularity verification is also coNP-complete for friend-oriented hedonic games.

3.2.2 Personal Contribution and Preceding Versions

A preliminary version of this paper has been accepted for publication at AAMAS'22 [87].

All technical results of the paper are my contribution. The writing and polishing was done jointly with Jörg Rothe.

3.2.3 Publication

The full article [88] is appended here.

Popularity and Strict Popularity in Average-Based and Minimum-Based Altruistic Hedonic Games

Anna Maria Kerkmann
Heinrich-Heine-Universität Düsseldorf, Germany

Jörg Rothe
Heinrich-Heine-Universität Düsseldorf, Germany

Abstract

We consider average- and min-based altruistic hedonic games and study the problem of verifying popular and strictly popular coalition structures. While strict popularity verification has been shown to be coNP-complete in min-based altruistic hedonic games, this problem has been open for equal- and altruistic-treatment average-based altruistic hedonic games. We solve these two open cases of strict popularity verification and then provide the first complexity results for popularity verification in (average- and min-based) altruistic hedonic games, where we cover all three degrees of altruism.

2012 ACM Subject Classification Theory of computation → Algorithmic game theory; Theory of computation → Solution concepts in game theory; Theory of computation → Problems, reductions and completeness

Keywords and phrases Cooperative Game Theory, Coalition Formation, Hedonic Game, Popularity, Altruism, Computational Complexity

Digital Object Identifier 10.4230/LIPIcs.CVIT.2016.23

Related Version *Preliminary 2-page extended abstract without any proofs*: [15]

Funding This work was supported in part by DFG grants RO 1202/14-2 and RO 1202/21-1.
Anna Maria Kerkmann: Anna is member of the PhD-programme "Online Participation," supported by the North Rhine-Westphalian funding scheme "Forschungskollegs."

1 Introduction

Much work has been done in recent years to study *hedonic games*, coalition formation games where players express their preferences over those coalitions that contain them. Drèze and Greenberg [9] were the first to propose hedonic games and Bogomolnaia and Jackson [4] and Banerjee *et al.* [3] formally defined and investigated them. For more background and the rich literature on hedonic games, we refer to the book chapters by Aziz and Savani [2] and Elkind and Rothe [10] and the survey by Woeginger [20].

We focus on *altruistic* hedonic games (AHGs) that, based on the friend-and-enemy encoding of the players' preferences due to Dimitrov *et al.* [8], were introduced by Nguyen *et al.* [16]. Schlueter and Goldsmith [18] generalized them to *"super AHGs,"* using ideas of the *"social distance games"* due to Brânzei and Larson [6]. Bullinger and Kober [7] introduced the related notion of *loyalty in hedonic games*. Nguyen *et al.* [16] defined three degrees of altruism depending on the order in which players take their own or their friends' preferences into account. They chose to model players' utilities by taking the *average* of these friends' valuations in the same coalition. Wiechers and Rothe [19] studied the same three degrees of altruism for *minimum-based* utilities, and Kerkmann and Rothe [14] applied the original model to *coalition formation games* in general. For an overview of various notions of altruism in game theory, we refer to the survey by Rothe [17].

We study both average- and min-based AHGs. For these two classes of games (and for hedonic games in general), many stability notions have been studied, including stability based on single-player deviations (such as Nash stability) or on deviations by groups of players (such as core stability) (see, e.g., Aziz and Savani [2], Elkind and Rothe [10], and Woeginger [20]). By contrast, for *popularity* and *strict popularity* we look at entire coalition structures (i.e., partitions of the players into coalitions) and ask—similarly to the notion of (weak) Condorcet winner in voting—whether a (strict) majority of players prefer a given coalition structure to every other coalition structure. Previous literature on popularity in hedonic games is, e.g., due to Aziz *et al.* [1], Brandt and Bullinger [5], and Kerkmann *et al.* [13]. We study the complexity of verifying (strictly) popular coalition structures in AHGs. While strict popularity verification is known to be coNP-complete in all three degrees of min-based AHGs [19] and for selfish-first average-based AHGs [16], its complexity was open for the other two degrees of average-based altruism.

We solve these two missing cases via technically rather involved constructions in Section 3. In addition, in Section 4 we provide the first complexity results for popularity verification in average- and min-based AHGs, covering for both all three degrees of altruism. We show that the problem in all cases is coNP-complete. Having closed all open problems for (strict) popularity verification in all these models, we conclude our work and give some future work directions in Section 5.

2 Preliminaries

We consider a set $N = [n]$ of n *players* (or *agents*), where subsets of the players can form *coalitions* (and we use the notation $[n] = \{1, \ldots, n\}$ for any integer n). For any player $i \in N$, $\mathcal{N}^i = \{C \subseteq N \mid i \in C\}$ denotes the set of coalitions containing i. A *coalition structure* is a partition $\Gamma = \{C_1, \ldots, C_k\}$ of the players into coalitions (i.e., $\bigcup_{i=1}^{k} C_i = N$ and $C_i \cap C_j = \emptyset$ for all $i, j \in [k]$ with $i \neq j$), where the coalition containing player i is denoted by $\Gamma(i)$. $\mathcal{C}_N$ is the set of all coalition structures for a set of agents N.

A *coalition formation game* is a pair $(N, \succeq)$, where N is a set of agents, $\succeq$ is a profile of preferences, and every preference $\succeq_i \subseteq \mathcal{C}_N \times \mathcal{C}_N$ is a complete weak order over all coalition structures. For coalition structures $\Gamma, \Delta \in \mathcal{C}_N$, we say that agent i *weakly prefers* Γ *to* Δ if $\Gamma \succeq_i \Delta$, that i *prefers* Γ *to* Δ ($\Gamma \succ_i \Delta$) if $\Gamma \succeq_i \Delta$ but not $\Delta \succeq_i \Gamma$, and that i is *indifferent between* Γ *and* Δ ($\Gamma \sim_i \Delta$) if $\Gamma \succeq_i \Delta$ and $\Delta \succeq_i \Gamma$.

A *hedonic game* is a coalition formation game $(N, \succeq)$ where the preference $\succeq_i$ of any agent $i \in N$ only depends on the coalitions that she is part of. This means that i is indifferent between any two coalition structures $\Gamma, \Delta \in \mathcal{C}_N$ as long as her coalition is the same, i.e., $\Gamma(i) = \Delta(i) \implies \Gamma \sim_i \Delta$. i's preference can then be represented by a complete weak order over the set $\mathcal{N}^i$ of coalitions containing i. For $A, B \in \mathcal{N}^i$, we say that player i *weakly prefers* A *to* B if $A \succeq_i B$ (and analogously for (strict) preference and indifference).

2.1 Altruistic Hedonic Games

Nguyen *et al.* [16] used the friends-and-enemies encoding by Dimitrov *et al.* [8] when first introducing altruistic hedonic games (AHGs). Under this encoding, each player i partitions the other players into a set of friends F_i and a set of enemies E_i, and assigns the following friend-oriented value to a coalition $A \in \mathcal{N}^i$:

$$v_i(A) = n|A \cap F_i| - |A \cap E_i|.$$

The friendship relations, which are assumed to be mutual, can then be represented by a *network of friends*, an undirected graph where two players are connected by an edge if and only if they are friends of each other.

Nguyen *et al.* [16] introduced altruism into an agent's preference by incorporating the average of her friends' valuations (of the friends that are in the same coalition) into her utility. Wiechers and Rothe [19] vary this model by considering the minimum instead. For any $A \in \mathcal{N}^i$, we use:

$$\mathrm{avg}_i^F(A) = \sum_{a \in A \cap F_i} \frac{v_a(A)}{|A \cap F_i|}; \qquad \mathrm{avg}_i^{F+}(A) = \sum_{a \in (A \cap F_i) \cup \{i\}} \frac{v_a(A)}{|(A \cap F_i) \cup \{i\}|}; \tag{1}$$

$$\min_i^F(A) = \min_{a \in A \cap F_i} v_a(A); \text{ and} \qquad \min_i^{F+}(A) = \min_{a \in (A \cap F_i) \cup \{i\}} v_a(A), \tag{2}$$

where the minimum of the empty set is defined as zero. We also define these values for coalition structures $\Gamma \in \mathcal{C}_N$, e.g., by $\mathrm{avg}_i^F(\Gamma) = \mathrm{avg}_i^F(\Gamma(i))$. The three *degrees of altruism*, introduced by Nguyen *et al.* [16], are the following. For a constant $M \geq n^5$, agent i's

- *selfish-first (SF) preference* is defined by $A \succeq_i^{SF} B \iff u_i^{SF}(A) \geq u_i^{SF}(B)$, with the SF utility function $u_i^{SF}(A) = M \cdot v_i(A) + \mathrm{avg}_i^F(A)$;
- *equal-treatment (EQ) preference* is defined by $A \succeq_i^{EQ} B \iff u_i^{EQ}(A) \geq u_i^{EQ}(B)$, with the EQ utility function $u_i^{EQ}(A) = \mathrm{avg}_i^{F+}(A)$; and
- *altruistic-treatment (AL) preference* is defined by $A \succeq_i^{AL} B \iff u_i^{AL}(A) \geq u_i^{AL}(B)$, with the AL utility function $u_i^{AL}(A) = v_i(A) + M \cdot \mathrm{avg}_i^F(A)$.

The constant factor $M \geq n^5$ ensures that the SF preference is first determined by the agent's own valuation for her coalition while the AL preference is first determined by her friends' valuations (see [16, Theorems 1 & 2]). The min-based altruistic preferences are defined analogously, using the minimum (see (2)) instead of the average. They will be denoted by $\succeq^{minSF}$, $\succeq^{minEQ}$, and $\succeq^{minAL}$.

A pair $(N, \succeq)$, where $\succeq$ is a profile of preferences defined by one of the average-based degrees of altruism, is called an *altruistic hedonic game* (AHG) with *average-based altruistic preferences* $\succeq$. A game $(N, \succeq^{\min})$ with *min-based altruistic preferences* $\succeq^{\min}$ is said to be a *min-based altruistic hedonic game* (MBAHG). Based on the degree of altruism, we call, say, an AHG with SF preferences an *SF AHG*, etc.

2.2 Popularity

We now define popularity, which is based on the pairwise comparison of coalition structures. For a hedonic game $(N, \succeq)$ and two coalition structures $\Gamma, \Delta \in \mathcal{C}_N$, let $\#_{\Gamma \succ \Delta} = |\{i \in N \mid \Gamma \succ_i \Delta\}|$ be the number of players that prefer Γ to Δ. A coalition structure $\Gamma \in \mathcal{C}_N$ is *popular* (respectively, *strictly popular*) if, for every other coalition structure $\Delta \in \mathcal{C}_N, \Delta \neq \Gamma$, it holds that $\#_{\Gamma \succ \Delta} \geq \#_{\Delta \succ \Gamma}$ (respectively, $\#_{\Gamma \succ \Delta} > \#_{\Delta \succ \Gamma}$). Define the problems:

P-Veri	
Given:	A hedonic game $(N, \succeq)$ and a coalition structure Γ.
Question:	Is Γ popular in $(N, \succeq)$?

P-Exi	
Given:	A hedonic game $(N, \succeq)$.
Question:	Is there a popular coalition structure in $(N, \succeq)$?

The strict variants of the problems, SP-Veri and SP-Exi, are defined analogously.

In the following two sections, we will solve the two missing cases of Nguyen *et al.* [16] by showing that SP-Veri is coNP-complete for EQ and AL AHGs, and we will also show that P-Veri is coNP-complete as well for all three degrees of altruism in AHGs and MBAHGs.

It is easy to see that all these verification problems are in coNP (cf. Nguyen *et al.* [16, Theorem 12]). To show their coNP-hardness, we reduce from the complement of the following NP-complete problem [11, 12]:

Restricted Exact Cover by 3-Sets (RX3C)	
Given:	A set $B = \{1, \dots, 3k\}$ (for some integer $k \geq 2$) and a collection $\mathscr{S} = \{S_1, \dots, S_{3k}\}$ of 3-element subsets of B, where each element of B occurs in exactly three sets in $\mathscr{S}$.
Question:	Does there exist an exact cover of B in $\mathscr{S}$, i.e., a subset $\mathscr{S}' \subseteq \mathscr{S}$ of size k such that every element of B occurs in exactly one set in $\mathscr{S}'$?

Specifically, to prove coNP-hardness of (strict) popularity verification, we construct from an RX3C instance $(B, \mathscr{S})$ the network of friends of a hedonic game $(N, \succeq)$ and a coalition structure Γ and show that Γ is *not* (strictly) popular under the considered model if and only if there exists an exact cover of B in $\mathscr{S}$.

3 Strict Popularity in AHGs

We start with strict popularity. While Wiechers and Rothe [19] showed that SP-Veri is coNP-complete for all three degrees of altruism in MBAHGs, Nguyen *et al.* [16] showed the same result only for SF AHGs. We solve the two missing cases (i.e., for EQ and AL) in Theorems 3 and 9. In their proofs, we will use the following two observations. The first observation says that, under EQ and AL, a player i prefers adding a friend's friend k to her current coalition provided that k is not her friend.

▶ **Observation 1.** *For any $D \in \mathcal{N}^i$, $j \in F_i \cap D$, and $k \in (F_j \setminus F_i) \setminus D$, it holds that $D \cup \{k\} \succ_i^{EQ} D$ and $D \cup \{k\} \succ_i^{AL} D$.*

Proof. It holds that

$$
\begin{aligned}
u_i^{EQ}(D \cup \{k\}) = \mathrm{avg}_i^{F+}(D \cup \{k\}) &= \sum_{a \in (D \cap F_i) \cup \{i\}} \frac{v_a(D \cup \{k\})}{|(D \cap F_i) \cup \{i\}|} \\
&= \frac{1}{|(D \cap F_i) \cup \{i\}|} \cdot \Big(v_i(D \cup \{k\}) + v_j(D \cup \{k\}) + \sum_{a \in (D \cap F_i) \setminus \{j\}} v_a(D \cup \{k\}) \Big) \\
&\geq \frac{1}{|(D \cap F_i) \cup \{i\}|} \cdot \Big(v_i(D) - 1 + v_j(D) + n + \sum_{a \in (D \cap F_i) \setminus \{j\}} (v_a(D) - 1) \Big) \\
&= \frac{1}{|(D \cap F_i) \cup \{i\}|} \cdot \Big(\sum_{a \in (D \cap F_i) \cup \{i\}} v_a(D) + n - |D \cap F_i| \Big) \\
&> \mathrm{avg}_i^{F+}(D) = u_i^{EQ}(D)
\end{aligned}
$$

and, by similar transformations of equations, we obtain that $\mathrm{avg}_i^F(D \cup \{k\}) > \mathrm{avg}_i^F(D)$. Thus $D \cup \{k\} \succ_i^{EQ} D$ and $D \cup \{k\} \succ_i^{AL} D$. ❏

With the help of Observation 1, the following is implied.

▶ **Observation 2.** *If player $i \in N$ has only one friend j (i.e., $F_i = \{j\}$), then $C = \{j\} \cup F_j$ is i's unique most preferred coalition under EQ and AL.*

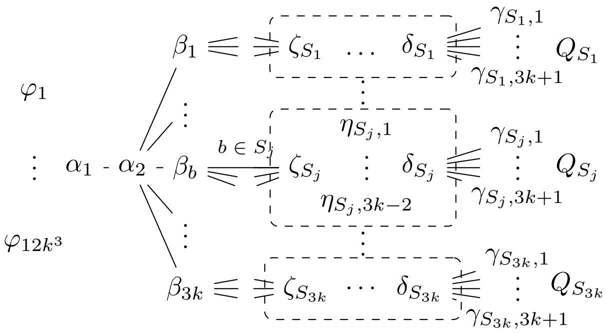


Figure 1 Network of friends in the proof of Theorem 3. A dashed rectangle indicates that all players inside are friends of each other.

Proof. The proof is the same for $\succeq_i^{EQ}$ and $\succeq_i^{AL}$. We will simply use u_i for which either u_i^{EQ} or u_i^{AL} can be substituted. Assume that $D \neq C$ is one of i's most preferred coalitions. Then $u_i(D) \geq u_i(C)$. It is obvious that $D \subseteq C$ because every player in $N \setminus C$ is an enemy of i's and j's and can thus only decrease i's utility. Further, since j is i's only friend, it is clear that $j \in D$ (otherwise, we would have $u_i(D) \leq 0 < u_i(C)$). Then, by Observation 1, it follows that D contains all friends of j's. Hence, $D = C$, which is a contradiction. ❑

We are now ready to solve the two problems that Nguyen *et al.* [16] left open for the complexity of SP-Veri, namely for EQ AHGs and AL AHGs. We start with the former.

▶ **Theorem 3.** SP-Veri *is* coNP-*complete for EQ AHGs.*

Proof. Given an instance of $(B, \mathscr{S})$ of RX3C, with $B = \{1, \ldots, 3k\}$ and $\mathscr{S} = \{S_1, \ldots, S_{3k}\}$, we define the set of players $N = P \cup A \cup \bigcup_{S \in \mathscr{S}} Q_S$ with $P = \{\varphi_1, \ldots, \varphi_{12k^3}\}$, $A = \{\alpha_1, \alpha_2\} \cup \{\beta_b \mid b \in B\}$, and $Q_S = \{\zeta_S, \eta_{S,j}, \delta_S, \gamma_{S,\ell} \mid j \in [3k-2], \ell \in [3k+1]\}$ for every $S \in \mathscr{S}$. We then construct the network of friends shown in Figure 1 and define $\Gamma = \{\{\varphi_1\}, \ldots, \{\varphi_{12k^3}\}, A, Q_{S_1}, \ldots, Q_{S_{3k}}\}$. It holds that $n = |N| = 12k^3 + 2 + 3k + 3k(6k+1) = 12k^3 + 18k^2 + 6k + 2$.

Specifically, the friendship relationships are as follows:

- α_2 is friends with α_1 and every $\beta_b, b \in B$.
- For $S \in \mathscr{S}$, ζ_S is friends with the three β_b with $b \in S$.
- For $S \in \mathscr{S}$, all players in $\{\zeta_S, \eta_{S,j}, \delta_S \mid j \in [3k-2]\}$ are friends of each other.
- For $S \in \mathscr{S}$, δ_S is friends with every $\gamma_{S,\ell}, \ell \in [3k+1]$.

The idea of this proof is to show that there can only be a coalition structure Δ that is equally popular to Γ if and only if there is an exact cover for B. We start by stating some useful claims. The first two claims are direct consequences of Observation 2 and the third claim is obvious as the φ-players do not have any friends.

▷ Claim 4. α_1 prefers A to every other coalition.

▷ Claim 5. For every $S \in \mathscr{S}$ and $\ell \in [3k+1]$, $\gamma_{S,\ell}$ prefers Q_S to every other coalition.

▷ Claim 6. For $h \in [12k^3]$, φ_h prefers $\{\varphi_h\}$ to every other coalition.

We further need the following two claims whose proofs are deferred to Section A.1 of the appendix. Note that the proof of Claim 7 is technically rather involved.

▷ Claim 7. For $S \in \mathscr{S}$, ζ_S prefers Q_S to every other coalition.

▷ Claim 8. If β_b with $b \in B$ prefers Δ to Γ, then $\zeta_S \in \Delta(\beta_b)$ for some $S \in \mathscr{S}$ with $b \in S$.

Now, using these claims, we will show that there is an exact cover of B if and only if Γ is not strictly popular.

Only if: Assume there exists an exact cover $\mathscr{S}' \subseteq \mathscr{S}$ of B. Then, for the coalition structure $\Delta = \{\{\varphi_1\}, \dots, \{\varphi_{12k^3}\}, A \cup \bigcup_{S \in \mathscr{S}'} Q_S\} \cup \{Q_S \mid S \in \mathscr{S} \setminus \mathscr{S}'\}$, we can show that Δ and Γ are equally popular: All players in Q_S with $S \in \mathscr{S} \setminus \mathscr{S}'$ and all $\varphi_h, h \in [12k^3]$, are obviously indifferent between Γ and Δ. By Claims 4, 5, and 7, α_1, $\gamma_{S,\ell}$, and ζ_S, with $S \in \mathscr{S}'$ and $\ell \in [3k+1]$, prefer Γ to Δ. The remaining players prefer Δ to Γ: For α_2, it holds that

$$
\begin{aligned}
u^{EQ}_{\alpha_2}(\Delta) &= \frac{1}{3k+2} \cdot \Big(\sum_{a \in \{\alpha_1, \alpha_2, \beta_1, \dots, \beta_{3k}\}} v_a(\Delta) \Big) \\
&= \frac{1}{3k+2} \cdot \Big(v_{\alpha_1}(\Gamma) - k(6k+1) + v_{\alpha_2}(\Gamma) - k(6k+1) + \sum_{b \in [3k]} (v_{\beta_b}(\Gamma) + n - k(6k+1) + 1) \Big) \\
&= \frac{1}{3k+2} \cdot \Big(\sum_{a \in \{\alpha_1, \alpha_2, \beta_1, \dots, \beta_{3k}\}} v_a(\Gamma) \Big) + \frac{1}{3k+2} \cdot \Big(-(3k+2)k(6k+1) + 3k(n+1) \Big) \\
&= u^{EQ}_{\alpha_2}(\Gamma) + \frac{k}{3k+2} \cdot \Big(-(3k+2)(6k+1) + 3(n+1) \Big) \\
&= u^{EQ}_{\alpha_2}(\Gamma) + \frac{k}{3k+2} \cdot \Big(36k^3 + 36k^2 + 3k + 7 \Big) > u^{EQ}_{\alpha_2}(\Gamma).
\end{aligned}
$$

For any $b \in [3k]$, β_b is part of exactly one $S \in \mathscr{S}'$. Thus there is exactly one ζ_S in $\Delta(\beta_b)$ that is her friend and we have

$$
\begin{aligned}
u^{EQ}_{\beta_b}(\Delta) &= \frac{1}{3}\Big(v_{\beta_b}(\Delta) + v_{\alpha_2}(\Delta) + v_{\zeta_S}(\Delta) \Big) = \frac{1}{3}\Big(2n - (3k + k(6k+1) - 1) \\
&\quad + n(3k+1) - k(6k+1) + n(3k+2) - (k(6k+1) - 1) \Big) \\
&= \frac{1}{3}\Big(n(6k+5) - (3k(6k+1) + 3k - 2) \Big) = n(2k + \tfrac{5}{3}) - (k(6k+2) - \tfrac{2}{3}) \\
&> n\left(\tfrac{3}{2} \cdot k + 1\right) - \tfrac{3}{2} \cdot k \\
&= \frac{1}{2}\big(n - 3k + n(3k+1) \big) = \frac{1}{2}\big(v_{\beta_b}(\Gamma) + v_{\alpha_2}(\Gamma) \big) = u^{EQ}_{\beta_b}(\Gamma).
\end{aligned}
$$

For $\eta_{S,j}$ and δ_S with $S \in \mathscr{S}'$ and $j \in [3k-2]$, we can similarly compute that $u^{EQ}_{\eta_{S,j}}(\Delta) = u^{EQ}_{\eta_{S,j}}(\Gamma) + 6k^2 + 20k + 5 + \frac{3}{k}$ and $u^{EQ}_{\delta_S}(\Delta) = u^{EQ}_{\delta_S}(\Gamma) + \frac{60k^2+14k+8}{6k+1}$.[1]

Overall, we have $\#_{\Delta \succ \Gamma} = |\{\alpha_2, \beta_1, \dots, \beta_{3k}\} \cup \{\eta_{S,j}, \delta_S \mid S \in \mathscr{S}', j \in [3k-2]\}| = 1 + 3k + k(3k-1) = 1 + k(3k+2) = |\{\alpha_1\} \cup \{\zeta_S, \gamma_{S,\ell} \mid S \in \mathscr{S}', \ell \in [3k+1]\}| = \#_{\Gamma \succ \Delta}$.

Hence, Γ is not strictly popular.

If: Assume that Γ is not strictly popular under EQ, i.e., there is a coalition structure $\Delta \neq \Gamma$ with $\#_{\Delta \succ \Gamma} \geq \#_{\Gamma \succ \Delta}$. Let $k' = |\{S \in \mathscr{S} \mid Q_S \notin \Delta\}$ be the number of sets Q_S that are not a coalition in Δ. Then, by Claims 5 and 7, all $\gamma_{S,\ell}$ and ζ_S from these k' sets Q_S prefer Γ to Δ. Further, no φ_h can ever prefer Δ to Γ, and all players in the $3k - k'$ sets $Q_S \in \Delta$ are indifferent between Γ and Δ.

First, observe that $k' \geq 1$. If $k' = 0$ then, for every $S \in \mathscr{S}$, Q_S is a coalition in Δ. Then, by Claim 8, no β_b prefers Δ to Γ and, obviously, β_b can only be indifferent between Γ and Δ

[1] Detailed calculations are given in Section A.2 of the appendix.

if $\Delta(\beta_b) = A$. It follows that $A \in \Delta$ because otherwise all β_b would prefer Γ to Δ and there would thus be more players who prefer Γ to Δ than vice versa. However, this means that $\Delta = \Gamma$, which is a contradiction.

Second, observe that A is not a coalition in Δ. If this were the case, all players in A were indifferent between Γ and Δ. Then, $\#_{\Gamma \succ \Delta} \geq |\{\zeta_S, \gamma_{S,1}, \ldots, \gamma_{S,3k+1} \mid S \in \mathscr{S}'\}| = k' \cdot (3k+2)$ and $\#_{\Delta \succ \Gamma} \leq |\{\eta_{S,1}, \ldots, \eta_{S,3k-2}, \delta_S \mid S \in \mathscr{S}'\}| = k' \cdot (3k-1)$. With $k' \geq 1$, this contradicts $\#_{\Delta \succ \Gamma} \geq \#_{\Gamma \succ \Delta}$.

Third, observe that $k' \leq k$. For a contradiction, assume that $k' > k$. Since $A \notin \Delta$, we know by Claim 4 that α_1 prefers Γ to Δ. So, we have $\#_{\Gamma \succ \Delta} \geq |\{\alpha_1\} \cup \{\zeta_S, \gamma_{S,1}, \ldots, \gamma_{S,3k+1} \mid S \in \mathscr{S}'\}| = 1 + k' \cdot (3k+2)$ and $\#_{\Delta \succ \Gamma} \leq |\{\alpha_2, \beta_1, \ldots, \beta_{3k}\} \cup \{\eta_{S,1}, \ldots, \eta_{S,3k-2}, \delta_S \mid S \in \mathscr{S}'\}| = 1 + 3k + k' \cdot (3k-1)$. Overall, $\#_{\Delta \succ \Gamma} \leq 1 + 3k + k' \cdot (3k-1) < 1 + 3k' + k' \cdot (3k-1) = 1 + k' \cdot (3k+2) = \#_{\Gamma \succ \Delta}$, which is a contradiction.

Finally, observe that $k' \geq k$. For a contradiction, assume that $k' < k$. Because of Claim 8 we then know that at most $3k'$ β-players prefer Δ to Γ. The remaining $3k - 3k'$ β_b do not have any ζ_S with $b \in S$ in their coalitions in Δ. Together with $A \notin \Delta$, it follows that these $3k - 3k'$ β-players prefer Γ to Δ. Hence, $\#_{\Gamma \succ \Delta} \geq 1 + 3k - 3k' + k' \cdot (3k+2) = 1 + 3k + k' \cdot (3k-1) > 1 + 3k' + k' \cdot (3k-1) \geq \#_{\Delta \succ \Gamma}$. This contradicts $\#_{\Delta \succ \Gamma} \geq \#_{\Gamma \succ \Delta}$.

Since we have $k' = k$, for $\#_{\Delta \succ \Gamma} \geq \#_{\Gamma \succ \Delta}$ to hold, every β_b needs to prefer Δ to Γ. By Claim 8, this is only possible if every β_b has a ζ_S with $b \in S$ in her coalition in Δ. This implies that $\{S \in \mathscr{S} \mid Q_S \notin \Delta\}$ is an exact cover of B. ❑

For strict popularity in AL AHGs, we can use the same construction but have to modify our arguments appropriately.

▶ **Theorem 9.** SP-VERI *is* coNP*-complete for AL AHGs.*

Proof. Consider the construction from Theorem 3 again with the network of friends shown in Figure 1. Only some details in the proof of correctness are different when considering AL instead of EQ. We again start our proof by stating some claims. Claims 4, 5, 6, and 8 from the proof of Theorem 3 also hold for AL: α_1 prefers A to every other coalition; for every $S \in \mathscr{S}$ and $\ell \in [3k+1]$, $\gamma_{S,\ell}$ prefers Q_S to every other coalition; for every $h \in [12k^3]$, φ_h prefers $\{\varphi_h\}$ to every other coalition; and if β_b prefers Δ to Γ, then $\zeta_S \in \Delta(\beta_b)$ for some $S \in \mathscr{S}$ with $b \in S$. In addition, we have the following claim whose proof is deferred to Section A.3 due space constraints.

▷ Claim 10. For $S \in \mathscr{S}$, ζ_S prefers $\{\zeta_S, \delta_S, \gamma_{S,1}, \ldots, \gamma_{S,3k+1}\}$ and every coalition $\{\zeta_S, \delta_S, \gamma_{S,1}, \ldots, \gamma_{S,3k+1}\} \setminus \{\gamma_{S,\ell}\}$, $\ell \in [3k+1]$, to Q_S, and ζ_S prefers Q_S to every other coalition.

We now show that there is an exact cover of B if and only if Γ is not strictly popular under AL.

Only if: Assume that there is an exact cover $\mathscr{S}' \subseteq \mathscr{S}$ of B. As in the proof of Theorem 3, let $\Delta = \{\{\varphi_1\}, \ldots, \{\varphi_{12k^3}\}, A \cup \bigcup_{S \in \mathscr{S}'} Q_S\} \cup \{Q_S \mid S \in \mathscr{S} \setminus \mathscr{S}'\}$. We will show that Δ and Γ are equally popular under AL.

Omitting the detailed calculations, we have $\mathrm{avg}^F_{\alpha_2}(\Delta) = \mathrm{avg}^F_{\alpha_2}(\Gamma) + \frac{3k(n+1)-(3k+1)k(6k+1)}{3k+1}$, $\mathrm{avg}^F_{\eta_{S,j}}(\Delta) = \mathrm{avg}^F_{\eta_{S,j}}(\Gamma) + \frac{3n+3-(3k-1)\left(3k+2+(k-1)(6k+1)\right)}{3k-1}$, and $\mathrm{avg}^F_{\delta_S}(\Delta) = \mathrm{avg}^F_{\delta_S}(\Gamma) + \frac{3n+3-6k\left(3k+2+(k-1)(6k+1)\right)}{6k}$ for $S \in \mathscr{S}'$ and $j \in [3k-2]$. Also using the preceding claims, it then follows that $\#_{\Delta \succ \Gamma} = |\{\alpha_2, \beta_1, \ldots, \beta_{3k}\} \cup \{\eta_{S,1}, \ldots, \eta_{S,3k-2}, \delta_S \mid S \in \mathscr{S}'\}| = 1 + 3k + k(3k-1) = 1 + k(3k+2) = |\{\alpha_1\} \cup \{\zeta_S, \gamma_{S,1}, \ldots, \gamma_{S,3k+1} \mid S \in \mathscr{S}'\}| = \#_{\Gamma \succ \Delta}$.

Hence, Γ is not strictly popular.

If: Assume that Γ is not strictly popular, i.e., that there is a coalition structure $\Delta \neq \Gamma$ with $\#_{\Delta \succ \Gamma} \geq \#_{\Gamma \succ \Delta}$.

First, note that there is no $S \in \mathscr{S}$ with $\{\zeta_S, \delta_S, \gamma_{S,1}, \dots, \gamma_{S,3k+1}\}$ in Δ or $\{\zeta_S, \delta_S, \gamma_{S,1}, \dots, \gamma_{S,3k+1}\} \setminus \{\gamma_{S,\ell}\}$ in Δ for any $\ell \in [3k+1]$: For the sake of contradiction, assume that there is such an $S \in \mathscr{S}$. Then, ζ_S prefers Δ to Γ; and $\eta_{S,j}$, δ_S, and $\gamma_{S,\ell}$ for $j \in [3k-2]$ and $\ell \in [3k+1]$ prefer Γ to Δ. For ζ_S and $\gamma_{S,\ell}$, this follows from the preceding claims. For $\eta_{S,j}$ and δ_S, this can be shown by direct calculations (see Section A.4 of the appendix). Then $\#^{\text{in } Q_S}_{\Delta \succ \Gamma} = 1$ and $\#^{\text{in } Q_S}_{\Gamma \succ \Delta} = 6k$. For all other $S' \in \mathscr{S}$, it holds that $\#^{\text{in } Q_{S'}}_{\Delta \succ \Gamma} \leq 3k$ and $\#^{\text{in } Q_{S'}}_{\Gamma \succ \Delta} \geq 3k+1$ if $Q_{S'} \notin \Delta$; and $\#^{\text{in } Q_{S'}}_{\Delta \succ \Gamma} = \#^{\text{in } Q_{S'}}_{\Gamma \succ \Delta} = 0$ if $Q_{S'} \in \Delta$. This means that, in $Q_{S'}$, at least as many players prefer Γ to Δ as the other way around. Then only the players in $\{\alpha_2, \beta_1, \dots, \beta_{3k}\}$ could prefer Δ to Γ. However, since $\#^{\text{in } Q_S}_{\Gamma \succ \Delta} = 6k$, this means that $\#_{\Gamma \succ \Delta} > \#_{\Delta \succ \Gamma}$, a contradiction. The remainder of the proof proceeds identically to the *If*-part in the proof of Theorem 3. ❑

From Theorems 3 and 9, we get the following corollary.

▶ **Corollary 11.** SP-Exi *is* coNP*-hard for EQ and AL AHGs.*

Proof. We use the same reduction as in the proof of Theorem 3 but do not give any coalition structure as a part of the instance. It then holds that there exists a strictly popular coalition structure for the defined game if and only if there is no exact cover of B. The correctness of this equivalence follows from the proofs of Theorems 3 and 9. Indeed, Γ as defined in the proof of Theorem 3 is strictly popular under EQ and AL preferences if there is no exact cover. If, on the other hand, there does exist an exact cover, then Δ as defined in the proof of Theorem 3 is as popular as Γ while there is still no coalition structure that is more popular than Γ. Hence, no strictly popular coalition structure can exist in this case. ❑

4 Popularity in AHGs and MBAHGs

Now, we provide the first complexity results for P-Veri in AHGs and MBAHGs, and we cover for both all three degrees of altruism. As mentioned earlier, Nguyen *et al.* [16, Theorem 12] showed that SP-Veri is coNP-complete for SF AHGs and Wiechers and Rothe [19, Theorem 4] showed the same result for SF MBAHGs. We modify their proofs to establish the same results for P-Veri.

▶ **Theorem 12.** P-Veri *is* coNP*-complete for SF AHGs and SF MBAHGs.*

Proof. The proof of this theorem, which is the same for SF AHGs and SF MBAHGs, is inspired by the proofs of Nguyen *et al.* [16, Theorem 12] and Wiechers and Rothe [19, Theorem 4] for SP-Veri. Given an instance $(B, \mathscr{S})$ of RX3C, with $B = \{1, \dots, 3k\}$ and $\mathscr{S} = \{S_1, \dots, S_{3k}\}$, we construct the network of friends shown in Figure 2 with the set of players $N = \{\alpha\} \cup \{\beta_b \mid b \in B\} \cup \bigcup_{S \in \mathscr{S}} Q_S$, where $Q_S = \{\zeta_S, \eta_{S,j} \mid j \in [4]\}$ for $S \in \mathscr{S}$, and we define the coalition structure $\Gamma = \{\{\alpha, \beta_1, \dots, \beta_{3k}\}, Q_{S_1}, \dots, Q_{S_{3k}}\}$. Specifically, the friendship relationships are:

- All players in $\{\alpha\} \cup \{\beta_b \mid b \in B\}$ are friends.
- For $S \in \mathscr{S}$, ζ_S is friends with α and all β_b with $b \in S$.
- For $S \in \mathscr{S}$, all players in Q_S are friends of each other.

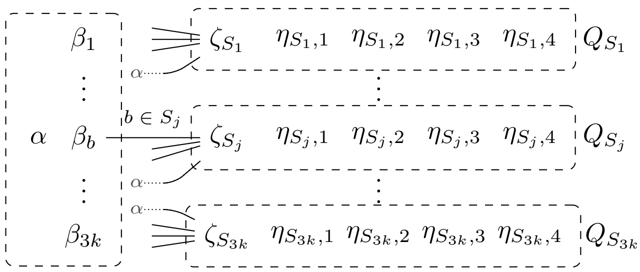


Figure 2 Network of friends in the proof of Theorem 12. A dashed rectangle indicates that all players inside are friends of each other.

We show that Γ is not popular if and only if there exists an exact cover of B in $\mathscr{S}$.

If: Assume that there exists an exact cover $\mathscr{S}' \subseteq \mathscr{S}$ of B. Then, the coalition structure $\Delta = \{\{\alpha, \beta_1, \dots, \beta_{3k}\} \cup \bigcup_{S \in \mathscr{S}'} Q_S\} \cup \{Q_S \mid S \in \mathscr{S} \setminus \mathscr{S}'\}$ is more popular than Γ:

$$
\begin{aligned}
\#_{\Delta \succ \Gamma} &= |\{\alpha, \beta_1, \dots, \beta_{3k}\} \cup \{\zeta_S \mid S \in \mathscr{S}'\}| = 1 + 3k + k \\
&> 4k = |\{\eta_{S,j} \mid S \in \mathscr{S}', j \in [4]\}| = \#_{\Gamma \succ \Delta}.
\end{aligned}
$$

Hence, Γ is not popular.

Only if: Assume that Γ is not popular, so there is a coalition structure $\Delta \neq \Gamma$ with $\#_{\Delta \succ \Gamma} > \#_{\Gamma \succ \Delta}$. First observe, for any $S \in \mathscr{S}$ and $j \in [4]$, that Q_S is $\eta_{S,j}$'s unique most preferred coalition, as it contains all of her friends and none of her enemies. Thus $\eta_{S,j}$ prefers Γ to Δ if $Q_S \notin \Delta$, and is indifferent if $Q_S \in \Delta$.

Now, let $k' = |\{S \in \mathscr{S} \mid Q_S \notin \Delta\}$ be the number of sets Q_S that are not a coalition in Δ.

Assume that $k' > k$. Then

$$
\begin{aligned}
&\#_{\Gamma \succ \Delta} \geq |\{\eta_{S,j} \mid Q_S \notin \Delta, j \in [4]\}| = 4k' \text{ and} \\
&\#_{\Delta \succ \Gamma} \leq |\{\alpha, \beta_1, \dots, \beta_{3k}\} \cup \{\zeta_S \mid Q_S \notin \Delta\}| = 3k + 1 + k' < 4k' + 1.
\end{aligned}
$$

Since $\#_{\Delta \succ \Gamma}$ is integral, this implies $\#_{\Gamma \succ \Delta} \geq 4k' \geq \#_{\Delta \succ \Gamma}$, a contradiction. Hence, $k' \leq k$.

Next, assume that $k' < k$. For any $b \in B$, observe that $\Gamma(\beta_b) = \{\alpha, \beta_1, \dots, \beta_{3k}\}$ is a clique. Hence, β_b can only prefer Δ to Γ if there are at least $3k+1$ of her friends in $\Delta(\beta_b)$, i.e., there is at least one ζ_S with $b \in S$ in $\Delta(\beta_b)$. Since there are k' ζ_S available (with $Q_S \notin \Delta$), there thus are at most $3k'$ β-players who prefer Δ to Γ. All other β-players (at least $3k - 3k'$) prefer Γ to Δ. Note that they are not indifferent between the two coalition structures: They would only be indifferent if $\{\alpha, \beta_1, \dots, \beta_{3k}\} \in \Delta$. However, this is not possible as it would imply that Δ is not more popular than Γ. We now have

$$
\begin{aligned}
\#_{\Gamma \succ \Delta} &\geq |\{\eta_{S,j} \mid Q_S \notin \Delta, j \in [4]\} \cup \{\beta_b | \text{ there is no } Q_S \notin \Delta \text{ with } b \in S\}| \\
&= 4k' + 3k - 3k' = 3k + k' > 4k' \text{ and} \\
\#_{\Delta \succ \Gamma} &\leq |\{\alpha\} \cup \{\beta_b | \text{ there is an } Q_S \notin \Delta \text{ with } b \in S\} \cup \{\zeta_S \mid Q_S \notin \Delta\}| \\
&= 1 + 3k' + k' = 4k' + 1.
\end{aligned}
$$

Since $\#_{\Gamma \succ \Delta}$ is integral, this implies $\#_{\Gamma \succ \Delta} \geq 4k' + 1 \geq \#_{\Delta \succ \Gamma}$, which is a contradiction. Thus we have $k' = k$.

Now, since exactly $4k$ η-players prefers Γ to Δ and because of $\#_{\Delta \succ \Gamma} > \#_{\Gamma \succ \Delta}$, $4k + 1$ players need to prefer Δ to Γ. Thus α, all β_b with $b \in B$, and all ζ_S with $Q_S \notin \Delta$ prefer Δ to Γ. As observed earlier, this means that every β_b has a ζ_S with $b \in S$ in her coalition in Δ.

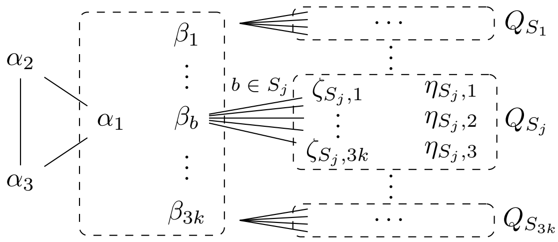


Figure 3 Network of friends in the proof of Theorem 15. A dashed rectangle indicates that all players inside are friends of each other.

Since there are $3k$ β-players and k ζ_S with $Q_S \notin \Delta$, this implies that $\{S \in \mathscr{S} \mid Q_S \notin \Delta\}$ is an exact cover of B. ❑

Since the altruistic tie-breaker is never used in the construction of Theorem 12, we get the following corollary.

▶ **Corollary 13.** P-VERI *is* coNP-*complete for friend-oriented hedonic games.*

With a slight adaptation of the construction in the proof of Theorem 3 we can show the following theorem.

▶ **Theorem 14.** P-VERI *is* coNP-*complete for EQ AHGs and AL AHGs.*

Proof. Consider the same construction as in the proof of Theorem 3 but delete player α_1 who under EQ and AL preferred Γ to the equally popular coalition structure Δ. Then $\Gamma' = \{\{\varphi_1\}, \dots, \{\varphi_{12k^3}\}, A \setminus \{\alpha_1\}, Q_{S_1}, \dots, Q_{S_{3k}}\}$ is not popular if and only if there is an exact cover of B.[2] The proof is analogous to the proofs of Theorems 3 and 9. ❑

Wiechers and Rothe [19] showed that SP-VERI is coNP-complete for EQ MBAHGs. We substantially modify their proof to establish the same result for P-VERI.

▶ **Theorem 15.** P-VERI *is* coNP-*complete for EQ MBAHGs.*

Proof. The proof of this theorem is inspired by proofs of Wiechers and Rothe [19, Theorem 4] and Kerkmann and Rothe [14, Theorem 7]. Given an instance of $(B, \mathscr{S})$ of RX3C, with $B = \{1, \dots, 3k\}$ and $\mathscr{S} = \{S_1, \dots, S_{3k}\}$, we construct the network of friends shown in Figure 3 with the set of players $N = \{\alpha_1, \alpha_2, \alpha_3\} \cup \{\beta_b \mid b \in B\} \cup \bigcup_{S \in \mathscr{S}} Q_S$, where $Q_S = \{\zeta_{S,\ell}, \eta_{S,j} \mid \ell \in [3k], j \in [3]\}$ for $S \in \mathscr{S}$, and we define the coalition structure $\Gamma = \{\{\alpha_2, \alpha_3\}, \{\alpha_1, \beta_1, \dots, \beta_{3k}\}, Q_{S_1}, \dots, Q_{S_{3k}}\}$. The friendship relationships are as follows:

- α_1, α_2, and α_3 are friends of each other.
- All players in $\{\alpha_1\} \cup \{\beta_b \mid b \in B\}$ are friends.
- For $S \in \mathscr{S}$ and $\ell \in [3k]$, $\zeta_{S,\ell}$ is friends with the three β_b with $b \in S$.
- For $S \in \mathscr{S}$, all players in Q_S are friends of each other.

We show that Γ is not popular if and only if there exists an exact cover of B in $\mathscr{S}$.

If: Assume that there exists an exact cover $\mathscr{S}' \subseteq \mathscr{S}$ of B. Then, for the coalition structure $\Delta = \{\{\alpha_1, \alpha_2, \alpha_3\}\} \cup \{\{\beta_b \mid b \in S\} \cup \{\zeta_{S,1}, \dots, \zeta_{S,3k}\} \mid S \in \mathscr{S}'\} \cup \{\{\eta_{S,1}, \eta_{S,2}, \eta_{S,3}\} \mid$

[2] Specifically, $\Delta' = \{\{\varphi_1\}, \dots, \{\varphi_{12k^3}\}, (A \setminus \{\alpha_1\}) \cup \bigcup_{S \in \mathscr{S}'} Q_S\} \cup \{Q_S \mid S \in \mathscr{S} \setminus \mathscr{S}'\}$ is more popular than Γ' (by one player) if there is an exact cover $\mathscr{S}'$ for B.

$S \in \mathscr{S}'\} \cup \{Q_S \mid S \in \mathscr{S} \setminus \mathscr{S}'\}$, it holds that α_2 and α_3 prefer Δ to Γ as they are in a clique of size three in Δ but in a clique of size two in Γ; all β_b with $b \in B$ prefer Δ to Γ as they are in a clique of size $3k+3$ in Δ but in a clique of size $3k+1$ in Γ; α_1 prefers Γ to Δ as she is in a clique of size $3k+1$ in Γ but in a clique of size three in Δ; all $\eta_{S,j}$ with $S \in \mathscr{S}'$ and $j \in [3]$ prefer Γ to Δ as they are in a clique of size $3k+3$ in Γ but in a clique of size three in Δ; and all remaining players are indifferent between Γ and Δ as they are in cliques of the same size in both coalition structures. So, we have $\#_{\Delta \succ \Gamma} = |\{\alpha_2, \alpha_3, \beta_1, \dots, \beta_{3k}\}| = 3k+2 > 3k+1 = 3|\mathscr{S}'|+1 = |\{\alpha_1\} \cup \{\eta_{S,1}, \eta_{S,2}, \eta_{S,3} \mid S \in \mathscr{S}'\}| = \#_{\Gamma \succ \Delta}$. Hence, Γ is not popular.

Only if: Assume that there is a coalition structure $\Delta \neq \Gamma$ with $\#_{\Delta \succ \Gamma} > \#_{\Gamma \succ \Delta}$. Then the following four claims hold; their easy proofs are given in Section A.5 of the appendix.

▷ Claim 16. For $S \in \mathscr{S}$ and $j \in [3]$, $\eta_{S,j}$ prefers Q_S to every other coalition.

▷ Claim 17. For $S \in \mathscr{S}$ and $\ell \in [3k]$, $\zeta_{S,\ell}$ has exactly two most preferred coalitions: Q_S and $\{\beta_b \mid b \in S\} \cup \{\zeta_{S,\ell} \mid \ell \in [3k]\}$.

▷ Claim 18. $\{\alpha_1, \beta_1, \dots, \beta_{3k}\}$ is α_1's unique most preferred coalition.

Note that Claims 16, 17, and 18 imply that there is no coalition structure that any of $\eta_{S,j}$, $\zeta_{S,\ell}$, or α_1 prefers to Γ. The β-players, however, prefer some coalition structures to Γ.

▷ Claim 19. For any $b \in B$, if β_b prefers Δ to Γ, then $\Delta(\beta_b) = \{\beta_a \mid a \in S\} \cup \{\zeta_{S,\ell} \mid \ell \in [3k]\}$ for some $S \in \mathscr{S}$ with $b \in S$.

Now, since $\#_{\Delta \succ \Gamma} > \#_{\Gamma \succ \Delta}$, there is a player $i \in N$ who prefers Δ to Γ. We distinguish the following two cases.

Case 1: $i = \beta_c$ for some $c \in B$. Let $k' = |\{S \in \mathscr{S} \mid Q_S \notin \Delta\}|$. Then, by Claim 16, there are $3k'$ η-players who prefer Γ to Δ and the remaining η-players are indifferent. Since β_c prefers Δ to Γ, we know by Claim 19 that $\Delta(\beta_c) = \{\beta_a \mid a \in S\} \cup \{\zeta_{S,\ell} \mid \ell \in [3k]\}$ for some $S \in \mathscr{S}$ with $c \in S$. Thus, by Claim 18, α_1 prefers Γ to Δ.

We will now see that $k' = k$.

First, assume that $k' > k$. Then $\#_{\Gamma \succ \Delta} \geq |\{\alpha_1\} \cup \{\eta_{S,j} \mid Q_S \notin \Delta, j \in [3]\}| = 3k'+1 > 3k+1$ and $\#_{\Delta \succ \Gamma} \leq |\{\alpha_2, \alpha_3, \beta_1, \dots, \beta_{3k}\}| = 3k+2$. Hence, $\#_{\Gamma \succ \Delta} \geq 3k+2 \geq \#_{\Delta \succ \Gamma}$, which contradicts $\#_{\Delta \succ \Gamma} > \#_{\Gamma \succ \Delta}$.

Second, assume that $k' < k$. Per one $Q_S \notin \Delta$, there are at most three β_b with $b \in S$ who prefer Δ to Γ (see Claim 19). Hence, $\#_{\Delta \succ \Gamma} \leq |\{\alpha_2, \alpha_3\} \cup \{\beta_b \mid b \in S, Q_S \notin \Delta\}| = 3k'+2 < 3k+2$. All remaining $3k - 3k'$ β_b do not have any $\zeta_{S,\ell}$ with $b \in S$ in $\Delta(\beta_b)$ and thus prefer Γ to Δ. We get $\#_{\Gamma \succ \Delta} \geq |\{\alpha_1\} \cup \{\eta_{S,j} \mid Q_S \notin \Delta, j \in [3]\} \cup \{\beta_b \mid Q_S \in \Delta \text{ for all } S \in \mathscr{S} \text{ with } b \in S\}| \geq 1 + 3k' + 3k - 3k' = 3k+1$. Hence, $\#_{\Gamma \succ \Delta} \geq 3k+1 \geq \#_{\Delta \succ \Gamma}$, which again is a contradiction.

It follows that $k' = k$ and thus $\#_{\Gamma \succ \Delta} \geq |\{\alpha_1\} \cup \{\eta_{S,j} \mid Q_S \notin \Delta, j \in [3]\}| = 3k+1$. Hence, since $\#_{\Delta \succ \Gamma} > \#_{\Gamma \succ \Delta}$, there are at least $3k+2$ players preferring Δ to Γ, which can only be α_2, α_3, and all β_b, $b \in B$. Then, by Claim 19, every β_b is in a coalition $\Delta(\beta_b) = \{\beta_a \mid a \in S\} \cup \{\zeta_{S,\ell} \mid \ell \in [3k]\}$ for some $S \in \mathscr{S}$ with $b \in S$. This implies that $\{S \in \mathscr{S} \mid Q_S \notin \Delta\}$ is an exact cover of B.

Case 2: $i = \alpha_2$ or $i = \alpha_3$. Since α_2 or α_3 prefer Δ to Γ, it follows that $\{\alpha_1, \alpha_2, \alpha_3\} \subseteq \Delta(\alpha_2)$. Then, considering only the α-players, we have $\#_{\Delta \succ \Gamma} \geq 2$ and $\#_{\Gamma \succ \Delta} \geq 1$. If at least one β_b prefers Δ to Γ, we are in Case 1 and an exact cover of B is already implied. Hence, assume that there is no β_b that prefers Δ to Γ. Then $\#_{\Delta \succ \Gamma} = 2$ because, by Claims 16 and 17, no other player can prefer Δ to Γ. With $\#_{\Delta \succ \Gamma} > \#_{\Gamma \succ \Delta}$, it follows that no player β_b, $\zeta_{S,\ell}$, nor

$\eta_{S,j}$ prefers Γ to Δ. Hence, by Claim 16, $Q_S \in \Delta$ for every $S \in \mathscr{S}$. However, this implies that all β_b prefer Γ to Δ, which is a contradiction. ❏

Wiechers and Rothe [19, Theorem 4] showed that SP-Veri is coNP-complete for AL MBAHGs. We extensively modify their proof to establish the same result for P-Veri but defer the proof of Theorem 20 to Section A.6 of the appendix.

▶ **Theorem 20.** P-Veri *is* coNP-*complete for AL MBAHGs.*

Finally, we turn to P-Exi. Note that we cannot simply modify the preceding theorems in order to show the hardness of P-Exi (similarly to how we used Theorems 3 and 9 to obtain Corollary 11) because a tie between two most popular coalition structures would not suffice to show the nonexistence of a popular coalition structure. However, for both AHGs and MBAHGs and all three degrees of altruism, there exist examples where no popular coalition structures exist (see Section A.7 of the appendix) and we suspect that P-Exi is hard for all considered models.

5 Conclusions and Future Research

We have solved the two remaining open problems regarding the complexity of strict popularity verification in AHGs, namely for equal treatment (Theorem 3) and altruistic treatment (Theorem 9). The proofs of these results required new ideas and are technically demanding. The corresponding results for MBAHGs have already been established by Wiechers and Rothe [19, Thm. 4]. In addition, we have provided the first complexity results for popularity verification in AHGs and MBAHGs, covering for both all three degrees of altruism (Theorems 12, 14, 15, and 20). Hence, the complexity of popularity verification and strict popularity verification is now settled in AHGs and MBAHGs; the picture is complete.

Moreover, we have seen that our hardness result for popularity verification (Theorem 12) extends to friend-oriented hedonic games. Additionally, we get some implications for classes of hedonic games that generalize AHGs. For instance, since the "super AHGs" by Schlueter and Goldsmith [18] generalize SF AHGs, all our hardness results for SF AHGs extend to this class as well. Also, all our results for EQ MBAHGs carry over to the "loyal variant of symmetric friend-oriented hedonic games" by Bullinger and Kober [7].

An interesting field for future research could be altruistic hedonic games in which agents may dynamically change their degree of altruism. In such a model, the agents' degree of altruism might depend on the well-being of others. For instance, they might act more altruistically when others are doing worse than themselves, while they are more selfish when others are doing better than themselves. Also, their degree of altruism might depend on the global level of welfare. While a global well-being might not evoke a strong degree of altruism, a severe suffering of their friends might do so.

References

1 H. Aziz, F. Brandt, and H. Seedig. Computing desirable partitions in additively separable hedonic games. *Artificial Intelligence*, 195:316–334, 2013.

2 H. Aziz and R. Savani. Hedonic games. In F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. Procaccia, editors, *Handbook of Computational Social Choice*, chapter 15, pages 356–376. Cambridge University Press, 2016.

3 S. Banerjee, H. Konishi, and T. Sönmez. Core in a simple coalition formation game. *Social Choice and Welfare*, 18(1):135–153, 2001.

4 A. Bogomolnaia and M. Jackson. The stability of hedonic coalition structures. *Games and Economic Behavior*, 38(2):201–230, 2002.

5 F. Brandt and M. Bullinger. Finding and recognizing popular coalition structures. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems*, pages 195–203. IFAAMAS, 2020.

6 S. Brânzei and K. Larson. Social distance games. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence*, pages 91–96. AAAI Press/IJCAI, July 2011.

7 M. Bullinger and S. Kober. Loyalty in cardinal hedonic games. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence*, pages 66–72. ijcai.org, July 2021.

8 D. Dimitrov, P. Borm, R. Hendrickx, and S. Sung. Simple priorities and core stability in hedonic games. *Social Choice and Welfare*, 26(2):421–433, 2006.

9 J. Drèze and J. Greenberg. Hedonic coalitions: Optimality and stability. *Econometrica*, 48(4):987–1003, 1980.

10 E. Elkind and J. Rothe. Cooperative game theory. In J. Rothe, editor, *Economics and Computation. An Introduction to Algorithmic Game Theory, Computational Social Choice, and Fair Division*, Springer Texts in Business and Economics, chapter 3, pages 135–193. Springer-Verlag, 2015.

11 M. Garey, D. Johnson, and L. Stockmeyer. Some simplified NP-complete graph problems. *Theoretical Computer Science*, 1:237–267, 1976.

12 T. Gonzalez. Clustering to minimize the maximum intercluster distance. *Theoretical Computer Science*, 38:293–306, 1985.

13 A. Kerkmann, J. Lang, A. Rey, J. Rothe, H. Schadrack, and L. Schend. Hedonic games with ordinal preferences and thresholds. *Journal of Artificial Intelligence Research*, 67:705–756, 2020.

14 A. Kerkmann and J. Rothe. Altruism in coalition formation games. In *Proceedings of the 29th International Joint Conference on Artificial Intelligence*, pages 347–353. ijcai.org, 7 2020.

15 A. Kerkmann and J. Rothe. Popularity and strict popularity in altruistic hedonic games and minimum-based altruistic hedonic games (extended abstract). In *Proceedings of the 21th International Conference on Autonomous Agents and Multiagent Systems*. IFAAMAS, May 2022. To appear.

16 N. Nguyen, A. Rey, L. Rey, J. Rothe, and L. Schend. Altruistic hedonic games. In *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems*, pages 251–259. IFAAMAS, May 2016.

17 J. Rothe. Thou shalt love thy neighbor as thyself when thou playest: Altruism in game theory. In *Proceedings of the 35th AAAI Conference on Artificial Intelligence*, pages 15070–15077. AAAI Press, February 2021.

18 J. Schlueter and J. Goldsmith. Super altruistic hedonic games. In *Proceedings of the 33rd International Florida Artificial Intelligence Research Society Conference*, pages 160–165. AAAI Press, May 2020.

19 A. Wiechers and J. Rothe. Stability in minimization-based altruistic hedonic games. In *Proceedings of the 9th European Starting AI Researchers' Symposium*, volume 2655. CEUR-WS.org, August 2020.

20 G. Woeginger. Core stability in hedonic coalition formation. In *Proceedings of the 39th International Conference on Current Trends in Theory and Practice of Computer Science*, pages 33–50. Springer-Verlag *Lecture Notes in Computer Science #7741*, January 2013.

A Appendix

In this appendix, we give all proof details that were omitted from the paper due to space constraints.

A.1 Proofs of Claims 7 and 8 in the Proof of Theorem 3

▷ Claim 7. For every $S \in \mathscr{S}$, ζ_S prefers Q_S to every other coalition.

Proof of Claim 7. It holds that

$$
\begin{aligned}
u^{EQ}_{\zeta_S}(Q_S) &= \frac{v_{\zeta_S}(Q_S) + (3k-2)v_{\eta_{S,1}}(Q_S) + v_{\delta_S}(Q_S)}{3k} \\
&= \frac{(3k-1)(n(3k-1)-(3k+1)) + n(6k)}{3k} \\
&= \frac{n(9k^2+1)-(9k^2-1)}{3k} \\
&= n\left(3k + \frac{1}{3k}\right) - \left(3k - \frac{1}{3k}\right).
\end{aligned}
$$

Hence, ζ_S and her friends have more than $3k$ friends in Q_S on average. Now, assume that there is a coalition $D \neq Q_S$ that ζ_S weakly prefers to every other coalition. It is clear that $D \subseteq Q_S \cup \{\beta_b \mid b \in S\} \cup \{\alpha_2\} \cup \{\zeta_{S'} | S' \in \mathscr{S}, S \cap S' \neq \emptyset\}$ because all other players are enemies of ζ_S's and of all her friends.

Assume that there is some β_b in D. We will show that ζ_S prefers $D \setminus \{\beta_b\}$ to D, which is a contradiction. It holds that ζ_S prefers $D \setminus \{\beta_b\}$ to D if and only if $u^{EQ}_{\zeta_S}(D \setminus \{\beta_b\}) > u^{EQ}_{\zeta_S}(D)$. Let $x = |D \cap F_{\zeta_S}|$, $t = \sum_{a \in ((D \setminus \{\beta_b\}) \cap F_{\zeta_S}) \cup \{\zeta_S\}} v_a(D \setminus \{\beta_b\})$, $v = \sum_{a \in (D \cap F_{\zeta_S}) \cup \{\zeta_S\}} v_a(D)$, and $w = v - t$. Then

$$
\begin{aligned}
u^{EQ}_{\zeta_S}(D \setminus \{\beta_b\}) > u^{EQ}_{\zeta_S}(D) \quad &\Leftrightarrow \quad \frac{v-w}{x} - \frac{v}{x+1} > 0 \\
&\Leftrightarrow \quad \frac{(x+1)(v-w) - xv}{x(x+1)} > 0 \\
&\Leftrightarrow \quad (x+1)(v-w) - xv > 0 \\
&\Leftrightarrow \quad xv + v - (x+1)w - xv > 0 \\
&\Leftrightarrow \quad v - (x+1)w > 0 \\
&\Leftrightarrow \quad \frac{v}{x+1} > w \\
&\Leftrightarrow \quad u^{EQ}_{\zeta_S}(D) > w.
\end{aligned}
$$

For w, we have

$$
\begin{aligned}
w &= \sum_{a \in (D \cap F_{\zeta_S}) \cup \{\zeta_S\}} \Big(v_a(D)\Big) - t \\
&= \sum_{a \in (D \setminus \{\beta_b\}) \cap F_{\zeta_S}} \Big(v_a(D)\Big) + v_{\beta_b}(D) + v_{\zeta_S}(D) - t \\
&= \sum_{a \in (D \setminus \{\beta_b\}) \cap F_{\zeta_S}} \Big(v_a(D \setminus \{\beta_b\}) - 1\Big) + v_{\beta_b}(D) + v_{\zeta_S}(D \setminus \{\beta_b\}) + n - t \\
&= \sum_{a \in ((D \setminus \{\beta_b\}) \cap F_{\zeta_S}) \cup \{\zeta_S\}} \Big(v_a(D \setminus \{\beta_b\})\Big) - |(D \setminus \{\beta_b\}) \cap F_{\zeta_S}| + \ v_{\beta_b}(D) + n - t \\
&= -|(D \setminus \{\beta_b\}) \cap F_{\zeta_S}| + v_{\beta_b}(D) + n \\
&\leq -|(D \setminus \{\beta_b\}) \cap F_{\zeta_S}| + 4n + n \\
&< 5n \ < \ u^{EQ}_{\zeta_S}(Q_S) \ \leq \ u^{EQ}_{\zeta_S}(D).
\end{aligned}
$$

Hence, ζ_S prefers $D \setminus \{\beta_b\}$ to D; a contradiction that implies that there is no β_b in D. It then follows that $D \subseteq Q_S$.

Since ζ_S has at most $3k-1$ friends in D and $u^{EQ}_{\zeta_S}(D) \geq u^{EQ}_{\zeta_S}(Q_S) = n(3k+\frac{1}{3k})-(3k-\frac{1}{3k})$, there has to be at least one friend of ζ_S in D who has at least $3k+1$ friends in D. The only player for which this is possible is δ_S, so $\delta_S \in D$. Then, by Observation 1, it holds that $\gamma_{S,1}, \ldots, \gamma_{S,3k+1} \in D$. Thus $\{\zeta_S, \delta_S, \gamma_{S,1}, \ldots, \gamma_{S,3k+1}\} \subseteq D \subseteq Q_S$.

Now, let $y = |D \cap \{\eta_{S,1}, \ldots, \eta_{S,3k-2}\}|$ be the number of η-players in D. Then

$$
\begin{aligned}
u^{EQ}_{\zeta_S}(D) &= \frac{v_{\zeta_S}(D) + y \cdot v_{\eta_{S,j}}(D) + v_{\delta_S}(D)}{y+2} \\
&= \frac{(y+1) \cdot (n(y+1) - (3k+1)) + n(y+1+3k+1)}{y+2} \\
&= n\left(\frac{y^2+3y+3k+3}{y+2}\right) - \frac{(y+1)(3k+1)}{y+2}.
\end{aligned}
$$

We know that $u^{EQ}_{\zeta_S}(D) \geq u^{EQ}_{\zeta_S}(Q_S)$ holds, for which we get the following equivalences:

$$
\begin{aligned}
& u^{EQ}_{\zeta_S}(D) \geq u^{EQ}_{\zeta_S}(Q_S) \\
\Leftrightarrow\ & n\left(\frac{y^2+3y+3k+3}{y+2}\right) - \frac{(y+1)(3k+1)}{y+2} \geq n\left(\frac{9k^2+1}{3k}\right) - \frac{9k^2-1}{3k} \\
\Leftrightarrow\ & 0 \geq n\left(\frac{9k^2+1}{3k} - \frac{y^2+3y+3k+3}{y+2}\right) - \left(\frac{(3k-1)(3k+1)}{3k} - \frac{(y+1)(3k+1)}{y+2}\right) \\
\Leftrightarrow\ & 0 \geq n\left(\frac{(9k^2+1)(y+2) - (y^2+3y+3k+3)(3k)}{(3k)(y+2)}\right) \\
& \quad - \left(\frac{(3k-1)(3k+1)(y+2) - (y+1)(3k+1)(3k)}{(3k)(y+2)}\right) \\
\Leftrightarrow\ & 0 \geq n\left(\frac{9k^2y + y + 18k^2 + 2 - 3ky^2 - 9ky - 9k^2 - 9k}{(3k)(y+2)}\right) \\
& \quad - \left(\frac{(3k+1)\Big((3k-1)(y+2) - (y+1)(3k)\Big)}{(3k)(y+2)}\right) \\
\Leftrightarrow\ & 0 \geq n\left(\frac{(3k-y-2)(3ky+3k-1)}{(3k)(y+2)}\right) - \left(\frac{(3k+1)(3k-y-2)}{(3k)(y+2)}\right). \qquad (3)
\end{aligned}
$$

This implies

$$0 \geq \frac{(3k-y-2)(3ky+3k-1)}{(3k)(y+2)}.$$

For a contradiction, assume that $0 < \frac{(3k-y-2)(3ky+3k-1)}{(3k)(y+2)}$. Since $3ky+3k-1>0$, $3k>0$, and $y+2>0$, it then follows that $3k-y-2>0$, i.e., $y<3k-2$. Then, for $0 \leq y < 3k-2$ and $k \geq 2$, the minimum of $\frac{(3k-y-2)(3ky+3k-1)}{(3k)(y+2)}$ is reached for $y=3k-3$, namely $\frac{9k^2-6k-1}{3k(3k-1)}$. But even for the minimum we have $n\left(\frac{9k^2-6k-1}{3k(3k-1)}\right) - \frac{(3k+1)(3k-y-2)}{(3k)(y+2)} > 0$, which is a contradiction to Equation (3).

Since $3ky+3k-1>0$, $3k>0$, and $y+2>0$, it follows that $3k-y-2 \leq 0$. Thus $y \geq 3k-2$. Hence, all η-players are in D, so $D = Q_S$. This is a contradiction and completes the proof. ❑ Claim 7

▷ Claim 8. If β_b with $b \in B$ prefers Δ to Γ, then $\zeta_S \in \Delta(\beta_b)$ for some $S \in \mathscr{S}$ with $b \in S$.

Proof of Claim 8. Assume that there is no ζ_S with $b \in S$ in $\Delta(\beta_b)$. Then α_2 is β_b's only remaining friend that could be in $\Delta(\beta_b)$. By Observation 1, β_b gets the most utility from Δ if $\Delta(\beta_b) = A$. This means that β_b does not prefer Δ to Γ. ❑ Claim 8

A.2 Detailed Calculation of Utilities in the Proof of Theorem 3

For $\eta_{S,j}$ with $S \in \mathscr{S}'$ and $j \in [3k-2]$, we have

$$\begin{aligned}
u^{EQ}_{\eta_{S,j}}(\Delta) &= \frac{1}{3k}\Big(v_{\zeta_S}(\Delta) + (3k-2)v_{\eta_{S,j}}(\Delta) + v_{\delta_S}(\Delta)\Big) \\
&= \frac{1}{3k}\Big(v_{\zeta_S}(\Gamma) + 3n - (3k-1) - (k-1)(6k+1) \\
&\quad + (3k-2)\big(v_{\eta_{S,j}}(\Gamma) - (3k+2) - (k-1)(6k+1)\big) \\
&\quad + v_{\delta_S}(\Gamma) - (3k+2) - (k-1)(6k+1)\Big) \\
&= \frac{v_{\zeta_S}(\Gamma) + (3k-2)v_{\eta_{S,j}}(\Gamma) + v_{\delta_S}(\Gamma)}{3k} \\
&\quad + \frac{1}{3k}\Big(3n + 3 - (3k)\big(3k+2+(k-1)(6k+1)\big)\Big) \\
&= u^{EQ}_{\eta_{S,j}}(\Gamma) + 6k^2 + 20k + 5 + {}^{3}\!/\!_{k} \\
&> u^{EQ}_{\eta_{S,j}}(\Gamma).
\end{aligned}$$

For δ_S with $S \in \mathscr{S}'$, we have

$$\begin{aligned}
u^{EQ}_{\delta_S}(\Delta) &= \frac{v_{\zeta_S}(\Delta) + (3k-2)v_{\eta_{S,j}}(\Delta) + v_{\delta_S}(\Delta) + (3k+1)v_{\gamma_{S,\ell}}(\Delta)}{6k+1} \\
&= \frac{v_{\zeta_S}(\Gamma) + (3k-2)v_{\eta_{S,j}}(\Gamma) + v_{\delta_S}(\Gamma) + (3k+1)v_{\gamma_{S,\ell}}(\Gamma)}{6k+1} \\
&\quad + \frac{3n + 3 - (6k+1)\big(3k+2+(k-1)(6k+1)\big)}{6k+1} \\
&= u^{EQ}_{\delta_S}(\Gamma) + \frac{60k^2+14k+8}{6k+1} \\
&> u^{EQ}_{\delta_S}(\Gamma).
\end{aligned}$$

A.3 Proof of Claim 10 in the Proof of Theorem 9

▷ Claim 10. For $S \in \mathscr{S}$, ζ_S prefers $\{\zeta_S, \delta_S, \gamma_{S,1}, \ldots, \gamma_{S,3k+1}\}$ and every coalition $\{\zeta_S, \delta_S, \gamma_{S,1}, \ldots, \gamma_{S,3k+1}\} \setminus \{\gamma_{S,\ell}\}$, $\ell \in [3k+1]$, to Q_S, and ζ_S prefers Q_S to every other coalition.

Proof of Claim 10. For $C = \{\zeta_S, \delta_S, \gamma_{S,1}, \ldots, \gamma_{S,3k+1}\}$, it holds that $\mathrm{avg}^F_{\zeta_S}(C) = v_{\delta_S}(C) = n(3k+2)$; for $C_\ell = C \setminus \{\gamma_{S,\ell}\}$ with $\ell \in [3k+1]$, it holds that $\mathrm{avg}^F_{\zeta_S}(C_\ell) = v_{\delta_S}(C_\ell) = n(3k+1)$; and for Q_S, we have

$$\begin{aligned}
\mathrm{avg}^F_{\zeta_S}(Q_S) &= \frac{(3k-2)v_{\eta_{S,j}}(Q_S) + v_{\delta_S}(Q_S)}{3k-1} \\
&= \frac{(3k-2)(n(3k-1)-(3k+1)) + n(6k)}{3k-1} \\
&= \frac{n(9k^2-3k+2) - (9k^2-3k-2)}{3k-1} \\
&= n\left(3k + \frac{2}{3k-1}\right) - \left(3k - \frac{2}{3k-1}\right).
\end{aligned}$$

Thus ζ_S prefers C to every C_ℓ, and every C_ℓ to Q_S.

Now, let D with $D \neq C$ and $D \neq C_\ell$ for $\ell \in [3k+1]$ be a coalition that ζ_S weakly prefers to every coalition except for C and C_ℓ. We will show that $D = Q_S$. Similarly as in the proof Claim 7, it follows that $\{\zeta_S, \delta_S\} \subseteq D \subseteq Q_S$. (We omit the details because this proof is very similar.) Now, let $x = |D \cap \{\gamma_{S,1}, \ldots, \gamma_{S,3k+1}\}|$ be the number of γ-players in D and let $y = |D \cap \{\eta_{S,1}, \ldots, \eta_{S,3k-2}\}|$ be the number of η-players in D.

First, assume $y = 0$. Then $\mathrm{avg}^F_{\zeta_S}(D) = v_{\delta_S}(D) = n(x+1)$. Since ζ_S weakly prefers D to Q_S, we know that $\mathrm{avg}^F_{\zeta_S}(D) \geq \mathrm{avg}^F_{\zeta_S}(Q_S)$, i.e., $n(x+1) \geq n\left(3k + \frac{2}{3k-1}\right) - \left(3k - \frac{2}{3k-1}\right)$. This implies $x \geq 3k$. This is a contradiction because it implies that $D = C$ or $D = C_\ell$ for some $\ell \in [3k+1]$. Thus we have $y \geq 1$.

By Observation 1, $\{\gamma_{S,1}, \ldots, \gamma_{S,3k+1}\} \subseteq D$; otherwise, ζ_S would prefer $D' = D \cup \{\gamma_{S,1}, \ldots, \gamma_{S,3k+1}\}$ to D. This would be a contradiction to ζ_S weakly preferring D to every coalition except for C and C_ℓ. (Note that $D' \neq C$ and $D' \neq C_\ell$ because of $y \geq 1$.) It then holds that

$$\begin{aligned}
\mathrm{avg}^F_{\zeta_S}(D) &= \frac{y v_{\eta_{S,j}}(D) + v_{\delta_S}(D)}{y+1} \\
&= \frac{y\big(n(y+1) - (3k+1)\big) + n(y+1+3k+1)}{y+1} \\
&= n\left(\frac{y^2+2y+3k+2}{y+1}\right) - \frac{y(3k+1)}{y+1}.
\end{aligned}$$

Now, rearranging $\mathrm{avg}^F_{\zeta_S}(D) \geq \mathrm{avg}^F_{\zeta_S}(Q_S)$, the difference

$$n\left(\frac{(3k-y-2)(3ky-y-2)}{(y+1)(3k-1)}\right) - \frac{(3k-y-2)(3k+1)}{(y+1)(3k-1)}$$

cannot be positive. It then follows that

$$0 \geq \frac{(3k-y-2)(3ky-y-2)}{(y+1)(3k-1)}.$$

Since $k \geq 2$ and $y \geq 1$, this implies that $0 \geq 3k-y-2$, i.e., $y \geq 3k-2$. Hence, we have $D = Q_S$. ❑ Claim 10

A.4 Additional Details for the Proof of Theorem 9

In the *If*-part of the proof of Theorem 9, we state that there is no $S \in \mathscr{S}$ with $C = \{\zeta_S, \delta_S, \gamma_{S,1}, \dots, \gamma_{S,3k+1}\}$ in Δ or $C_\ell = \{\zeta_S, \delta_S, \gamma_{S,1}, \dots, \gamma_{S,3k+1}\} \setminus \{\gamma_{S,\ell}\}$ in Δ for any $\ell \in [3k+1]$. If there were such an $S \in \mathscr{S}$, then any $\eta_{S,j}$ and δ_S with $j \in [3k-2]$ would prefer Γ to Δ. The corresponding utilities, which are omitted above, are as follows.

For δ_S, we have

$$\begin{aligned}
\mathrm{avg}^F_{\delta_S}(\Gamma) &= \frac{v_{\zeta_S}(\Gamma) + (3k-2)v_{\eta_{S,j}}(\Gamma) + (3k+1)v_{\gamma_{S,\ell}}(\Gamma)}{6k} \\
&= \frac{(3k-1)(n(3k-1)-(3k+1)) + (3k+1)(n-(6k-1))}{6k} \\
&= \frac{n(9k^2-3k+2)-(27k^2+3k-2)}{6k} \\
&= n(\frac{3}{2}\cdot k - \frac{1}{2} + \frac{1}{3k}) - (\frac{27}{6}\cdot k + \frac{1}{2} - \frac{1}{3k}); \\
\mathrm{avg}^F_{\delta_S}(C) &= \frac{v_{\zeta_S}(C) + (3k+1)v_{\gamma_{S,\ell}}(C)}{3k+2} = \frac{n-(3k+1)+(3k+1)(n-(3k+1))}{3k+2} \\
&= n-(3k+1) < \mathrm{avg}^F_{\delta_S}(\Gamma); \text{ and} \\
\mathrm{avg}^F_{\delta_S}(C_\ell) &= \frac{v_{\zeta_S}(C_\ell) + (3k)v_{\gamma_{S,\ell}}(C_\ell)}{3k+1} = \frac{n-3k+(3k)(n-3k)}{3k+1} = n-3k < \mathrm{avg}^F_{\delta_S}(\Gamma).
\end{aligned}$$

For any $\eta_{S,j}$ with $j \in [3k-2]$, we have

$$\mathrm{avg}^F_{\eta_{S,j}}(\Gamma) = \mathrm{avg}^F_{\zeta_S}(\Gamma) = n\left(3k + \frac{2}{3k-1}\right) - \left(3k - \frac{2}{3k-1}\right).$$

If C or C_ℓ is in Δ, then the best coalition that could form for $\eta_{S,j}$ is $\{\eta_{S,1}, \dots, \eta_{S,3k-2}\}$. Hence,

$$\mathrm{avg}^F_{\eta_{S,j}}(\Delta) \le \frac{(3k-3)(n(3k-3))}{3k-3} = n(3k-3) < \mathrm{avg}^F_{\eta_{S,j}}(\Gamma).$$

It follows that all δ_S and $\eta_{S,j}$ with $j \in [3k-2]$ prefer Γ to Δ.

A.5 Proofs of the Claims for Theorem 15

We now give the proofs of Claims 16, 17, 18, and 19 from the proof of Theorem 15.

▷ Claim 16. For $S \in \mathscr{S}$ and $j \in [3]$, $\eta_{S,j}$ prefers Q_S to every other coalition.

Proof of Claim 16. Since Q_S is a clique of size $3k+3$, it holds that $u^{minEQ}_{\eta_{S,j}}(Q_S) = n(3k+2)$. As every $\eta_{S,j}$ has only $3k+2$ friends in total, Q_S is the only clique of size $3k+3$ that can reach this utility for $\eta_{S,j}$. Every other coalition $C \in \mathcal{N}_{\eta_{S,j}}$ either contains fewer friends or more enemies of $\eta_{S,j}$'s than Q_S, which leads to a decrease in utility for $\eta_{S,j}$. ❑ Claim 16

▷ Claim 17. For $S \in \mathscr{S}$ and $\ell \in [3k]$, $\zeta_{S,\ell}$ has exactly two most preferred coalitions: Q_S and $\{\beta_b \mid b \in S\} \cup \{\zeta_{S,\ell} \mid \ell \in [3k]\}$.

Proof of Claim 17. Since the coalitions $A = Q_S$ and $B = \{\beta_b \mid b \in S\} \cup \{\zeta_{S,\ell} \mid \ell \in [3k]\}$ are cliques of size $3k+3$, it holds that $u^{minEQ}_{\zeta_{S,\ell}}(A) = u^{minEQ}_{\zeta_{S,\ell}}(B) = n(3k+2)$. For a contradiction, assume that there is another coalition C with $u^{minEQ}_{\zeta_{S,\ell}}(C) \ge n(3k+2)$.

In case of $u^{minEQ}_{\zeta_{S,\ell}}(C) = n(3k+2)$, C would be a clique of size $3k+3$. However, there are no other cliques of size $3k+3$ containing $\zeta_{S,\ell}$ besides A and B.

In case of $u^{minEQ}_{\zeta_{S,\ell}}(C) > n(3k+2)$, $\zeta_{S,\ell}$ and all her friends each need to have at least $3k+3$ friends in C. Each η-player has only $3k+2$ friends in total and thus cannot be part of C. However, without the η-players, $\zeta_{S,\ell}$ has only $3k+2$ friends in total. Thus we have a contradiction. ❑ Claim 17

▷ Claim 18. $\{\alpha_1, \beta_1, \ldots, \beta_{3k}\}$ is α_1's unique most preferred coalition.

Proof of Claim 18. For coalition $A = \{\alpha_1, \beta_1, \ldots, \beta_{3k}\}$, it holds that $u^{minEQ}_{\alpha_1}(A) = n3k$. If there were another coalition $B \neq A$ with $u^{minEQ}_{\alpha_1}(B) \geq u^{minEQ}_{\alpha_1}(A) = n3k$, α_1 would have at least $3k$ friends in B and all these friends would also have at least $3k$ friends in B. Since α_2 and α_3 have only two friends in total, it holds that $\alpha_2 \notin B$ and $\alpha_3 \notin B$. However, α_1's remaining $3k$ friends are β-players, which implies $A = \{\alpha_1, \beta_1, \ldots, \beta_{3k}\} \subseteq B$. Since any additional ζ- or η-player in B would contradict $u^{minEQ}_{\alpha_1}(B) \geq n3k$, we get $A = B$, which also is a contradiction. ❑ Claim 18

▷ Claim 19. For any $b \in B$, if β_b prefers Δ to Γ, then $\Delta(\beta_b) = \{\beta_a \mid a \in S\} \cup \{\zeta_{S,\ell} \mid \ell \in [3k]\}$ for some $S \in \mathscr{S}$ with $b \in S$.

Proof of Claim 19. Assume that $\Delta \succ^{minEQ}_{\beta_b} \Gamma$. Since $\Gamma(\beta_b) = \{\alpha_1, \beta_1, \ldots, \beta_{3k}\}$ is a clique, it follows that β_b has a friend in $\Delta(\beta_b)$ that is not in $\Gamma(\beta_b)$. Hence, there is a $\zeta_{S,\ell}$ with $b \in S$ in $\Delta(\beta_b)$.

Now assume that $\Delta(\beta_b) \neq \{\beta_a \mid a \in S\} \cup \{\zeta_{S,\ell} \mid \ell \in [3k]\}$. Then $\zeta_{S,\ell} \in \Delta(\beta_b)$ together with Claim 17 implies that all $\zeta_{S,\ell'}$ with $\ell' \in [3k]$ prefer Γ to Δ. Further, Claim 16 implies that $\eta_{S,j}$, $j \in [3]$, prefer Γ to Δ. Hence, we have $\#_{\Gamma\succ\Delta} \geq 3k+3$. From Claims 16, 17, and 18 we know that no η, ζ, or α_1 can prefer Δ to Γ. Hence, $\#_{\Delta\succ\Gamma} \leq |\{\alpha_2, \alpha_3, \beta_1, \ldots, \beta_{3k}\}| = 3k+2$. We get $\#_{\Gamma\succ\Delta} > \#_{\Delta\succ\Gamma}$, which is a contradiction. ❑ Claim 19

A.6 Proof of Theorem 20

We now give the proof of Theorem 20.

▶ **Theorem 20.** P-Veri *is* coNP*-complete for AL MBAHGs.*

Proof. We use the same set of players and network of friends as in the proof of Theorem 15 that is shown in Figure 3. We again consider coalition structure $\Gamma = \{\{\alpha_1, \beta_1, \ldots, \beta_{3k}\}, \{\alpha_2, \alpha_3\}, Q_{S_1}, \ldots, Q_{S_{3k}}\}$ and show that Γ is not popular under AL if and only if there exists an exact cover of B in $\mathscr{S}$.

If: Assume that there exists an exact cover $\mathscr{S}' \subseteq \mathscr{S}$ of B. Then $\Delta = \{\{\alpha_1, \alpha_2, \alpha_3\}\} \cup \{\{\beta_b \mid b \in S\} \cup \{\zeta_{S,\ell} \mid \ell \in [3k]\} \mid S \in \mathscr{S}'\} \cup \{\{\eta_{S,1}, \eta_{S,2}, \eta_{S,3}\} \mid S \in \mathscr{S}'\} \cup \{Q_S \mid S \in \mathscr{S} \setminus \mathscr{S}'\}$ is more popular than Γ: All players in Q_S with $S \in \mathscr{S} \setminus \mathscr{S}'$ are obviously indifferent between Γ and Δ because their coalitions stay the same. The utilities of the other players are shown in Table 1. Hence, Γ is not popular because

$$\begin{aligned}\#_{\Delta\succ\Gamma} &= |\{\alpha_2, \alpha_3, \beta_1, \ldots, \beta_{3k}\}| = 3k+2 > 3k+1 = 3|\mathscr{S}'|+1 \\ &= |\{\alpha_1\} \cup \{\eta_{S,j} \mid S \in \mathscr{S}', j \in [3]\}| = \#_{\Gamma\succ\Delta}.\end{aligned}$$

Only if: Assume that there is a coalition structure $\Delta \neq \Gamma$ with $\#_{\Delta\succ\Gamma} > \#_{\Gamma\succ\Delta}$. Then we can iteratively show the following claims.

Table 1 Utilities of the players in $N \setminus \bigcup_{S \in \mathscr{S} \setminus \mathscr{S}'} Q_S$ for the proof of Theorem 20

player i	$u_i^{minAL}(\Gamma)$		$u_i^{minAL}(\Delta)$
α_1	$M \cdot n3k + n3k$	$>$	$M \cdot 2n + 2n$
α_2, α_3	$M \cdot n + n$	$<$	$M \cdot 2n + 2n$
β_b, $b \in B$	$M \cdot n3k + n3k$	$<$	$M \cdot n(3k+2) + n(3k+2)$
$\zeta_{S,\ell}$, $S \in \mathscr{S}'$, $\ell \in [3k]$	$M \cdot n(3k+2) + n(3k+2)$	$=$	$M \cdot n(3k+2) + n(3k+2)$
$\eta_{S,j}$, $S \in \mathscr{S}'$, $j \in [3]$	$M \cdot n(3k+2) + n(3k+2)$	$>$	$M \cdot 2n + 2n$

▷ Claim 21. For any $S \in \mathscr{S}$ and $\ell \in [3k]$, if $\zeta_{S,\ell}$ prefers Δ to Γ, then $\Delta(\zeta_{S,\ell})$ contains no $\eta_{S,j}$ with $j \in [3]$, no $\zeta_{S,\ell'}$ with $\ell' \in [3k]$ and $\ell' \neq \ell$, at least one β_b with $b \in S$, and $3k+2$ other friends of β_b's.

Proof of Claim 21. Assume that $\zeta_{S,\ell}$ prefers Δ to Γ and let $\Delta(\zeta_{S,\ell}) = D$. As $\Gamma(\zeta_{S,\ell}) = Q_S$ is a clique of size $3k+3$, we have $u_{\zeta_{S,\ell}}^{minAL}(\Delta) > u_{\zeta_{S,\ell}}^{minAL}(\Gamma) = n(3k+2) + Mn(3k+2)$. Thus D contains at least one friend of $\zeta_{S,\ell}$'s and every friend of $\zeta_{S,\ell}$'s in D has at least $3k+3$ friends in D. Since the players $\eta_{S,j}$, $j \in [3]$, each have only $3k+2$ friends in total, they cannot be part of D. By omitting these players, all $\zeta_{S,\ell'}$, $\ell' \neq \ell$, only have $3k+2$ friends left and cannot be part of D either. Hence, D contains at least one β_b with $b \in S$, and $3k+2$ other friends of β_b's. ❑ Claim 21

▷ Claim 22. For any $S \in \mathscr{S}$ and $\ell \in [3k]$, if $\zeta_{S,\ell}$ prefers Δ to Γ then at least $3k$ other players in Q_S prefer Γ to Δ.

Proof of Claim 22. Assume that $\zeta_{S,\ell}$ prefers Δ to Γ.

Since there are only three β_b with $b \in S$, we know by Claim 21 that at most two other $\zeta_{S,\ell'}$ with $\ell' \in [3k]$ and $\ell' \neq \ell$ can prefer Δ to Γ at the same time. All other players from Q_S obviously prefer Γ to Δ because they can only stay among themselves in Δ. Thus at least $3k+3-3 = 3k$ players in Q_S prefer Γ to Δ. ❑ Claim 22

▷ Claim 23. For any $S \in \mathscr{S}$ and $\ell \in [3k]$, there are exactly two coalitions $A \subseteq N$ with $v_{\zeta_{S,\ell}}^{minAL}(A) = Mn(3k+2) + n(3k+2)$, namely Q_S and $\{\beta_b \mid b \in S\} \cup \{\zeta_{S,\ell} \mid \ell \in [3k]\}$.

Proof of Claim 23. Since Q_S and $\{\beta_b \mid b \in S\} \cup \{\zeta_{S,\ell} \mid \ell \in [3k]\}$ are cliques of size $3k+3$, the statement is clearly true for them. Every other coalition C with the same valuation would also have to be a clique of size $3k+3$ containing $\zeta_{S,\ell}$. However, such a clique C does not exist in the given network of friends. ❑ Claim 23

▷ Claim 24. For every $\eta_{S,j}$ with $S \in \mathscr{S}$ and $j \in [3]$, there is no coalition that is in a tie with Q_S.

Proof of Claim 24. Let $C \subseteq N$ be a coalition containing $\eta_{S,j}$ and satisfying $u_{\eta_{S,j}}^{minAL}(C) = u_{\eta_{S,j}}^{minAL}(Q_S) = M(3k+2) + n(3k+2)$. Then it has to contain exactly $3k+2$ friends of $\eta_{S,j}$'s who are all friends with each other. Since $\eta_{S,j}$ has exactly $3k+2$ friends, this is clearly determined as Q_S. ❑ Claim 24

▷ Claim 25. For any $S \in \mathscr{S}$ and $j \in [3]$, if $\eta_{S,j}$ prefers Δ to Γ then all other players in Q_S prefer Γ to Δ.

Proof of Claim 25. Assume that $\eta_{S,j}$ prefers Δ to Γ and let $\Delta(\eta_{S,j}) = D$. Then every friend of $\eta_{S,j}$'s in D has at least $3k+3$ friends in D. Thus $\eta_{S,j'} \notin D$ for $j' \in [3]$, $j' \neq j$, since they only have $3k+2$ friends in total. Now there only remain the players $\zeta_{S,\ell}, \ell \in [3k]$ which (after omitting the $\eta_{S,j'}$) have exactly $3k+3$ friends left. Thus $\{\beta_b \mid b \in S\} \cup \{\zeta_{S,\ell} \mid \ell \in [3k]\} \subseteq D$. As $\eta_{S,j}$ has fewer friends in D than in Q_S, it follows that every $\zeta_{S,\ell}, \ell \in [3k]$ prefers Q_S to D. Since the $\eta_{S,j'}$ with, $j' \in [3], j' \neq j$, only have themselves left as friends, they clearly also prefer Γ to Δ. ❏ Claim 25

▷ Claim 26. If α_2 or α_3 prefer Δ to Γ then α_1 prefers Γ to Δ.

Proof of Claim 26. Assume that α_2 prefers Δ to Γ. Then α_2 has at least one friend in $\Delta(\alpha_2)$ and every friend of α_2 in $\Delta(\alpha_2)$ has at least two friends in $\Delta(\alpha_2)$. Hence, $\{\alpha_1, \alpha_2, \alpha_3\} \subseteq \Delta(\alpha_2)$ or $\{\alpha_1, \alpha_2, \beta_b\} \subseteq \Delta(\alpha_2)$ for some $b \in B$. In both cases, $u^{minAL}_{\alpha_1}(\Delta) \leq M \cdot v_{\alpha_2}(\Delta) + v_{\alpha_1}(\Delta) \leq M \cdot 2n + v_{\alpha_1}(\Delta) < M \cdot n3k + n3k = u^{minAL}_{\alpha_1}(\Gamma)$. Thus α_1 prefers Γ to Δ. Due to symmetry, the same arguments work if α_3 prefers Δ to Γ. ❏ Claim 26

▷ Claim 27. No $\zeta_{S,\ell}$ with $S \in \mathscr{S}$ and $\ell \in [3k]$ prefers Δ to Γ.

Proof of Claim 27. Assume that some $\zeta_{S,\ell}$ prefers Δ to Γ. Then, by Claim 22, $3k$ players from Q_S prefer Γ to Δ. Further, by Claim 21, $\Delta(\zeta_{S,\ell})$ does not contain any other player from Q_S but does contain a β_b with $b \in S$ and $3k+2$ friends of β_b that are not in Q_S. Then $u^{minAL}_{\beta_b}(\Delta) \leq M \cdot v_{\zeta_{S,\ell}}(\Delta) + v_{\beta_b}(\Delta) \leq M \cdot 3n + v_{\beta_b}(\Delta) < M \cdot n3k + n3k = u^{minAL}_{\beta_b}(\Gamma)$. Hence, β_b prefers Γ to Δ. Summing up, we have $\#_{\Gamma \succ \Delta} \geq 3k+1$. With $\#_{\Delta \succ \Gamma} > \#_{\Gamma \succ \Delta}$, this implies $\#_{\Delta \succ \Gamma} \geq 3k+2$. Thus there have to be $3k+1$ players, besides $\zeta_{S,\ell}$, who prefer Δ to Γ. Since there are only $3k-1$ β-players left who might prefer Δ to Γ, there have to be at least two other players who prefer Δ to Γ. Because of Claim 26, there can only be two α-players who prefer Δ to Γ if there is also one α-player who prefers Γ to Δ. Hence, in any case, there has to be at least one additional player i of the form $i = \zeta_{S',\ell'}$ or $i = \eta_{S',j'}$ who prefers Δ to Γ. If $i = \zeta_{S,\ell'}$ for some $\ell' \in [3k], \ell' \neq \ell$, then with the same arguments as for $\zeta_{S,\ell}$ there has to be an additional $\beta_{b'}$ who prefers Γ to Δ. If i is from another $Q_{S'}$, $S' \neq S$, then again, by Claim 22 or 25, at least $3k$ further players prefer Γ to Δ. Both cases again imply that there have to be some more ζ- and η-players who prefer Δ to Γ. Inductively, it follows that there are more players who prefer Γ to Δ than vise versa. This is a contradiction. ❏ Claim 27

▷ Claim 28. No $\eta_{S,j}$ with $S \in \mathscr{S}$ and $j \in [3]$ prefers Δ to Γ.

Proof of Claim 28. Assume that some $\eta_{S,j}$ prefers Δ to Γ. Then, by Claim 25, the other $3k+2$ players in Q_S prefer Γ to Δ. Hence, $\#_{\Delta \succ \Gamma} > \#_{\Gamma \succ \Delta}$ implies $\#_{\Delta \succ \Gamma} \geq 3k+3$. Since no $\zeta_{S,\ell}$ prefers Δ to Γ by Claim 27 and since not all β- and all α-players can prefer Δ to Γ at the same time (see Claim 26), there is another player $\eta_{S',j'}$ with $S' \neq S$, $j' \in [3]$ who prefers Δ to Γ. However, this again implies $3k+2$ players from $Q_{S'}$ who prefer Γ to Δ. Inductively, there are always more players who prefer Γ to Δ than vise versa, which is a contradiction. ❏ Claim 28

▷ Claim 29. α_1 prefers Γ to Δ.

Figure 4 Networks of friends in Example 31 (left) and Example 32 (right)

Proof of Claim 29. First, if α_1 prefers Δ to Γ, by Claim 26, α_2 and α_3 do not prefer Δ to Γ. Moreover, $u^{minAL}_{\alpha_1}(\Delta) > u^{minAL}_{\alpha_1}(\Gamma) = M \cdot n3k + n3k$, which means that all friends of α_1's in $\Delta(\alpha_1)$ have at least $3k+1$ friends in $\Delta(\alpha_1)$. Clearly, $\alpha_2 \notin \Delta(\alpha_1)$ and $\alpha_3 \notin \Delta(\alpha_1)$ but there is at least one β_b in $\Delta(\alpha_1)$. Since this β_b needs $3k+1$ friends in $\Delta(\alpha_1)$, there is at least one $\zeta_{S,\ell}$ with $b \in S$ in $\Delta(\alpha_1)$. With Claims 23, 24, 27, and 28, it follows that all $3k+3$ players from Q_S prefer Γ to Δ. Hence, $\#_{\Gamma \succ \Delta} \geq |Q_S| = 3k + 3$ and $\#_{\Delta \succ \Gamma} \leq |\{\alpha_1\} \cup \{\beta_1, \dots, \beta_{3k}\}| = 3k + 1$, contradicting $\#_{\Delta \succ \Gamma} > \#_{\Gamma \succ \Delta}$.

Second, if α_1 is indifferent between Γ and Δ, then $u^{minAL}_{\alpha_1}(\Delta) = M \cdot n3k + n3k$, which means that α_1 has exactly $3k$ friends in $\Delta(\alpha_1)$ and all these friends have exactly $3k$ friends in $\Delta(\alpha_1)$. This implies $\Delta(\alpha_1) = \{\alpha_1, \beta_1, \dots, \beta_{3k}\}$. However, this is a contradiction because there is no player left who could prefer Δ to Γ. ❑ Claim 29

▷ Claim 30. For every $S \in \mathscr{S}$, either $Q_S \in \Delta$ or $\{\beta_b \mid b \in S\} \cup \{\zeta_{S,\ell} \mid \ell \in [3k]\} \in \Delta$.

Proof of Claim 30. Assume that the statement does not hold for some $S \in \mathscr{S}$. Then, by Claims 23 and 24, no player in Q_S is indifferent between Γ and Δ. By Claims 27 and 28, no player in Q_S prefers Δ to Γ. Thus all $3k + 3$ players from Q_S prefer Γ to Δ. Hence, $\#_{\Gamma \succ \Delta} \geq |Q_S| = 3k+3$ and $\#_{\Delta \succ \Gamma} \leq |\{\alpha_2, \alpha_3, \beta_1, \dots, \beta_{3k}\}| = 3k+2$, which is a contradiction to $\#_{\Delta \succ \Gamma} > \#_{\Gamma \succ \Delta}$. ❑ Claim 30

Now, we use all these claims to show that the existence of Δ implies the existence of an exact cover of B. Let $k' = |\{S \in \mathscr{S} \mid Q_S \notin \Delta\}$ (or, equivalently, $k' = |\{S \in \mathscr{S} \mid \{\beta_b \mid b \in S\} \cup \{\zeta_{S,\ell} \mid \ell \in [3k]\} \in \Delta\}$). It is clear that $k' \geq 1$ because otherwise Δ could not be more popular than Γ. We show that $k' = k$. First, assume that $k' > k$. Then, by the preceding claims, we have $\#_{\Gamma \succ \Delta} \geq |\{\eta_{S,j} \mid Q_S \notin \Delta, j \in [3]\} \cup \{\alpha_1\}| = 3k' + 1 > 3k + 1$ and $\#_{\Delta \succ \Gamma} \leq |\{\alpha_2, \alpha_3, \beta_1, \dots, \beta_{3k}\}| = 3k + 2$. This contradicts $\#_{\Delta \succ \Gamma} > \#_{\Gamma \succ \Delta}$. Second, assume that $k' < k$. All $3k'$ β-players that are in one of the k' coalitions of the form $\{\beta_b \mid b \in S\} \cup \{\zeta_{S,\ell} \mid \ell \in [3k]\}$ prefer Δ to Γ. However, all other $3k - 3k'$ β-players have no ζ-players in their coalitions and thus prefer Γ to Δ. Hence, $\#_{\Gamma \succ \Delta} \geq 3k' + 1 + (3k - 3k') = 3k + 1 > 3k' + 1$ and $\#_{\Delta \succ \Gamma} \leq 3k' + 2$. This again contradicts $\#_{\Delta \succ \Gamma} > \#_{\Gamma \succ \Delta}$. Thus $k' = k$. Now, since there are $k' = k$ sets $S \in \mathscr{S}$ such that each $\{\beta_b \mid b \in S\} \cup \{\zeta_{S,\ell} \mid \ell \in [3k]\}$ contains three distinct β_b, $\{S \in \mathscr{S} \mid Q_S \notin \Delta\}$ is an exact cover of B of size k. ❑

A.7 Popularity Existence

Finally, we give examples where no popular coalition structures exist for AHGs and MBAHGs. They were verified using brute force.

▶ **Example 31.** For the left network of friends in Figure 4, there is no popular coalition structure under all three degrees of altruism in AHGs.

▶ **Example 32.** For the right network of friends in Figure 4, there is no popular coalition structure under all three degrees of altruism in MBAHGs.

3.3 Altruism in Coalition Formation Games

The following article studies altruism in the more general scope of coalition formation games.

Publication (Kerkmann et al. [90])

A. Kerkmann, S. Cramer, and J. Rothe. "Altruism in Coalition Formation Games". Submitted to the *Annals of Mathematics and Artificial Intelligence*. 2022

3.3.1 Summary

Inspired by the altruistic hedonic games by Nguyen et al. [107], this work introduces *altruism in general coalition formation games*. While extending the framework of Nguyen et al. [107], we model agents to behave altruistically to *all their friends*, not only to the friends in their current coalitions (as it is the case for altruistic hedonic games). The model is grounded on the *friends-and-enemies encoding* by Dimitrov et al. [50] where players can be represented by the vertices of an undirected graph with the edges representing mutual friendship relations. We then consider the *friend-oriented valuations* of the agents and distinguish between the three *degrees of altruism* introduced by Nguyen et al. [107]: selfish first, equal treatment, and altruistic treatment. We further distinguish between a sum-based and minimum-based aggregation of valuations. We show that our resulting altruistic models satisfy some desirable properties and argue that it is not reasonable to exclude any of an agent's friends from her altruistic behavior. We show that all our models lead to *unanimous* preferences while the altruistic hedonic games by Nguyen et al. [107] (and the min-based altruistic hedonic games by Wiechers and Rothe [145]) can lead to equal-treatment and altruistic-treatment preferences that are not unanimous. Moreover, we show that our models also fulfill some basic properties introduced by Nguyen et al. [107] but our models fulfill more types of *monotonicity* than the altruistic hedonic models. After completing the axiomatic study of altruistic coalition formation games, we consider some common stability notions from the context of hedonic games. We extend the notions to the more general context of our work and study the computational complexity of the associated *verification* and *existence problems*. We obtain broad results for the case of selfish-first preferences and initiate the study for the other two degrees of altruism. In particular, we show that the verification and existence problems are in P for individual rationality, Nash stability, and individual stability in all our altruistic models (all three degrees of altruism and both aggregation functions). For core stability, popularity, and strict popularity verification, we obtain coNP-completeness results for the selfish-first models. Core stability and strict core stability existence are trivial for selfish-first altruistic coalition formation games as there always exist strictly core stable coalition structures in these games. Furthermore, we obtain several upper bounds on the complexity of perfectness verification and existence.

3.3.2 Personal Contribution and Preceding Versions

Me and Jörg Rothe published a work about sum-based altruism in coalition formation games at IJCAI'20 [84]. This journal article merges the IJCAI'20 paper [84] with a Bachelor's thesis about min-based altruism by Simon Cramer [46] and further results that were partly presented by me and Jörg Rothe at COMSOC'21 (with nonarchival proceedings [86]). Parts of this work were also presented at the 16th and 17th International Symposium on Artificial Intelligence and Mathematics (ISAIM'20 with nonarchival proceedings [85] and ISAIM'22 without any proceedings).

The model that we present in this work extends a model introduced by Nguyen et al. [107]. The presented extension of the model to sum-based altruistic coalition formation games and all technical results concerning sum-based altruistic coalition formation games are my contribution. Furthermore, I contributed all axiomatic results from Section 3 and extended some results for sum-based altruistic coalition formation games to the min-based case (viz., Example 3, Theorem 5, Corollary 1, and Proposition 8).

The writing and polishing of the paper was done jointly with all co-authors.

3.3.3 Publication

The full article [90] is appended here.

Altruism in Coalition Formation Games

Anna Maria Kerkmann[1*], Simon Cramer[1] and Jörg Rothe[1]

[1]Institut für Informatik, Heinrich-Heine-Universität Düsseldorf, Universitätsstr. 1, Düsseldorf, 40225, Germany.

*Corresponding author(s). E-mail(s): anna.kerkmann@hhu.de; Contributing authors: simon.cramer@hhu.de; rothe@hhu.de;

Abstract

Nguyen *et al.* [1] introduced altruistic hedonic games in which agents' utilities depend not only on their own preferences but also on those of their friends in the same coalition. We propose to extend their model to coalition formation games in general, considering also the friends in other coalitions. Comparing our model to altruistic hedonic games, we argue that excluding some friends from the altruistic behavior of an agent is a major disadvantage that comes with the restriction to hedonic games. After introducing our model and showing some desirable properties, we additionally study some common stability notions and provide a computational analysis of the associated verification and existence problems.

Keywords: Coalition formation, Hedonic game, Altruism, Cooperative game theory

1 Introduction

We consider coalition formation games where agents have to form coalitions based on their preferences. Among other compact representations of hedonic coalition formation games, Dimitrov *et al.* [2] in particular proposed the *friends-and-enemies encoding with friend-oriented preferences* which involves a *network of friends*: a (simple) undirected graph whose vertices are the players and where two players are connected by an edge exactly if they are friends of each other. Players not connected by an edge consider each other as enemies. Under friend-oriented preferences, player i prefers a coalition C to a coalition D if C contains more of i's friends than D, or C and D have the same number of i's friends but C contains fewer enemies of i's than D. This is a special case of the *additive encoding* [3]. For more background on these two compact representations, see Section 2 and the book chapter by Aziz and Savani [4].

$$1 \text{ —— } 2 \text{ —— } 3 \text{ —— } 4$$

Fig. 1: Network of friends for Example 1

Based on friend-oriented preferences, Nguyen *et al.* [1] introduced *altruistic hedonic games* where agents gain utility not only from their own satisfaction but also from their friends' satisfaction. However, Nguyen *et al.* [1] specifically considered hedonic games only, which require that an agent's utility only depends on her own coalition. In their interpretation of altruism, the utility of an agent is composed of the agent's own valuation of her coalition and the valuation of all this agent's friends *in this coalition*. While Nguyen *et al.* [1] used the average when aggregating some agents' valuations, Wiechers and Rothe [5] proposed a variant of altruistic hedonic games where some agents' valuations are aggregated by taking the minimum.

Inspired by the idea of altruism, we extend the model of altruism in hedonic games to coalition formation games in general. That is, we propose a model where agents behave altruistically to *all their friends*, not only to the friends in the same coalition. Not restricting to hedonic games, we aim to capture a more natural notion of altruism where none of an agent's friends is excluded from her altruistic behavior.

Example 1 To become acquainted with this idea of altruism, consider the coalition formation game that is represented by the *network of friends* in Figure 1. For the coalition structures $\Gamma = \{\{1,2,3\},\{4\}\}$ and $\Delta = \{\{1,2,4\},\{3\}\}$, it is clear that player 1 is indifferent between coalitions $\{1,2,3\}$ and $\{1,2,4\}$ under friend-oriented preferences, as both coalitions contain 1's only friend (player 2) and one of 1's enemies (either 3 or 4). Under altruistic hedonic preferences [1], however, player 1 behaves altruistically to her friend 2 (who is friends with 3 but not with 4) and therefore prefers $\{1,2,3\}$ to $\{1,2,4\}$. Now, consider the slightly modified coalition structures $\Gamma' = \{\{1\},\{2,3\},\{4\}\}$ and $\Delta' = \{\{1\},\{2,4\},\{3\}\}$. Intuitively, one would still expect 1 to behave altruistically to her friend 2. However, under any *hedonic* preference (which requires the players' preferences to depend *only* on their own coalitions), player 1 (being in the same coalition for both Γ' and Δ') must be indifferent between Γ' and Δ'.

In order to model altruism globally, we release the restriction to hedonic games and introduce *altruistic coalition formation games* where agents behave altruistically to all their friends, independently of their current coalition.

1.1 Related Work

Coalition formation games, as considered here, are closely related to the subclass of *hedonic games* which has been broadly studied in the literature, addressing the issue of compactly representing preferences, conducting axiomatic analyses, dealing with different notions of stability, and investigating the computational complexity of the associated problems (see, e.g., the book chapter by Aziz and Savani [4]).

Closest related to our work are the *altruistic hedonic games* by Nguyen *et al.* [1] (see also the related minimization-based variant by Wiechers and Rothe [5]), which we modify to obtain our more general models of altruism. Based on the model due to Nguyen *et al.* [1], Schlueter and Goldsmith [6] defined *super altruistic hedonic* *games* where friends have a different impact on an agent based on their distances in the underlying network of friends. More recently, Bullinger and Kober [7] introduced *loyalty in cardinal hedonic games* where agents are loyal to all agents in their so-called loyalty set. In their model, the utilities of the agents in the loyalty set are aggregated by taking the minimum. They then study the loyal variants of common classes of cardinal hedonic games such as additively separable and friend-oriented hedonic games.[1]

Altruism has also been studied for *noncooperative* games. Most prominently, Ashlagi *et al.* [8] introduced *social context games* where a social context is applied to a strategic game and the costs in the resulting game depend on the original costs and a graph of neighborhood. Their so-called *MinMax collaborations* (where players seek to minimize the maximal cost of their own and their neighbors) are related to our minimization-based equal-treatment model. Still, the model of Ashlagi *et al.* [8] differs from ours in that they consider *non*cooperative games. Other work considering noncooperative games with social networks is due to Bilò *et al.* [9] who study social context games for other underlying strategic games than Ashlagi *et al.* [8], Hoefer *et al.* [10] who study *considerate equilibria* in strategic games, and Anagnostopoulos *et al.* [11] who study *altruism* and *spite* in strategic games. Further work studying altruism in noncooperative games without social networks is due to Hoefer and Skopalik [12], Chen *et al.* [13], Apt and Schäfer [14], and Rahn and Schäfer [15].

1.2 Our Contribution

Conceptually, we extend the models of altruism proposed by Nguyen *et al.* [1] and Wiechers and Rothe [5] from hedonic games to general coalition formation games. We argue how this captures a more global notion of altruism and show that our models fulfill some desirable properties that are violated by the previous models. We then study the common stability concepts in this model and analyze the associated verification and existence problems in terms of their computational complexity.

This work extends a preliminary version that appeared in the proceedings of the *29th International Joint Conference on Artificial Intelligence* (IJCAI'20) [16]. Parts of this work were also presented at the 16th and 17th International Symposium on Artificial Intelligence and Mathematics (ISAIM'20 and ISAIM'22) and at the *8th International Workshop on Computational Social Choice* (COMSOC'21), each with nonarchival proceedings.

2 The Model

In *coalition formation games*, players divide into groups based on their preferences. Before introducing altruism, we now give some foundations of such games.

[1]Note that their loyal variant of symmetric friend-oriented hedonic games is equivalent to the minimization-based altruistic hedonic games under equal treatment introduced by Wiechers and Rothe [5].

2.1 Coalition Formation Games

Let $N = \{1, \dots, n\}$ be a set of *agents* (or *players*). Each subset of N is called a *coalition*. A *coalition structure* Γ is a partition of N, and we denote the set of all possible coalition structures for N by $\mathscr{C}_N$. For a player $i \in N$ and a coalition structure $\Gamma \in \mathscr{C}_N$, $\Gamma(i)$ denotes the unique coalition in Γ containing i. Now, a *coalition formation game (CFG)* is a pair $(N, \succeq)$, where $N = \{1, \dots, n\}$ is a set of agents, $\succeq = (\succeq_1, \dots, \succeq_n)$ is a profile of preferences, and every preference $\succeq_i \in \mathscr{C}_N \times \mathscr{C}_N$ is a complete weak order over all possible coalition structures. Given two coalition structures Γ, $\Delta \in \mathscr{C}_N$, we say that i *weakly prefers* Γ to Δ if $\Gamma \succeq_i \Delta$. When $\Gamma \succeq_i \Delta$ but not $\Delta \succeq_i \Gamma$, we say that i *prefers* Γ to Δ (denoted by $\Gamma \succ_i \Delta$), and we say that i is *indifferent* between Γ and Δ (denoted by $\Gamma \sim_i \Delta$) if $\Gamma \succeq_i \Delta$ and $\Delta \succeq_i \Gamma$.

Note that *hedonic games* are a special case of coalition formation games where the agents' preference relations only depend on the coalitions containing themselves. In a hedonic game $(N, \succeq)$, agent $i \in N$ is indifferent between any two coalition structures Γ and Δ as long as her coalition is the same, i.e., $\Gamma(i) = \Delta(i) \implies \Gamma \sim_i \Delta$. Therefore, the preference order of any agent $i \in N$ in a hedonic game $(N, \succeq)$ is usually represented by a complete weak order over the set of coalitions containing i.

2.2 The "Friends and Enemies" Encoding

Since $|\mathscr{C}_N|$, the number of all possible coalition structures, is extremely large in the number of agents,[2] it is not reasonable to ask every agent for her complete preference over $\mathscr{C}_N$. Instead, we are looking for a way to compactly represent the agents' preferences. In the literature, many such representations have been proposed for hedonic games, such as the *additive encoding* [3, 19, 20], the *singleton encoding* due to Cechlárová and Romero-Medina [21] and further studied by Cechlárová and Hajduková [22], the *friends-and-enemies encoding* due to Dimitrov *et al.* [2], and FEN-hedonic games due to Kerkmann *et al.* [23] and also used by Rothe *et al.* [24]. Here, we use the *friends-and-enemies encoding* due to Dimitrov *et al.* [2]. We focus on their friend-oriented model and will later adapt it to our altruistic model.

In the friend-oriented model, the preferences of the agents in N are given by a network of friends, i.e., a (simple) undirected graph $G = (N, A)$ whose vertices are the players and where two players $i, j \in N$ are connected by an edge $\{i, j\} \in A$ exactly if they are each other's friends. Agents not connected by an edge consider each other as enemies. For an agent $i \in N$, we denote the set of i's friends by $F_i = \{j \in N \mid \{i, j\} \in A\}$ and the set of i's enemies by $E_i = N \setminus (F_i \cup \{i\})$. Under *friend-oriented preferences* as defined by Dimitrov *et al.* [2], between any two coalitions players prefer the coalition with more friends, and if there are equally many friends in both coalitions, they prefer the coalition with fewer enemies:

$$C \succeq_i^F D \iff |C \cap F_i| > |D \cap F_i| \text{ or } (|C \cap F_i| = |D \cap F_i| \text{ and } |C \cap E_i| \leq |D \cap E_i|).$$

This can also be represented additively. Assigning a value of n to each friend and a value of -1 to each enemy, agent $i \in N$ values coalition C containing herself with

[2] The number of possible partitions of a set with n elements equals the n-th Bell number [17, 18], defined as $B_n = \sum_{k=0}^{n-1} \binom{n-1}{k} B_k$ with $B_0 = B_1 = 1$. For example, for $n = 10$ agents, we have $B_{10} = 115{,}975$ possible coalition structures.

$v_i(C) = n|C \cap F_i| - |C \cap E_i|$. Note that $-(n-1) \leq v_i(C) \leq n(n-1)$, and $v_i(C) > 0$ if and only if there is at least one friend of i's in C. For a given coalition structure $\Gamma \in \mathscr{C}_N$, we also write $v_i(\Gamma)$ for player i's value of $\Gamma(i)$.

Furthermore, we denote the sum of the values of i's friends by $\mathrm{sum}_i^F(\Gamma) = \sum_{f \in F_i} v_f(\Gamma)$. Analogously, we also define $\mathrm{sum}_i^{F+}(\Gamma) = \sum_{f \in F_i \cup \{i\}} v_f(\Gamma)$, $\min_i^F(\Gamma) = \min_{f \in F_i} v_f(\Gamma)$, and $\min_i^{F+}(\Gamma) = \min_{f \in F_i \cup \{i\}} v_f(\Gamma)$.

2.3 Three Degrees of Altruism

When we now define altruistic coalition formation games based on the friend-oriented preference model, we consider the same three degrees of altruism that Nguyen *et al.* [1] introduced for altruistic hedonic games. However, we adapt them to our model, extending the agents' altruism to *all* their friends, not only to their friends in the same coalition.

- **Selfish First (SF):** Agents first rank coalition structures based on their own valuations. Only in the case of a tie between two coalition structures, their friends' valuations are considered as well.
- **Equal Treatment (EQ):** Agents treat themselves and their friends the same. That means that an agent $i \in N$ and all of i's friends have the same impact on i's utility for a coalition structure.
- **Altruistic Treatment (AL):** Agents first rank coalition structures based on their friends' valuations. They only consider their own valuations in the case of a tie.

We further distinguish between a sum-based and a min-based aggregation of some agents' valuations. Formally, for an agent $i \in N$ and a coalition structure $\Gamma \in \mathscr{C}_N$, we denote i's sum-based utility for Γ under SF by $u_i^{sumSF}(\Gamma)$, under EQ by $u_i^{sumEQ}(\Gamma)$, and under AL by $u_i^{sumAL}(\Gamma)$, and her min-based utility for Γ under SF by $u_i^{minSF}(\Gamma)$, under EQ by $u_i^{minEQ}(\Gamma)$, and under AL by $u_i^{minAL}(\Gamma)$. For a constant $M \geq n^3$, they are defined as

$$\begin{aligned}
u_i^{sumSF}(\Gamma) &= M \cdot v_i(\Gamma) + \mathrm{sum}_i^F(\Gamma); & u_i^{minSF}(\Gamma) &= M \cdot v_i(\Gamma) + \min_i^F(\Gamma);\\
u_i^{sumEQ}(\Gamma) &= \mathrm{sum}_i^{F+}(\Gamma); & u_i^{minEQ}(\Gamma) &= \min_i^{F+}(\Gamma);\\
u_i^{sumAL}(\Gamma) &= v_i(\Gamma) + M \cdot \mathrm{sum}_i^F(\Gamma); & u_i^{minAL}(\Gamma) &= v_i(\Gamma) + M \cdot \min_i^F(\Gamma).
\end{aligned}$$

In the case of $F_i = \emptyset$, we define the minimum of the empty set to be zero.

For any coalition structures $\Gamma, \Delta \in \mathscr{C}_N$, agent i's sum-based SF preference is then defined by $\Gamma \succeq_i^{sumSF} \Delta \iff u_i^{sumSF}(\Gamma) \geq u_i^{sumSF}(\Delta)$. Her other altruistic preferences ($\succeq_i^{sumEQ}$; $\succeq_i^{sumAL}$; $\succeq_i^{minSF}$; $\succeq_i^{minEQ}$; and $\succeq_i^{minAL}$) are defined analogously, using the respective utility functions. The factor M, which is used for the SF and AL models, ensures that an agent's utility is first determined by the agent's own valuation in the SF model and first determined by the friends' valuations in the AL model. Similarly as Nguyen *et al.* [1] prove the corresponding properties in hedonic games, we can show that for $M \geq n^3$, $v_i(\Gamma) > v_i(\Delta)$ implies $\Gamma \succ_i^{sumSF} \Delta$ and $\Gamma \succ_i^{minSF} \Delta$, and for $M \geq n^2$, $\mathrm{sum}_i^F(\Gamma) > \mathrm{sum}_i^F(\Delta)$ implies $\Gamma \succ_i^{sumAL} \Delta$ while $\min_i^F(\Gamma) > \min_i^F(\Delta)$ implies $\Gamma \succ_i^{minAL} \Delta$. An *altruistic coalition formation game (ACFG)* is a coalition formation

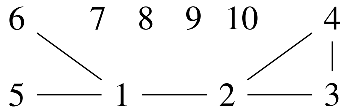


Fig. 2: Network of friends for Example 2

Table 1: Values for the game in Example 2 with the network of friends in Figure 2

	v_1	v_2	v_5	v_6	sum_1^F	sum_1^{F+}	$\min_1^F$	$\min_1^{F+}$
Γ	10	10	0	0	10	20	0	0
Δ	16	20	5	5	30	46	5	5

game where the agents' preferences were obtained by a network of friends via one of these cases of altruism. Hence, we distinguish between sum-based SF, sum-based EQ, sum-based AL, min-based SF, min-based EQ, and min-based AL ACFGs. For any ACFG, the players' utilities can obviously be computed in polynomial time.

3 Monotonicity and Other Properties in ACFGs

Nguyen *et al.* [1] focus on altruism in hedonic games where an agent's utility only depends on her own coalition. As we have already seen in Example 1, there are some aspects of altruistic behavior that cannot be realized by hedonic games. The following example shows that our model crucially differs from the models due to Nguyen *et al.* [1] and Wiechers and Rothe [5].

Example 2 Consider an ACFG $(N, \succeq)$ with the network of friends in Figure 2 and the coalition structures $\Gamma = \{\{1,2\},\{3\},\{4\},\dots,\{10\}\}$ and $\Delta = \{\{1,5,\dots,10\},\{2,3,4\}\}$. We will now compare agent 1's preferences for these two coalition structures under our altruistic models to 1's preferences under the altruistic hedonic models [1, 5]. Table 1 shows all relevant values that are needed to compute the utilities of agent 1.

One can observe that agent 1 and all her friends assign a greater value to Δ than to Γ. Consequently, also the aggregations of the friends' values (sum_1^F, sum_1^{F+}, $\min_1^F$, $\min_1^{F+}$) are greater for Δ. Hence, 1 prefers Δ to Γ under all our sum-based and min-based altruistic preferences.

The hedonic models due to Nguyen *et al.* [1] and Wiechers and Rothe [5], however, are blind to the fact that agent 1 and all her friends are better off in Δ than in Γ. Under their altruistic hedonic preferences, player 1 compares the two coalition structures Γ and Δ only based on her own coalitions $\Gamma(1) = \{1,2\}$ and $\Delta(1) = \{1,5,\dots,10\}$. She then only considers her friends that are in the same coalition, i.e., player 2 for Γ and players 5 and 6 for Δ. This leads to 1 preferring $\Gamma(1)$ to $\Delta(1)$ under altruistic hedonic EQ and AL preferences. In particular, the average (and minimum) valuation of 1's friends in $\Gamma(1)$ is 10 while the average (and minimum) valuation of 1's friends in $\Delta(1)$ is 5. Also considering 1's own value for EQ, the average (and minimum) in $\Gamma(1)$ is 10 while the average (respectively, minimum) value in $\Delta(1)$ is $8.\overline{6}$ (respectively, 5).

3.1 Some Basic Properties

As we have seen in Example 2, altruistic *hedonic* games [1, 5] allow for players that prefer coalition structures that make themselves and all their friends worse off. To avoid this kind of unreasonable behavior, we focus on general coalition formation games. In fact, all our altruistic *coalition-formation* preferences fulfill unanimity: For an ACFG $(N, \succeq)$ and a player $i \in N$, we say that $\succeq_i$ is *unanimous* if, for any two coalition structures $\Gamma, \Delta \in \mathscr{C}_N$, $v_a(\Gamma) > v_a(\Delta)$ for each $a \in F_i \cup \{i\}$ implies $\Gamma \succ_i \Delta$.

This property crucially distinguishes our preference models from the corresponding altruistic *hedonic* preferences, which are not unanimous under EQ or AL preferences, as Example 2 shows. Note that Nguyen *et al.* [1] define a restricted version of unanimity in altruistic hedonic games by considering only the agents' own coalitions. Other desirable properties that were studied by Nguyen *et al.* [1] for altruistic hedonic preferences can be generalized to coalition formation games. We show that these desirable properties also hold for our models. First, we collect some basic observations:

Observation 1 Consider any ACFG $(N, \succeq)$ with an underlying network of friends G.

1. All preferences $\succeq_i$, $i \in N$, are reflexive and transitive.
2. For any player $i \in N$ and any two coalition structures $\Gamma, \Delta \in \mathscr{C}_N$, it can be decided in polynomial time (in the number of agents) whether $\Gamma \succeq_i \Delta$.
3. The preferences $\succeq_i$, $i \in N$, only depend on the structure of G.

Note that the third statement of Observation 1 implies that the properties that Nguyen *et al.* [1] call *anonymity* and *symmetry* are both satisfied in ACFGs. Another desirable property they consider is called *sovereignty of players* and inspired by the axiom of *"citizens' sovereignty"* from social choice theory:[3] Given a set of agents N, a coalition structure $\Gamma \in \mathscr{C}_N$, and an agent $i \in N$, we say that *sovereignty of players* is satisfied if there is a network of friends G on N such that Γ is i's most preferred coalition structure in any ACFG induced by G.

Proposition 1 *ACFGs satisfy sovereignty of players under all sum-based and min-based SF, EQ, and AL altruistic preferences.*

Proof. Sovereignty of players in ACFGs can be shown with an analogous construction as in the proof of Nguyen *et al.* [1, Theorem 5]: For a given set of players N, player $i \in N$, and coalition structure $\Gamma \in \mathscr{C}_N$, we construct a network of friends where all players in $\Gamma(i)$ are friends of each other while there are no other friendship relations. Then Γ is i's (nonunique) most preferred coalition structure under all sum-based and min-based SF, EQ, and AL altruistic preferences. □

[3] Informally stated, a voting rule satisfies *citizens' sovereignty* if every candidate can be made a winner of an election for a suitably chosen preference profile.

3.2 Monotonicity

The next property describes the monotonicity of preferences and further distinguishes our models from altruistic hedonic games. In fact, Nguyen *et al.* [1] define two types of monotonicity, which we here adapt to our setting.

Definition 1 Consider any ACFG $(N, \succeq)$, agents $i, j \in N$ with $j \in E_i$, and coalition structures $\Gamma, \Delta \in \mathscr{C}_N$. Let further $\succeq'_i$ be the preference relation resulting from $\succeq_i$ when j turns from being i's enemy to being i's friend (all else remaining equal). We say that $\succeq_i$ is

- *type-I-monotonic* if (1) $\Gamma \succ_i \Delta$, $j \in \Gamma(i) \cap \Delta(i)$, and $v_j(\Gamma) \geq v_j(\Delta)$ implies $\Gamma \succ'_i \Delta$, and (2) $\Gamma \sim_i \Delta$, $j \in \Gamma(i) \cap \Delta(i)$, and $v_j(\Gamma) \geq v_j(\Delta)$ implies $\Gamma \succeq'_i \Delta$;
- *type-II-monotonic* if (1) $\Gamma \succ_i \Delta$, $j \in \Gamma(i) \setminus \Delta(i)$, and $v_j(\Gamma) \geq v_j(\Delta)$ implies $\Gamma \succ'_i \Delta$, and (2) $\Gamma \sim_i \Delta$, $j \in \Gamma(i) \setminus \Delta(i)$, and $v_j(\Gamma) \geq v_j(\Delta)$ implies $\Gamma \succeq'_i \Delta$.

Theorem 2 *Let $(N, \succeq)$ be an ACFG.*

1. *If $(N, \succeq)$ is sum-based, its preferences satisfy type-I- and type-II-monotonicity.*
2. *If $(N, \succeq)$ is min-based, its preferences satisfy type-II- but not type-I-monotonicity.*

Proof. Let $(N, \succeq)$ be an ACFG with an underlying network of friends $G = (N, H)$. Consider $i \in N$, $\Gamma, \Delta \in \mathscr{C}_N$, and $j \in E_i$ and denote with $G' = (N, H \cup \{\{i, j\}\})$ the network of friends resulting from G when j turns from being i's enemy to being i's friend (all else being equal). Let $(N, \succeq')$ be the ACFG induced by G'. For any agent $a \in N$ and coalition structure $\Gamma \in \mathscr{C}_N$, denote a's value for Γ in G' by $v'_a(\Gamma)$, a's preference relation in $(N, \succeq')$ by $\succeq'_a$, and a's friends and enemies in $(N, \succeq')$ by F'_a and E'_a, respectively. That is, we have $F'_i = F_i \cup \{j\}$, $E'_i = E_i \setminus \{j\}$, $F'_j = F_j \cup \{i\}$, and $E'_j = E_j \setminus \{i\}$. Further, v'_i, v'_j, and $\succeq'_i$ might differ from v_i, v_j, and $\succeq_i$, while the friends, enemies, and values of all other players stay the same, i.e., $F'_a = F_a$, $E'_a = E_a$, and $v'_a = v_a$ for all $a \in N \setminus \{i, j\}$.

Type-I-monotonicity under sum-based preferences.

Let $j \in \Gamma(i) \cap \Delta(i)$ and $v_j(\Gamma) \geq v_j(\Delta)$. It then holds that

$$v'_i(\Gamma) = n|\Gamma(i) \cap F'_i| - |\Gamma(i) \cap E'_i| = n|\Gamma(i) \cap F_i| + n - |\Gamma(i) \cap E_i| + 1 = v_i(\Gamma) + n + 1.$$

Equivalently, $v'_i(\Delta) = v_i(\Delta) + n + 1$, $v'_j(\Gamma) = v_j(\Gamma) + n + 1$, and $v'_j(\Delta) = v_j(\Delta) + n + 1$. Furthermore,

$$\begin{aligned} \mathrm{sum}_i^{F'}(\Gamma) &= \sum_{a \in F'_i} v'_a(\Gamma) = \sum_{a \in F_i \cup \{j\}} v'_a(\Gamma) = \sum_{a \in F_i} v_a(\Gamma) + v'_j(\Gamma) \\ &= \mathrm{sum}_i^F(\Gamma) + v_j(\Gamma) + n + 1 \quad \text{and} \qquad (1) \\ \mathrm{sum}_i^{F'}(\Delta) &= \mathrm{sum}_i^F(\Delta) + v_j(\Delta) + n + 1. \qquad (2) \end{aligned}$$

(1) sumSF: If $\Gamma \succ_i^{sumSF} \Delta$ then either (i) $v_i(\Gamma) = v_i(\Delta)$ and $\mathrm{sum}_i^F(\Gamma) > \mathrm{sum}_i^F(\Delta)$, or (ii) $v_i(\Gamma) > v_i(\Delta)$.

In case (i), $v_i(\Gamma) = v_i(\Delta)$ implies $v_i'(\Gamma) = v_i'(\Delta)$. Applying $\mathrm{sum}_i^F(\Gamma) > \mathrm{sum}_i^F(\Delta)$ and $v_j(\Gamma) \geq v_j(\Delta)$ to (1) and (2), we get $\mathrm{sum}_i^{F\prime}(\Gamma) > \mathrm{sum}_i^{F\prime}(\Delta)$. This together with $v_i'(\Gamma) = v_i'(\Delta)$ implies $\Gamma \succ_i^{sumSF\prime} \Delta$.

In case (ii), $v_i(\Gamma) > v_i(\Delta)$ implies $v_i'(\Gamma) > v_i'(\Delta)$. Hence, $\Gamma \succ_i^{sumSF\prime} \Delta$.

If $\Gamma \sim_i^{sumSF} \Delta$ then $v_i(\Gamma) = v_i(\Delta)$ and $\mathrm{sum}_i^F(\Gamma) = \mathrm{sum}_i^F(\Delta)$. $v_i(\Gamma) = v_i(\Delta)$ implies $v_i'(\Gamma) = v_i'(\Delta)$. Applying $\mathrm{sum}_i^F(\Gamma) = \mathrm{sum}_i^F(\Delta)$ and $v_j(\Gamma) \geq v_j(\Delta)$ to (1) and (2), we get $\mathrm{sum}_i^{F\prime}(\Gamma) \geq \mathrm{sum}_i^{F\prime}(\Delta)$. This together with $v_i'(\Gamma) = v_i'(\Delta)$ implies $\Gamma \succeq_i^{sumSF\prime} \Delta$.

(2) sumEQ: If $\Gamma \succ_i^{sumEQ} \Delta$ then $\mathrm{sum}_i^F(\Gamma) + v_i(\Gamma) > \mathrm{sum}_i^F(\Delta) + v_i(\Delta)$. Using (1), (2), $v_i'(\Gamma) = v_i(\Gamma) + n + 1$, $v_i'(\Delta) = v_i(\Delta) + n + 1$, and $v_j(\Gamma) \geq v_j(\Delta)$, this implies $\mathrm{sum}_i^{F\prime}(\Gamma) + v_i'(\Gamma) > \mathrm{sum}_i^{F\prime}(\Delta) + v_i'(\Delta)$. Hence, $\Gamma \succ_i^{sumEQ\prime} \Delta$.

If $\Gamma \sim_i^{sumEQ} \Delta$, using the same equations, $\Gamma \succeq_i^{sumEQ\prime} \Delta$ is implied.

(3) sumAL: If $\Gamma \succ_i^{sumAL} \Delta$ then either (i) $\mathrm{sum}_i^F(\Gamma) = \mathrm{sum}_i^F(\Delta)$ and $v_i(\Gamma) > v_i(\Delta)$, or (ii) $\mathrm{sum}_i^F(\Gamma) > \mathrm{sum}_i^F(\Delta)$.

In case (i), $\mathrm{sum}_i^F(\Gamma) = \mathrm{sum}_i^F(\Delta)$ together with (1), (2), and $v_j(\Gamma) \geq v_j(\Delta)$ implies $\mathrm{sum}_i^{F\prime}(\Gamma) \geq \mathrm{sum}_i^{F\prime}(\Delta)$. Further, $v_i(\Gamma) > v_i(\Delta)$ together with $v_i'(\Gamma) = v_i(\Gamma) + n + 1$ and $v_i'(\Delta) = v_i(\Delta) + n + 1$ implies $v_i'(\Gamma) > v_i'(\Delta)$. Altogether, this implies $\Gamma \succ_i^{sumAL\prime} \Delta$.

In case (ii), $\mathrm{sum}_i^{F\prime}(\Gamma) > \mathrm{sum}_i^{F\prime}(\Delta)$ is implied and $\Gamma \succ_i^{sumAL\prime} \Delta$ follows.

If $\Gamma \sim_i^{sumAL} \Delta$ then $\mathrm{sum}_i^F(\Gamma) = \mathrm{sum}_i^F(\Delta)$ and $v_i(\Gamma) = v_i(\Delta)$. Using the same equations as before, $\Gamma \succeq_i^{sumAL\prime} \Delta$ is implied.

Type-II-monotonicity under sum-based and min-based preferences.

Let $j \in \Gamma(i) \setminus \Delta(i)$ and $v_j(\Gamma) \geq v_j(\Delta)$. It follows that $v_i'(\Gamma) = v_i(\Gamma) + n + 1$, $v_i'(\Delta) = v_i(\Delta)$, $v_j'(\Gamma) = v_j(\Gamma) + n + 1$, and $v_j'(\Delta) = v_j(\Delta)$. Furthermore,

$$\mathrm{sum}_i^{F\prime}(\Gamma) = \mathrm{sum}_i^F(\Gamma) + v_j(\Gamma) + n + 1, \tag{3}$$

$$\mathrm{sum}_i^{F\prime}(\Delta) = \mathrm{sum}_i^F(\Delta) + v_j(\Delta), \tag{4}$$

$$\min_i^{F\prime}(\Gamma) = \min\big(\min_i^F(\Gamma), v_j(\Gamma) + n + 1\big), \tag{5}$$

$$\min_i^{F\prime}(\Delta) = \min\big(\min_i^F(\Delta), v_j(\Delta)\big), \tag{6}$$

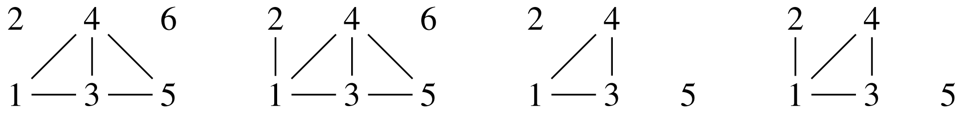
$$\min_i^{F+\prime}(\Gamma) = \min\big(\min_i^F(\Gamma), v_j(\Gamma) + n + 1, v_i(\Gamma) + n + 1\big), \quad \text{and} \tag{7}$$

$$\min_i^{F+\prime}(\Delta) = \min\big(\min_i^F(\Delta), v_j(\Delta), v_i(\Delta)\big). \tag{8}$$

(1) sumSF and minSF: If $\Gamma \succeq_i^{SF} \Delta$ then $v_i(\Gamma) \geq v_i(\Delta)$. Hence, $v_i'(\Gamma) = v_i(\Gamma) + n + 1 \geq v_i(\Delta) + n + 1 > v_i(\Delta) = v_i'(\Delta)$, which implies $\Gamma \succ_i^{SF\prime} \Delta$.

(2) sumEQ: If $\Gamma \succeq_i^{sumEQ} \Delta$ then $\mathrm{sum}_i^F(\Gamma) + v_i(\Gamma) \geq \mathrm{sum}_i^F(\Delta) + v_i(\Delta)$. Together with (3), (4), and $v_j(\Gamma) \geq v_j(\Delta)$ this implies $\mathrm{sum}_i^{F\prime}(\Gamma) + v_j'(\Gamma) > \mathrm{sum}_i^{F\prime}(\Delta) + v_j'(\Delta)$. Hence, $\Gamma \succ_i^{sumEQ\prime} \Delta$.

(3) sumAL: If $\Gamma \succeq_i^{sumAL} \Delta$ then $\mathrm{sum}_i^F(\Gamma) \geq \mathrm{sum}_i^F(\Delta)$. Together with (3), (4), and $v_j(\Gamma) \geq v_j(\Delta)$ this implies $\mathrm{sum}_i^{F\prime}(\Gamma) > \mathrm{sum}_i^{F\prime}(\Delta)$, so $\Gamma \succ_i^{sumAL\prime} \Delta$.



(a) Network of $\mathscr{G}_1$ (b) Network of $\mathscr{G}_1'$ (c) Network of $\mathscr{G}_2$ (d) Network of $\mathscr{G}_2'$

Fig. 3: Networks of friends in the proof of Theorem 2

(4) minEQ: First, assume that $\Gamma \succ_i^{minEQ} \Delta$. We then have $\min\big(\min_i^F(\Gamma), v_i(\Gamma)\big) > \min\big(\min_i^F(\Delta), v_i(\Delta)\big)$. It follows that $\Gamma \succ_i^{minEQ\prime} \Delta$ because

$$\begin{aligned} \min_i^{F+\prime}(\Gamma) &= \min\big(\min_i^F(\Gamma), v_j(\Gamma)+n+1, v_i(\Gamma)+n+1\big) \\ &> \min\big(\min_i^F(\Delta), v_j(\Gamma), v_i(\Delta)\big) \geq \min\big(\min_i^F(\Delta), v_j(\Delta), v_i(\Delta)\big) = \min_i^{F+\prime}(\Delta). \end{aligned} \tag{9}$$

Second, assume $\Gamma \sim_i^{minEQ} \Delta$. Then $\min\big(\min_i^F(\Gamma), v_i(\Gamma)\big) = \min\big(\min_i^F(\Delta), v_i(\Delta)\big)$. Similarly as in (9), it follows that $\min_i^{F+\prime}(\Gamma) \geq \min_i^{F+\prime}(\Delta)$. Hence, $\Gamma \succeq_i^{minEQ\prime} \Delta$.

(5) minAL: First, assume $\Gamma \succ_i^{minAL} \Delta$. Then either (i) $\min_i^F(\Gamma) > \min_i^F(\Delta)$, or (ii) $\min_i^F(\Gamma) = \min_i^F(\Delta)$ and $v_i(\Gamma) > v_i(\Delta)$.

In case of (i), we get $\Gamma \succ_i^{minAL\prime} \Delta$ because

$$\begin{aligned} \min_i^{F\prime}(\Gamma) &= \min\big(\min_i^F(\Gamma), v_j(\Gamma)+n+1\big) \geq \min\big(\min_i^F(\Gamma), v_j(\Delta)+n+1\big) \\ &> \min\big(\min_i^F(\Delta), v_j(\Delta)\big) = \min_i^{F\prime}(\Delta). \end{aligned} \tag{10}$$

In case of (ii), similarly as in (10), we get $\min_i^{F\prime}(\Gamma) \geq \min_i^{F\prime}(\Delta)$. Furthermore, $v_i(\Gamma) > v_i(\Delta)$ implies $v_i'(\Gamma) > v_i'(\Delta)$. Hence, $\Gamma \succ_i^{minAL\prime} \Delta$.

Second, assume that $\Gamma \sim_i^{minAL} \Delta$. Then $\min_i^F(\Gamma) = \min_i^F(\Delta)$ and $v_i(\Gamma) = v_i(\Delta)$. Similarly as in (10), we get $\min_i^{F\prime}(\Gamma) \geq \min_i^{F\prime}(\Delta)$. Furthermore, $v_i(\Gamma) = v_i(\Delta)$ implies $v_i'(\Gamma) > v_i'(\Delta)$. Hence, $\Gamma \succ_i^{minAL\prime} \Delta$.

Type-I-monotonicity under min-based preferences.

To see that $\succeq^{minSF}$ is not type-I-monotonic, consider the game $\mathscr{G}_1$ with the network of friends in Figure 3a. Furthermore, consider the coalition structures $\Gamma = \{\{1,2\},\{3,4,5\},\{6\}\}$ and $\Delta = \{\{1,2\},\{3,4,5,6\}\}$ and players $i = 1$ and $j = 2$ with $2 \in \Gamma(1) \cap \Delta(1)$, and $v_2(\Gamma) = -1 = v_2(\Delta)$. It holds that $v_1(\Gamma) = v_1(\Delta) = -1$, $\min_1^F(\Gamma) = 2n$, and $\min_1^F(\Delta) = 2n-1$. Hence, $\Gamma \succ_1^{minSF} \Delta$.

Now, making 2 a friend of 1's leads to the game $\mathscr{G}_1'$ with the network of friends in Figure 3b. For this game, we have $v_1'(\Gamma) = v_1'(\Delta) = n$ and $\min_1^{F\prime}(\Gamma) = \min_1^{F\prime}(\Delta) = n$. This implies $\Gamma \sim_1^{minSF\prime} \Delta$, which contradicts type-I-monotonicity.

To see that $\succeq^{minEQ}$ and $\succeq^{minAL}$ are not type-I-monotonic, consider the game $\mathscr{G}_2$ with the network of friends in Figure 3c. Consider the coalition structures $\Gamma = \{\{1,2,3,4\},\{5\}\}$ and $\Delta = \{\{1,2,3,4,5\}\}$ and players $i = 1$ and $j = 2$ with $2 \in \Gamma(1) \cap \Delta(1)$, and $v_2(\Gamma) = -3 > -4 = v_2(\Delta)$. It holds that $\min_1^{F+}(\Gamma) = \min_1^F(\Gamma) = 2n-1$, and $\min_1^{F+}(\Delta) = \min_1^F(\Delta) = 2n-2$. Hence, $\Gamma \succ_1^{minEQ} \Delta$ and $\Gamma \succ_1^{minAL} \Delta$.

Now, making 2 a friend of 1's leads to the game $\mathscr{G}_2'$ with the network of friends in Figure 3d. For this game, we have $\min_1^{F+\prime}(\Gamma) = \min_1^{F\prime}(\Gamma) = n$ and $\min_1^{F+\prime}(\Delta) = \min_1^{F\prime}(\Delta) = n$. This implies $\Gamma \sim_1^{minEQ\prime} \Delta$ and $\Gamma \sim_1^{minAL\prime} \Delta$, contradicting type-I-monotonicity and completing the proof. □

Note that the hedonic models of altruism [1, 5] violate both type-I- and type-II-monotonicity for EQ and AL. Hence, it is quite remarkable that all three degrees of our extended sum-based model of altruism satisfy both types of monotonicity.

4 Stability in ACFGs

The main question in coalition formation games is which coalition structures might form. There are several stability concepts that are well-studied for hedonic games, each indicating whether a given coalition structure would be accepted by the agents or if there are other coalition structures that are more likely to form. Although we consider more general coalition formation games, we can easily adapt these definitions to our framework.

Let $(N, \succeq)$ be an ACFG with preferences $\succeq = (\succeq_1, \dots, \succeq_n)$ obtained from a network of friends via one of the three degrees of altruism and with either sum-based or min-based aggregation of the agents' valuations. We use the following notation. For a coalition structure $\Gamma \in \mathscr{C}_N$, a player $i \in N$, and a coalition $C \in \Gamma \cup \{\emptyset\}$, $\Gamma_{i\to C}$ denotes the coalition structure that arises from Γ when moving i to C, i.e.,

$$\Gamma_{i\to C} = \Gamma \setminus \{\Gamma(i), C\} \cup \{\Gamma(i) \setminus \{i\}, C \cup \{i\}\}.$$

In addition, we use $\Gamma_{C\to\emptyset}$, with $C \subseteq N$, to denote the coalition structure that arises from Γ when all players in C leave their respective coalition and form a new one, i.e.,

$$\Gamma_{C\to\emptyset} = \Gamma \setminus \{\Gamma(j) \mid j \in C\} \cup \{\Gamma(j) \setminus C \mid j \in C\} \cup \{C\}.$$

Finally, for any two coalition structures $\Gamma, \Delta \in \mathscr{C}_N$, let $\#_{\Gamma \succ \Delta} = |\{i \in N \mid \Gamma \succ_i \Delta\}|$ be the number of players that prefer Γ to Δ. Now, we are ready to define the common stability notions.

Definition 2 A coalition structure Γ is said to be

- *Nash stable* if no player prefers moving to another coalition:

$$(\forall i \in N)(\forall C \in \Gamma \cup \{\emptyset\})[\Gamma \succeq_i \Gamma_{i\to C}];$$

- *individually rational* if no player would prefer being alone:

$$(\forall i \in N)[\Gamma \succeq_i \Gamma_{i\to\emptyset}];$$

- *individually stable* if no player prefers moving to another coalition and could deviate to it without harming any player in that coalition:

$$(\forall i \in N)(\forall C \in \Gamma \cup \{\emptyset\})\left[\Gamma \succeq_i \Gamma_{i \to C} \vee (\exists j \in C)[\Gamma \succ_j \Gamma_{i \to C}]\right];$$

- *contractually individually stable* if no player prefers another coalition and could deviate to it without harming any player in the new or the old coalition:

$$(\forall i \in N)(\forall C \in \Gamma \cup \{\emptyset\})\left[\Gamma \succeq_i \Gamma_{i \to C} \vee (\exists j \in C)[\Gamma \succ_j \Gamma_{i \to C}] \vee (\exists k \in \Gamma(i))[\Gamma \succ_k \Gamma_{i \to C}]\right];$$

- *totally individually stable* if no player prefers another coalition and could deviate to it without harming any other player:

$$(\forall i \in N)(\forall C \in \Gamma \cup \{\emptyset\})\left[\Gamma \succeq_i \Gamma_{i \to C} \vee (\exists l \in N \setminus \{i\})[\Gamma \succ_l \Gamma_{i \to C}]\right];$$

- *core stable* if no nonempty coalition blocks Γ:

$$(\forall C \subseteq N, C \neq \emptyset)(\exists i \in C)[\Gamma \succeq_i \Gamma_{C \to \emptyset}];$$

- *strictly core stable* if no coalition weakly blocks Γ:

$$(\forall C \subseteq N)(\exists i \in C)[\Gamma \succ_i \Gamma_{C \to \emptyset}] \vee (\forall i \in C)[\Gamma \sim_i \Gamma_{C \to \emptyset}];$$

- *popular* if for every other coalition structure Δ, at least as many players prefer Γ to Δ as there are players who prefer Δ to Γ:

$$(\forall \Delta \in \mathscr{C}_N, \Delta \neq \Gamma)\left[\#_{\Gamma \succ \Delta} \geq \#_{\Delta \succ \Gamma}\right];$$

- *strictly popular* if for every other coalition structure Δ, more players prefer Γ to Δ than there are players who prefer Δ to Γ:

$$(\forall \Delta \in \mathscr{C}_N, \Delta \neq \Gamma)\left[\#_{\Gamma \succ \Delta} > \#_{\Delta \succ \Gamma}\right];$$

- *perfect* if no player prefers any coalition structure to Γ:

$$(\forall i \in N)(\forall \Delta \in \mathscr{C}_N)[\Gamma \succeq_i \Delta].$$

Note that "totally individual stability" is a new notion which we introduce here. It strengthens the notion of contractually individual stability and makes sense in the context of coalition formation games because players' preferences may also be influenced by coalitions they are not part of.

We now study the associated *verification* and *existence problems* in terms of their computational complexity. We assume the reader to be familiar with the complexity classes P (deterministic polynomial time), NP (nondeterministic polynomial time)

Table 2: Complexity results in sum-based and min-based SF ACFGs

Stability notion α	α-VERIFICATION	α-EXISTENCE
Individual rationality	in P[1]	trivial[1]
Nash stability	in P[1]	trivial[1]
Individual stability	in P[1]	trivial[1]
Core stability	coNP-complete[2]	trivial
Strict core stability	in coNP[2]	trivial
Popularity	coNP-complete[2]	not trivial[1]
Strict popularity	coNP-complete[2]	coNP-hard
Perfectness	in P[2]	in P[3]

[1] also holds for sum-based and min-based EQ and AL ACFGs
[2] is in coNP for any ACFG
[3] is in coNP for sum-based EQ ACFGs

and coNP (the class of complements of NP sets). For more background on computational complexity, we refer to, e.g., the textbooks by Garey and Johnson [25] and Rothe [26]. Given a stability concept α, we define:

- α-VERIFICATION: Given an ACFG $(N, \succeq)$ and a coalition structure $\Gamma \in \mathscr{C}_N$, does Γ satisfy α?
- α-EXISTENCE: Given an ACFG $(N, \succeq)$, does there exist a coalition structure $\Gamma \in \mathscr{C}_N$ that satisfies α?

Table 2 summarizes the results for these problems under sum-based and min-based SF preferences. We will also give results for EQ and AL in this section. In Table 2, however, we only mark if the results for EQ and AL match those for SF.

4.1 Individual Rationality

Verifying individual rationality is easy: We just need to iterate over all agents and compare two coalition structures in each iteration. Since players' utilities can be computed in polynomial time, individual rationality can be verified in time polynomial in the number of agents. The existence problem is trivial, since $\Gamma = \{\{1\}, \dots, \{n\}\}$ is always individually rational. Furthermore, we give the following characterization.

Theorem 3 *Given an ACFG $(N, \succeq)$, a coalition structure $\Gamma \in \mathscr{C}_N$ is individually rational*

1. *under sum-based SF, sum-based EQ, sum-based AL, min-based SF, or min-based AL preferences if and only if it holds for all players $i \in N$ that $\Gamma(i)$ contains a friend of i's or i is alone, formally: $(\forall i \in N)[\Gamma(i) \cap F_i \neq \emptyset \vee \Gamma(i) = \{i\}]$;*
2. *under min-based EQ preferences if and only if for all players $i \in N$, $\Gamma(i)$ contains a friend of i's or i is alone or there is a friend of i's whose valuation of Γ is less than or equal to i's valuation of Γ, formally: $(\forall i \in N)[\Gamma(i) \cap F_i \neq \emptyset \vee \Gamma(i) = \{i\} \vee (\exists j \in F_i)[v_j(\Gamma) \leq v_i(\Gamma)]]$.*

Proof. 1. To show the implication from left to right, if Γ is individually rational, we assume for the sake of contradiction that $\Gamma(i) \cap F_i = \emptyset$ and $\Gamma(i) \neq \{i\}$ for some player $i \in N$. First, we observe that for all $j \in F_i$ we have $v_j(\Gamma) = v_j(\Gamma_{i\to\emptyset})$, as their respective coalition is not affected by i's move. It directly follows that, for all considered models of altruism, player i's utilities for Γ and $\Gamma_{i\to\emptyset}$ only depend on her own valuation, which is greater for $\Gamma_{i\to\emptyset}$ than for Γ (since there are enemies in $\Gamma(i)$ but not in $\Gamma_{i\to\emptyset}(i)$). Hence, i prefers $\Gamma_{i\to\emptyset}$ to Γ, so Γ is not individually rational. This is a contradiction.

The implication from right to left is obvious for all considered models of altruism.

2. From left to right, we have that Γ is individually rational and, for the sake of contradiction, we assume that there is a player $i \in N$ with $\Gamma(i) \cap F_i = \emptyset$ and $\Gamma(i) \neq \{i\}$ and for all $j \in F_i$ we have $v_j(\Gamma) > v_i(\Gamma)$. Since i is the least satisfied player in $F_i \cup \{i\}$, we have $u_i^{minEQ}(\Gamma) = v_i(\Gamma)$. With $v_j(\Gamma_{i\to\emptyset}) = v_j(\Gamma) > v_i(\Gamma)$ for all $j \in F_i$ and $v_i(\Gamma_{i\to\emptyset}) = 0 > v_i(\Gamma)$, we immediately obtain $u_i^{minEQ}(\Gamma_{i\to\emptyset}) > u_i^{minEQ}(\Gamma)$ and $\Gamma_{i\to\emptyset} \succ_i^{minEQ} \Gamma$. This is a contradiction to Γ being individually rational.

From right to left, we have to consider two cases. First, if $\Gamma(i) \cap F_i \neq \emptyset$ or $\Gamma(i) = \{i\}$ for some $i \in N$, we obviously have $\Gamma \succeq_i^{minEQ} \Gamma_{i\to\emptyset}$. Second, if $\Gamma(i) \cap F_i = \emptyset$ and $\Gamma(i) \neq \{i\}$, we know that there is at least one $j \in F_i$ with $v_j(\Gamma) \leq v_i(\Gamma) < 0$. Let j' denote a least satisfied friend of i's in Γ (pick one randomly if there are more than one). Since $\Gamma(i) \cap F_i = \emptyset$, it holds that $\Gamma(j) = \Gamma_{i\to\emptyset}(j)$ for all $j \in F_i$. Consequently, j' is i's least satisfied friend in both coalition structures and we have $u_i^{minEQ}(\Gamma) = v_{j'}(\Gamma) = v_{j'}(\Gamma_{i\to\emptyset}) = u_i^{minEQ}(\Gamma_{i\to\emptyset})$. Hence, $\Gamma \sim_i^{minEQ} \Gamma_{i\to\emptyset}$, so Γ is individually rational. □

4.2 Nash Stability

Since there are at most $|N|$ coalitions in a coalition structure $\Gamma \in \mathscr{C}_N$, we can verify Nash stability in polynomial time: We just iterate over all agents $i \in N$ and all the (at most $|N|+1$) coalitions $C \in \Gamma \cup \{\emptyset\}$ and check whether $\Gamma \succeq_i \Gamma_{i\to C}$. Since we can check a player's altruistic preferences over any two coalition structures in polynomial time and since we have at most a quadratic number of iterations ($|N| \cdot (|N|+1)$), Nash stability verification is in P for any ACFG.

Nash stability existence is trivially in P for any ACFG; indeed, the same example that Nguyen *et al.* [1] gave for altruistic hedonic games works here as well. Specifically, for $C = \{i \in N \mid F_i = \emptyset\} = \{c_1, \dots, c_k\}$ the coalition structure $\{\{c_1\}, \dots, \{c_k\}, N \setminus C\}$ is Nash stable.

4.3 Individual Stability

For individual stability, contractually individual stability, and totally individual stability, existence is also trivially in P. Nash stability implies all these three concepts, hence, the Nash stable coalition structure given above is also (contractually; totally) individually stable.

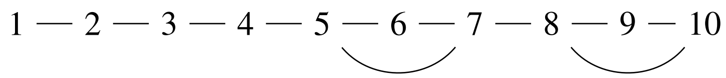


Fig. 4: Networks of friends for Example 3

Verification is also in P for these stability concepts. Similarly to Nash stability, we can iterate over all players and all coalitions and check the respective conditions in polynomial time.

4.4 Core Stability and Strict Core Stability

We now turn to core stability and state some results for sum-based and min-based SF ACFGs. We first show that (strict) core stability existence is trivial for SF ACFGs.

Theorem 4 *Let* $(N, \succeq^{SF})$ *be a (sum-based or min-based) SF ACFG with the underlying network of friends* G. *Let further* $C_1, \dots, C_k$ *be the vertex sets of the connected components of* G. *Then* $\Gamma = \{C_1, \dots, C_k\}$ *is strictly core stable (and thus core stable).*

Proof. For the sake of contradiction, assume that Γ were not strictly core stable, i.e., that there is a coalition $D \neq \emptyset$ that weakly blocks Γ. Consider some player $i \in D$. Since i weakly prefers deviating from $\Gamma(i)$ to D, there have to be at least as many friends of i's in D as in $\Gamma(i)$. Since $\Gamma(i)$ contains all of i's friends, D also has to contain all friends of i's. Then all these friends of i's also have all their friends in D for the same reason, and so on. Consequently, D contains all players from the connected component $\Gamma(i)$, i.e., $\Gamma(i) \subseteq D$.

Since D weakly blocks Γ, D cannot be equal to $\Gamma(i)$ and thus needs to contain some $\ell \notin \Gamma(i)$. Yet, this is a contradiction, as ℓ is an enemy of i's and i would prefer Γ to $\Gamma_{D \to \emptyset}$ if D contains the same number of friends as $\Gamma(i)$ but more enemies than $\Gamma(i)$. □

However, the coalition structure from Theorem 4 is not necessarily core stable under EQ and AL preferences.

Example 3 Let $N = \{1, \dots, 10\}$ and consider the network of friends G shown in Figure 4. Consider the coalition structure consisting of the connected component of G (i.e., of only the grand coalition: $\Gamma = \{N\}$) and the coalition $C = \{8, 9, 10\}$. C blocks Γ under sum-based and min-based EQ and AL preferences. To see this, consider how players 7, 8, 9, and 10 value Γ and $\Gamma_{C \to \emptyset}$:

$$v_7(\Gamma) = v_8(\Gamma) = 30 - 6 = 24, \qquad v_7(\Gamma_{C \to \emptyset}) = 20 - 4 = 16,$$
$$v_9(\Gamma) = v_{10}(\Gamma) = 20 - 7 = 13, \qquad v_8(\Gamma_{C \to \emptyset}) = v_9(\Gamma_{C \to \emptyset}) = v_{10}(\Gamma_{C \to \emptyset}) = 20.$$

We then obtain

- $\mathrm{sum}_8^{F+}(\Gamma) = 74 < 76 = \mathrm{sum}_8^{F+}(\Gamma_{C \to \emptyset})$ and $\mathrm{sum}_9^{F+}(\Gamma) = \mathrm{sum}_{10}^{F+}(\Gamma) = 50 < 60 = \mathrm{sum}_9^{F+}(\Gamma_{C \to \emptyset}) = \mathrm{sum}_{10}^{F+}(\Gamma_{C \to \emptyset})$, so $\Gamma_{C \to \emptyset} \succ_i^{sumEQ} \Gamma$ for all $i \in C$;
- $\mathrm{sum}_8^{F}(\Gamma) = 50 < 56 = \mathrm{sum}_8^{F}(\Gamma_{C \to \emptyset})$ and $\mathrm{sum}_9^{F}(\Gamma) = \mathrm{sum}_{10}^{F}(\Gamma) = 37 < 40 = \mathrm{sum}_9^{F}(\Gamma_{C \to \emptyset}) = \mathrm{sum}_{10}^{F}(\Gamma_{C \to \emptyset})$, so $\Gamma_{C \to \emptyset} \succ_i^{sumAL} \Gamma$ for all $i \in C$;

- $\min_8^{F+}(\Gamma) = \min_8^F(\Gamma) = 13 < 16 = \min_8^{F+}(\Gamma_{C\to\emptyset}) = \min_8^F(\Gamma_{C\to\emptyset})$ and $\min_9^{F+}(\Gamma) = \min_9^F(\Gamma) = \min_{10}^{F+}(\Gamma) = \min_{10}^F(\Gamma) = 13 < 20 = \min_9^{F+}(\Gamma_{C\to\emptyset}) = \min_9^F(\Gamma_{C\to\emptyset}) = \min_{10}^{F+}(\Gamma_{C\to\emptyset}) = \min_{10}^F(\Gamma_{C\to\emptyset})$, which implies $\Gamma_{C\to\emptyset} \succ_i^{minEQ} \Gamma$ and $\Gamma_{C\to\emptyset} \succ_i^{minAL} \Gamma$ for all $i \in C$.

Thus C blocks Γ under sum-based and min-based EQ and AL preferences.

Turning to (strict) core stability verification, we can show that this problem is hard under SF preferences, and we suspect that this hardness also extends to EQ and AL.

Theorem 5 *Strict core stability verification and core stability verification are in* coNP *for any ACFG. For (sum-based and min-based) SF ACFGs, core stability verification is even* coNP-*complete.*

Proof. To see that strict core stability verification and core stability verification are in coNP, consider any coalition structure $\Gamma \in \mathscr{C}_N$ in an ACFG $(N, \succeq)$. Γ is not (strictly) core stable if there is a coalition $C \subseteq N$ that (weakly) blocks Γ. Hence, we nondeterministically guess a coalition $C \subseteq N$ and check whether C (weakly) blocks Γ. This can be done in polynomial time since the preferences of the agents in C for the coalition structures Γ and $\Gamma_{C\to\emptyset}$ can be verified in polynomial time for all our altruistic models.

To show coNP-hardness of core stability verification under min-based SF ACFGs, we use RX3C, which is a restricted variant of Exact Cover by 3-Sets and known to be NP-complete [25, 27]. We provide a polynomial-time many-one reduction from RX3C to the complement of our verification problem. Let $(B, \mathscr{S})$ be an instance of RX3C, consisting of a set $B = \{1, \dots, 3k\}$ and a collection $\mathscr{S} = \{S_1, \dots, S_{3k}\}$ of 3-element subsets of B, where each element of B occurs in exactly three sets in $\mathscr{S}$. The question is whether there exists an exact cover for B in $\mathscr{S}$, i.e., a subset $\mathscr{S}' \subseteq \mathscr{S}$ with $|\mathscr{S}'| = k$ and $\bigcup_{S \in \mathscr{S}'} S = B$. We assume that $k > 4$.

From $(B, \mathscr{S})$ we now construct the following ACFG. The set of players is $N = \{\beta_b \mid b \in B\} \cup \{\zeta_S, \alpha_{S,1}, \alpha_{S,2}, \alpha_{S,3}, \delta_{S,1}, \dots, \delta_{S,4k-3} \mid S \in \mathscr{S}\}$ and we define the sets

$$\begin{aligned} Beta &= \{\beta_b \mid b \in B\}, \\ Zeta &= \{\zeta_S \mid S \in \mathscr{S}\}, \text{ and} \\ Q_S &= \{\zeta_S, \alpha_{S,1}, \alpha_{S,2}, \alpha_{S,3}, \delta_{S,1}, \dots, \delta_{S,4k-3}\} \text{ for each } S \in \mathscr{S}. \end{aligned}$$

Figure 5 shows the network of friends, where a dashed rectangle around a group of players means that all these players are friends of each other:

- All players in *Beta* are friends of each other.
- For every $S \in \mathscr{S}$, ζ_S is friend with every β_b with $b \in S$ and with $\alpha_{S,1}$, $\alpha_{S,2}$, and $\alpha_{S,3}$.
- For every $S \in \mathscr{S}$, $\alpha_{S,1}$, $\alpha_{S,2}$, $\alpha_{S,3}$, and $\delta_{S,1}$ are friends of each other.
- For every $S \in \mathscr{S}$, all players in $\{\delta_{S,1}, \dots, \delta_{S,4k-3}\}$ are friends of each other.

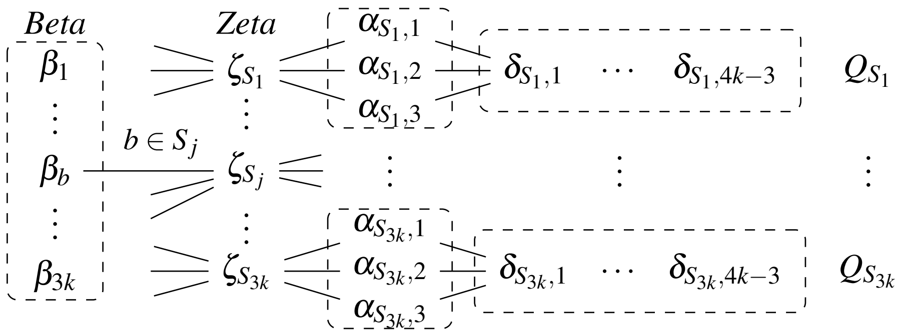


Fig. 5: Network of friends in the proof of Theorem 5 that is used to show coNP-hardness of core stability verification in min-based SF ACFGs. A dashed rectangle around a group of players indicates that all these players are friends of each other.

Furthermore, consider the coalition structure $\Gamma = \{Beta, Q_{S_1}, \dots, Q_{S_{3k}}\}$. We will now show that $\mathscr{S}$ contains an exact cover for B if and only if Γ is not core stable under the min-based SF model.

Only if: Assume that there is an exact cover $\mathscr{S}' \subseteq \mathscr{S}$ for B. Then $|\mathscr{S}'| = k$. Consider coalition $C = Beta \cup \{\zeta_S \mid S \in \mathscr{S}'\}$. C blocks Γ, i.e., $\Gamma_{C\to\emptyset} \succ_i^{minSF} \Gamma$ for all $i \in C$, because (a) every $\beta_b \in Beta$ has $3k$ friends in C but only $3k-1$ friends in *Beta* and (b) every ζ_S with $S \in \mathscr{S}'$ has 3 friends and $4k-4$ enemies in C but 3 friends and $4k-3$ enemies in Q_S.

If: Assume that Γ is not core stable and let $C \subseteq N$ be a coalition that blocks Γ. Then $\Gamma_{C\to\emptyset} \succ_i^{minSF} \Gamma$ for all $i \in C$. First, observe that every $i \in C$ needs to have at least as many friends in C as in $\Gamma(i)$. So, if any $\alpha_{S,j}$ or $\delta_{S,j}$ is in C, it follows quite directly that $Q_S \subseteq C$. However, since Q_S is a coalition in Γ and since every other player (from $N \setminus Q_S$) is an enemy of all δ-players, any coalition C with $Q_S \subseteq C$ cannot be a blocking coalition for Γ. This contradiction implies that no $\alpha_{S,j}$ or $\delta_{S,j}$ is in C.

We now have $C \subseteq Beta \cup Zeta$. Since any $\beta_b \in C$ has $3k-1$ friends and no enemies in $\Gamma(\beta_b)$ and prefers $\Gamma_{C\to\emptyset}$ to Γ, one of the following holds: (a) β_b has at least $3k$ friends in C or (b) β_b has $3k-1$ friends and no enemies in C and β_b's friends assign a higher value to $\Gamma_{C\to\emptyset}$ than to Γ. For a contradiction, assume that (b) holds for some $\beta_b \in C$. First, observe that there are exactly $3k$ players in C (namely, β_b and β_b's $3k-1$ friends). We now distinguish two cases:

Case 1: *All the $3k-1$ friends of β_b's are β-players.* Then C consists of all β-players, i.e., $C = Beta$. This is a contradiction, as *Beta* is already a coalition in Γ.

Case 2: *There are some ζ-players in C that are β_b's friends.* Since β_b has three ζ-friends in total and no enemies in C, there are between one and three ζ-players in C. Hence, there are between $3k-3$ and $3k-1$ β-players in C. Then one of the β-players has no ζ-friend in C. (The at most three ζ-players are friends with at most nine β-players, but $3k-3 > 9$ for $k > 4$.) Consequently, this β-player has only the other (at most $3k-2$) β-players as friends in C and does not prefer $\Gamma_{C\to\emptyset}$ to Γ. This is a contradiction.

Hence, option (a) holds for each $\beta_b \in C$. In total, each β_b has exactly three ζ-friends and $3k-1$ β-friends. Thus at least $3k-3$ of β_b's friends in C are β-players and at least one of β_b's friends in C is a ζ-player. Also counting β_b herself, there are at least $3k-2$ β-players in C. Since all of these $3k-2$ β-players have at least one ζ-friend in C, there are at least k ζ-players in C. (Note that $k-1$ ζ-players are friends with at most $3k-3$ β-players.)

Consider some $\zeta_S \in C$. Since ζ_S has three friends and $4k-3$ enemies in Q_S, at most three friends in C, and prefers $\Gamma_{C\to\emptyset}$ to Γ, ζ_S has exactly three friends and at most $4k-3$ enemies in C. Hence, C contains at most $4k-3+3+1=4k+1$ players.

So far we know that there are at least $3k-2$ β-players in C. If C contains exactly $3k-2$ (or $3k-1$) β-players then each of this players has only $3k-3$ (or $3k-2$) β-friends in C and additionally needs at least three (or two) ζ-friends in C. Hence, we have at least $(3k-2)\cdot 3 = 9k-6$ (or $6k-2$) edges between the β- and ζ-players in C. Then there are at least $3k-2$ (or $2k$) ζ-players in C. Thus there are at least $(3k-2)+(3k-2)=6k-4$ (or $5k-1$) players in C which is a contradiction because there are at most $4k+1$ players in C. Hence, there are exactly $3k$ β-players in C.

Summing up, there are exactly $3k$ β-players, at least k ζ-players, and at most $4k+1$ players in C. Hence, there are k or $k+1$ ζ-players in C. For the sake of contradiction, assume that there are $k+1$ ζ-players in C. Then each $\zeta_S \in C$ has $4k-3$ enemies in C. Since ζ_S prefers $\Gamma_{C\to\emptyset}$ to Γ, this implies that ζ_S has exactly three friends and $4k-3$ enemies in C and the minimal value assigned to $\Gamma_{C\to\emptyset}$ by ζ_S's friends is higher than the minimal value assigned to Γ by ζ_S's friends. In both coalition structures, the minimal value is given by ζ_S's α-friends. However, since these α-players lose ζ_S as a friend when ζ_S deviates to C, the minimal value assigned to Γ is higher than for $\Gamma_{C\to\emptyset}$. This is a contradiction. Hence, there are exactly k ζ-players in C. Finally, since every of the $3k$ $\beta_b \in C$ has one of the k $\zeta_S \in C$ as a friend, it holds that $\{S \mid \zeta_S \in C\}$ is an exact cover for B. This completes the coNP-hardness proof for min-based SF ACFGs.

For sum-based SF ACFGs, coNP-hardness of core stability verification can be shown by a similar construction. Again, given an instance $(B,\mathscr{S})$ of RX3C, with $B=\{1,\dots,3k\}$, $\mathscr{S}=\{S_1,\dots,S_{3k}\}$, and $k>8$, we construct the following ACFG. The set of players is $N=\{\beta_b \mid b\in B\}\cup\{\zeta_S,\alpha_{S,1},\alpha_{S,2},\alpha_{S,3},\delta_{S,1},\dots,\delta_{S,4k-3} \mid S\in\mathscr{S}\}$. We define the sets $Beta=\{\beta_b \mid b\in B\}$ and $Q_S=\{\alpha_{S,1},\alpha_{S,2},\alpha_{S,3},\delta_{S,1},\dots,\delta_{S,4k-3}\}$ for each $S\in\mathscr{S}$. The network of friends is given in Figure 6, where a dashed rectangle around a group of players means that all these players are friends of each other:

- All players in $Beta$ are friends of each other.
- For every $S\in\mathscr{S}$, all players in Q_S are friends of each other.
- For every $S\in\mathscr{S}$, ζ_S is friend with $\alpha_{S,1}$, $\alpha_{S,2}$, and $\alpha_{S,3}$ and with every β_b with $b\in S$.

Similar arguments as above show that the coalition structure $\Gamma=\{Beta\}\cup\{\{\zeta_S\}\cup Q_S \mid S\in\mathscr{S}\}$ is not core stable under sum-based SF preferences if and only if $\mathscr{S}$ contains an exact cover for B. $\square$

4.5 Popularity and Strict Popularity

Now we take a look at popularity and strict popularity. For all considered models of altruism, there are games for which no (strictly) popular coalition structure exists.

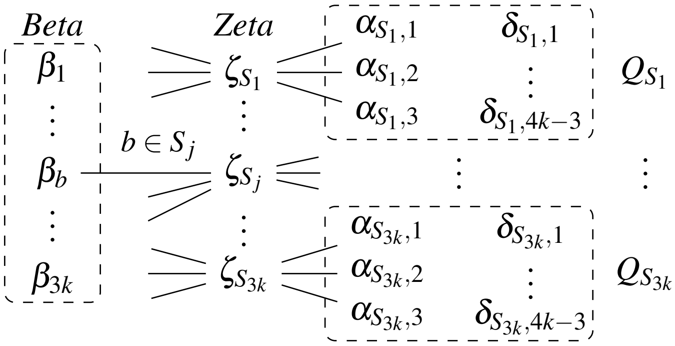


Fig. 6: Network of friends in the proof of Theorem 5 that is used to show coNP-hardness of core stability verification in sum-based SF ACFGs. A dashed rectangle around a group of players indicates that all these players are friends of each other.

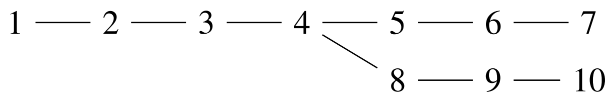


Fig. 7: Network of friends for Example 4

Example 4 Let $N = \{1, \dots, 10\}$ and consider the network of friends shown in Figure 7. Then there is no strictly popular and no popular coalition structure for any of the sum-based or min-based degrees of altruism. Since perfectness implies popularity, there is also no perfect coalition structure for this ACFG.

Recall from Footnote 2 that there are 115,975 possible coalition structures for this game with ten players, which we all tested for this example by brute force.

We now show that, under sum-based and min-based SF preferences, it is hard to verify if a given coalition structure is popular or strictly popular, and it is also hard to decide whether there exists a strictly popular coalition structure for a given SF ACFG.

Theorem 6 *Popularity verification and strict popularity verification are in* coNP *for any ACFG. For (sum-based and min-based) SF ACFGs, popularity verification and strict popularity verification are* coNP*-complete and strict popularity existence is* coNP*-hard.*

Proof. First, we observe that the verification problems are in coNP: To verify that a given coalition structure Γ is not (strictly) popular, we can nondeterministically guess a coalition structure Δ, compare both coalition structures in polynomial time, and accept exactly if Δ is more popular than (or at least as popular as) Γ.

To show coNP-hardness of strict popularity verification for min-based SF ACFGs, we again employ a polynomial-time many-one reduction from RX3C. Let $(B, \mathscr{S})$ be an instance of RX3C, consisting of a set $B = \{1, \dots, 3k\}$ and a collection $\mathscr{S} = \{S_1, \cdots, S_{3k}\}$ of 3-element subsets of B. Recall that every element of B occurs in exactly three sets in $\mathscr{S}$ and the question is whether there is an exact cover $\mathscr{S}' \subseteq \mathscr{S}$ of B.

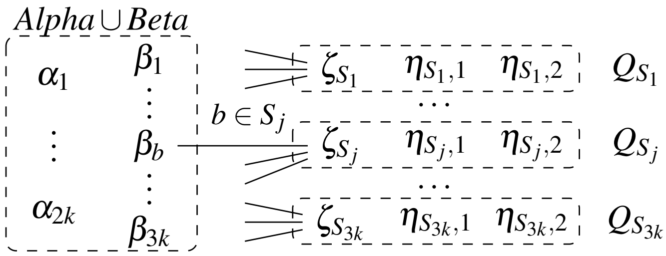


Fig. 8: Network of friends in the proof of Theorem 6 that is used to show coNP-hardness of strict popularity verification in min-based SF ACFGs. A dashed rectangle around a group of players indicates that all these players are friends of each other.

We now construct a network of friends based on this instance. The set of players is given by $N = \{\alpha_1, \dots, \alpha_{2k}\} \cup \{\beta_b \,|\, b \in B\} \cup \{\zeta_S, \eta_{S,1}, \eta_{S,2} \,|\, S \in \mathscr{S}\}$, so in total we have $n = 14k$ players. For convenience, we define $Alpha = \{\alpha_1, \dots, \alpha_{2k}\}$, $Beta = \{\beta_b \,|\, b \in B\}$, and $Q_S = \{\zeta_S, \eta_{S,1}, \eta_{S,2} \,|\, S \in \mathscr{S}\}$ for $S \in \mathscr{S}$. The network of friends is shown in Figure 8, where a dashed square around a group of players means that all these players are friends of each other: All players in $Alpha \cup Beta$ are friends of each other; for every $S \in \mathscr{S}$, all players in Q_S are friends of each other; and ζ_S is a friend of every β_b with $b \in S$.

We consider the coalition structure $\Gamma = \{Alpha \cup Beta\} \cup \{Q_S \,|\, S \in \mathscr{S}\}$ and will now show that $\mathscr{S}$ contains an exact cover for B if and only if Γ is not strictly popular under min-based SF preferences.

Only if: Assuming that there is an exact cover $\mathscr{S}' \subset \mathscr{S}$ for B, we define the coalition structure $\Delta = \{Alpha \cup Beta \cup \bigcup_{S \in \mathscr{S}'} Q_S\} \cup \{Q_S \,|\, S \in \mathscr{S} \setminus \mathscr{S}'\}$. We will now show that Δ is as popular as Γ under min-based SF preferences.

First, all $2k$ α-players prefer Γ to Δ, since they only add enemies to their coalition in Δ. Second, the $3k$ β-players prefer Δ to Γ, as each β-player gains a ζ-friend and then has $5k$ friends instead of $5k-1$. Next, we consider the Q_S-groups for $S \in \mathscr{S}'$, i.e., the groups that were added to $Alpha \cup Beta$ in Δ. We observe that every ζ_S-player in these Q_S-groups prefers Δ to Γ, since ζ_S gains three additional β-friends. For the η-players, on the other hand, the new coalition only contains more enemies, so the η-players prefer Γ to Δ. Since we have $|\mathscr{S}'| = k$, this means k ζ-players prefer Δ to Γ, and $2k$ η-players prefer Γ to Δ. Finally, we consider the remaining Q_S-groups with $S \in \mathscr{S} \setminus \mathscr{S}'$. Here, the coalition containing these players is the same in Γ and Δ. Hence, for each player $p \in Q_S$, we have $v_p(\Gamma) = v_p(\Delta)$. Thus the players have to ask their friends for their valuations. For $\zeta_S \in Q_S$ with $S \in \mathscr{S} \setminus \mathscr{S}'$, the minimum value of her friends is in both structures given by an η-friend, since $\eta_{S,1}$ and $\eta_{S,2}$ value Γ and Δ both with $n \cdot 2$, while the β-friends of ζ_S assign values $n \cdot (5k-1)$ to Γ and $n \cdot 5k - (3k-1)$ to Δ. So we have $u^{minSF}_{\zeta_S}(\Gamma) = u^{minSF}_{\zeta_S}(\Delta)$ and, therefore, $2k$ ζ-players that are indifferent. The η-players in Q_S, $S \in \mathscr{S} \setminus \mathscr{S}'$, are also indifferent, as all their friends value Γ and Δ the same. In total, $\#_{\Delta \succ \Gamma} = |Beta \cup \{\zeta_S \,|\, S \in \mathscr{S}'\}| = 4k = |Alpha \cup \{\eta_{S,1}, \eta_{S,2} \,|\, S \in \mathscr{S}'\}| = \#_{\Gamma \succ \Delta}$ and, therefore, Δ is exactly as popular as Γ, so Γ is not strictly popular.

If: Assuming that Γ is not strictly popular, there is some coalition structure $\Delta \in \mathscr{C}_N$ with $\Delta \neq \Gamma$ such that Δ is at least as popular as Γ under min-based SF preferences. We will now show that this implies the existence of an exact cover for B in $\mathscr{S}$.

First of all, we observe that all α-players' most preferred coalition is $Alpha \cup Beta$, as it contains all their friends and no enemies. Thus we have $\Gamma \succ_{\alpha}^{minSF} \Delta$ if $Alpha \cup Beta \notin \Delta$ and $\Gamma \sim_{\alpha}^{minSF} \Delta$ if $Alpha \cup Beta \in \Delta$.

For the sake of contradiction, we assume that $Alpha \cup Beta \in \Delta$. As $\Delta \neq \Gamma$, the players in the Q_S-groups have to be partitioned differently. However, that would not increase any player's valuation since every player in Q_S can only lose friends and gain enemies. That means that no β-player prefers Δ to Γ, as they are in the same coalition as in Γ and their friends are not more satisfied. We also have at least three players of a Q_S-group that are no longer in the same coalition, so they prefer Γ to Δ. This is a contradiction, as we assumed that Δ is at least as popular as Γ. Thus we have $Alpha \cup Beta \notin \Delta$.

Now consider the η-players. For every $S \in \mathscr{S}$, we know that Q_S is the best valued coalition for $\eta_{S,1}$ and $\eta_{S,2}$. So again, $\eta_{S,1}$ and $\eta_{S,2}$ prefer Γ to Δ if and only if $Q_S \notin \Delta$, and they are indifferent otherwise. Define $k' = |\{S \in \mathscr{S} \mid Q_S \notin \Delta\}|$. So $2k'$ is the number of η-players that prefer Γ to Δ, and the remaining $6k - 2k'$ η-players are indifferent between Γ and Δ. We first collect some observations:

1. All $2k$ α-players prefer Γ to Δ.
2. $2k'$ η-players prefer Γ to Δ, and $6k - 2k'$ η-players are indifferent.
3. $3k - k'$ ζ-players are in the same coalition in both coalition structures, so their utilities depend on their friends' valuations. In Γ, the minimum value of their friends is given by an η-player. Since this η-player is also in the same coalition in Δ and thus assigns the same value, it is not possible that the minimum value of the friends is higher in Δ than in Γ. So $3k - k'$ ζ-players are indifferent or prefer Γ to Δ.
4. We have $14k$ players in total, so we can have at most $14k - 2k - 2k' - (6k - 2k') - (3k - k') = 3k + k'$ players that prefer Δ to Γ.

Next, we show that $k' = k$. First, assume that $k' > k$: We have $\#_{\Gamma \succ \Delta} \geq 2k + 2k'$, and since $k' > k$, we have $2k + 2k' > 3k + k' \geq \#_{\Delta \succ \Gamma}$. This is a contradiction to $\#_{\Gamma \succ \Delta} \leq \#_{\Delta \succ \Gamma}$, so we obtain $k' \leq k$.

Second, let us assume $k' < k$: Since every ζ-player has three β-friends and there are k' ζ-players that are not in their respective Q_S coalition in Δ, there are at most $3k'$ β-players that gain a ζ-friend in Δ. The $3k - 3k'$ other β-players have at most $5k - 1$ friends in Δ, namely all other α- and β-players. But as $Alpha \cup Beta \notin \Delta$, they would also gain at least one enemy, so we have $3k - 3k'$ β-players that prefer Γ. That means we have $\#_{\Gamma \succ \Delta} \geq 2k + 2k' + 3k - 3k' = 5k - k'$ and $\#_{\Delta \succ \Gamma} \leq 3k + k' - (3k - 3k') = 4k'$. Since $k' < k$, we have $5k - k' > 5k - k = 4k > 4k'$, and therefore, $\#_{\Gamma \succ \Delta} > \#_{\Delta \succ \Gamma}$, which again is a contradiction. Thus we conclude that $k' \geq k$ and, in total, $k' = k$.

Consequently, we know that $4k$ players prefer Γ to Δ, namely all α-players and the $2k$ η-players that are not in Q_S anymore. Subtracting all the indifferent players, we observe that all other players have to prefer Δ to Γ in order to ensure $\#_{\Gamma \succ \Delta} \leq \#_{\Delta \succ \Gamma}$. These other players are the $3k$ β-players and the k ζ-players that are not in

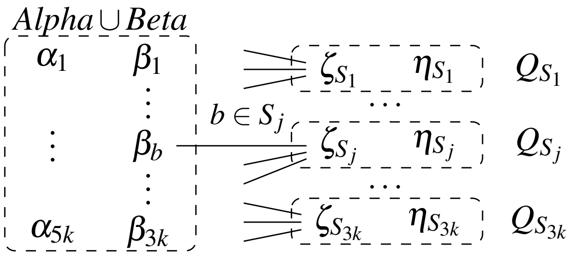


Fig. 9: Network of friends in the proof of Theorem 6 that is used to show coNP-hardness of strict popularity verification in sum-based SF ACFGs. A dashed rectangle around a group of players indicates that all these players are friends of each other.

Q_S anymore. Finally, that is only possible if every β-player gains a ζ-friend in Δ. Hence each one of those k ζ-players has to be friends with three different β-players. Therefore, the set $\{S \in \mathscr{S} \mid Q_S \notin \Delta\}$ is an exact cover for B.

To show coNP-hardness of strict popularity verification for sum-based SF ACFGs, we use a similar construction. For an instance $(B, \mathscr{S})$ of RX3C with $B = \{1, \dots, 3k\}$ and $\mathscr{S} = \{S_1, \dots, S_{3k}\}$, where each element of B occurs in exactly three sets in $\mathscr{S}$, we construct the following ACFG. The set of players is given by $N = \{\alpha_1, \dots, \alpha_{5k}\} \cup \{\beta_b \mid b \in B\} \cup \{\zeta_S, \eta_S \mid S \in \mathscr{S}\}$. Let $Alpha = \{\alpha_1, \dots, \alpha_{5k}\}$, $Beta = \{\beta_b \mid b \in B\}$, and $Q_S = \{\zeta_S, \eta_S\}$ for each $S \in \mathscr{S}$. The network of friends is given in Figure 9, where a dashed rectangle around a group of players means that all these players are friends of each other: All players in $Alpha \cup Beta$ are friends of each other and, for every $S \in \mathscr{S}$, ζ_S is friends with η_S and every β_b with $b \in S$.

Consider the coalition structure $\Gamma = \{Alpha \cup Beta, Q_{S_1}, \dots, Q_{S_{3k}}\}$. We show that $\mathscr{S}$ contains an exact cover for B if and only if Γ is not strictly popular.

Only if: Assuming that there is an exact cover $\mathscr{S}' \subseteq \mathscr{S}$ for B and considering coalition structure $\Delta = \{Alpha \cup Beta \cup \bigcup_{S \in \mathscr{S}'} Q_S\} \cup \{Q_S \mid S \in \mathscr{S} \setminus \mathscr{S}'\}$, it can be shown with similar arguments as before that $\#_{\Delta \succ \Gamma} = |\{\beta_1, \dots, \beta_{3k}, \zeta_{S_1}, \dots, \zeta_{S_{3k}}\}| = 6k = |\{\alpha_1, \dots, \alpha_{5k}\} \cup \{\eta_S \mid S \in \mathscr{S}'\}| = \#_{\Gamma \succ \Delta}$. Hence, Δ and Γ are equally popular.

If: Assuming that Γ is not strictly popular, i.e., that there is a coalition structure $\Delta \in \mathscr{C}_N$, $\Delta \neq \Gamma$, with $\#_{\Gamma \succ \Delta} \leq \#_{\Delta \succ \Gamma}$, it can be shown similarly as before that the set $\{S \in \mathscr{S} \mid Q_S \notin \Delta\}$ is an exact cover for B.

The results for strict popularity existence and popularity verification can be shown by slightly modifying the above reductions.

To show that strict popularity existence is coNP-hard for min-based and sum-based SF ACFGs, we consider the same two reductions as before but the coalition structures Γ are not given as a part of the problem instances. Then, there is an exact cover for B if and only if there is no strictly popular coalition structure. In particular, if there is an exact cover for B, Γ and Δ as defined in the proofs above are in a tie and every other coalition structure is beaten by Γ. And if there is no exact cover for B then Γ beats every other coalition structure and thus is strictly popular.

Popularity verification for min-based and sum-based SF ACFGs can be shown to be coNP-complete by using the same constructions as for strict popularity verification (see Figure 8 and 9) but reducing the numbers of α-players to $2k-1$ and $5k-1$, respectively. Then there is an exact cover for B if and only if Γ, as defined above, is not popular. □

4.6 Perfectness

Turning now to perfectness, we start with the SF model.

Theorem 7 *For any sum-based or min-based SF ACFG $(N,\succeq)$ with an underlying network of friends G, a coalition structure $\Gamma \in \mathscr{C}_N$ is perfect if and only if it consists of the connected components of G and all of them are cliques.*

Proof. From left to right, assume that the coalition structure $\Gamma \in \mathscr{C}_N$ is perfect. It then holds for all agents $i \in N$ and all coalition structures $\Delta \in \mathscr{C}_N$, $\Delta \neq \Gamma$, that i weakly prefers Γ to Δ. It follows that $v_i(\Gamma) \geq v_i(\Delta)$ for all $\Delta \in \mathscr{C}_N$, $\Delta \neq \Gamma$, and $i \in N$. Hence, every agent $i \in N$ has the maximal valuation $v_i(\Gamma) = n \cdot |F_i|$ and is together with all of her friends and none of her enemies. This implies that each coalition in Γ is a connected component and a clique.

The implication from right to left is obvious. □

Since it is easy to check this characterization, perfect coalition structures can be verified in polynomial time for sum-based and min-based SF ACFGs. It follows directly from Theorem 7 that the corresponding existence problem is also in P.

Corollary 1 For any sum-based or min-based SF ACFG $(N,\succeq)$ with an underlying network of friends G, there exists a perfect coalition structure if and only if all connected components of G are cliques.

We further get the following upper bounds.

Proposition 8 *For any ACFG, perfectness verification is in* coNP.

Proof. Consider any ACFG $(N,\succeq)$. A coalition structure $\Gamma \in \mathscr{C}_N$ is not perfect if and only if there is an agent $i \in N$ and a coalition structure $\Delta \in \mathscr{C}_N$ such that $\Delta \succ_i \Gamma$. Hence, we can nondeterministically guess an agent $i \in N$ and a coalition structure $\Delta \in \mathscr{C}_N$ and verify in polynomial time whether $\Delta \succ_i \Gamma$. □

Furthermore, we initiate the characterization of perfectness in ACFGs. The *diameter* of a connected graph component is the greatest distance between any two of its vertices. For sum-based EQ ACFGs, we get the following implication.

Proposition 9 *For any sum-based EQ ACFG with an underlying network of friends G, it holds that if a coalition structure Γ is perfect for it, then Γ consists of the connected components of G and all these components have a diameter of at most two.*

Proof. We first show that, in a perfect coalition structure, all agents have to be together with all their friends. For the sake of contradiction, assume that Γ is perfect but there are $i, j \in N$ with $j \in F_i$ and $j \notin \Gamma(i)$. We distinguish two cases.

Case 1: *All* $f \in F_i \cap \Gamma(i)$ *have a friend in* $\Gamma(j)$. Consider the coalition structure Δ that results from the union of $\Gamma(i)$ and $\Gamma(j)$, i.e., $\Delta = \Gamma \setminus \{\Gamma(i), \Gamma(j)\} \cup \{\Gamma(i) \cup \Gamma(j)\}$. It holds that i and all friends of i's either gain an additional friend in Δ or their coalition stays the same: First, i keeps all friends from $\Gamma(i)$ and gets j as an additional friend. Hence, i has at least one friend more in Δ than in Γ and we have $v_i(\Delta) > v_i(\Gamma)$. Second, all friends $f \in F_i \cap \Gamma(i)$ have a friend in $\Gamma(j)$ and therefore also gain at least one additional friend from the union of the two coalitions. Hence, $v_f(\Delta) > v_f(\Gamma)$ for all $f \in F_i \cap \Gamma(i)$. Third, all friends $f \in F_i \cap \Gamma(j)$ have i as friend. Hence, they also gain one friend from the union. Thus $v_f(\Delta) > v_f(\Gamma)$ for all $f \in F_i \cap \Gamma(j)$. Finally, all $f \in F_i$ who are not in $\Gamma(i)$ or $\Gamma(j)$ value Γ and Δ the same because their coalition is the same in both coalition structures. Hence, $v_f(\Delta) = v_f(\Gamma)$ for all $f \in F_i$ with $f \notin \Gamma(j)$ and $f \notin \Gamma(i)$. Summing up, we have $u_i^{sumEQ}(\Delta) > u_i^{sumEQ}(\Gamma)$, so i prefers Δ to Γ, which is a contradiction to Γ being perfect.

Case 2: *There is an* $f \in F_i \cap \Gamma(i)$ *who has no friends in* $\Gamma(j)$. Consider the coalition structure Δ that results from j moving to $\Gamma(i)$, i.e., $\Delta = \Gamma_{j \to \Gamma(i)}$. Let $k \in F_i \cap \Gamma(i)$ be one of the agents who have no friends in $\Gamma(j)$. Then $v_k(\Delta) = v_k(\Gamma) - 1$; $v_i(\Delta) = v_i(\Gamma) + n$; for all $f \in F_k \cap \Gamma(i)$, $f \neq i$, we have $v_f(\Delta) \geq v_f(\Gamma) - 1$; and for all $f \in F_k, f \notin \Gamma(i)$ (and $f \notin \Gamma(j)$), we have $v_f(\Delta) = v_f(\Gamma)$. Hence,

$$
\begin{aligned}
u_k^{sumEQ}(\Delta) &= \sum_{a \in F_k \cup \{k\}} v_a(\Delta) = \sum_{a \in F_k \cap \Gamma(i), a \neq i} v_a(\Delta) + \sum_{a \in F_k \setminus \Gamma(i)} v_a(\Delta) + v_k(\Delta) + v_i(\Delta) \\
&\geq \sum_{a \in F_k \cap \Gamma(i), a \neq i} v_a(\Gamma) - 1 + \sum_{a \in F_k \setminus \Gamma(i)} v_a(\Gamma) + v_k(\Gamma) - 1 + v_i(\Gamma) + n \\
&= \sum_{a \in F_k \cup \{k\}} v_a(\Gamma) - (|F_k \cap \Gamma(i)| - 1) - 1 + n \\
&= u_k^{sumEQ}(\Gamma) - \underbrace{|F_k \cap \Gamma(i)|}_{<n} + n > u_k^{sumEQ}(\Gamma).
\end{aligned}
$$

Therefore, k prefers Δ to Γ, which again is a contradiction to Γ being perfect.

Next, assume that Γ is perfect but there is a coalition C in Γ that has a diameter greater than two. Then there are agents $i, j \in C$ with a distance greater than two. Thus j is an enemy of i's and an enemy of all of i's friends. It follows that i prefers coalition structure $\Gamma_{j \to \emptyset}$ to Γ, which is a contradiction to Γ being perfect.

Summing up, in a perfect coalition structure Γ for a sum-based EQ ACFG every agent is together with all her friends and every coalition in Γ has a diameter of at most two. Together this implies that Γ consists of the connected components of G and all these components have a diameter of at most two. □

From Propositions 8 and 9, we get the following corollary.

Corollary 2 For sum-based EQ ACFGs, perfectness existence is in coNP.

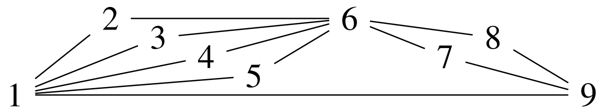


Fig. 10: Network of friends for Example 5

However, Proposition 9 is not an equivalence. The converse does not hold, as the following example shows.

Example 5 Consider the sum-based EQ ACFG $(N, \succeq^{sumEQ})$ with the network of friends G in Figure 10. The coalition structure $\Gamma = \{N\}$ consists of the only connected component of G, which has a diameter of two. However, agent 1 prefers $\Delta = \{\{1, \dots, 6\}, \{7, 8, 9\}\}$ to Γ because

$$\begin{aligned} u_1^{sumEQ}(\Gamma) &= v_1(\Gamma) + \dots + v_5(\Gamma) + v_9(\Gamma) = (9 \cdot 5 - 3) + 4 \cdot (9 \cdot 2 - 6) + (9 \cdot 3 - 5) = 112 \\ &< 113 = (9 \cdot 4 - 1) + 4 \cdot (9 \cdot 2 - 3) + (9 \cdot 2 - 0) = v_1(\Delta) + \dots + v_5(\Delta) + v_9(\Delta) \\ &= u_1^{sumEQ}(\Delta). \end{aligned}$$

Hence, Γ is not perfect.

5 Conclusions and Open Problems

We have proposed to extend the models of altruistic hedonic games due to Nguyen *et al.* [1] and Wiechers and Rothe [5] to coalition formation games in general. We have compared our more general models to altruism in hedonic games and have motivated our work by removing some crucial disadvantages that come with the restriction to hedonic games. In particular, we have shown that all degrees of our general altruistic preferences are unanimous while this is not the case for all altruistic hedonic preferences. Furthermore, all our sum-based degrees of altruism fulfill two types of monotonicity that are violated by the corresponding hedonic equal- and altruistic-treatment preferences.

We have furthermore studied the common stability notions and have initiated a computational analysis of the associated verification and existence problems (see Table 2 for an overview of our results). We also gave characterizations for some of the stability notions, using graph-theoretical properties of the underlying network of friends. For future work, we propose to complete this analysis, close all gaps between complexity-theoretic upper and lower bounds, and get a full characterization for all stability notions.

Acknowledgments. We thank the anonymous IJCAI'20 reviewers for helpful comments. This work was supported in part by DFG grants RO 1202/14-2 and RO 1202/21-1. The first and third author have been supported in part by the research project "Online Participation" within the North Rhine-Westphalian funding scheme "Forschungskollegs."

Declarations: Conflict of Interest/Competing Interest Statement

Funding: This work was supported in part by Deutsche Forschungsgemeinschaft under grants RO 1202/14-2 and RO 1202/21-1 and the NRW project "Online Participation."

Non-financial interests: Author Jörg Rothe is or has been on the following editorial boards of scientific journals:

- *Annals of Mathematics and Artificial Intelligence* (AMAI), Associate Editor, since 01/2020,
- *Journal of Artificial Intelligence Research* (JAIR), Associate Editor, since 09/2017,
- *Journal of Universal Computer Science* (J.UCS), Editorial Board, since 01/2005,
- *Mathematical Logic Quarterly* (MLQ – Wiley), Editorial Board, 01/2008–12/2019, and
- *MDPI Algorithms*, Editorial Board, since 04/2021.

References

[1] Nguyen, N., Rey, A., Rey, L., Rothe, J., Schend, L.: Altruistic hedonic games. In: Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems, pp. 251–259. IFAAMAS, Singapore (2016)

[2] Dimitrov, D., Borm, P., Hendrickx, R., Sung, S.-C.: Simple priorities and core stability in hedonic games. Social Choice and Welfare **26**(2), 421–433 (2006)

[3] Aziz, H., Brandt, F., Seedig, H.G.: Computing desirable partitions in additively separable hedonic games. Artificial Intelligence **195**, 316–334 (2013)

[4] Aziz, H., Savani, R.: Hedonic games. In: Brandt, F., Conitzer, V., Endriss, U., Lang, J., Procaccia, A. (eds.) Handbook of Computational Social Choice, pp. 356–376. Cambridge University Press, Cambridge (2016). Chap. 15

[5] Wiechers, A., Rothe, J.: Stability in minimization-based altruistic hedonic games. In: Proceedings of the 9th European Starting AI Researchers' Symposium, vol. 2655, p. 3. CEUR-WS.org, Santiago Compostela (2020)

[6] Schlueter, J., Goldsmith, J.: Super altruistic hedonic games. In: Proceedings of the 33rd International Florida Artificial Intelligence Research Society Conference, pp. 160–165. AAAI Press, Miami (2020)

[7] Bullinger, M., Kober, S.: Loyalty in cardinal hedonic games. In: Proceedings of the 30th International Joint Conference on Artificial Intelligence, pp. 66–72. ijcai.org, Montreal (2021)

[8] Ashlagi, I., Krysta, P., Tennenholtz, M.: Social context games. In: Proceedings

of the 4th International Workshop on Internet & Network Economics, pp. 675–683. Springer, Shanghai (2008)

[9] Bilò, V., Celi, A., Flammini, M., Gallotti, V.: Social context congestion games. Theoretical Computer Science **514**, 21–35 (2013)

[10] Hoefer, M., Penn, M., Polukarova, M., Skopalik, A., Vöcking, B.: Considerate equilibrium. In: Proceedings of the 22nd International Joint Conference on Artificial Intelligence, pp. 234–239. AAAI Press/IJCAI, Barcelona (2011)

[11] Anagnostopoulos, A., Becchetti, L., B. de Keijzer, Schäfer, G.: Inefficiency of games with social context. In: Proceedings of the 6th International Symposium on Algorithmic Game Theory, pp. 219–230. Springer, Aachen (2013)

[12] Hoefer, M., Skopalik, A.: Altruism in atomic congestion games. In: Proceedings of the 17th Annual European Symposium on Algorithms, pp. 179–189. Springer, Copenhagen (2009)

[13] Chen, P., Keijzer, B.d., Kempe, D., Schäfer, G.: The robust price of anarchy of altruistic games. In: Proceedings of the 7th International Workshop on Internet & Network Economics, pp. 383–390. Springer, Singapore (2011)

[14] Apt, K., Schäfer, G.: Selfishness level of strategic games. Journal of Artificial Intelligence Research **49**, 207–240 (2014)

[15] Rahn, M., Schäfer, G.: Bounding the inefficiency of altruism through social contribution games. In: Proceedings of the 9th International Workshop on Internet & Network Economics, pp. 391–404. Springer, Cambridge (2013)

[16] Kerkmann, A., Rothe, J.: Altruism in coalition formation games. In: Proceedings of the 29th International Joint Conference on Artificial Intelligence, pp. 347–353. ijcai.org, Yokohama (2020)

[17] Bell, E.T.: The iterated exponential integers. Annals of Mathematics, 539–557 (1938)

[18] Rota, G.-C.: The number of partitions of a set. The American Mathematical Monthly **71**(5), 498–504 (1964)

[19] Sung, S.-C., Dimitrov, D.: Computational complexity in additive hedonic games. European Journal of Operational Research **203**(3), 635–639 (2010)

[20] Woeginger, G.: A hardness result for core stability in additive hedonic games. Mathematical Social Sciences **65**(2), 101–104 (2013)

[21] Cechlárová, K., Romero-Medina, A.: Stability in coalition formation games. International Journal of Game Theory **29**(4), 487–494 (2001)

[22] Cechlárová, K., Hajduková, J.: Computational complexity of stable partitions with B-preferences. International Journal of Game Theory **31**(3), 353–364 (2003)

[23] Kerkmann, A.M., Lang, J., Rey, A., Rothe, J., Schadrack, H., Schend, L.: Hedonic games with ordinal preferences and thresholds. Journal of Artificial Intelligence Research **67**, 705–756 (2020)

[24] Rothe, J., Schadrack, H., Schend, L.: Borda-induced hedonic games with friends, enemies, and neutral players. Mathematical Social Sciences **96**, 21–36 (2018)

[25] Garey, M., Johnson, D.: Computers and Intractability: A Guide to the Theory of NP-Completeness. W. H. Freeman and Company, New York, USA (1979)

[26] Rothe, J.: Complexity Theory and Cryptology. An Introduction to Cryptocomplexity. EATCS Texts in Theoretical Computer Science. Springer, Heidelberg and Berlin, Germany (2005)

[27] Gonzalez, T.: Clustering to minimize the maximum intercluster distance. Theoretical Computer Science **38**, 293–306 (1985)

Chapter 4

Local Fairness in Hedonic Games via Individual Threshold Coalitions

This chapter summarizes the following journal article in which we introduce and study three local fairness notions for hedonic games:

Publication (Kerkmann et al. [93])

> A. Kerkmann, N. Nguyen, and J. Rothe. "Local Fairness in Hedonic Games via Individual Threshold Coalitions". In: *Theoretical Computer Science* 877 (2021), pp. 1–17

4.1 Summary

In this work, we introduce and study three notions of *local fairness* in hedonic games. Previous literature by Bogomolnaia and Jackson [21], Aziz et al. [10], Wright and Vorobeychik [148], and Peters [114, 115] considers *envy-freeness* as a notion of fairness in hedonic games. However, this notion requires agents to inspect other coalitions than their own. In contrast to this notion, our *local* fairness notions can be decided solely based on the agents' own coalitions and their individual preferences.

We define the three local fairness notions *min-max fairness*, *grand-coalition fairness*, and *max-min fairness* based on three different threshold coalitions. For each agent, these thresholds are solely defined on her individual preference. Moreover, a coalition structure is fair for an agent if she weakly prefers her coalition in this coalition structure to her threshold coalition.

After introducing the three local fairness notions, we show that they form a strict hierarchy: max-min fairness implies grand-coalition fairness which in turn implies min-max fairness. We also relate the three notions to other stability notions that are known from the literature, such as Nash stability, core stability, envy-freeness by replacement, and individual rationality. We then study the problem of computing the fairness thresholds and determine the complexity of this problem in the context of additively separable hedonic games. We also determine subclasses of hedonic games where fair coalition structures are guaranteed to exist. Since this does not hold for general additively separable hedonic games, we also ask for the complexity of determining whether a fair coalition structure exists in a given additively separable hedonic game.

Afterwards, we study the minimum and maximum price of local fairness which describe the best-case and worst-case loss of social welfare of a coalition structure that satisfies fairness compared to the coalition structure with maximum utilitarian social welfare. In doing so, we concentrate on min-max fairness which is the weakest of our three local fairness notions and constrains the set of possible coalition structures less than the other two notions. For symmetric additively separable hedonic games, we show that the maximum price of min-max fairness is not bounded by a constant but the minimum price of min-max fairness is always one.

Finally, we discuss an alternative fairness notion and argue that there is no local fairness notion stronger than individual rationality such that fair coalition structures exist for every hedonic game.

4.2 Personal Contribution and Preceding Versions

This journal publication extends a preliminary conference version by Nhan-Tam Nguyen and Jörg Rothe [105] that was also presented at CoopMAS'16 [106]. Contents that I contributed to our work are additional writing and improved presentation throughout the paper (e.g., the reordering of definitions in Section 2, the improvement of Figure 1, the reorganization of Sections 3.1 and 3.3, additional Footnotes 3, 4, 5, and 6, the extension of Definition 8, and the remark after Corollary 6), additional related work in Section 1.3, the examples and explanations in Propositions 1 and 2, the first example and explanation in Proposition 3, Section 4.1 (where a preliminary version of Theorem 4 was contained in [105]), the first part of Theorem 6 that shows the membership of Min-Max-Exist in NP, Observation 6, and the argumentation for Proposition 4.

4.3 Publication

The full article [93] is appended here.

Theoretical Computer Science 877 (2021) 1–17

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Local fairness in hedonic games via individual threshold coalitions ☆

Anna Maria Kerkmann, Nhan-Tam Nguyen, Jörg Rothe *

Institut für Informatik, Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany

A R T I C L E   I N F O

Article history:
Received 21 September 2020
Received in revised form 27 January 2021
Accepted 24 March 2021
Available online 26 March 2021

Keywords:
Coalition formation
Hedonic game
Fairness
Price of local fairness
Cooperative game theory
Computational complexity

A B S T R A C T

Hedonic games are coalition formation games where players only specify preferences over coalitions they are part of. We introduce and systematically study three local fairness notions in hedonic games called max-min fairness, grand-coalition fairness, and min-max fairness. To this end, we define suitable threshold coalitions for these three concepts. A coalition structure (i.e., a partition of the players into coalitions) is considered locally fair if all players' coalitions in this structure are each at least as good as their threshold coalitions. Based on this approach, we then introduce three specific notions of local fairness by suitably adapting fairness notions from fair division. We show that they form a proper hierarchy and how they are related to previously studied solution concepts in hedonic games. We also study the computational aspects of finding threshold coalitions and of deciding whether fair coalition structures exist in additively separable hedonic games, and we investigate the related price of local fairness.

1. Introduction

Coalition formation plays a crucial role in multiagent systems when agents have to cooperate. A commonly studied model of coalition formation is the model of hedonic games. These are coalition formation games with nontransferable utility, which were first introduced by Drèze and Greenberg [34] and studied later on by Banerjee et al. [9], Bogomolnaia and Jackson [21], and others (see, e.g., the book chapter by Aziz and Savani [7] for an overview). A key feature of hedonic games is that the players' preferences depend only on coalitions they are part of. Since players specify their preferences over an exponential-size domain (in the number of players), various compact representations have been proposed, which either are fully expressive but may still have an exponential size in the worst case or are always succinct but may restrict the preference domain one way or the other [8,37,21,33,4,49,61,29]. Most of these studies are concerned with stability issues. Intuitively, these capture incentives of (groups of) players to deviate from their current coalition and joining a different coalition so as to increase their individual utility values. Thus stability-related questions address a decentralized aspect of hedonic games.

A more recent approach to hedonic games is social welfare maximization [24,5,6]. This idea is different because social welfare maximization usually presupposes a central authority guiding the maximization process by eliciting preferences and suggesting or enforcing an optimal solution. This enforcement may be necessary because the optimality of a solution is

☆ A preliminary version appeared in the proceedings of the *15th International Conference on Autonomous Agents and Multiagent Systems* (AAMAS 2016, [54]) and was also presented at the *7th Workshop on Cooperative Games in Multiagent Systems* (CoopMAS 2016).

* Corresponding author.
E-mail addresses: anna.kerkmann@hhu.de (A.M. Kerkmann), nhan-tam.nguyen@hhu.de (N.-T. Nguyen), rothe@hhu.de (J. Rothe).

https://doi.org/10.1016/j.tcs.2021.03.027
0304-3975/

determined by a global criterion, such as utilitarian or egalitarian social welfare, and may affect some players' utility values negatively compared with the status quo.

1.1. Motivation

We will focus on the concept of local fairness. Fairness is an important aspect besides stability and social welfare maximization (see the related work section and, e.g., the work of de Jong et al. [32] for a discussion of fairness in multiagent systems and the book chapters by Bouveret et al. [22] and Lang and Rothe [50] for fair division of indivisible goods). Fairness is related to both centralized and decentralized approaches. On the one hand, the center may want to ensure a certain utility level for each player. This goal can be achieved by a global fairness condition. However, fairness does not per se presuppose the existence of a center. On the other hand, players may not consider their current coalition fair, given their individual preferences. While we agree with Bogomolnaia and Jackson [21] that stability has a "'restricted fairness' flavor," we add that one can also take the complementary view that lack of fairness can be a major cause of instability.

To make this more concrete, consider a situation where all players except a single player in some coalition consider this coalition their favorite one, yet for that single player this coalition is actually only marginally better than being alone. However, because everyone else prefers this coalition and thus is much better off than that player, she rejects this coalition. This can be considered an unfair situation and is comparable to an ultimatum game situation where the first player's (the proposer's) proposal is very imbalanced and so the second player (responder) rejects the proposal because it is below her fair share (see, again, [32]). Note that we would have to contrast the single player's utility to the other players' utility values in order to explain the predicament. This approach of balanced utility values and inequality reduction either requires a center that knows all players' utility values or it requires that players look at coalitions (and even other players' well-being) outside of their own.[1] To some extent, however, this is at odds with the idea of hedonic game because players in such games should only be interested in their own coalition.

Works considering fairness in hedonic games are due to Bogomolnaia and Jackson [21], Aziz et al. [5], Wright and Vorobeychik [65], and Peters [56,57]. They consider envy-freeness as a notion of fairness. This traditional fairness notion says that a partition of the agents is fair if no agent envies another agent for her coalition. However, this notion requires players to inspect other coalitions. If there is a large number of coalitions, that is something we would like to avoid. Therefore, we propose and study notions of *local* fairness—restricted fairness notions with the additional constraint that players only compare their current coalition to some bound that *solely* depends on their individual preferences.[2] We feel that this is in the general spirit of the decentralized aspect of hedonic games.

1.2. Contribution

In order to achieve this goal of specifying local fairness criteria for hedonic games, we introduce notions that are inspired by ideas from the field of fair division of indivisible goods. Our main contributions are the following:

1. We introduce the idea of local fairness and three specific such fairness notions in hedonic games. We show that these concepts form a (strict) hierarchy, and we relate them to previously studied concepts. Perhaps somewhat surprisingly, the hierarchy strikingly differs from the corresponding scale proposed by Bouveret and Lemaître [23] in the context of fair division of indivisible goods.
2. We systematically study the complexity of finding "threshold coalitions" and of determining whether a fair coalition structure exists in a given additively separable hedonic game. We also find that two of our notions coincide in such games.
3. We initiate the study of the price of local fairness in hedonic games.

A preliminary version of this paper was presented at the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'16) [54]. Since then we have completed the computational complexity analysis in Section 4 by strengthening some complexity results (Theorems 5, 6, and 7), added a new section about the relations among the fairness and stability notions in additively separable hedonic games (Section 4.1), and added a new section showing that there is no local fairness notion, stronger than individual rationality, such that there always exists a coalition structure satisfying this fairness notion (Section 6).

[1] On a related note, in *altruistic hedonic games* introduced by Nguyen et al. [53], players care not only about their own utility when deciding which coalition to join but also about their friends' utilities in these coalitions. Wiechers and Rothe [64] study *minimization-based altruistic hedonic games* in which an agent's utility for a coalition depends on her own utility and that of a friend in the same coalition that is worst off. An even broader approach is taken by Kerkmann and Rothe [47] who introduce *altruistic coalition formation games*, taking into account the utilities of *all* friends, not only of those in the same coalition. Rothe [60] surveys various approaches to altruism in game theory including the papers mentioned above.

[2] In case of envy-freeness, by contrast, the bound would also depend on the whole coalition structure.

1.3. Related work

Surveys and book chapters on hedonic games are, for example, due to Aziz and Savani [7], Elkind and Rothe [36], and Hajdukova [43]. Bogomolnaia and Jackson [21] already mention envy-freeness in their work, but they focus on studying stability notions. Aziz et al. [5] study the complexity of determining the existence of stable coalition structures in additively separable hedonic games (first introduced by Bogomolnaia and Jackson [21]). They also consider the social welfare maximization approach and the notion of envy-freeness.

The work by Wright and Vorobeychik [65] is somewhat related to ours. They study hedonic games under the perspective of mechanism design and propose mechanisms for solving the team formation problem. A key difference is that they consider additively separable hedonic games with nonnegative values only. Since in this case the grand coalition is most preferred by every player, Wright and Vorobeychik introduce cardinality constraints on feasible coalition sizes. They also consider *envy bounded by a single teammate*, which for the aforementioned reasons is not suitable for our goals. In addition, they introduce the *maximin share guarantee for team formation*, which is based on the idea of replacing players. This, however, leads to a provably different notion than ours (see Theorem 4). More recently, Ueda [63] introduced the notion of *justified envy-freeness*, which is a weakening of envy-freeness. *Weakly justified envy-freeness*, due to Barrot and Yokoo [11], in fact is a stronger version of justified envy-freeness and is thus situated between envy-freeness and justified envy-freeness. They study the implications between (weakly) justified envy-freeness and several stability notions. Furthermore, they consider the conjunction of justified envy-freeness with various stability notions and study the existence of coalition structures that satisfy these conjunctions.

Recently introduced stability notions include *strong Nash stability*, proposed by Karakaya [46], and *strictly strong Nash stability*, due to Aziz and Brandl [3]. Brânzei and Larson [24] study *social welfare maximization* and *core stability* in additively separable hedonic games. Moreover, they consider the so-called stability gap. Bilò et al. [18,19] study similar notions in fractional hedonic games. *Pareto optimality* can be considered a notion of stability [51] as well. Elkind et al. [35] investigate the *price of Pareto optimality* in various representations of hedonic games. Peters and Elkind [58] give conditions on when certain classes of hedonic games admit intractable problems regarding the existence of stable coalition structures. Peters [57] considers restrictions of hedonic games that admit fast algorithms and he models allocating indivisible goods as a hedonic game.

Fair division of indivisible goods and hedonic games are closely related because both fields deal with partitions of sets. In fair division, a set of goods needs to be partitioned into n subsets, where n is the number of agents. Usually, it is assumed that goods cannot be shared. This is a departure point from hedonic games because the number of coalitions in a partition is only bounded above by n. The no-externality assumption, however, is prevalent (or even defining) in both fields: This assumption requires the agents' utilities to depend only on the subset of goods that they receive (fair division) or the coalition that they are part of (hedonic games).

Surveys and book chapters on fair division are due, for example, to Chevaleyre et al. [30], Nguyen et al. [55], Lang and Rothe [50], and Bouveret et al. [22]. Three main fairness criteria considered in the context of fair division are *equitability*, *proportionality*, and *envy-freeness*. Gourvès et al. [42] further introduce *jealousy-freeness*, which is a relaxation of equitability. Envy-freeness, which was first introduced by Foley [38] in resource allocation, is also considered in hedonic games (see Section 1.1). A relaxation of this notion, *envy-freeness up to one good*, was introduced by Budish [26] and further studied by, e.g., Caragiannis et al. [28], Bilò et al. [17], Benabbou et al. [14], and Bérczi et al. [15]. Informally speaking, an allocation of goods to agents is *envy-free up to one good* if each agent's envy towards any other agent can be eliminated by removing one single good from the envied agent's bundle.

Some authors consider fair division with an underlying network of neighborhood among the agents. For instance, Abebe et al. [1] say an allocation is *locally envy-free*[3] if no agents envy any of their neighbors. Here, the expression "local" has another meaning than in our notion of *local fairness*: "Local envy-freeness" refers to envy occurring only *locally* in the agents' neighborhood, whereas in "local fairness" agents can decide about fairness *locally* by only considering their own coalition and their own preferences.

Inspired by the cut-and-choose protocol from cake cutting, Budish [26] furthermore introduced the *max-min fair share* criterion. Since then, it has been studied by, e.g., Procaccia and Wang [59], Kurokawa et al. [48], Amanatidis et al. [2], Heinen et al. [44,45], Nguyen et al. [52], Barman and Murthy [10], Caragiannis et al. [28], Gourvès and Monnot [41], Suksompong [62], Biswas and Barman [20], and Bei et al. [12]. Bouveret and Lemaître [23] introduce *min-max fair share* and propose a scale of even more demanding fairness criteria. Bertsimas et al. [16] and Caragiannis et al. [27] study the *price of fairness* in fair division.

1.4. Organization of the paper

In Section 2, we formally define hedonic games and relevant notions of stability. In Section 3, we introduce our notions of local fairness and relate them to other stability, fairness, and optimality concepts. In Section 4, we study our notions in additively separable hedonic games under computational aspects. The price of local fairness is considered in Section 5 and

[3] This notion is also called *graph envy-freeness* [31,25] or *envy-freeness on networks* [13].

an alternative fairness notion is briefly discussed in Section 6, followed by a discussion of our findings and the conclusions in Section 7.

2. Preliminaries

We denote by $N = \{1, \dots, n\}$ the set of *players* (or *agents*). A *coalition* is a subset of N and a *coalition structure* π is a partition of N. The set of all coalition structures over N is $\Pi(N)$. We denote by $\pi(i)$ the unique coalition with player i in coalition structure π and by $\mathscr{N}_i = \{C \subseteq N \mid i \in C\}$ all coalitions that player i is a member of. Every player i has a weak and complete preference order $\succeq_i$ over $\mathscr{N}_i$. For $A, B \in \mathscr{N}_i$, we write $A \succeq_i B$ if player i *weakly prefers* coalition A to B; we write $A \succ_i B$ if player i *(strictly) prefers* coalition A to B (i.e., $A \succeq_i B$ but not $B \succeq_i A$); and we write $A \sim_i B$ if i is indifferent between A and B (i.e., $A \succeq_i B$ and $B \succeq_i A$). We denote by $\mathrm{top}(\succeq_i)$ player i's most preferred coalition (with arbitrary tie-breaking). Let $\succeq$ be the collection of all $\succeq_i$, $i \in N$. A *hedonic game* is a pair $(N, \succeq)$. It is an *additively separable hedonic game* (ASHG) if for every $i \in N$, there is a valuation function $v_i : N \to \mathbb{Q}$ such that $\sum_{j \in A} v_i(j) \geq \sum_{j \in B} v_i(j) \iff A \succeq_i B$. We write (N, v) for an additively separable hedonic game, where v is the collection of all v_i, $i \in N$. We assume normalization of the valuation functions, that is, $v_i(i) = 0$. We overload the notation to mean $v_i(A) = \sum_{j \in A} v_i(j)$ for each coalition $A \in \mathscr{N}_i$.

Now we define previously studied notions of stability that are relevant for this work. Stability in hedonic games refers to players not having an incentive to deviate from their current coalitions in a given coalition structure. We distinguish, as is common, between group deviations, individual deviations, and other notions.

We consider the following notions of group deviations.

Definition 1. Let $(N, \succeq)$ be a hedonic game.

1. A nonempty coalition $C \subseteq N$ *blocks* a coalition structure π if $C \succ_i \pi(i)$ for every $i \in C$. A coalition structure π is *core stable* (CS) if no coalition blocks π.
2. A coalition $C \subseteq N$ *weakly blocks* a coalition structure π if $C \succeq_i \pi(i)$ for every $i \in C$ and there is some $j \in C$ with $C \succ_j \pi(j)$. A coalition structure π is *strictly core stable* (SCS) if no coalition weakly blocks π.
3. Given a coalition $H \subseteq N$, coalition structure π' is *reachable from coalition structure $\pi \neq \pi'$ by coalition H* if for all $i, j \in N \setminus H$, we have $\pi(i) = \pi(j) \iff \pi'(i) = \pi'(j)$. A nonempty coalition $H \subseteq N$ *weakly Nash-blocks* coalition structure π if there exists some coalition structure π' that is reachable from π by coalition H such that $\pi'(i) \succeq_i \pi(i)$ for every $i \in H$ and there is some $j \in H$ with $\pi'(j) \succ_j \pi(j)$. We say π is *strictly strong Nash stable* (SSNS) if there is no coalition that weakly Nash-blocks π.
4. A coalition structure π' *Pareto-dominates* coalition structure π if $\pi'(i) \succeq_i \pi(i)$ for every $i \in N$ and there is some $j \in N$ with $\pi'(j) \succ_j \pi(j)$. A coalition structure π is *Pareto-optimal* (PO) if no coalition structure Pareto-dominates it.

Note that a coalition structure π is Pareto-optimal exactly if N does not weakly Nash-block π.

As to individual deviations, we need the following stability notions.

Definition 2. Let $(N, \succeq)$ be a hedonic game.

1. A coalition $C \in \mathscr{N}_i$ is *acceptable for $i \in N$* if $C \succeq_i \{i\}$. A coalition structure π is *individually rational* (IR) if $\pi(i)$ is acceptable for every $i \in N$.
2. A coalition structure π is *Nash stable* (NS) if no player would like to deviate to another coalition, i.e., if $\pi(i) \succeq_i C \cup \{i\}$ for every $i \in N$ and $C \in \pi \cup \{\emptyset\}$.
3. A coalition structure π is *contractually individually stable* (CIS) if for every $i \in N$, the existence of a coalition $C \in \pi \cup \{\emptyset\}$ with $C \cup \{i\} \succ_i \pi(i)$ implies that there exists some $j \in C$ such that $C \succ_j C \cup \{i\}$ or there exists some $k \in \pi(i)$ such that $\pi(k) \succ_k \pi(k) \setminus \{i\}$.

Of the remaining notions we need the following.

Definition 3. Let $(N, \succeq)$ be a hedonic game.

1. A coalition structure π is *perfect* (PF) if every player is in one of her most preferred coalitions, i.e., if $\pi(i) \succeq_i C$ for every $i \in N$ and $C \in \mathscr{N}_i$.
2. A coalition structure π is *envy-free by replacement* (EF-R) if no player i would prefer to replace another agent j, i.e., if $\pi(i) \succeq_i (\pi(j) \setminus \{j\}) \cup \{i\}$ for every $i, j \in N$, $i \neq j$.

We furthermore define two notions that are based on the maximization of social welfare and are defined for ASHGs only.

Definition 4. Let (N, v) be an ASHG. A coalition structure $\pi \in \Pi(N)$ *maximizes*

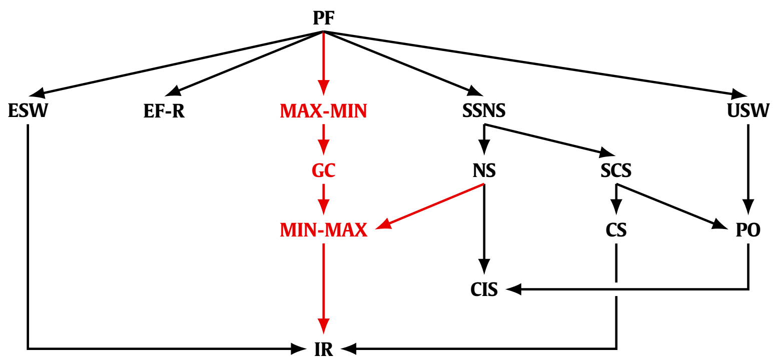


Fig. 1. Relations among stability and other notions defined for hedonic games. An arrow from notion A to notion B means that every coalition structure that is A is also B. For example, every strictly core stable (SCS) coalition structure is core stable (CS) and Pareto-optimal (PO). If notion B is not reachable from notion A by any directed path then A does not imply B. The three local fairness notions that we introduce and study in this paper and their relations among each other and to the previously known notions are colored in red. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

1. *utilitarian social welfare* (USW) if for every $\pi' \in \Pi(N)$, $\sum_{i \in N} v_i(\pi(i)) \geq \sum_{i \in N} v_i(\pi'(i))$;
2. *egalitarian social welfare* (ESW) if for every $\pi' \in \Pi(N)$, $\min_{i \in N} v_i(\pi(i)) \geq \min_{i \in N} v_i(\pi'(i))$.

For the notions from Definition 4, we make the common assumption (see, for example, [5,24]) in coalition formation that values are interpersonally comparable.[4]

We also say that a coalition structure π *satisfies some notion* X if π is X or maximizes X.

Fig. 1 shows the relationships between these notions (colored in black) and the three local fairness notions (colored in red) that we will introduce in the next section. For more explanations of the above definitions and their interrelations we refer to the surveys, book chapters, and papers mentioned in the related work section. The notions are chosen such that our separation results in the next section also apply to intermediate notions such as *contractual strict core stability* and *individual stability*.

We assume the reader to be familiar with the fundamental notions of computational complexity theory, including the complexity classes P (deterministic polynomial time) and NP (nondeterministic polynomial time), the notion of polynomial-time many-one reducibility, and the notions of NP-hardness and NP-completeness based on this reducibility. In addition, we will also consider NP-hardness and -completeness in the "strong sense" [39]. Note that, unless P = NP, strongly NP-hard problems do not even have a fully polynomial-time approximation scheme and they cannot have pseudo-polynomial-time algorithms.

3. Local fairness in hedonic games

Formally, a local fairness notion is a function f that maps a preference order $\succeq_i$ to a coalition $f(\succeq_i) \in \mathscr{N}_i$. A coalition structure π is f*-fair* if

$$\pi(i) \succeq_i f(\succeq_i)$$

for every $i \in N$, i.e., if all agents are at least as happy with their current coalition in π as with the coalition specified for them by f.

We now introduce our specific fairness criteria, starting with the weakest one. IR is the most basic notion of stability. It is also the weakest fairness criterion. Similarly to the example in the introduction, a player who is in a coalition not acceptable to her is exploited by the other players in that coalition if this coalition is acceptable to them. In other words, this player is forced to be in a disliked coalition just for other players to benefit. In this case, a coalition structure consisting of singletons is more preferable for this player. Note that IR ($f(\succeq_i) = \{i\}$) and PF ($f(\succeq_i) = \mathrm{top}(\succeq_i)$) are examples of local fairness criteria that only propose a threshold coalition a player has to be a member of. They are two opposing extreme cases as IR demands a threshold that can be situated rather low in the agents' preference orders and PF demands the highest threshold possible (up to tie-breaking among equally liked top coalitions). In a sense, we look for criteria situated between these two notions. Because all fairness criteria have to satisfy IR necessarily,[5] we consider such fairness criteria only.

[4] Informally, this means that agents that draw the same happiness from another agent also assign the same valuation to that agent. Or, if an agent draws twice as much happiness from an agent as another agent, this agent's valuation is twice the valuation of the other agent, and so on. In this sense, a certain valuation always represents the same amount of happiness independent of the considered agent.

[5] In particular, if we were to allow local fairness notions based on threshold coalitions ranked strictly below $\{i\}$ in $\succeq_i$, we would still consider it "fair" for agent i to be forced to be in a coalition that i likes less than being alone. Clearly, that would be counterintuitive, as in any reasonable notion of fairness agents should be allowed to leave coalitions disliked to such an extent.

3.1. Min-Max fairness

Before we formally define the min-max threshold, we illustrate it with the following situation: A player is arriving late and all other players have already formed a coalition structure without her (where the specific form of the coalition structure is irrelevant for this argument). Because the player could not participate in the coalition formation process, the player is allowed to join any coalition. Clearly, this player joins her most preferred coalition. This describes a fairness criterion because someone who was neglected should be allowed to adapt to the situation in the best possible way.

Definition 5. The *min-max threshold of player* $i \in N$ is defined as

$$\mathrm{MinMax}_i = \min_{\pi \in \Pi(N \setminus \{i\})} \max_{C \in \pi \cup \{\emptyset\}} C \cup \{i\},$$

where minimization and maximization are with respect to $\succeq_i$. A coalition structure π satisfies *min-max fairness* (MIN-MAX) if

$$\pi(i) \succeq_i \mathrm{MinMax}_i$$

for every $i \in N$.

This notion is the hedonic-games analogue of min-max fair share, originally proposed by Bouveret and Lemaître [23] in fair division. We relate min-max fairness to previously known notions of stability in hedonic games. By definition, min-max fairness satisfies IR. Clearly, since USW does not imply IR,[6] it cannot satisfy min-max fairness. Similarly, EF-R, PO, and CIS cannot imply min-max fairness.

Observation 1.

1. *Every min-max fair coalition structure is IR.*
2. *An USW, EF-R, PO, or CIS coalition structure does not necessarily satisfy min-max fairness.*

Later (in Section 3.3 on max-min fairness) we will see that min-max fairness is independent of most stability notions in the sense that it does not imply them. Now, we check which stability notions except for PF imply min-max fairness.

Theorem 1. *Every NS coalition structure satisfies min-max fairness.*

Proof. Let π be a NS coalition structure and $i \in N$. Then $\pi(i) \succeq_i C \cup \{i\}$ for every $C \in \pi \cup \{\emptyset\}$. Since MinMax_i is a best coalition in a *worst* coalition structure for i, $\pi(i) \succeq_i \mathrm{MinMax}_i$. □

Since SSNS and PF imply NS, it follows from Theorem 1 that these two notions also imply min-max fairness. However, we will show that no other of the above defined stability or welfare optimization notions implies min-max fairness, starting with SCS and ESW.

Proposition 1. *A coalition structure that is SCS or maximizes ESW does not necessarily satisfy min-max fairness.*

Proof. Consider the following hedonic game:

$$\{1,3\} \succ_1 \{1,2,3\} \succ_1 \{1\} \succ_1 \{1,2\},$$
$$\{2,3\} \succ_2 \{2\} \succ_2 \{1,2,3\} \succ_2 \{1,2\},$$
$$\{1,2,3\} \succ_3 \{2,3\} \succ_3 \{1,3\} \succ_3 \{3\}.$$

It is additively separable as it can also be represented via the following values $v_i(j)$:

$i \backslash j$	1	2	3
1	0	−1	2
2	−2	0	1
3	1	2	0

[6] See Figure 1 in the paper by Aziz et al. [5, p. 319] where what we call USW is referred to as MaxUtil.

We compute each player's min-max threshold coalition. Player 1 is faced with coalition structures $\pi = \{\{2\}, \{3\}\}$ and $\pi' = \{\{2, 3\}\}$ over $N \setminus \{1\} = \{2, 3\}$. The best acceptable coalition for player 1 to join with respect to π is $\{1, 3\}$, and with respect to π' is $\{1, 2, 3\}$. Thus $\mathrm{MinMax}_1 = \{1, 2, 3\}$. Analogously, $\mathrm{MinMax}_2 = \{2\}$ and $\mathrm{MinMax}_3 = \{2, 3\}$. Therefore, in every min-max fair coalition structure, player 1 has to be in coalition $\{1, 3\}$ or $\{1, 2, 3\}$, player 2 in $\{2, 3\}$ or $\{2\}$, and player 3 in $\{1, 2, 3\}$ or $\{2, 3\}$. Hence, there is no min-max fair coalition structure. However, coalition structure $\pi'' = \{\{1\}, \{2, 3\}\}$ is SCS and maximizes ESW.

For the sake of contradiction, assume π'' were not SCS. Then there is a coalition $C \in \{\{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{1, 2, 3\}\}$ that weakly blocks π''. First observe that player 2 cannot be part of C because 2 is already in her unique most preferred coalition. Hence, $C \in \{\{3\}, \{1, 3\}\}$. However, agent 3 prefers her current coalition to both $\{3\}$ and $\{1, 3\}$, which is a contradiction to the assumption that there is a weakly blocking coalition.

Now assume that π'' does not maximize ESW. Then there is another coalition structure $\pi''' \in \Pi(N)$ with an ESW greater than $\min_{i\in\{1,2,3\}} v_i(\pi''(i)) = \min\{0, 1, 2\} = 0$. However, the ESW of π''' can only be positive if agent 2 is in $\{2, 3\}$, as 2 has a nonpositive valuation for all other coalitions. This is a contradiction, as it implies $\pi''' = \{\{1\}, \{2, 3\}\} = \pi''$. $\square$

The example given in the proof of Proposition 1 further shows that min-max fair coalition structures do not always exist (which is to be expected from any reasonable notion of fairness; envy-freeness is a classic fairness condition in fair division of indivisible goods, but in conjunction with completeness or Pareto optimality such partitions do not always exist either).

Since SCS implies IR and CS, we get the following corollary.

Corollary 1. *An IR or CS coalition structure does not necessarily satisfy min-max fairness.*

3.2. Grand-coalition fairness

Bogomolnaia and Jackson [21] proposed the grand coalition as a notion of fairness. We recover their idea in the context of local fairness: It can be seen as an analogue of the fair-division notion of proportionality in the setting of hedonic games. As IR ($f(\succeq_i) = \{i\}$) is the minimal fairness criterion and we want our notions to be situated above IR, we set $\{i\}$ as a minimal bound in the next definition.

Definition 6. The *grand-coalition threshold of player $i \in N$* is defined as

$$\mathrm{GC}_i = \max\{\{i\}, N\},$$

where we maximize with respect to $\succeq_i$. A coalition structure π satisfies *grand-coalition fairness* (GC) if

$$\pi(i) \succeq_i \mathrm{GC}_i$$

for every $i \in N$.

Grand-coalition fairness is a notion of fairness because the grand coalition can be interpreted as an average: Every player has to face both her friends and her enemies. Note that a proportionality threshold is typically defined as the ratio of the valuation for the whole to the number of players. Since players "share" their coalitions, it is not clear which number the valuation of the whole should be compared to. Comparing to the number of coalitions in a coalition structure, however, violates our locality requirement: thresholds should only depend on a player's own preference. Therefore, we forego using ratios in the definition of grand-coalition fairness.

First, we show that grand-coalition fairness is strictly stronger a requirement than min-max fairness.

Theorem 2. *Every grand-coalition fair coalition structure satisfies min-max fairness, yet a min-max fair coalition structure does not necessarily satisfy grand-coalition fairness.*

Proof. Let $i \in N$. Every coalition structure serves as an upper bound of MinMax_i. Consider the coalition structure $\{N\}$. Then $\max\{\{i\}, N\} \succeq_i \mathrm{MinMax}_i$.

Conversely, consider the following hedonic game:

$$\{1, 2\} \succ_1 \{1\} \succ_1 \{1, 2, 3\} \succ_1 \{1, 3\},$$
$$\{1, 2, 3\} \succ_2 \{2, 3\} \succ_2 \{1, 2\} \succ_2 \{2\},$$
$$\{1, 2, 3\} \succ_3 \{2, 3\} \succ_3 \{1, 3\} \succ_3 \{3\}.$$

The players' min-max threshold coalitions are $\mathrm{MinMax}_1 = \{1\}$, $\mathrm{MinMax}_2 = \{2, 3\}$, and $\mathrm{MinMax}_3 = \{2, 3\}$. Thus $\{\{1\}, \{2, 3\}\}$ satisfies min-max fairness as each player is in her min-max threshold coalition, but it does not satisfy grand-coalition fairness as players 2 and 3 prefer the grand coalition $\{1, 2, 3\}$ to $\{2, 3\}$. $\square$

Since none of USW, ESW, EF-R, PO, CIS, SCS, CS, or IR implies min-max fairness, we get the following corollary.

Corollary 2. *A coalition structure that satisfies USW, ESW, EF-R, PO, CIS, SCS, CS, or IR does not necessarily satisfy grand-coalition fairness.*

Later we will see that grand-coalition fairness is independent of all other considered notions except for PF. For now we show that these notions do not imply grand-coalition fairness.

Proposition 2. *A SSNS coalition structure does not necessarily satisfy grand-coalition fairness.*

Proof. Consider the following hedonic game:

$$\{1,2\} \succ_1 \{1\} \succ_1 \{1,2,3\} \succ_1 \{1,3\},$$
$$\{1,2,3\} \succ_2 \{1,2\} \succ_2 \{2,3\} \succ_2 \{2\},$$
$$\{2,3\} \succ_3 \{3\} \succ_3 \{1,2,3\} \succ_3 \{1,3\}.$$

Coalition structure $\pi = \{\{1,2\},\{3\}\}$ is SSNS but is not grand-coalition fair.

For the sake of contradiction, assume π were not SSNS. Then there is a coalition H and a coalition structure π' that is reachable from π by H such that $\pi'(i) \succeq_i \pi(i)$ for every $i \in H$ and there is some $j \in H$ with $\pi'(j) \succ_j \pi(j)$. Since π' is different from π, $\{1,2\}$ cannot be a coalition in π'. This directly implies that agent 1 is not part of H because 1 prefers $\{1,2\}$ to every other coalition. If agent 2 is in H, then with $\{1,2\} \notin \pi'$ and $\pi'(2) \succeq_2 \pi(2) = \{1,2\}$ we can conclude that $\pi' = \{\{1,2,3\}\}$. However, this implies that 3 is in H which is in contradiction with $\pi(3) = \{3\} \succ_3 \{1,2,3\} = \pi'(3)$. Thus $2 \notin H$, which implies $H = \{3\}$. However, there is no coalition structure that is reachable from π by $\{3\}$ that agent 3 prefers. This is the final contradiction.

To see that π is not grand-coalition fair, consider agent 2 who is in coalition $\{1,2\}$ according to π but would prefer to be in the grand coalition. □

Since SSNS implies NS, we get the following corollary.

Corollary 3. *A NS coalition structure does not necessarily satisfy grand-coalition fairness.*

3.3. Max-Min fairness

We motivate the next fairness notion with the following situation: Suppose some player is allowed to partition all players excluding herself but does not know which coalition she will be part of in the end. Since she had the right to choose a partition, she has to live with all possible consequences. In other words, she could end up in any of these coalitions, even the worst. The only choice she will always be allowed to make in the end is to leave the coalition she is assigned to and be on her own instead. Therefore, a player would partition all remaining players so that the worst coalition among them is as good as possible for her.

Definition 7. The *max-min threshold of player $i \in N$* is defined as

$$\mathrm{MaxMin}_i = \max_{\pi \in \Pi(N\setminus\{i\})} \max\{\{i\}, \min_{C \in \pi} C \cup \{i\}\},$$

where maximization and minimization are with respect to $\succeq_i$. A coalition structure π satisfies *max-min fairness* (MAX-MIN) if

$$\pi(i) \succeq_i \mathrm{MaxMin}_i$$

for every $i \in N$.

Max-min fairness is the hedonic-games analogue of max-min fair share due to Budish [26]. We show that max-min fairness is strictly stronger than grand-coalition fairness.

Theorem 3. *Every max-min fair coalition structure satisfies grand-coalition fairness, yet a grand-coalition fair coalition structure does not necessarily satisfy max-min fairness.*

Proof. Let $i \in N$. The coalition structure π consisting of the grand coalition without i is the one where $\max_{C\in\pi\cup\{\emptyset\}} C \cup \{i\}$ and $\max\{\{i\}, \min_{C\in\pi} C \cup \{i\}\}$ become equal. Since every coalition structure gives a lower bound for MaxMin_i and an upper bound for MinMax_i, we have

$$\text{MaxMin}_i \succeq_i \text{GC}_i \succeq_i \text{MinMax}_i.$$

Conversely, consider the following hedonic game:

$$\{1,2\} \succ_1 \{1,3\} \succ_1 \{1,2,3\} \succ_1 \{1\},$$
$$\{1,2,3\} \succ_2 \{1,2\} \succ_2 \{2,3\} \succ_2 \{2\},$$
$$\{2,3\} \succ_3 \{1,2,3\} \succ_3 \{3\} \succ_3 \{1,3\}.$$

Coalition structure $\{\{1,2,3\}\}$ satisfies grand-coalition fairness but not max-min fairness because player 1's max-min threshold coalition is $\{1,3\}$. $\square$

Theorems 2 and 3 give additional motivation of grand-coalition fairness: It is strictly situated between max-min and min-max fairness. It follows from Theorem 3 and our previous results for grand-coalition fairness that max-min fairness is not implied by any of the considered notions except for PF.

Proposition 3. *A max-min fair coalition structure does not necessarily satisfy CIS, CS, EF-R, or ESW.*

Proof. For CIS, CS, and EF-R, consider the following hedonic game:

$$\{1,2\} \succ_1 \{1\} \succ_1 \{1,2,3\} \succ_1 \{1,3\},$$
$$\{1,2\} \succ_2 \{2\} \succ_2 \{1,2,3\} \succ_2 \{2,3\},$$
$$\{2,3\} \succ_3 \{3\} \succ_3 \{1,2,3\} \succ_3 \{1,3\}.$$

The max-min threshold coalitions are $\text{MaxMin}_i = \{i\}$, $i \in \{1,2,3\}$. For example, player 1 can choose between the coalition structures $\{\{2\},\{3\}\}$ and $\{\{2,3\}\}$. In $\{\{2\},\{3\}\}$, the worst coalition that she might end up in is $\{1,3\}$ and, in $\{\{2,3\}\}$, it is $\{1,2,3\}$. Since she always has the right to be on her own, which she prefers in both cases, her max-min threshold is $\text{MaxMin}_1 = \max\{\max\{\{1\},\{1,3\}\}, \max\{\{1\},\{1,2,3\}\}\} = \{1\}$.

Thus coalition structures $\pi = \{\{1\},\{2\},\{3\}\}$ and $\pi' = \{\{1,2\},\{3\}\}$ satisfy max-min fairness. However, π is neither CS nor CIS: Coalition $\{1,2\}$ blocks π and agent 2 could deviate to coalition $\{1\}$ without decreasing anyone's valuation. On the other hand, π' is not EF-R, since agent 3 envies agent 1 and would like to replace her.

For ESW, finally, consider the following additively separable hedonic game, defined via the values $v_i(j)$:

$i \backslash j$	1	2	3	4
1	0	1	−2	0
2	1	0	−2	0
3	−1	−1	0	1
4	−1	−1	1	0

We again have $\text{MaxMin}_i = \{i\}$, $i \in \{1,2,3,4\}$. Hence, coalition structure $\{\{1,2\},\{3\},\{4\}\}$ satisfies max-min fairness. But it does not maximize ESW since coalition structure $\{\{1,2\},\{3,4\}\}$ has a higher ESW. $\square$

Corollary 4.

1. *A grand-coalition fair or min-max fair coalition structure does not necessarily satisfy CIS, CS, EF-R, or ESW.*
2. *A max-min fair, grand-coalition fair, or min-max fair coalition structure does not necessarily satisfy NS, PO, SSNS, SCS, USW, or PF.*

See Fig. 1 for a summary of the results of this section.

4. Local fairness in ASHGs

In this section, we study the existence of fair coalition structures, the complexity of computing fairness thresholds and of deciding whether a hedonic game admits a fair coalition structure. Since additively separable hedonic games are a well-studied class of hedonic games (see, e.g., the book chapter by Aziz and Savani [7] and the references therein), we will focus on this class. In addition, it will be easier to compare our complexity results to some results in fair division with additive utility functions. Before starting with the computational analysis, we first concretize our results from Section 3 in regard to ASHGs.

4.1. Additional relations in ASHGs

In Section 3, we introduced the three notions of local fairness and related them to each other and to well-known stability notions, considering general hedonic games (recall Fig. 1). Considering ASHGs only, however, some more implications than those shown in Fig. 1 might hold. Indeed, for ASHGs, grand-coalition fairness and max-min fairness coincide:

Theorem 4. *In additively separable hedonic games (N, v), for every $i \in N$ we have*

$$v_i(\mathrm{MaxMin}_i) = v_i(\mathrm{GC}_i).$$

Proof. By Theorem 3, it remains to show that $\max\{0, v_i(N)\} = v_i(\mathrm{GC}_i) \geq v_i(\mathrm{MaxMin}_i)$ for each $i \in N$.

If $v_i(N) < 0$, then $v_i(\mathrm{GC}_i) = 0$. Suppose that $v_i(\mathrm{MaxMin}_i) > 0$. Then there is some $\pi \in \Pi(N \setminus \{i\})$ such that $\min_{C \in \pi} v_i(C \cup \{i\}) > 0$. Then $v_i(C \cup \{i\}) > 0$ for every coalition $C \in \pi$. This, however, implies

$$v_i(N) = \sum_{C \in \pi} v_i(C \cup \{i\}) > 0,$$

contradicting $v_i(N) < 0$, so $v_i(\mathrm{MaxMin}_i) \leq 0 = v_i(\mathrm{GC}_i)$.

If $v_i(N) \geq 0$, then $v_i(\mathrm{GC}_i) = v_i(N)$. Suppose that $v_i(\mathrm{MaxMin}_i) > v_i(N)$. Then there is some $\pi \in \Pi(N \setminus \{i\})$ such that $\min_{C \in \pi} v_i(C \cup \{i\}) > v_i(N)$. Thus $v_i(C \cup \{i\}) > v_i(N)$ for every coalition $C \in \pi$, which implies

$$v_i(N) = \sum_{C \in \pi} v_i(C \cup \{i\}) > v_i(N),$$

again a contradiction. Thus $v_i(\mathrm{MaxMin}_i) \leq v_i(N) = v_i(\mathrm{GC}_i)$. This completes the proof. □

All other relations, stated in Observation 1, Theorems 1 and 2, Propositions 1, 2, and 3, and Corollaries 1, 2, 3, and 4, remain valid for ASHGs. This can easily be seen because, first, all relations between the previously studied notions (colored in black in Fig. 1) also hold for ASHGs, second, all implications stated in the above mentioned theorems and propositions are shown for general hedonic games and thus do hold for ASHGs in particular, and third, all hedonic games stated as counterexamples (proving that certain implications do not hold) are additively separable.

4.2. Min-Max fairness

We begin our computational complexity analysis with min-max fairness and start by computing min-max thresholds. Since we have valuation functions in ASHGs, we can compare to the *value* of threshold coalitions. In particular, we consider the following decision problem:

Min-Max-Threshold	
Given:	A set N of players, a player i's valuation function v_i, and a rational number k.
Question:	Does it hold that $v_i(\mathrm{MinMax}_i) \geq k$?

By considering coalition structures consisting of either the grand coalition or only of singletons, we have the following observations that show that Min-Max-Threshold is easy to solve for certain restricted valuation functions.

Observation 2. *If $v_i(N) \leq 0$, then $v_i(\mathrm{MinMax}_i) = 0$.*

Observation 3. *If $v_i(j) \geq 0$ for every $j \in N$, then $v_i(\mathrm{MinMax}_i) = \max_{j \in N} v_i(j)$.*

For general valuation functions, however, we have this result:

Theorem 5. *For general ASHGs,* Min-Max-Threshold *is strongly* coNP*-complete.*

Proof. We consider the complementary problem, which for the same input asks whether $v_i(\mathrm{MinMax}_i) < k$ holds. Membership of this problem in NP follows from guessing a coalition structure π and comparing the maximum value of a coalition in π to k.

To show strong NP-hardness, we reduce from a restricted variant of 3-Partition, which is known to be strongly NP-complete (see, e.g., the book by Garey and Johnson [39]). A 3-Partition instance consists of a set $X = \{x_1, \dots, x_{3m}\}$ and nonnegative integers $s(x)$ for each $x \in X$ such that $\sum_{x \in X} s(x) = mB$ and $B/4 < s(x) < B/2$ for each $x \in X$. The question is whether X can be partitioned into m disjoint sets $S_1, \dots, S_m$ such that $\sum_{x \in S_i} s(x) = B$, $1 \leq i \leq m$. Note that we can multiply each size by $4m$, a polynomial, while still maintaining strong NP-completeness. Then B is divisible by 4 and $\frac{B}{4m} \geq 1$. Also, we can assume that $m \geq 2$ (the case of $m = 1$ is trivial).

The reduction works as follows. Set $N = \{x_1, \dots, x_{3m}, d_1, \dots, d_m, a\}$, where the d_i are referred to as *dummy players*. Consider valuation function v_a, which is defined as follows: $v_a(x_i) = s(x_i)$, $1 \leq i \leq 3m$, $v_a(d_j) = -\frac{3}{4}B$, $1 \leq j \leq m$. Set $k = \frac{B}{4} + 1$. We show that there is a 3-partition if and only if $v_a(\text{MinMax}_a) < k$.

From left to right: Suppose $S_1, \dots, S_m$ is a 3-partition. Consider the coalition structure $\pi = \{S_i \cup \{d_i\}\}_{1 \leq i \leq m}$. We have $\max_{C \in \pi \cup \{\emptyset\}} v_a(C \cup \{a\}) = \max\{B - \frac{3}{4}B, 0\} = \frac{B}{4}$. Therefore, $v_a(\text{MinMax}_a) \leq \frac{B}{4} < k$.

From right to left: Suppose we have a no-instance for 3-partition. Let $\pi \in \Pi(N \setminus \{a\})$ be an arbitrary coalition structure. Denote by ℓ the number of coalitions in π. We will show that for each ℓ, $1 \leq \ell \leq 4m$, there is a coalition in π that is valued at least $\frac{B}{4} + 1$.

Case 1: $\ell = 1$. The value of the grand coalition is $mB - \frac{3}{4}mB = \frac{1}{4}mB \geq \frac{B}{4} + 1$, if $m \geq 2$.

Case 2: $2 \leq \ell < m - 1$. There is a coalition that is valued at least $\frac{mB}{4\ell} > \frac{mB}{4(m-1)}$. The right-hand side is at least $\frac{B}{4} + 1$ if and only if $B \geq 4m - 4$, which holds because of $\frac{B}{4m} \geq 1$.

Case 3: $\ell = m - 1$. Suppose for the sake of contradiction that the maximum value of all coalitions in π is strictly less than $\frac{B}{4} + 1$. Then $v(N) < (m-1)(\frac{B}{4} + 1) = \frac{1}{4}mB + m - \frac{B}{4} - 1 \leq \frac{1}{4}mB$. The last inequality holds because of $\frac{B}{4m} \geq 1$.

Case 4: $\ell = m$. If each coalition contains three distinct players x_i, then there is a coalition S_i that is valued more than B (when dummy players in S_i are ignored) because we have started from a no-instance. If there is no dummy player, we are done. If there is more than one dummy player, then by the pigeonhole principle there is a coalition with three distinct players x_i, yet without a dummy player. Again, we are done. If there is exactly one dummy player in S_i, $v(S_i) > B - \frac{3}{4}B = \frac{B}{4}$. By integrality, $v(S_i) \geq \frac{B}{4} + 1$. If there is a coalition S_i that contains at least four distinct players x_i, we can argue similarly and use integrality: If there is one dummy player, $v(S_i) > 4 \cdot \frac{B}{4} - \frac{3}{4}B = \frac{1}{4}B$. If there is no dummy player, we are done. If there is more than one dummy player, there is a coalition that contains no dummy player, but each x_i is valued strictly more than $\frac{1}{4}B$.

Case 5: $m + 1 \leq \ell \leq 4m$. By the pigeonhole principle there is a coalition that contains no dummy player. Since each x_i is valued more than $\frac{B}{4}$, we have that such a coalition has value at least $\frac{B}{4} + 1$ by integrality.

Since all coalition structures have a coalition which is valued at least $\frac{B}{4} + 1$, we have $v_a(\text{MinMax}_a) \geq k$. □

Turning now to the question of whether fair coalition structures exist, we first define the decision problem that we study.

Min-Max-Exist	
Given:	An additively separable hedonic game (N, v).
Question:	Does there exist a coalition structure $\pi \in \Pi(N)$ that is min-max fair?

We first observe that Min-Max-Exist is trivial in the case of nonnegative valuations.

Observation 4. *If $v_i(j) \geq 0$ for every $i, j \in N$, then $\{N\}$ satisfies min-max fairness, i.e., min-max fair coalition structures always exist.*

We say an additively separable hedonic game is *symmetric* if $v_i(j) = v_j(i)$ for every $i, j \in N$. Since there always exist NS coalition structures in symmetric ASHGs [21] and since NS implies min-max fairness by Theorem 1, we have

Corollary 5. *Symmetric ASHGs always admit min-max fair coalition structures.*

For general additively separable hedonic games, however, there do not always exist min-max fair coalition structures. In fact, the problem of deciding whether there exists a min-max fair coalition structure for a given additively separable hedonic game (not restricted to be symmetric) is NP-complete.

Theorem 6. *For general ASHGs,* Min-Max-Exist *is* NP*-complete.*

Proof. First, we show that Min-Max-Exist is in NP. We nondeterministically guess a coalition structure $\pi \in \Pi(N)$ and n further coalition structures $\pi^i \in \Pi(N \setminus \{i\}), i \in N$. Then, for each $i \in N$, we check if i weakly prefers the coalition she is assigned to in π to every coalition she could be assigned to in π^i, i.e., if $\pi(i) \succeq_i C \cup \{i\}$ for every $C \in \pi^i \cup \{\emptyset\}$. We do this by checking whether $v_i(\pi(i)) \geq v_i(\max_{C \in \pi^i \cup \{\emptyset\}} C \cup \{i\})$, where maximization is with respect to $\succeq_i$. If the latter equation is true for all $i \in N$, then π is min-max fair. Note that $(\pi, \pi^1, \dots, \pi^n)$ is a witness (of length $O(n^2)$) for the existence of a min-max fair coalition structure and that verifying the min-max fairness of π can then be done in polynomial time by checking the above mentioned inequalities. Hence, membership of Min-Max-Exist in NP follows.

For NP-hardness, we reduce from Monotone-One-in-Three-3SAT (see, e.g., the comment of Garey and Johnson [39, p. 259] on One-in-Three-3SAT): Given a boolean formula φ that contains only clauses with three positive literals, does there exist a satisfying assignment such that each clause has exactly one true literal?

Now we describe the reduction. Let $\ell_1, \dots, \ell_n$ be the variables, $C_1, \dots, C_m$ the clauses of φ, and r_i the number of distinct clauses ℓ_i appears in. For every variable ℓ_i that appears in some clause, we introduce three *variable players*, a_i, b_i, and c_i, and for every clause C_k, we add three *clause players*, D_k, E_k, and F_k.

The valuation functions are defined as follows:

Variable players of type 1, a_i, $1 \le i \le n$, value every other variable player except for b_i and c_i with $-3m-1$. They value b_i with $3r_i$, c_i with 0, all clause players D_k, E_k, and F_k for which $\ell_i \in C_k$ with -1, and all remaining players with 0. Variable players of type 2, b_i, $1 \le i \le n$, value c_i with $3r_i$, all clause players D_k, E_k, and F_k for which $\ell_i \in C_k$ with 1, and all remaining players with 0. Variable players of type 3, c_i, $1 \le i \le n$, value every clause player with -1 and all remaining players with 0.

Clause players of type 1, D_k, $1 \le k \le m$, value E_k with -10, F_k with 15, and all remaining players with 0. Clause players of type 2, E_k, $1 \le k \le m$, value D_k with -20, F_k with 21, all a_i for which $\ell_i \in C_k$ with 20, and all remaining players with 0. Clause players of type 3, F_k, $1 \le k \le m$, value D_k with 10, E_k with 20, and all remaining players with 0.

We compute the min-max thresholds before showing the required equivalence. The min-max threshold of type-1 variable players a_i is 0, of type-2 variable players b_i is $3r_i$, of type-3 variable players c_i is 0, of type-1 clause players D_k is 5, of type-2 clause players E_k is 20 (consider the coalition structure where D_k and F_k are together and all remaining players are in single coalitions), and of type-3 clause players F_k is 20. We show that the given formula is satisfied by an assignment with exactly one true literal per clause if and only if there is a coalition structure satisfying min-max fairness.

From left to right: Suppose there is a satisfying assignment τ with the above property. Denote by $\ell_1, \dots, \ell_o$ all variables which are true under τ and by $\ell_{o+1}, \dots, \ell_n$ all remaining variables. Denote by $T(\ell_i)$ the set of clauses that become true under τ via ℓ_i and let $\mathscr{C}(\ell_i) = \{D_k, E_k, F_k \mid C_k \in T(\ell_i)\}$. Consider the following coalition structure:

$$\pi = \{\mathscr{C}(\ell_1) \cup \{a_1, b_1\}, \{c_1\}, \dots, \mathscr{C}(\ell_o) \cup \{a_o, b_o\}, \{c_o\}, \{a_{o+1}, b_{o+1}, c_{o+1}\}, \dots, \{a_n, b_n, c_n\}\}.$$

In words, for each clause satisfied by some literal, we put into one coalition the corresponding type-1 and type-2 variable players and all three corresponding clause players. In this case, the corresponding type-3 variable player stays alone. If a variable satisfies no clause, then the corresponding variable players are in a coalition that only consists of them. Since each clause is satisfied by exactly one literal, no player is in multiple coalitions simultaneously.

We compute every player's value. Type-1 clause players D_k are in the same coalition as E_k and F_k. Therefore, they have a value of 5. Similarly, type-3 clause players F_k have a value of 30. For variables ℓ_i that are true under τ, since type-1 variable players a_i never share coalitions with other variable players except for b_i and all clause players corresponding to clauses that contain ℓ_i, these a_i have a value of $-3r_i + 3r_i = 0$. For the same reason, type-2 variable players b_i have a value of $3r_i$ and type-2 clause players E_k have a value of $-20 + 21 + 20 = 21$. Type-3 variable players c_i are alone and, hence, have value 0. For variables ℓ_j that are false under τ, the coalition consisting of all variable players corresponding to such a variable gives the corresponding type-1 and type-2 variable players, a_j and b_j, a value of $3r_j$ and the corresponding type-3 variable player c_j a value of 0. Overall, every player achieves her min-max threshold, so coalition structure π satisfies min-max fairness.

From right to left: Suppose there is some coalition structure π satisfying min-max fairness. Let F_k be the type-3 clause player for some clause C_k. Since $v_{F_k}(\pi(F_k)) \ge 20$, $E_k \in \pi(F_k)$. Because $F_k \in \pi(F_k)$ and $v_{D_k}(\pi(D_k)) \ge 5$, $D_k \in \pi(F_k)$. Thus $\{D_k, E_k, F_k\} \subseteq \pi(F_k)$. Because $v_{E_k}(\pi(F_k)) \ge 20$, some variable player a_i corresponding to a literal occurring in clause C_k has to be in $\pi(F_k)$; otherwise, $v_{E_k}(\pi(F_k)) = 1$. If variable players different from b_i or c_i are in $\pi(F_k)$, no such variable player can satisfy the min-max threshold. Since variable player a_i joins $\{D_k, E_k, F_k\}$, $b_i \in \pi(F_k)$. Because $v_{b_i}(\pi(F_k)) \ge 3r_i$, all clause players that are valued positively by b_i are in $\pi(F_k)$. Otherwise, $c_i \in \pi(F_k)$, but then $u_{c_i}(\pi(F_k)) < 0$ because clause players are in $\pi(F_k)$. If clause players that are valued 0 by a_i are in the same coalition, other variable players have to be in that coalition as well, but then min-max fairness is not satisfied. Therefore, every $\pi(F_k)$ contains exactly one a_i. Then we can construct a satisfying assignment with the one-in-three property by making all variables ℓ_i true where a_i is in a coalition with clause players. $\square$

4.3. Grand-coalition and Max-Min fairness

Due to Theorem 4, we can consider grand-coalition and max-min fairness at the same time. We define the threshold and existence problems for grand-coalition and max-min fairness analogously to Min-Max-Threshold and Min-Max-Exist. Since computing the value of the grand coalition is easy in additively separable hedonic games, we have the following result.

Observation 5. *For general ASHGs,* Max-Min-Threshold *and* Grand-Coalition-Threshold *are in* P.

Considering the existence problem, we can make a similar observation as for min-max fairness.

Observation 6. *If $v_i(j) \ge 0$ for every $i, j \in N$, then $\{N\}$ satisfies grand-coalition and max-min fairness, i.e., grand-coalition and max-min fair coalition structures always exist.*

However, checking whether there exists a grand-coalition fair or max-min fair coalition structure is hard for general valuation functions.

Theorem 7. *For general ASHGs, the two problems* GRAND-COALITION-EXIST *and* MAX-MIN-EXIST *are strongly* NP*-complete.*

Proof. Membership in NP follows from guessing and checking. Checking works in polynomial time because of Observation 5. We reduce from a restricted variant of the problem EXACT-COVER-BY-THREE-SETS (X3C) that is known to be strongly NP-complete [40]: Let $(B, \mathscr{S})$ be an instance of X3C with $B = \{b_1, \dots, b_{3m}\}$, a collection $\mathscr{S} = \{S_1, \dots, S_n\}$ with $S_i \subseteq B$ and $|S_i| = 3$ for each i, and $n > m$, where every element of B appears in exactly three triplets from $\mathscr{S}$. The question is whether there exists an exact cover of B, i.e., a subcollection $\mathscr{S}' \subseteq \mathscr{S}$ of size m such that $B = \bigcup_{S_i \in \mathscr{S}'} S_i$.

Let $C(n,m)$ and $D(n,m)$ be functions of n and m to be determined later. For easier notation, we will omit (n,m) and will simply write C and D instead. Consider the additively separable hedonic game with players $N = \{b_1, \dots, b_{3m}, S_1, \dots, S_n, a_1, \dots, a_{n-m}\}$ and with the following valuation functions:

- $v_{b_i}(b_j) = 0$, $v_{b_i}(S_k) = C$ if $b_i \in S_k$, and $v_{b_i}(a_l) = \frac{-2C}{n-m}$, for $1 \le i, j \le 3m$, $1 \le k \le n$, and $1 \le l \le n-m$.
- For $1 \le k \le n$, if $b_i \in S_k$, set $v_{S_k}(b_i) = C$; otherwise, set $v_{S_k}(b_i) = -D$.
- Furthermore, we have $v_{S_k}(S_o) = -D$, $k \neq o$, and $v_{S_k}(a_l) = 3C$ for $1 \le l \le n-m$.
- The valuation function for a_l, $1 \le l \le n-m$, maps b_i, $1 \le i \le 3m$, and a_p, $l \neq p$, to $-(n+1)C$; and it maps S_k, $1 \le k \le n$, to C.

We want the grand-coalition threshold of S_k, $1 \le k \le n$, to be $3C$. Therefore, we need

$$3C - D(3(m-1) + (n-1)) + 3C(n-m) = 3C,$$

which is equivalent to

$$D = \frac{3C(n-m)}{3(m-1) + (n-1)}.$$

Setting $C = 3(m-1) + (n-1)$ gives $D = 3(n-m)$. Since each b_i appears in exactly three triplets, we have

$$\mathrm{GC}_{b_i} = 3C + (n-m)\frac{-2C}{n-m} = C.$$

The grand-coalition threshold of a_l, $1 \le l \le n-m$, is 0. We show that there is an exact cover if and only if there is a coalition structure satisfying grand-coalition (and thus, equivalently in ASHGs, max-min) fairness.

From left to right: Suppose $\mathscr{S}'$ is an exact cover with index set I. Then the coalition structure

$$\{\{S_i, b_x, b_y, b_z\}_{i \in I}, \{S_j, a_{\pi(j)}\}_{j \notin I}\}$$

satisfies grand-coalition fairness, where $S_i = \{b_x, b_y, b_z\}$ and $\pi : \{1, \dots, n\} \setminus I \to \{1, \dots, n-m\}$ is some permutation: Each b_i, $1 \le i \le 3m$, gets a value of C, each S_k, $1 \le k \le n$, a value of $3C$, and each a_l, $1 \le l \le n-m$, a value of C.

From right to left: Suppose π satisfies grand-coalition fairness. Then $v_{b_i}(\pi(b_i)) \ge C$. This implies that there is at least one S_j in $\pi(b_i)$. Suppose S_j and S_k with $j \neq k$ are in $\pi(b_i)$. Note that $v_{a_l}(\pi(a_l)) \ge 0$. Therefore, no a_l, $1 \le l \le n-m$, is in $\pi(b_i)$. Furthermore, the fact that $v_{S_j}(\pi(S_j)) \ge 3C$ and $v_{S_k}(\pi(S_k)) \ge 3C$ means that S_j and S_k only derive value from b-players. Since S_j and S_k are distinct and each S-player only values exactly three b-players, there is some b-player in $\pi(b_i)$ that has value $-D$ for S_j or S_k, which implies a value less than $3C$ for that S-player (contradiction). Hence, we have found an exact cover because each b_i is exactly with one S_j. □

5. Price of local fairness

Now we study the price of local fairness in additively separable hedonic games. Informally, the maximum (minimum) price of fairness captures the loss in social welfare of a worst (best) coalition structure that satisfies some fairness criterion. We denote by $\mathrm{SW}_G(\pi)$ the USW of coalition structure π in an additively separable hedonic game $G = (N, v)$, that is, $\mathrm{SW}_G(\pi) = \sum_{i \in N} v_i(\pi(i))$. We omit G when it is clear from the context.

Definition 8. Let $G = (N, v)$ be an additively separable hedonic game and let π^* denote a coalition structure maximizing USW. Define the *maximum price of min-max fairness in G* by

- $\text{Max-PoMMF}(G) = \max_{\pi \in \Pi(N) \text{ is min-max fair}} \frac{\mathrm{SW}(\pi^*)}{\mathrm{SW}(\pi)}$ if there is some min-max fair $\pi \in \Pi(N)$ and $\mathrm{SW}(\pi) > 0$ for all min-max fair $\pi \in \Pi(N)$;
- $\text{Max-PoMMF}(G) = 1$ if $\mathrm{SW}(\pi^*) = 0$ (and thus $\mathrm{SW}(\pi) = 0$ for all min-max fair $\pi \in \Pi(N)$); and
- $\text{Max-PoMMF}(G) = +\infty$ otherwise.

Define the *minimum price of min-max fairness in G* by

- Min-PoMMF$(G) = \min_{\pi \in \Pi(N) \text{ is min-max fair}} \frac{\mathrm{SW}(\pi^*)}{\mathrm{SW}(\pi)}$ if there is a min-max fair $\pi \in \Pi(N)$ with $\mathrm{SW}(\pi) > 0$;
- Min-PoMMF$(G) = 1$ if $\mathrm{SW}(\pi^*) = 0$ (and thus $\mathrm{SW}(\pi) = 0$ for all min-max fair $\pi \in \Pi(N)$); and
- Min-PoMMF$(G) = +\infty$ otherwise.

Note that we have $\mathrm{SW}(\pi^*) \geq 0$ and $\mathrm{SW}(\pi) \geq 0$, where π^* maximizes USW and π is min-max fair.

Because the grand coalition maximizes USW under nonnegative valuation functions, the minimum and maximum price of grand-coalition fairness is one. Since this bound is not really informative, we now make some suitable assumptions to strengthen our results. We will only consider min-max fairness, the weakest fairness notion, in order to constrain the set of coalition structures as little as possible. In addition, just as Elkind et al. [35] focus on Pareto optimality because such coalition structures always exist, we will restrict our study to symmetric additively separable hedonic games so as to guarantee the existence of min-max fair coalition structures.

Unfortunately, the maximum price of min-max fairness is not bounded by a constant value even for nonnegative valuation functions.

Theorem 8. *Let $G = (N, v)$ be a symmetric ASHG of n players with $v_i(j) \geq 0$ for every $i, j \in N$. Then*

$$\text{Max-PoMMF}(G) \leq n - 1,$$

and this bound is tight.

Proof. If $\mathrm{SW}(\pi^*) = 0$, then Max-PoMMF$(G) = 1$. Otherwise, there are $i, j \in N$, $i \neq j$, such that $v_i(j) > 0$. We can upper-bound $\mathrm{SW}(\pi^*)$ by $\sum_{i \in N} v_i(N)$. By Observation 3, we can lower-bound the value of every player i by $\max_{j \in N} v_i(j)$. Thus

$$\begin{aligned} \text{Max-PoMMF}(G) &\leq \frac{\sum_{i \in N} v_i(N)}{\sum_{i \in N} \max_{j \in N} v_i(j)} \\ &\leq \frac{\sum_{i \in N} (n-1) \max_{j \in N} v_i(j)}{\sum_{i \in N} \max_{j \in N} v_i(j)} \\ &= n - 1. \end{aligned}$$

To see that this bound is tight, consider a game with n players, where n is even. Every player values every other player with $a > 0$. Thus the min-max threshold of every player is a. Therefore, the coalition structure that consists of $n/2$ pairs satisfies min-max fairness and has minimum USW of $n \cdot a$ among all min-max fair coalition structures. The coalition structure consisting of the grand coalition that maximizes USW, however, has a USW of $n(n-1)a$. □

To obtain a meaningful bound in the above result, we need the existence of a min-max fair coalition structure, which is guaranteed in symmetric ASHGs. We will use the next result when turning to Min-PoMMF.

In symmetric ASHGs, every coalition structure that maximizes USW is NS (see the proof of Proposition 2 in the paper by Bogomolnaia and Jackson [21, p. 213]) and, for general hedonic games, NS implies min-max fairness (see Theorem 1). Hence, we get the following proposition.

Proposition 4. *In symmetric ASHGs, any coalition structure that maximizes USW is min-max fair.*

From Proposition 4 we immediately have the following corollary.

Corollary 6. *Let G be a symmetric ASHG. Then*

$$\text{Min-PoMMF}(G) = 1.$$

Note that Proposition 4 is more than we need to prove Corollary 6. It is also sufficient to notice that there always exists a coalition structure that is min-max fair and maximizes USW, namely $\{N\}$. Since for Min-PoMMF we minimize $\frac{\mathrm{SW}(\pi^*)}{\mathrm{SW}(\pi)}$ over all min-max fair coalition structures π, this minimum is always 1 for $\pi = \{N\}$.

6. An alternative notion

One might wonder whether it is possible to define a local fairness notion that always exists. For individual rationality this is the case. Now we consider a slightly modified notion that returns a coalition that is ranked just above the singleton coalition if this is possible; otherwise, it returns the singleton coalition. However, the following hedonic game shows that a coalition structure that satisfies this fairness notion does not always exist:

$$\{1,2\} \succ_1 \{1\} \succ_1 \{1,2,3\} \succ_1 \{1,3\},$$
$$\{2,3\} \succ_2 \{2\} \succ_2 \{1,2,3\} \succ_2 \{1,2\},$$
$$\{1,3\} \succ_3 \{3\} \succ_3 \{1,2,3\} \succ_3 \{2,3\}.$$

Under additively separable hedonic games, this notion corresponds to the question of whether it is possible to guarantee a positive value to each agent (or zero if no positively valued coalition exists). The existence problem is still NP-complete. We present a short proof sketch. The reduction is again from the restricted variant of X3C where each b_i appears in exactly three S_j (see the proof of Theorem 7). Map an instance $(B, \mathscr{S})$ with $B = \{b_1, \ldots, b_{3m}\}$ and $\mathscr{S} = \{S_1, \ldots, S_n\}$ to an ASHG (N, v) with $N = B \cup \mathscr{S} \cup \{h_1, \ldots, h_{n-m}\}$ and valuation function v defined as follows:

- $v_{b_i}(b_j) = -1$ for each $b_i, b_j \in B$ with $b_i \neq b_j$, and $v_{b_i}(S_p) = 3$ if $b_i \in S_p$;
- $v_{S_p}(b_i) = 1$ if $b_i \in S_p$, and $v_{S_p}(h_k) = 1$, $1 \le k \le n-m$;
- $v_{h_i}(S_p) = 1$, $1 \le p \le n$; and
- all other entries are $-cn$ for some sufficiently large constant c.

If there is an exact cover, put the subset and its members into the same coalition. The remaining subsets form coalitions of size two with an h_i each. This coalition structure satisfies our alternative notion of fairness.

Conversely, if there is a coalition structure according to the alternative notion of fairness, then h_i has positive value and is with some S_p. Since S_p cannot be with other players of the same type, each h_i is with a unique S_p. The remaining m subsets S_p have to be with some b_i, which in turn have to be in a coalition with an S_p. Each b_i can share the coalition with at most two other b_j. Since there are $3m$ players b_i and m remaining subsets S_p, there is an exact cover.

7. Discussion and conclusion

We have introduced three new notions of (local) fairness in hedonic games and studied the connection with previously studied notions. Our notions themselves form a strict hierarchy: Every max-min fair coalition structure is grand-coalition fair (but not vice versa), and every grand-coalition fair coalition structure is min-max fair (but not vice versa). Although our local fairness criteria are inspired by notions from the field of fair division, our results are very different. Bouveret and Lemaître's scale of fairness criteria for additive utility functions [23] says that an envy-free partition of goods satisfies min-max fair share, which in turn implies proportionality, which in turn implies max-min fair share. So our strongest notion of local fairness is the weakest notion in fair division of indivisible goods (according to this scale). In addition, in additively separable hedonic games, we have seen that grand-coalition fairness and max-min fairness coincide. This is not the case in fair division (if one equates grand-coalition fairness with proportionality). Also note that Nash stability (or, equivalently, a definition of envy-freeness based on joining) implies min-max fairness but none of the stronger notions. In the setting of allocating indivisible goods, envy-freeness even implies min-max fair share. So it is one of the strongest notions there. We consider these results surprising, as the intuition from fair division of indivisible goods is no longer valid in this different context. The main reasons are the already mentioned difference between the number of allowed subsets of a partition and that players can "share" coalitions. This missing intuition is also a reason of why we have checked in detail whether any known stability notions in hedonic games imply one of our local fairness notions or are implied by them.

Then we have studied the complexity of computing threshold coalitions and of deciding whether an additively separable hedonic game admits a locally fair coalition structure. Although nearly all of these problems are intractable, our local fairness criteria still have some meaning. They give additional motivation to notions of stability, such as Nash stability. Moreover, in a decentralized setting the hardness of a problem can be "distributed" (of course, the intractability cannot simply disappear). Giving players a yardstick for fairness that only depends on their own preferences reduces the amount of communication that is necessary to check whether a coalition structure is fair. Our complexity results are also comparable to the results by Bouveret and Lemaître [23] and Heinen et al. [45] with the exception that no computational lower bound is known for deciding whether a max-min fair share allocation exists, whereas in ASHGs we know that the corresponding problem is strongly NP-complete. Also note that with min-max fairness we have found a notion that is strictly stronger than individual rationality, but is still satisfied by every coalition structure maximizing utilitarian social welfare in symmetric additively separable hedonic games.

At last, making the fact that coalition structures satisfying one of our three local fairness notions are not always guaranteed to exist perhaps a bit more bearable, we showed that there is no local fairness notion stronger than individual rationality such that there always exists a coalition structure satisfying it. Furthermore, we have initiated the study of price of local fairness in hedonic games. Our results here are unsatisfactory in the sense that either the price is unbounded or not very informative. Therefore, we consider finding suitable restrictions to players' valuation functions such that the maximum price of min-max fairness is bounded by a nontrivial constant an interesting research question for future work. Interesting future work would also be identifying (other) sufficient conditions that imply the existence of a fair coalition structure, determining the complexity of searching for a min-max fair coalition structure in symmetric additively separable hedonic games, and investigating the complexity of computing thresholds and of verifying the existence of locally fair coalition structures for other classes of hedonic games such as, e.g., fractional hedonic games.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank the anonymous AAMAS 2016, CoopMAS 2016, and TCS reviewers for their helpful comments. The first author is a member of the PhD-programme "Online Participation," supported by the North Rhine-Westphalian funding scheme "Forschungskollegs." This work was supported in part by Deutsche Forschungsgemeinschaft under grants RO-1202/14-2 and RO-1202/21-1.

References

[1] R. Abebe, J. Kleinberg, D. Parkes, Fair division via social comparison, in: Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems, IFAAMAS, May 2017, pp. 281–289.
[2] G. Amanatidis, E. Markakis, A. Nikzad, A. Saberi, Approximation algorithms for computing maximin share allocations, in: Proceedings of the 42nd International Colloquium on Automata, Languages, and Programming, in: Lecture Notes in Computer Science, vol. 9134, Springer-Verlag, July 2015, pp. 39–51.
[3] H. Aziz, F. Brandl, Existence of stability in hedonic coalition formation games, in: Proceedings of the 11th International Joint Conference on Autonomous Agents and Multiagent Systems, IFAAMAS, June 2012, pp. 763–770.
[4] H. Aziz, F. Brandt, P. Harrenstein, Fractional hedonic games, in: Proceedings of the 13th International Joint Conference on Autonomous Agents and Multiagent Systems, IFAAMAS, May 2014, pp. 5–12.
[5] H. Aziz, F. Brandt, H. Seedig, Computing desirable partitions in additively separable hedonic games, Artif. Intell. 195 (2013) 316–334.
[6] H. Aziz, S. Gaspers, J. Gudmundsson, J. Mestre, H. Täubig, Welfare maximization in fractional hedonic games, in: Proceedings of the 24th International Joint Conference on Artificial Intelligence, AAAI Press/IJCAI, July 2015, pp. 461–467.
[7] H. Aziz, R. Savani, Hedonic games, in: F. Brandt, V. Conitzer, U. Endriss, J. Lang, A. Procaccia (Eds.), Handbook of Computational Social Choice, Cambridge University Press, 2016, pp. 356–376, chapter 15.
[8] C. Ballester, NP-completeness in hedonic games, Games Econ. Behav. 49 (1) (2004) 1–30.
[9] S. Banerjee, H. Konishi, T. Sönmez, Core in a simple coalition formation game, Soc. Choice Welf. 18 (1) (2001) 135–153.
[10] S. Barman, S. Murthy, Approximation algorithms for maximin fair division, in: Proceedings of the 2017 ACM Conference on Economics and Computation, ACM, June 2017, pp. 647–664.
[11] N. Barrot, M. Yokoo, Stable and envy-free partitions in hedonic games, in: Proceedings of the 28th International Joint Conference on Artificial Intelligence, July 2019, pp. 67–73, ijcai.org.
[12] X. Bei, S. Liu, X. Lu, H. Wang, Maximin fairness with mixed divisible and indivisible goods, Technical Report, ACM Computing Research Repository, 2020, arXiv:2002.05245.
[13] X. Bei, Y. Qiao, S. Zhang, Networked fairness in cake cutting, in: Proceedings of the 26th International Joint Conference on Artificial Intelligence, AAAI Press, August 2017, pp. 3632–3638.
[14] N. Benabbou, M. Chakraborty, A. Igarashi, Y. Zick, Finding fair and efficient allocations when valuations don't add up, Technical Report, ACM Computing Research Repository, 2020, arXiv:2003.07060.
[15] K. Bérczi, E. Bérczi-Kovács, E. Boros, F. Gedefa, N. Kamiyama, T. Kavitha, Y. Kobayashi, K. Makino, Envy-Free Relaxations for Goods, Chores, and Mixed Items, Technical Report, ACM Computing Research Repository, 2020, arXiv:2006.04428.
[16] D. Bertsimas, V. Farias, N. Trichakis, The price of fairness, Oper. Res. 59 (1) (2011) 17–31.
[17] V. Bilò, I. Caragiannis, M. Flammini, A. Igarashi, G. Monaco, D. Peters, C. Vinci, W. Zwicker, Almost envy-free allocations with connected bundles, in: Proceedings of the 10th Conference on Innovations in Theoretical Computer Science, in: LIPIcs, vol. 124, Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2018, pp. 14:1–14:21.
[18] V. Bilò, A. Fanelli, M. Flammini, G. Monaco, L. Moscardelli, Nash stability in fractional hedonic games, in: Proceedings of the 10th International Workshop on Internet & Network Economics, in: Lecture Notes in Computer Science, vol. 8877, Springer-Verlag, December 2014, pp. 486–491.
[19] V. Bilò, A. Fanelli, M. Flammini, G. Monaco, L. Moscardelli, On the price of stability of fractional hedonic games, in: Proceedings of the 14th International Joint Conference on Autonomous Agents and Multiagent Systems, IFAAMAS, May 2015, pp. 1239–1247.
[20] A. Biswas, S. Barman, Fair division under cardinality constraints, in: Proceedings of the 27th International Joint Conference on Artificial Intelligence, 2018, pp. 91–97, ijcai.org.
[21] A. Bogomolnaia, M. Jackson, The stability of hedonic coalition structures, Games Econ. Behav. 38 (2) (2002) 201–230.
[22] S. Bouveret, Y. Chevaleyre, N. Maudet, Fair allocation of indivisible goods, in: F. Brandt, V. Conitzer, U. Endriss, J. Lang, A. Procaccia (Eds.), Handbook of Computational Social Choice, Cambridge University Press, 2016, pp. 284–310, chapter 12.
[23] S. Bouveret, M. Lemaître, Characterizing conflicts in fair division of indivisible goods using a scale of criteria, J. Auton. Agents Multi-Agent Syst. 30 (2) (2016) 259–290.
[24] S. Brânzei, K. Larson, Coalitional affinity games and the stability gap, in: Proceedings of the 21st International Joint Conference on Artificial Intelligence, AAAI Press/IJCAI, 2009, pp. 79–84.
[25] R. Bredereck, A. Kaczmarczyk, R. Niedermeier, Envy-free allocations respecting social networks, in: Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems, IFAAMAS, July 2018, pp. 283–291.
[26] E. Budish, The combinatorial assignment problem: approximate competitive equilibrium from equal incomes, J. Polit. Econ. 119 (6) (2011) 1061–1103.
[27] I. Caragiannis, C. Kaklamanis, P. Kanellopoulos, M. Kyropoulou, The efficiency of fair division, Theory Comput. Syst. 50 (4) (2012) 589–610.
[28] I. Caragiannis, D. Kurokawa, H. Moulin, A. Procaccia, N. Shah, J. Wang, The unreasonable fairness of maximum nash welfare, in: Proceedings of the 17th ACM Conference on Economics and Computation, ACM, July 2016, pp. 305–322.
[29] K. Cechlárová, A. Romero-Medina, Stability in coalition formation games, Int. J. Game Theory 29 (4) (2001) 487–494.

[30] Y. Chevaleyre, P. Dunne, U. Endriss, J. Lang, M. Lemaître, N. Maudet, J. Padget, S. Phelps, J. Rodríguez-Aguilar, P. Sousa, Issues in multiagent resource allocation, Informatica 30 (1) (2006) 3–31.
[31] Y. Chevaleyre, U. Endriss, N. Maudet, Distributed fair allocation of indivisible goods, Artif. Intell. 242 (2017) 1–22.
[32] S. de Jong, K. Tuyls, K. Verbeeck, Fairness in multi-agent systems, Knowl. Eng. Rev. 23 (2) (2008) 153–180.
[33] D. Dimitrov, P. Borm, R. Hendrickx, S. Sung, Simple priorities and core stability in hedonic games, Soc. Choice Welf. 26 (2) (2006) 421–433.
[34] J. Drèze, J. Greenberg, Hedonic coalitions: optimality and stability, Econometrica 48 (4) (1980) 987–1003.
[35] E. Elkind, A. Fanelli, M. Flammini, Price of Pareto optimality in hedonic games, in: Proceedings of the 30th AAAI Conference on Artificial Intelligence, AAAI Press, February 2016, pp. 475–481.
[36] E. Elkind, J. Rothe, Cooperative game theory, in: J. Rothe (Ed.), Economics and Computation. An Introduction to Algorithmic Game Theory, Computational Social Choice, and Fair Division, Springer-Verlag, 2015, pp. 135–193, chapter 3.
[37] E. Elkind, M. Wooldridge, Hedonic coalition nets, in: Proceedings of the 8th International Joint Conference on Autonomous Agents and Multiagent Systems, IFAAMAS, May 2009, pp. 417–424.
[38] D. Foley, Resource allocation and the public sector, Yale Econ. Essays 7 (1) (1967) 45–98.
[39] M. Garey, D. Johnson, Computers and Intractability: A Guide to the Theory of NP-Completeness, W. H. Freeman and Company, 1979.
[40] T. Gonzalez, Clustering to minimize the maximum intercluster distance, Theor. Comput. Sci. 38 (1985) 293–306.
[41] L. Gourvès, J. Monnot, On maximin share allocations in matroids, Theor. Comput. Sci. 754 (2019) 50–64.
[42] L. Gourvès, J. Monnot, L. Tlilane, Near fairness in matroids, in: Proceedings of the 21st European Conference on Artificial Intelligence, IOS Press, August 2014, pp. 393–398.
[43] J. Hajduková, Coalition formation games: a survey, Int. Game Theory Rev. 8 (4) (2006) 613–641.
[44] T. Heinen, N. Nguyen, T. Nguyen, J. Rothe, Approximation and complexity of the optimization and existence problems for maximin share, proportional share, and minimax share allocation of indivisible goods, J. Auton. Agents Multi-Agent Syst. 32 (6) (2018) 741–778.
[45] T. Heinen, N. Nguyen, J. Rothe, Fairness and rank-weighted utilitarianism in resource allocation, in: Proceedings of the 4th International Conference on Algorithmic Decision Theory, in: Lecture Notes in Artificial Intelligence, vol. 9346, Springer-Verlag, September 2015, pp. 521–536.
[46] M. Karakaya, Hedonic coalition formation games: a new stability notion, Math. Soc. Sci. 61 (3) (2011) 157–165.
[47] A. Kerkmann, J. Rothe, Altruism in coalition formation games, in: Proceedings of the 29th International Joint Conference on Artificial Intelligence, July 2020, pp. 347–353, ijcai.org.
[48] D. Kurokawa, A. Procaccia, J. Wang, When can the maximin share guarantee be guaranteed?, in: Proceedings of the 30th AAAI Conference on Artificial Intelligence, AAAI Press, February 2016, pp. 523–529.
[49] J. Lang, A. Rey, J. Rothe, H. Schadrack, L. Schend, Representing and solving hedonic games with ordinal preferences and thresholds, in: Proceedings of the 14th International Joint Conference on Autonomous Agents and Multiagent Systems, IFAAMAS, May 2015, pp. 1229–1238.
[50] J. Lang, J. Rothe, Fair division of indivisible goods, in: J. Rothe (Ed.), Economics and Computation. An Introduction to Algorithmic Game Theory, Computational Social Choice, and Fair Division, Springer-Verlag, 2015, pp. 493–550, chapter 8.
[51] T. Morrill, The roommates problem revisited, J. Econ. Theory 145 (5) (2010) 1739–1756.
[52] N. Nguyen, T. Nguyen, J. Rothe, Approximate solutions to max-min fair and proportionally fair allocations of indivisible goods, in: Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems, IFAAMAS, May 2017, pp. 262–271.
[53] N. Nguyen, A. Rey, L. Rey, J. Rothe, L. Schend, Altruistic hedonic games, in: Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems, IFAAMAS, May 2016, pp. 251–259.
[54] N. Nguyen, J. Rothe, Local fairness in hedonic games via individual threshold coalitions, in: Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems, IFAAMAS, May 2016, pp. 232–241.
[55] T. Nguyen, M. Roos, J. Rothe, A survey of approximability and inapproximability results for social welfare optimization in multiagent resource allocation, Ann. Math. Artif. Intell. 68 (1–3) (2013) 65–90.
[56] D. Peters, Complexity of hedonic games with dichotomous preferences, in: Proceedings of the 30th AAAI Conference on Artificial Intelligence, AAAI Press, February 2016, pp. 579–585.
[57] D. Peters, Graphical hedonic games of bounded treewidth, in: Proceedings of the 30th AAAI Conference on Artificial Intelligence, AAAI Press, February 2016, pp. 586–593.
[58] D. Peters, E. Elkind, Simple causes of complexity in hedonic games, in: Proceedings of the 24th International Joint Conference on Artificial Intelligence, AAAI Press/IJCAI, July 2015, pp. 617–623.
[59] A. Procaccia, J. Wang, Fair enough: guaranteeing approximate maximin shares, in: Proceedings of the 15th ACM Conference on Economics and Computation, ACM Press, 2014, pp. 675–692.
[60] J. Rothe, Thou shalt love thy neighbor as thyself when thou playest: altruism in game theory, in: Proceedings of the 35th AAAI Conference on Artificial Intelligence, AAAI Press, February 2021, in press.
[61] J. Rothe, H. Schadrack, L. Schend, Borda-induced hedonic games with friends, enemies, and neutral players, Math. Soc. Sci. 96 (2018) 21–36.
[62] Warut Suksompong, Approximate maximin shares for groups of agents, Math. Soc. Sci. 92 (2018) 40–47.
[63] Suguru Ueda, Envy based fairness in hedonic games, in: Proceedings of the Genetic and Evolutionary Computation Conference Companion, GECCO '18, ACM, 2018, pp. 1852–1853.
[64] A. Wiechers, J. Rothe, Stability in minimization-based altruistic hedonic games, in: Proceedings of the 9th European Starting AI Researchers' Symposium, Vol. 2655, August 2020, p. 3, CEUR-WS.org.
[65] M. Wright, Y. Vorobeychik, Mechanism design for team formation, in: Proceedings of the 29th AAAI Conference on Artificial Intelligence, AAAI Press, January 2015, pp. 1050–1056.

Chapter 5

Hedonic Games with Ordinal Preferences and Thresholds

In this chapter, we summarize the following journal article in which we introduce and study a new preference representation in hedonic games where agents submit ordinal rankings that are separated by two thresholds:

Publication (Kerkmann et al. [92])

A. Kerkmann, J. Lang, A. Rey, J. Rothe, H. Schadrack, and L. Schend. "Hedonic Games with Ordinal Preferences and Thresholds". In: *Journal of Artificial Intelligence Research* 67 (2020), pp. 705–756

5.1 Summary

In this work, we introduce and study a new class of hedonic games which we call *FEN-hedonic games*. In these games, the agents partition the other agents into friends, enemies, and players that they are neutral to. Additionally, they submit a weak order on their friends and on their enemies, respectively. The resulting preference representation is referred to as *weak ranking with double threshold*. Based on this representation, we then infer preferences over coalitions using the *responsive extension principle*. Since the resulting polarized responsive extensions are not always complete, we consider agents to *possibly* or *necessarily prefer* a coalition to another one if this preference holds for *at least one* or *all* completions of their polarized responsive extensions. Afterwards, we introduce so-called *optimistic* and *pessimistic preference extensions*.

Using these extensions, we then characterize stability in FEN-hedonic games. In addition, we study the problems of verifying stable coalition structures in FEN-hedonic games and of checking whether stable coalition structures exist. While doing so, we distinguish between *possible* and *necessary stability*, depending on whether there exists at least one extended preference profile that satisfies stability or whether all extended preference profiles satisfy stability. While these verification and existence problems for possible and necessary stability are in P for the strongest and weakest notion that we consider, namely for *perfectness* and *individual rationality*, we also show some hardness results for some other stability notions. For example, we show that possible and necessary Nash stability verification are in P, while possible and necessary Nash stability existence are NP-complete. We also show that possible verification is coNP-complete for core stability, strict core stability, Pareto optimality, popularity, and strict popularity. Also, necessary verification is coNP-complete for Pareto optimality, popularity, and strict popularity. Finally, we close our work with a short discussion and some directions for future work.

5.2 Personal Contribution and Preceding Versions

This journal paper largely extends two conference papers published by Jérôme Lang, Anja Rey, Jörg Rothe, Hilmar Schadrack, and Lena Schend at AAMAS'15 [96] and by me and Jörg Rothe at AAMAS'19 [89]. The modeling is due to the authors of the AAMAS'15 paper [96]. The writing of this journal paper was done jointly with all co-authors. The ideas of all technical results that also appear in the AAMAS'19 paper [89] are my contribution. Also some technical parts from the preceeding AAMAS'15 paper [96] that were revised for this journal paper are my contribution as well. Parts of the technical results of this journal paper have already appeared, in preliminary form, in my Master's Thesis [83]. However, their presentation and many of their proofs were improved in the journal paper.

5.3 Publication

The full article [92] is appended here.

Journal of Artificial Intelligence Research 67 (2020) 705–756 Submitted 06/2019; published 04/2020

Hedonic Games with Ordinal Preferences and Thresholds

Anna Maria Kerkmann ANNA.KERKMANN@HHU.DE
Heinrich-Heine-Universität Düsseldorf
40225 Düsseldorf, Germany

Jérôme Lang LANG@LAMSADE.DAUPHINE.FR
LAMSADE, CNRS, Université Paris-Dauphine, PSL
75775 Paris Cedex 16, France

Anja Rey ANJA.REY@TU-DORTMUND.DE
Technische Universität Dortmund
44221 Dortmund, Germany

Jörg Rothe ROTHE@HHU.DE
Hilmar Schadrack HILMAR.SCHADRACK@HHU.DE
Lena Schend LENA.SCHEND@GOOGLEMAIL.COM
Heinrich-Heine-Universität Düsseldorf
40225 Düsseldorf, Germany

Abstract

We propose a new representation setting for hedonic games, where each agent partitions the set of other agents into friends, enemies, and neutral agents, with friends and enemies being ranked. Under the assumption that preferences are monotonic (respectively, antimonotonic) with respect to the addition of friends (respectively, enemies), we propose a bipolar extension of the responsive extension principle, and use this principle to derive the (partial) preferences of agents over coalitions. Then, for a number of solution concepts, we characterize partitions that necessarily or possibly satisfy them, and we study the related problems in terms of their complexity.

1. Introduction

Hedonic games are cooperative games where agents form coalitions. Each agent has a preference relation over the set of all coalitions containing her. Various solution concepts—such as individual rationality, Nash stability, individual stability, core stability, popularity, and so on—have been proposed and studied for hedonic games. These solution concepts apply to coalition structures, that is, to partitions of the set of agents into disjoint coalitions. For instance, a coalition structure is individually rational if no agent prefers the coalition of which she is the only member to the coalition she is currently a member of, and it is Nash stable if no agent prefers to be integrated into another existing coalition than staying in her current coalition. (Other solution concepts will be explained later on.)

1.1 Hedonic Games: Standard Game-Theoretic vs. Engineering-Oriented Point of View

There are two different points of view under which we can study hedonic games. Under a standard game-theoretic point of view, a hedonic game is a model by which one can predict (or at least reason about) the coalitions that a set of agents, acting without the intervention of a central authority, may form given what we know their preferences. Under an engineering-oriented point of view, a game is the input of a problem whose output, computed by a central authority, is a set of coalitions, which should be as satisfactory as possible.

Both the origin and the description of the game differ under these two interpretations of hedonic games. Under the standard game-theoretic one, the game is written down by the modeler; under the engineering-oriented one, it is written down by the agents, each of which has to report their preferences on the possible coalitions they may end up forming.

The role of solution concepts is also different under these two interpretations. Under the standard game-theoretic one, solution concepts are assumed to be realistic models of what may happen; for instance, if a partition Γ of the agents is not individually rational (that is, at least one agent would be better leaving her coalition and be alone), and if agents are free to move out and form singleton coalitions, then it is not reasonable to expect that Γ will be the final outcome of the game; as another example, if agents can choose to leave their coalition and join any other existing coalition without asking for permission of the coalition they join, nor of the coalition they leave, then Nash stability is a relevant solution concept, and therefore, provided there exists at least a Nash stable coalition structure, we can expect that the resulting coalition structure will be Nash stable (and if there exists no Nash stable coalition structure, we can predict that the outcome of the game will be unstable).

Under the engineering-oriented interpretation, solution concepts are desiderata that one may impose on the outcome: For instance, individual rationality should be a hard constraint that the outcome must satisfy. *In this paper, we focus on this interpretation of hedonic games.* Therefore, our assumption will be that a central authority has first to elicit the agents' preferences about coalitions, and then to compute a desirable outcome. In the following paragraph, we will discuss the difficulties inherent to these two stages (elicitation and computation). Our contribution mostly consists in defining a new framework for hedonic games that comes with a representation language that offers a trade-off between expressivity and succinctness, and to study various stability notions in this setting.

1.2 Representing Preferences over Coalitions

Since each agent has to specify a preference relation over the set of all coalitions containing her, an important bottleneck is how the agents' preferences over the coalitions that contain them are expressed. As there are exponentially many (in the number of agents) coalitions containing agent i, it is not reasonable to expect agent i to express a ranking (or a utility function) over all these coalitions explicitly. This issue is often addressed by assuming that only a small part of the preference relation is expressed by the agent, and that this small part is then extended to a complete preference relation over coalitions using an appropriate extension principle. Various assumptions about the nature of the input (specifying what the agents are required to express) and the preference extension have been made in the literature (for recent surveys, see Aziz & Savani, 2016; Elkind & Rothe, 2015; Woeginger, 2013a):

1. The *individually rational encoding* (Ballester, 2004): Each agent explicitly ranks all coalitions she prefers to herself being alone, and only those ones.
2. *Hedonic coalition nets* (Elkind & Wooldridge, 2009): Each agent specifies her utility function over the set of all coalitions via (more or less) a set of weighted logical formulas.
3. The *singleton encoding* (Cechlárová & Romero-Medina, 2001; Cechlárová & Hajduková, 2003, 2004): Each agent ranks only single agents; under the optimistic (respectively, pes-

simistic) extension, X is preferred to Y if the best (respectively, worst) agent in X is preferred to the best (respectively, worst) agent in Y.

4. The *additive encoding* (Sung & Dimitrov, 2007, 2010; Aziz et al., 2013b; Woeginger, 2013b): Each agent gives a valuation (positive or negative) of each other agent; preferences are additively separable, and the extension principle is that the valuation of a set of agents, for agent i, is the sum of the valuations i gives to the agents in the set (and then the preference relation is derived from this valuation function).
5. *Fractional hedonic games* (Aziz et al., 2019; Bilò et al., 2014, 2015): Once again, each agent assigns a value to each other agent (and 0 to herself); an agent's utility of a coalition is the average value she assigns to the members of the coalition.
6. The *friends-and-enemies encoding* (Dimitrov et al., 2006; Sung & Dimitrov, 2007; Rey et al., 2016; Nguyen et al., 2016): Each agent partitions the set of other agents into two sets (her friends and her enemies); under the *friend-oriented preference extension*, coalition X is preferred to coalition Y if X contains more friends than Y, or as many friends as Y and fewer enemies than Y; under the *enemy-oriented preference extension*, X is preferred to Y if X contains fewer enemies than Y, or as many enemies as Y and more friends than Y.
7. The *anonymous encoding* (Ballester, 2004; Darmann et al., 2018): Each agent specifies only a preference relation over the number of agents in her coalition (ignoring who they are).
8. *Boolean hedonic games* (Aziz et al., 2016; Peters, 2016): Each agent partitions all coalitions into two subsets one of which she prefers to the other while being indifferent between the coalitions inside each of those two subsets. This partition is expressed compactly in propositional logic.

We can classify these various ways of specifying hedonic games according to two parameters:

- the nature of the output: *ordinal* (for each agent i, a preference relation over coalitions containing i), *cardinal* (a utility function over coalitions containing i), or *dichotomous* (a partition of coalitions containing i between good and bad ones);
- the nature of the language used for expressing the agent's preference over coalitions containing her: *explicit* (coalitions are listed in extension), *logical* (preferences are expressed using logical formulas or similar objects), *singleton-wise* (only single agents are ranked or given a value), or *anonymous* (preferences are expressed only on possible cardinalities of coalitions).

In Table 1 we classify each of our languages along these two parameters, where the numbers in the table correspond to the above enumeration of eight encodings of hedonic games.

Naturally, compact representation either does not avoid exponential-size representations in the worst case (Case 1 and, to a lesser extent, Case 2), or comes with a loss of expressivity, corresponding to a demanding domain restriction, such as separable preferences (Cases 3, 4, and 5), anonymous preferences (Case 7), or other domain restrictions that do not bear a specific name (Cases 6 and 8).

In Cases 2, 4, and 5, preferences are expressed numerically: Agents do explicitly express numbers. In all other cases, they are expressed ordinally. The difficulties with eliciting and aggregating numeric preferences have been long discussed in social choice (Sen, 1970), and for these reasons the community favors ordinal preferences.

Anonymity is a very demanding assumption, which does not allow to distinguish between agents. Even if it makes sense in some settings, such as in group activity selection (Darmann et al., 2018), it is unrealistic in most cases. The individually rational encoding is not compact in general.

	dichotomous	ordinal	cardinal
explicit		1	
logical	8		2
singleton-wise	6	3	4, 5
anonymous		7	

Table 1: Classification of some representation languages for hedonic games. They are referred to by their numbers in the enumeration.

1.3 Solution Concepts and Computation

While one difficulty we wish to overcome has to do with the space, time, and cognitive effort required from the agents for expressing their preferences, another difficulty consists in the complexity of computing coalition structures that satisfy some solution concept, or, when such a coalition structure is not guaranteed to exist, to check whether there exists a coalition structure that satisfies it. Because of the exponential number of possible coalition structures, there is no guarantee that these problems are easy to solve, and indeed in many cases they are hard. A lot of attention has been devoted to the computational complexity of the problems associated with various solution concepts under different representations; they are surveyed in Section 15.4 of the book chapter by Aziz and Savani (2016) and in Section 3.3.3 of the book chapter by Elkind and Rothe (2015).

1.4 Towards a More Satisfactory Representation

As discussed before, desiderata for hedonic game representations are expressivity, succinctness, and cognitive simplicity. We would like, ideally, a representation that satisfies the following three requirements, or that, at least, is a satisfactory trade-off between them:

(1) it should be reasonably expressive;

(2) it should be compact;

(3) it should be cognitively plausible, and easy to elicit from the agents.

Because of requirement (3), we want to stick to ordinal preferences which, among other advantages, are easier to elicit from the agents. Requirement (1) excludes the very demanding anonymity assumption, which does not allow to distinguish between agents. Requirement (2) excludes the individually rational encoding, which is not compact in general.

The only remaining representations are the friends-and-enemies and singleton encodings. However, we argue that they are insufficiently expressive and thus are poor on requirement (1). A problem with the friends-and-enemies encoding is that an agent cannot express preferences inside the friend set nor inside the enemy set: Preferences over individual agents are dichotomous (but preferences between coalitions are not, because they depend on the number of friends and enemies). A problem with the singleton encoding is that having simply a rank $\rhd_i$ for each agent i does not tell us which agents i would like to see in her coalition and which agents she would like not to see: For instance, if $\rhd_1$ is $2 \rhd_1 3 \rhd_1 4$, we know that 1 prefers 2 to 3 and 3 to 4, but nothing tells us whether 1

	dichotomous	ordinal	**polarized ordinal**	cardinal
explicit		1		
logical	8			2
singleton-wise	6	3	this work	4, 5
anonymous		7		

Table 2: Classification of some representation languages for hedonic games, *including ours*. The representation languages are again referred to by their numbers in the enumeration above.

prefers to be with 2 (respectively, with 3 and 4) to being alone, that is, whether the "absolute desirability" of 2, 3, and 4 is positive or negative (of course, if it is negative for 3, it is also negative for 4, etc.). Obviously, both ways are insufficiently informative: Specifying only a partition into favorable and disliked agents ("friends" and "enemies") does not tell which of her friends *i* prefers to which other agents, and which of her enemies she wants to avoid most. On the other hand, specifying a ranking over agents does not say which agents *i* prefers to be with rather than being alone.

Therefore, we propose a model that integrates the models of Cases 3 and 6 (as described in the list in Section 1.2 on pages 706–707): Each agent *i* first subdivides the other agents into three groups—her friends, her enemies, and an intermediate type of agent on which she has neither a positive nor a negative opinion—and then specifies a ranking of her friends and a ranking of her enemies. Our new representation of hedonic games with friends, enemies, and neutral players is called the model of *FEN-hedonic game*.

This representation is not purely ordinal: Along with the preference relation we have a threshold that indicates which coalitions are better than being alone, which ones are worse, and which ones are equally good. Such a structure, which we call *polarized ordinal*, is reminiscent of "approval-ranking" ballots in fallback voting (Brams & Sanver, 2009), where each voter ranks candidates along with indicating an approval threshold. Table 2 also shows where our representation method is located with respect to the two parameters mentioned earlier.

Based on this representation, we consider a natural extension of a player's preference, the *polarized responsive extension*, which is a partial order over coalitions containing the player.

Responsive preferences come from bipartite many-to-one matching markets (see, e.g., Roth, 1985; Roth & Sotomayor, 1992), and consider the comparison of one participant to another. In the context of many-to-one matching markets, an agent on the one side has *responsive preferences* over assignments of the agents on the other side if, for any two assignments that differ in only one agent, the assignment containing the most preferred agent is preferred. The responsive extension principle is sometimes called the *Bossong-Schweigert extension principle* (Bossong & Schweigert, 2006) (see also Delort et al., 2011).

How can we deal with incomparabilities within these partial orders? Our approach is to leave them open and define notions such as "possible" and "necessary" stability concepts. Questions of interest include appropriate characterizations of stability concepts and a computational study of the related problems in terms of their complexity.

1.5 Outline

This paper is a largely extended version of its conference predecessors (Lang, Rey, Rothe, Schadrack, & Schend, 2015; Kerkmann & Rothe, 2019). It contains all omitted proofs, as well as a broad spectrum of new insights, results, and discussions. This includes a more detailed discussion of our model and its advantages in comparison to existing models. We will also extend the axiomatic analysis of the different steps of the preference extensions, and use this analysis to further delimit our model to other models of hedonic game. On the other hand, our results on Borda-induced hedonic games from the conference version (Lang et al., 2015) are not contained here but have appeared in a separate journal article (Rothe, Schadrack, & Schend, 2018).

After introducing the needed notions and definitions in Section 2, explaining various representations of and stability concepts for hedonic games as well as some required terms from complexity theory, we turn in Section 3 to our new model of FEN-hedonic games, using ordinal preferences with double threshold and the polarized responsive principle. In Section 4, we study stability in FEN-hedonic games, in particular pinpointing the complexity of possible and necessary stability problems. In Section 5, we conclude our work and give a brief overview of open problems and possible future work.

2. Preliminaries

A *hedonic game* is a pair $(A,\succeq)$ consisting of a set of *players* (or *agents*) $A = \{1,2,\ldots,n\}$ and a profile of preference relations $\succeq = (\succeq_1,\succeq_2,\ldots,\succeq_n)$ defining for each player a weak preference order over all possible *coalitions* $C \subseteq A$ containing the player herself.[1] We denote the set of all coalitions containing player $i \in A$ with $\mathscr{A}_i$. For two coalitions $C,D \in \mathscr{A}_i$, we say that *i weakly prefers C to D* if $C \succeq_i D$; *i prefers C to D*, denoted by $C \succ_i D$, if $C \succeq_i D$ but not $D \succeq_i C$; and *i is indifferent between C and D*, denoted by $C \sim_i D$, if both $C \succeq_i D$ and $D \succeq_i C$. A *coalition structure* Γ for a given game $(A,\succeq)$ is a partition of A into disjoint coalitions, and for each player $i \in A$, $\Gamma(i)$ denotes the unique coalition in Γ containing i. We denote the set of all possible coalition structures for a hedonic game $(A,\succeq)$ by $\mathscr{C}_{(A,\succeq)}$. Occasionally, we may omit the hedonic game in this notion and just write $\mathscr{C}$ if this is clear from the context.

2.1 Some Known Representations of Hedonic Games

The need for a succinct representation of hedonic games calls for the definition of a compact representation language for preferences over coalitions containing a player. Specifically, using this language players should be required to express their preferences in a compact manner. At the same time, they should have the opportunity to express them in as much detail as possible. To address this issue, a number of sophisticated approaches have been proposed in the literature, and our new model to be introduced in Section 3 will draw on some of them. We list some of the known representations of hedonic games below.

We start with a very powerful class of hedonic games that was introduced by Banerjee et al. (2001). An *additively separable hedonic game* is given by a pair (A,w), where $A = \{1,2,\ldots,n\}$ is a set of players and $w = (w_1,w_2,\ldots,w_n)$, i.e., each player $i \in A$ has a *value function* $w_i : A \to \mathbb{R}$ by which she evaluates all players. Now, the players' preferences on coalitions containing them can be

[1] While we often stick to the convention that the players' names are numbers (as in $A = \{1,2,\ldots,n\}$), we will occasionally deviate from it for the sake of readability.

derived as follows, yielding the corresponding hedonic game $(A, \succeq)$: For each $i \in A$ and for any two coalitions $B, C \in \mathscr{A}_i$, it holds that $B \succeq_i C \iff \sum_{j \in B} w_i(j) \geq \sum_{j \in C} w_i(j)$.

Dimitrov et al. (2006) introduced a representation that is based on so-called friend- and enemy-oriented preference extensions and provides a subclass of the additively separable hedonic games. In their representation, each player $i \in A$ partitions the other players in a set of friends $F_i \subseteq A \setminus \{i\}$ and a set of enemies $E_i = A \setminus (F_i \cup \{i\})$, and if $B, C \in \mathscr{A}_i$, i's preference over these two coalitions is then determined by the number of friends and enemies she has in them as follows. In the *friend-oriented preference extension*, we define $B \succeq_i C$ if and only if $|B \cap F_i| > |C \cap F_i|$ or ($|B \cap F_i| = |C \cap F_i|$ and $|B \cap E_i| \leq |C \cap E_i|$), and in the *enemy-oriented preference extension*, we define $B \succeq_i C$ if and only if $|B \cap E_i| < |C \cap E_i|$ or ($|B \cap E_i| = |C \cap E_i|$ and $|B \cap F_i| \geq |C \cap F_i|$).

Both encodings can also be represented by additively separable hedonic games. To capture the friend-oriented encoding, each player i has to assign the value $|A|$ to her friends and the value -1 to her enemies. The enemy-oriented encoding results from value functions which assign to each friend the value 1 and to each enemy the value $-|A|$.

A different approach is taken by Cechlárová and Romero-Medina (2001) (see also Cechlárová & Hajduková, 2003, 2004), who expect the game to be given in the singleton encoding, i.e., each player $i \in A$ has to provide a complete ranking $\trianglerighteq_i$ over all players. For any coalition $B \in \mathscr{A}_i$, let $\mathscr{B}_i(B)$ be any *best* player $j \in B$ from i's view, i.e., $j \trianglerighteq_i k$ for each $k \in B$; and let $\mathscr{W}_i(B) = i$ if $B = \{i\}$, and otherwise let $\mathscr{W}_i(B)$ be any *worst* player $j \in B \setminus \{i\}$ from i's view, i.e., $k \trianglerighteq_i j$ for each $k \in B$. Now, for any $B, C \in \mathscr{A}_i$, we say *B is $\mathscr{B}$-preferred by i over C* if $\mathscr{B}_i(B) \triangleright_i \mathscr{B}_i(C)$ or ($\mathscr{B}_i(B) \sim_i \mathscr{B}_i(C)$ and $|B| < |C|$), and we say *B is $\mathscr{W}$-preferred by i over C* if $\mathscr{W}_i(B) \triangleright_i \mathscr{W}_i(C)$.

2.2 Stability Concepts

Important solution concepts for hedonic games are various notions of stability for coalition structures (see, e.g., Bogomolnaia & Jackson, 2002; Aziz et al., 2013b; Aziz, Brandt, & Harrenstein, 2013a; Aziz & Savani, 2016; Elkind & Rothe, 2015). We focus on concepts that deal with avoiding a player to deviate to another (possibly empty) existing coalition. Relatedly, other commonly studied concepts consider group deviations, such as core stability with the goal that there is no blocking coalition. A third group of stability concepts, such as Pareto optimality and popularity, is based on a relation comparing different coalition structures. For other restrictions of games and other properties, we refer, e.g., to the work of Banerjee et al. (2001).

The following properties are well-known, except for the last one (strict popularity), which is introduced here. A coalition structure Γ is called

- *perfect* if each player i weakly prefers $\Gamma(i)$ to every other coalition containing i;
- *individually rational* if each player $i \in A$ weakly prefers $\Gamma(i)$ to being alone in $\{i\}$;
- *Nash stable* if for each player $i \in A$ and for each coalition $C \in \Gamma \cup \{\emptyset\}$, $\Gamma(i) \succeq_i C \cup \{i\}$ (that is, no player wants to move to another coalition);
- *individually stable* if for each player $i \in A$ and for each coalition $C \in \Gamma \cup \{\emptyset\}$, it holds that $\Gamma(i) \succeq_i C \cup \{i\}$ or there exists a player $j \in C$ such that $C \succ_j C \cup \{i\}$ (that is, no player can move to another coalition without making some player in the new coalition worse off);
- *contractually individually stable* if for each player $i \in A$ and for each coalition $C \in \Gamma \cup \{\emptyset\}$, it holds that $\Gamma(i) \succeq_i C \cup \{i\}$, or there exists a player $j \in C$ such that $C \succ_j C \cup \{i\}$, or there exists a player $k \in \Gamma(i) \setminus \{i\}$ such that $\Gamma(i) \succ_k \Gamma(i) \setminus \{i\}$ (that is, no player can move to another coalition without making some player in the new coalition or in the old coalition worse off);

- *core stable* if for each nonempty coalition $C \subseteq A$, there exists a player $i \in C$ such that $\Gamma(i) \succeq_i C$ (that is, no coalition blocks Γ);
- *strictly core stable* if for each nonempty coalition $C \subseteq A$, there exists a player $i \in C$ such that $\Gamma(i) \succ_i C$ or for each player $i \in C$, we have $\Gamma(i) \sim_i C$ (that is, no coalition weakly blocks Γ);
- *Pareto optimal* if for each coalition structure $\Delta \neq \Gamma$, there exists a player $i \in A$ such that $\Gamma(i) \succ_i \Delta(i)$ or for each player $j \in A$, we have $\Gamma(j) \sim_j \Delta(j)$ (that is, no other coalition structure *Pareto-dominates* Γ);
- *popular* if for each coalition structure $\Delta \neq \Gamma$, the number of players i with $\Gamma(i) \succ_i \Delta(i)$ is at least as large as the number of players j with $\Delta(j) \succ_j \Gamma(j)$;
- *strictly popular* if for each coalition structure $\Delta \neq \Gamma$, the number of players i with $\Gamma(i) \succ_i \Delta(i)$ is larger than the number of players j with $\Delta(j) \succ_j \Gamma(j)$.[2]

2.3 Complexity Theory

When studying computational aspects of stability in hedonic games, there are two natural questions that arise. Let γ be a stability concept such as those defined in Section 2.2. How hard is it to decide whether a given solution for a given game is γ-stable and how hard is it for a given game to decide whether there exists a γ-stable outcome? The former question is the so-called *verification* variant, which we formally state as follows:

γ-VERIFICATION	
Given:	A hedonic game H and a coalition structure Γ.
Question:	Is Γ stable in the sense of γ in H?

The latter question, on the other hand, is referred to as the *existence* problem, defined as follows:

γ-EXISTENCE	
Given:	A hedonic game H.
Question:	Is there a coalition structure that is stable in the sense of γ in H?

We assume the reader to be familiar with the complexity classes P and NP. For each stability concept γ, whenever the problem γ-VERIFICATION is in P then γ-EXISTENCE is in NP, by simply guessing a coalition structure and then testing whether it satisfies γ. There are, however, no further direct connections between these two problems with respect to their complexity. We refer the reader to the interesting and detailed survey by Woeginger (2013a) for further information.

In Section 4, we will show the above problems to be NP-hard for several stability concepts in FEN-hedonic games, and we will do so by reductions from the following well-known NP-complete problems (see Garey & Johnson, 1979).

EXACT-COVER-BY-THREE-SETS (X3C)	
Given:	A set $B = \{b_1, b_2, \ldots, b_{3m}\}$, $m > 1$, and a collection $\mathscr{S} = \{S_1, S_2, \ldots, S_n\}$ of subsets $S_i \subseteq B$ with $\|S_i\| = 3$ for each i, $1 \leq i \leq n$.
Question:	Is there a subcollection $\mathscr{S}' \subseteq \mathscr{S}$ such that each element of B occurs in exactly one set in $\mathscr{S}'$?

[2] This notion is adapted from the voting-theoretic term of *Condorcet winner*: Such a candidate wins an election if and only if she beats each other candidate in pairwise comparison by a (strict) majority of votes.

Note that the problem X3C remains NP-complete even if each element in B occurs in at most three sets in $\mathscr{S}$ (see Garey & Johnson, 1979).

CLIQUE	
Given:	An undirected graph $G = (V, E)$ and a positive integer k.
Question:	Is there a clique (i.e., a subset $V' \subseteq V$ such that each two vertices in V' are connected by an edge) of size at least k in G?

Beyond that we will encounter problems from the second level of the polynomial hierarchy, namely from $\Sigma_2^p = \mathrm{NP}^{\mathrm{NP}}$ and $\Pi_2^p = \mathrm{coNP}^{\mathrm{NP}}$ (see Meyer & Stockmeyer, 1972; Stockmeyer, 1976). For more background on computational complexity, the reader is referred to the textbooks by Papadimitriou (1995) and Rothe (2005).

3. Polarized Responsive Preferences and FEN-Hedonic Games

We now introduce the class of FEN-hedonic games, defined via the following three steps:

1. Similarly to the singleton-encoding and to hedonic games with $\mathscr{W}$-preferences (recall Section 2 for the formal definitions), we start with the assumption that each player $i \in A$ has preferences over the remaining players in $A \smallsetminus \{i\}$. These preferences will be formally defined in Section 3.1 and denoted by $\trianglerighteq_i^{+0-}$ for each $i \in A$.
2. To obtain a hedonic game, we have to lift these preferences over players to preferences over coalitions. We will do so by applying a polarized version of the *responsive extension principle*, which we will formally define in Section 3.2; "PR-extension" in Figure 1 stands for *polarized responsive extension*.
3. These preferences, denoted by $\succeq_i^{+0-}$ for each $i \in A$, can be incomplete in the sense that there might be pairs of coalitions for which $\succeq_i^{+0-}$ does not determine which coalition player i prefers. By specifying these missing comparisons, we can extend each $\succeq_i^{+0-}$ to complete preferences, which we will collect in the set $\mathrm{Ext}\left(\succeq_i^{+0-}\right)$. With these complete preference extensions $\succeq_i \in \mathrm{Ext}\left(\succeq_i^{+0-}\right)$, we will then define the class of FEN-hedonic games in Section 3.3.

Figure 1 provides an overview of the just described tripartite procedure.

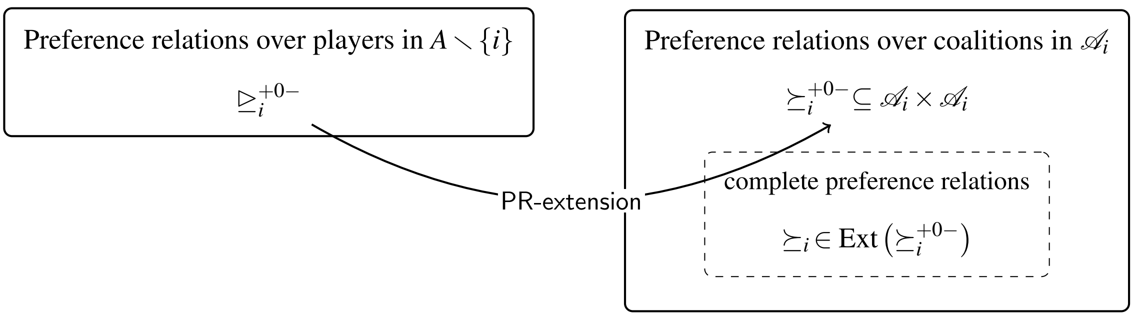


Figure 1: The process of defining the class of FEN-hedonic games for a fixed player $i \in A$

3.1 Preferences over Players: Ordinal Preferences with Double Threshold

Polarized responsive preferences are a combination of the singleton encoding and the friend- and enemy-oriented encoding with the additional degree of freedom that not all co-players have to be categorized as either friends or enemies. Furthermore, the players have the possibility to provide a ranking of their friends and of their enemies which allows a very fine-grained expression of their opinion. We formalize this intuition in the following definition.

Definition 1 (weak ranking with double threshold) *Let $A = \{1,2,\dots,n\}$ be a set of agents. For each $i \in A$, a* weak ranking with double threshold for agent i, *denoted by $\trianglerighteq_i^{+0-}$, consists of a partition of $A \setminus \{i\}$ into three sets:*

- *A_i^+ (i's friends), together with a weak order $\trianglerighteq_i^+$ over A_i^+,*
- *A_i^- (i's enemies), together with a weak order $\trianglerighteq_i^-$ over A_i^-, and*
- *A_i^0 (the neutral agents, i.e., the agents i does not care about).*

We also write $\trianglerighteq_i^{+0-}$ as $(\trianglerighteq_i^+ \,|\, A_i^0 \,|\, \trianglerighteq_i^-)$. Not having an order of the neutral agents can be interpreted as being indifferent about them all, so it holds that $j \sim_i k$ for all $j,k \in A_i^0$. Furthermore, we assume that each agent i strictly prefers all her friends to her neutral agents, and the neutral players to her enemies. The *weak order $\trianglerighteq_i$ induced by $\trianglerighteq_i^{+0-}$* is therefore defined as follows:

- $\trianglerighteq_i$ coincides with $\trianglerighteq_i^+$ on A_i^+;
- $f \rhd_i j$ for each $f \in A_i^+$ and $j \in A_i^0$;
- $j_1 \sim_i j_2 \sim_i \cdots \sim_i j_k$ for $A_i^0 = \{j_1, j_2, \dots, j_k\}$;
- $j \rhd_i e$ for each $j \in A_i^0$ and $e \in A_i^-$; and
- $\trianglerighteq_i$ coincides with $\trianglerighteq_i^-$ on A_i^-.

For a set $X = \{a_1, a_2, \dots, a_x\} \subseteq A$ in player i's preference, the shorthand $X_{\sim_i}$ denotes that player i is indifferent between all players in X, so $a_1 \sim_i a_2 \sim_i \cdots \sim_i a_x$. Occasionally, we will drop subscript i and simply write $X_\sim$ for $X_{\sim_i}$ when i is clear from the context. Whenever player i's set of friends or enemies is empty, we will slightly abuse notation and let $\emptyset$ denote the empty preference $\trianglerighteq_i^+$ or $\trianglerighteq_i^-$.

Example 2 *Let $A = \{1,2,\dots,11\}$ and let $\trianglerighteq_1^{+0-} = (2 \rhd_1 3 \sim_1 4 \,|\, \{5,6,7\} \,|\, 8 \rhd_1 9 \sim_1 10 \rhd_1 11)$ be a weak ranking with double threshold. This means that player 1 likes 2, 3, and 4 (and prefers 2 to both 3 and 4, and is indifferent between 3 and 4); 1 does not care about 5, 6, and 7 (and is indifferent between them); and 1 does not like 8, 9, 10, and 11 (but still prefers 8 to 9 and 10, is indifferent between 9 and 10, and prefers 9 and 10 to 11). The weak order $\trianglerighteq_1$ induced by $\trianglerighteq_1^{+0-}$ is* $2 \rhd_1 3 \sim_1 4 \rhd_1 5 \sim_1 6 \sim_1 7 \rhd_1 8 \rhd_1 9 \sim_1 10 \rhd_1 11$.

Note that here the preference between a friend and a neutral player is strict because we assume below that a coalition containing a friend instead of a neutral player is preferred. Analogously, the preference between a neutral player and an enemy is strict because a player does not care about having a neutral player in a coalition but is less happy with having an enemy in the coalition instead.

3.2 Preferences over Coalitions: Generalizing Responsive Preferences

Starting from a weak ranking with double threshold $\trianglerighteq_i^{+0-}$, which provides a ranking of the players in $A \setminus \{i\}$ from player i's perspective, we want to deduce player i's preferences over coalitions she is contained in. To do so, we suggest the following generalization of the *responsive extension principle*.

Definition 3 (extended preference order) *Let $\trianglerighteq_i^{+0-}$ be a weak ranking with double threshold for agent i. The* extended preference order $\succeq_i^{+0-}$ *is defined as follows. For every pair of coalitions $X,Y \in \mathscr{A}_i$, we have that $X \succeq_i^{+0-} Y$ if and only if the following two conditions hold: (1) There is an injective function σ from $Y \cap A_i^+$ to $X \cap A_i^+$ such that for every $y \in Y \cap A_i^+$, we have $\sigma(y) \trianglerighteq_i y$. (2) There is an injective function θ from $X \cap A_i^-$ to $Y \cap A_i^-$ such that for every $x \in X \cap A_i^-$, we have $x \trianglerighteq_i \theta(x)$. Finally, $X \succ_i^{+0-} Y$ if and only if $X \succeq_i^{+0-} Y$ and not $Y \succeq_i^{+0-} X$, and $X \sim^{+0-} Y$ if and only if $X \succeq^{+0-} Y$ and $Y \succeq^{+0-} X$.*

Intuitively speaking, for a fixed coalition, adding a further friend makes the coalition strictly more valuable, while adding an enemy causes the opposite. When exchanging two friends, the valuation of the coalition changes depending on the relation between the exchanged players (the same holds when two enemies are exchanged). When both a friend and an enemy are added or when they both are removed, the original and the new coalition are incomparable with respect to the responsive extension principle.

Thus, to construct the *polarized responsive extension* (PR-extension, for short) for a player *i*, we start with the coalition consisting of *i* and her friends (which is *i*'s most preferred coalition) and then construct all directly comparable coalitions by adding enemies, removing friends, or exchanging enemies or friends. For each newly obtained coalition, we repeat this procedure until we reach *i*'s least preferred coalition consisting of *i* and all of *i*'s enemies. Note that the elements of A_i^0 are disregarded, as their addition to or removal from a coalition does not change the coalition's value for *i*. The following examples illustrate the just presented extension principle.

Example 4 *For $A = \{1,2,\ldots,6\}$, consider the weak ranking with double threshold of player 1 given by $\trianglerighteq_1^{+0-} = (2 \triangleright_1 3 \sim_1 4 \,|\, \emptyset \,|\, 5 \triangleright_1 6)$. The graph in Figure 2 shows the partial order obtained from the polarized responsive extension of this preference, where an arc from coalition X to coalition Y implies that $X \succ_1^{+0-} Y$. Hence, any path leading from X' to Y' implies $X' \succ_1^{+0-} Y'$, whereas coalitions that are not connected by a path, such as $\{1,2,3\}$ and $\{1,2,3,4,5\}$, are incomparable. Note that if there were additional players $j > 6$ in A considered as neutral by player 1, the general picture would be the same with additional indifferences between any $C \subseteq \{2,\ldots,6\}$ and $\{1\} \cup C \cup N$ for any $N \subseteq A \smallsetminus \{1,\ldots,6\}$. These indifferences would occur at each level and for each coalition.*

Example 5 *For $A = \{1,2,3,4,5\}$, let the first player's preferences be $\trianglerighteq_1^{+0-} = (2 \triangleright_1 3 \mid \emptyset \mid 4 \triangleright_1 5)$. The graph in Figure 3 shows the partial order obtained from the polarized responsive extension of this preference using the same notation as in Example 4.*

Intuitively, the relation between two coalitions C and D ($C \succ_i^{+0-} D$, or $D \succ_i^{+0-} C$, or $C \sim_i^{+0-} D$, or C and D are incomparable) from player *i*'s point of view can be determined by the characterization given in Proposition 6, which is inspired by the work of Aziz et al. (2015) and of Bouveret et al. (2010) who show characterizations for the original responsive order in the context of fair division. Essentially, the characterization of Proposition 6 tells us how the extended order $\succeq_i^{+0-}$ induced by a weak ranking with double threshold $\trianglerighteq_i^{+0-}$ for agent *i* depends on the number of *i*'s friends and enemies in the coalitions and on the ranking of, respectively, friends and enemies contained in them.

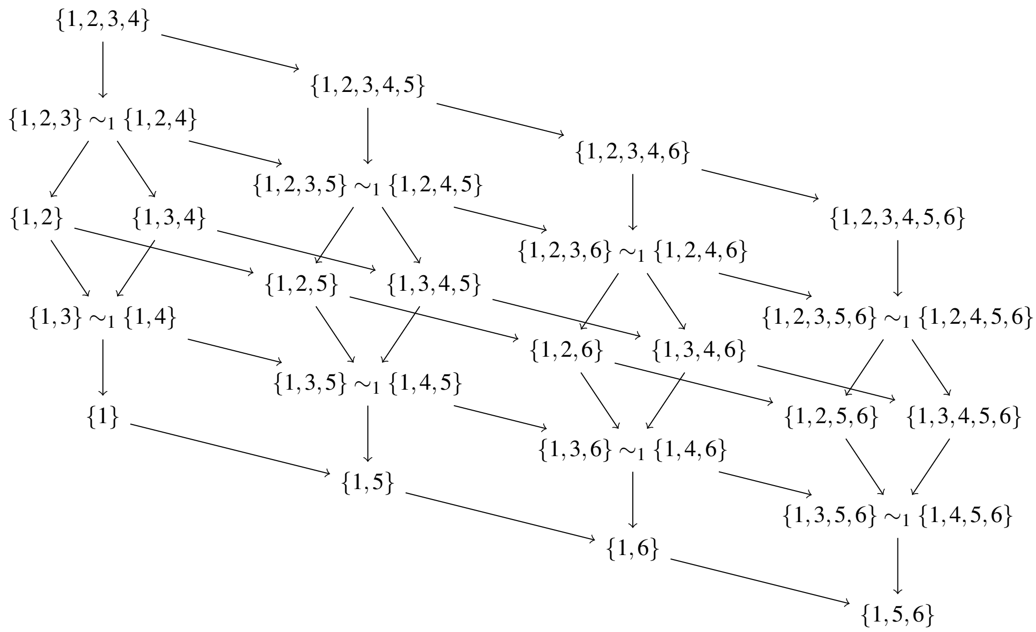


Figure 2: Partial order from the polarized responsive extension of $\trianglerighteq_1^{+0-} = (2 \rhd_1 3 \sim_1 4 \,|\, \emptyset \,|\, 5 \rhd_1 6)$

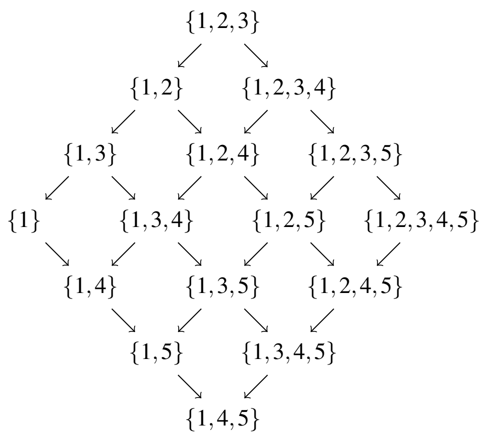


Figure 3: Partial order from the polarized responsive order of $\trianglerighteq_1^{+0-} = (2 \rhd_1 3 \,|\, \emptyset \,|\, 4 \rhd_1 5)$

Proposition 6 *Let $\trianglerighteq_i^{+0-}$ be a weak ranking with double threshold for agent i, and let C and D be any two coalitions containing i. Consider the orders $f_1 \trianglerighteq_i f_2 \trianglerighteq_i \cdots \trianglerighteq_i f_\mu$ with $\{f_1, f_2, \ldots, f_\mu\} = C \cap A_i^+$ and $f'_1 \trianglerighteq_i f'_2 \trianglerighteq_i \cdots \trianglerighteq_i f'_{\mu'}$ with $\{f'_1, f'_2, \ldots, f'_{\mu'}\} = D \cap A_i^+$, as well as $e_1 \trianglerighteq_i e_2 \trianglerighteq_i \cdots \trianglerighteq_i e_\nu$ with $\{e_1, e_2, \ldots, e_\nu\} = C \cap A_i^-$ and $e'_1 \trianglerighteq_i e'_2 \trianglerighteq_i \cdots \trianglerighteq_i e'_{\nu'}$ with $\{e'_1, e'_2, \ldots, e'_{\nu'}\} = D \cap A_i^-$. Then $C \succeq_i^{+0-} D$ if and only if*

(a) $\mu \geq \mu'$ *and* $\nu \leq \nu'$,

(b) *for each* k, $1 \le k \le \mu'$, *it holds that* $f_k \trianglerighteq_i f'_k$, *and*
(c) *for each* ℓ, $1 \le \ell \le \nu$, *it holds that* $e_{\nu-\ell+1} \trianglerighteq_i e'_{\nu'-\ell+1}$.

Proof. If (a) to (c) hold, the two injective functions $\sigma : D \cap A_i^+ \to C \cap A_i^+$ and $\theta : C \cap A_i^- \to D \cap A_i^-$ mapping $f'_k \mapsto f_k$ for each k, $1 \le k \le \mu'$, and $e_{\nu-\ell+1} \mapsto e'_{\nu'-\ell+1}$ for each ℓ, $1 \le \ell \le \nu$, satisfy $\sigma(f'_k) \trianglerighteq_i f'_k$ and $e_{\nu-\ell+1} \trianglerighteq_i \theta(e_{\nu-\ell+1})$, for the same range of k and ℓ. On the other hand, if there are two injective functions with the desired requirements, (a) holds. If there were some k with $f'_k \triangleright_i f_k$ (or some ℓ with $e'_{\nu'-\ell+1} \triangleright_i e_{\nu-\ell+1}$), this would imply $\sigma(f'_k) = f_j$ for some $j < k$ (respectively, $\theta(e_{\nu-\ell+1}) = e'_{\nu-j+1}$ for some $j > \ell$). This, however, implies that either a requirement is violated for f'_1 (or e_ν), or that σ (or θ) is not injective, a contradiction. ❑

3.3 The Class of FEN-Hedonic Games

Now we define hedonic games where each player has *friends*, *enemies*, and *neutral* co-players, and preferences over the former two sets such that we can derive each player's preference relation as introduced in the previous section. We call them *FEN-hedonic games* and define them formally as follows.

Definition 7 (FEN-hedonic game) *A* FEN-hedonic game *is a pair* $H = (A, (\trianglerighteq_1^{+0-}, \dots, \trianglerighteq_n^{+0-}))$, *where* $A = \{1, 2, \dots, n\}$ *is a set of players, and* $\trianglerighteq_i^{+0-}$ *gives the weak ranking with double threshold of player* $i \in A$ *as defined in Definition 1.*

To obtain the players' preferences over coalitions, we use the polarized responsive extension that we defined in Section 3.2. Since these preference relations $\succeq_i^{+0-}$ can be incomplete, we consider their extensions to complete relations, which have to preserve both already defined strict comparisons and indifferences.

Definition 8 (possible and necessary (weak) preference) *A preference relation* $\succeq_i$ *over* $\mathscr{A}_i$ extends $\succeq_i^{+0-}$ *if (1)* $C \succ_i^{+0-} D$ *implies* $C \succ_i D$ *for all coalitions* $C, D \in \mathscr{A}_i$*; and (2)* $C \sim_i^{+0-} D$ *implies* $C \sim_i D$ *for all coalitions* $C, D \in \mathscr{A}_i$.

Let $\mathrm{Ext}\left(\succeq_i^{+0-}\right)$ *be the set of all complete preference relations extending* $\succeq_i^{+0-}$. *We say*

- i possibly weakly prefers C to D *if* $C \succeq_i D$ *for some* $\succeq_i \in \mathrm{Ext}\left(\succeq_i^{+0-}\right)$;
- i possibly prefers C to D *if* $C \succ_i D$ *(i.e.,* $C \succeq_i D$ *and not* $D \succeq_i C$*) for some* $\succeq_i \in \mathrm{Ext}\left(\succeq_i^{+0-}\right)$;
- i necessarily weakly prefers C to D *if* $C \succeq_i D$ *for all* $\succeq_i \in \mathrm{Ext}\left(\succeq_i^{+0-}\right)$*; and*
- i necessarily prefers C to D *if* $C \succ_i D$ *(i.e.,* $C \succeq_i D$ *and not* $D \succeq_i C$*) for all* $\succeq_i \in \mathrm{Ext}\left(\succeq_i^{+0-}\right)$.

Equivalently, the above definitions can also be formulated as follows: *i possibly weakly prefers C to D* if $C \succeq_i^{+0-} D$ or C and D are incomparable with respect to $\succeq^{+0-}$; *i possibly prefers C to D* if $C \succ_i^{+0-} D$ or C and D are incomparable with respect to $\succeq^{+0-}$; *i necessarily weakly prefers C to D* if $C \succeq_i^{+0-} D$; and *i necessarily prefers C to D* if $C \succ_i^{+0-} D$.

We will see that weak rankings with double threshold can have various complete extensions.

Example 9 *Consider the FEN-hedonic game whose players in $A = \{1,2,3\}$ have the following weak rankings with double threshold: $\trianglerighteq_1^{+0-} = (2 \rhd_1 3 \mid \emptyset \mid \emptyset)$, $\trianglerighteq_2^{+0-} = (3 \mid \emptyset \mid 1)$, and $\trianglerighteq_3^{+0-} = (1 \mid \{2\} \mid \emptyset)$. The polarized responsive orders are*

$$\{1,2,3\} \succ_1^{+0-} \{1,2\} \succ_1^{+0-} \{1,3\} \succ_1^{+0-} \{1\}$$

for player 1,

$$\begin{tikzcd}
& \{2,3\} \arrow[dl, "\succ_2^{+0-}"'] \arrow[dr, "\succ_2^{+0-}"] & \\
\{2\} \arrow[dr, "\succ_2^{+0-}"'] & & \{1,2,3\} \arrow[dl, "\succ_2^{+0-}"] \\
& \{1,2\} &
\end{tikzcd}$$

for player 2, and

$$\{1,3\} \sim_3^{+0-} \{1,2,3\} \succ_3^{+0-} \{3\} \sim_3^{+0-} \{2,3\}$$

for player 3. So, two preferences are already complete, and there are three complete preferences extending $\succeq_2^{+0-}$, one setting $\{2\} \succ_2 \{1,2,3\}$, another setting $\{2\} \sim_2 \{1,2,3\}$, and the third setting $\{1,2,3\} \succ_2 \{2\}$, leaving all other relations the same.

3.4 Optimistic and Pessimistic Preference Extensions

We will now introduce two (generally incomplete) preference extensions. For a given coalition C containing player i, we consider the *optimistic extension* $\succeq_i^{+C}$ of player i's preference which ranks C as high as possible, and the *pessimistic extension* $\succeq_i^{-C}$ which ranks C as low as possible. We will make intensive use of these extensions for stating characterizations of stability in FEN-hedonic games, considering the stability concepts defined in Section 2.2, and for deriving polynomial-time algorithms in Section 4.

Definition 10 (optimistic and pessimistic extension) *Let $C \in \mathscr{A}_i$ be a coalition containing player $i \in A$ and $\succeq_i^{+0-}$ be i's preference relation. We define the following two relations:*

- $R_i^{+C} = \succeq_i^{+0-} \cup \{(C,C') \mid C' \not\succeq_i^{+0-} C,\ C' \in \mathscr{A}_i\}$ *and*
- $R_i^{-C} = \succeq_i^{+0-} \cup \{(C',C) \mid C \not\succeq_i^{+0-} C',\ C' \in \mathscr{A}_i\}$.

Let $\succeq_i^{+C}$ be the transitive closure of R_i^{+C} and $\succeq_i^{-C}$ be the transitive closure of R_i^{-C}. The strict preference relations $\succ_i^{+C}$ and $\succ_i^{-C}$ and indifference relations $\sim_i^{+C}$ and $\sim_i^{-C}$ are defined as usual.

Note that $\succeq_i^{+C}$ and $\succeq_i^{-C}$ extend $\succeq_i^{+0-}$ where $\succeq_i^{+C}$ ranks C as high as possible and $\succeq_i^{-C}$ ranks C as low as possible. Therefore, we call $\succeq_i^{+C}$ the optimistic *and $\succeq_i^{-C}$ the* pessimistic extension of $\succeq_i^{+0-}$ with respect to $C \in \mathscr{A}_i$.

The following characterization follows directly from the definitions of R_i^{+C} and R_i^{-C}.

Observation 11 *Consider any $C,D,E \in \mathscr{A}_i$.*

(1) $(D,E) \in R_i^{+C}$ *if and only if* $D \succeq_i^{+0-} E$ *or* $E \not\succeq_i^{+0-} D = C$.

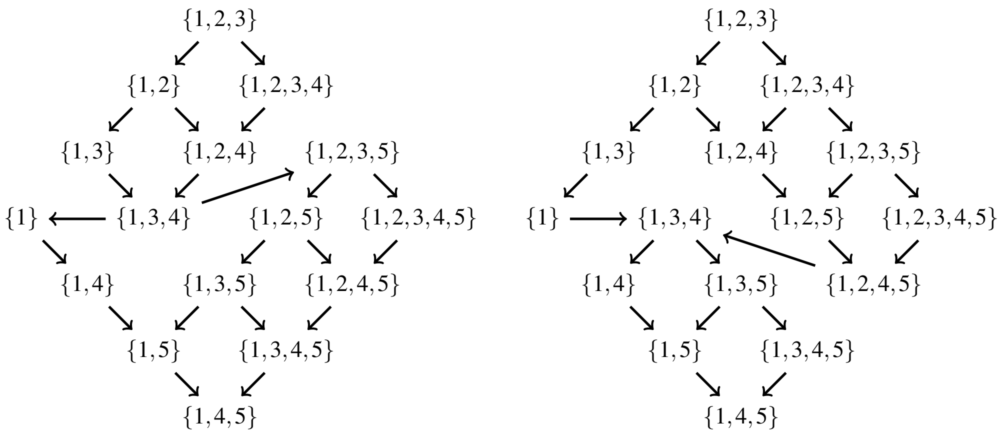


Figure 4: Optimistic extension $\succeq_1^{+\{1,3,4\}}$ (left) and pessimistic extension $\succeq_1^{-\{1,3,4\}}$ (right) of $\succeq_1^{+0-}$ from Example 5

(2) $(D,E) \in R_i^{-C}$ *if and only if* $D \succeq_i^{+0-} E$ *or* $C = E \not\succ_i^{+0-} D$.

Example 12 *Consider the FEN-hedonic game from Example 5 and let* $C = \{1,3,4\}$.

The relations $\succeq_1^{+C}$ *and* $\succeq_1^{-C}$ *are shown in Figure 4 where all arrows induce transitivity. As usual, we do not show the edges induced by reflexivity and transitivity.*

Proposition 13 *Let* $C, D, E \in \mathscr{A}_i$.

(1) $D \succeq_i^{+C} E$ *if and only if* $(D,E) \in R_i^{+C}$ *or* $(D,C),(C,E) \in R_i^{+C}$.

(2) $D \succeq_i^{-C} E$ *if and only if* $(D,E) \in R_i^{-C}$ *or* $(D,C),(C,E) \in R_i^{-C}$.

Proof. We only show (1); the proof for (2) is similar. If $C = D$ or $D = E$ or $C = E$, then (1) is obvious, so we asusme now that C, D and E are all different.

The implication from right to left is obvious.

To prove the implication from left to right, we show that $D \succeq_i^{+C} E$ and $(D,E) \notin R_i^{+C}$ imply $(D,C),(C,E) \in R_i^{+C}$. Assume that (a) $D \succeq_i^{+C} E$ and (b) $(D,E) \notin R_i^{+C}$. From (a) and the definition of $\succeq_i^{+C}$ as the transitive closure of R_i^{+C}, there are $m \geq 1$ and $C_1, \dots, C_m \in \mathscr{A}_i$ such that (c) $(D,C_1),(C_1,C_2),\dots,(C_m,E) \in R_i^{+C}$.

By taking the smallest such m, there is at most one $j \in \{1,\dots,m\}$ with $C = C_j$. Now, assume that $C \neq C_j$ for all $j \in \{1,\dots,m\}$. Because for $X \neq C$, $(X,Y) \in R_i^{+C}$ is equivalent to $X \succeq_i^{+0-} Y$, we have $D \succeq_i^{+0-} C_1 \succeq_i^{+0-} C_2 \succeq_i^{+0-} \dots \succeq_i^{+0-} C_m \succeq_i^{+0-} E$, hence $D \succeq_i^{+0-} E$, contradicting (b). Therefore, there must be exactly one $j \in \{1,\dots,m\}$ such that $C = C_j$.

If $j = 1$, then from (c) we directly get $(D,C) \in R_i^{+C}$. Since for $X \neq C$, $(X,Y) \in R_i^{+C}$ is equivalent to $X \succeq_i^{+0-} Y$, we furthermore get $C \succeq_i^{+0-} C_2 \succeq_i^{+0-} \dots \succeq_i^{+0-} C_m \succeq_i^{+0-} E$, hence $C \succeq_i^{+0-} E$, which implies $(C,E) \in R_i^{+C}$. Analogously, if $j = m$, we directly get $(C,E) \in R_i^{+C}$ and $D \succeq_i^{+0-} C_1 \succeq_i^{+0-} C_2 \succeq_i^{+0-} \dots \succeq_i^{+0-} C_{m-1} \succeq_i^{+0-} C$, which implies $(D,C) \in R_i^{+C}$. If $1 < j < m$, then (c) implies (d) $D \succeq_i^{+0-} C$, (e) $C_{j+1} \not\succeq_i^{+0-} C$, and (f) $C_{j+1} \succeq_i^{+0-} E$. (d) implies $(D,C) \in R_i^{+C}$. Assume now that $(C,E) \notin R_i^{+C}$, which is equivalent to $E \succ_i^{+0-} C$; together with (f) this implies $C_{j+1} \succ_i^{+0-} C$, contradicting (e). ❑

Proposition 14 *Consider any $C,D \in \mathscr{A}_i$.*

(1) $C \succeq_i^{+C} D$ if and only if $(C,D) \in R_i^{+C}$, which in turn holds if and only if $D \not\succ_i^{+0-} C$.

(2) $D \succeq_i^{+C} C$ if and only if $(D,C) \in R_i^{+C}$, which in turn holds if and only if $D \succeq_i^{+0-} C$.

(3) $C \succeq_i^{-C} D$ if and only if $(C,D) \in R_i^{-C}$, which in turn holds if and only if $C \succeq_i^{+0-} D$.

(4) $D \succeq_i^{-C} C$ if and only if $(D,C) \in R_i^{-C}$, which in turn holds if and only if $C \not\succ_i^{+0-} D$.

Proof. For each point the first equivalence follows by applying Proposition 13 and the second equivalence follows by applying Observation 11. ❑

Combining Observation 11, Proposition 13, and Proposition 14, we furthermore get:

Proposition 15 *Consider any $C,D,E \in \mathscr{A}_i$.*

(1) $D \succeq_i^{+C} E$ if and only if $D \succeq_i^{+0-} E$ or ($D \succeq_i^{+0-} C$ and $E \not\succ_i^{+0-} C$).

(2) $D \succeq_i^{-C} E$ if and only if $D \succeq_i^{+0-} E$ or ($C \not\succ_i^{+0-} D$ and $C \succeq_i^{+0-} E$).

Proof. We only show (1). (2) can be shown similarly.

By Proposition 13, we have $D \succeq_i^{+C} E$ if and only if $(D,E) \in R_i^{+C}$ or $(D,C),(C,E) \in R_i^{+C}$. By Observation 11, this is equivalent to $D \succeq_i^{+0-} E$ or $E \not\succ_i^{+0-} D = C$ or $(D,C),(C,E) \in R_i^{+C}$. By Proposition 14, this in turn is equivalent to $D \succeq_i^{+0-} E$ or $E \not\succ_i^{+0-} D = C$ or ($D \succeq_i^{+0-} C$ and $E \not\succ_i^{+0-} C$). Since $E \not\succ_i^{+0-} D = C$ implies ($D \succeq_i^{+0-} C$ and $E \not\succ_i^{+0-} C$), the condition can be shortened to $D \succeq_i^{+0-} E$ or ($D \succeq_i^{+0-} C$ and $E \not\succ_i^{+0-} C$). ❑

Proposition 16 *The strict relations $\succ_i^{+C}$ and $\succ_i^{-C}$ are acyclic.*

Proof. We give the proof for $\succ_i^{+C}$ only. The proof for $\succ_i^{-C}$ is similar.

Suppose that $\succ_i^{+C}$ contains a cycle $C_1 \succ_i^{+C} C_2 \succ_i^{+C} \cdots \succ_i^{+C} C_q \succ_i^{+C} C_1$. Because $\succeq_i^{+C}$ contains $\succ_i^{+C}$ and $\succeq_i^{+C}$ is the transitive closure of R_i^{+C}, R_i^{+C} contains a cycle of the form $C'_1, C'_2, \ldots, C'_k, C'_1$, with $C'_1 = C_1$ and $\{C_2, \ldots, C_q\} \subseteq \{C'_2, \ldots, C'_k\}$; in particular, $C_2 = C'_r$ for some $r \in \{2, \ldots, k\}$.

If $C \neq C'_j$ for all $j \in \{1, \ldots, k\}$, then because for $X \neq C$, $(X,Y) \in R_i^{+C}$ is equivalent to $X \succeq_i^{+0-} Y$, we have that $C_1 \sim_i^{+0-} C'_2 \sim_i^{+0-} \cdots \sim_i^{+0-} C'_k$, thus $C_1 \sim_i^{+0-} C'_r$, contradicting $C_1 \succ_i^{+C} C_2$.

Therefore, $C = C'_j$ and $C'_j \not\succeq_i^{+0-} C'_{j+1}$ for some $j \in \{1, \ldots, k\}$. Without loss of generality, let $j = 1$ be the only index with $C = C'_j$. Then $(C, C'_2), \ldots, (C'_k, C) \in R_i^{+C}$ implies $C \not\succeq_i^{+0-} C'_2$ and $C'_2 \succeq_i^{+0-} \cdots \succeq_i^{+0-} C'_k \succeq_i^{+0-} C$; however, by Proposition 14, $(C, C'_2) \in R_i^{+C}$ is equivalent to $C'_2 \not\succ_i^{+0-} C$, a contradiction. ❑

Proposition 17 *$\succeq_i^{+C}$ and $\succeq_i^{-C}$ are extensions of $\succeq_i^{+0-}$ with respect to Definition 8.*

Proof. We show that (1) and (2) from Definition 8 hold for $\succeq_i^{+C}$. The proof for $\succeq_i^{-C}$ is similar.

(1) Let $D \succ_i^{+0-} E$ for two coalitions $D, E \in \mathscr{A}_i$. This means that $D \succeq_i^{+0-} E$ and $E \not\succeq_i^{+0-} D$. By Proposition 15, $D \succeq_i^{+0-} E$ implies $D \succeq_i^{+C} E$. Assume, for the sake of contradiction, that $E \succeq_i^{+C} D$. Then, by Proposition 15, we have $E \succeq_i^{+0-} D$ or $(E \succeq_i^{+0-} C$ and $D \not\succ_i^{+0-} C)$. However, $E \succeq_i^{+0-} D$ does not hold because of $D \succ_i^{+0-} E$. $(E \succeq_i^{+0-} C$ and $D \not\succ_i^{+0-} C)$ does not hold either because $D \succ_i^{+0-} E \succeq_i^{+0-} C$ contradicts $D \not\succ_i^{+0-} C$. Hence, the assumption was false and $E \not\succeq_i^{+C} D$. With $D \succeq_i^{+C} E$ we have $D \succ_i^{+C} E$.

(2) Let $D \sim_i^{+0-} E$ for two coalitions $D, E \in \mathscr{A}_i$. This means that $D \succeq_i^{+0-} E$ and $E \succeq_i^{+0-} D$. Then, by Proposition 15, $D \succeq_i^{+0-} E$ implies $D \succeq_i^{+C} E$ and $E \succeq_i^{+0-} D$ implies $E \succeq_i^{+C} D$. Hence, $D \sim_i^{+C} E$. ❑

We furthermore observe that $\succeq_i^{+C}$ and $\succeq_i^{-C}$ are never undecided concerning C.

Observation 18 *For any two coalitions $C, D \in \mathscr{A}_i$, it holds that $C \succeq_i^{+C} D$ or $D \succeq_i^{+C} C$ and that $C \succeq_i^{-C} D$ or $D \succeq_i^{-C} C$.*

Proof. First, assume that $C \not\succeq_i^{+C} D$ and $D \not\succeq_i^{+C} C$. It follows by Proposition 14 that $D \succ_i^{+0-} C$ and that $D \not\succeq_i^{+0-} C$. This is a contradiction. Similarly, assume that $C \not\succeq_i^{-C} D$ and $D \not\succeq_i^{-C} C$. It then follows by Proposition 14 that $C \not\succeq_i^{+0-} D$ and $C \succ_i^{+0-} D$, which again is a contradiction. ❑

Proposition 19 *Consider any $C, D \in \mathscr{A}_i$.*

1. *There exists a complete extension $\succeq_i$ of $\succeq_i^{+0-}$ satisfying*
   - *(a) $C \succeq_i D$ if and only if $C \succeq_i^{+C} D$.*
   - *(b) $C \succ_i D$ if and only if $C \succ_i^{+C} D$.*
   - *(c) $D \succeq_i C$ if and only if $D \succeq_i^{-C} C$.*
   - *(d) $D \succ_i C$ if and only if $D \succ_i^{-C} C$.*
2. *All complete extensions $\succeq_i$ of $\succeq_i^{+0-}$ satisfy*
   - *(a) $C \succeq_i D$ if and only if $C \succeq_i^{-C} D$.*
   - *(b) $C \succ_i D$ if and only if $C \succ_i^{-C} D$.*
   - *(c) $D \succeq_i C$ if and only if $D \succeq_i^{+C} C$.*
   - *(d) $D \succ_i C$ if and only if $D \succ_i^{+C} C$.*

Proof. We only prove points 1(a) and (2)b. The proofs of all other points are similar.

First consider 1(a). From left to right, assume that $C \succeq_i^{+C} D$ does not hold. By Proposition 14, this implies $D \succ_i^{+0-} C$. Hence, every extension $\succeq_i$ of $\succeq_i^{+0-}$ satisfies $D \succ_i C$, which implies $C \not\succeq_i D$. From right to left, assume that $C \succeq_i^{+C} D$. Since $\succ_i^{+C}$ is acyclic by Proposition 16, $\succeq_i^{+C}$ can be extended to a complete preference relation. Consider such a complete extension $\succeq_i$ of $\succeq_i^{+C}$. It then holds by definition of the responsive extension that $C \succeq_i D$. Furthermore, since $\succeq_i^{+C}$ is an extension of $\succeq_i^{+0-}$ (and $\succeq_i$ is an extension of $\succeq_i^{+C}$), $\succeq_i$ is also an extension of $\succeq_i^{+0-}$.

Now consider 2(b). From left to right, assume that for all $\succeq_i \in \mathrm{Ext}\left(\succeq_i^{+0-}\right)$ we have $C \succ_i D$. Since $\succ_i^{-C}$ is acyclic by Proposition 16, $\succeq_i^{-C}$ can be extended to a complete preference relation.

Consider such an extension $\succeq_i \in \mathrm{Ext}\left(\succeq_i^{-C}\right)$. Then $\succeq_i$ is also an extension of $\succeq_i^{+0-}$. Hence, $C \succ_i D$ holds, which in turn means that $C \not\preceq_i D$. Since $\succeq_i$ is an extension of $\succeq_i^{-C}$, it follows that $C \not\preceq_i^{+C} D$. Since $\succeq_i^{-C}$ is never undecided concerning C (see Observation 18), it follows that $C \succ_i^{-C} D$. From right to left, assume that $C \succ_i^{-C} D$, i.e., $C \succeq_i^{-C} D$ and $D \not\succeq_i^{-C} C$. With Proposition 14(3) the latter implies $C \succ_i^{+0-} D$. Hence, for every extension $\succeq_i$ of $\succeq_i^{+0-}$, it holds that $C \succ_i D$. ❑

Finally, note that whether $(D,E) \in R_i^{+C}$ or $(D,E) \in R_i^{-C}$ holds for any three coalitions $C, D, E \in \mathscr{A}_i$ can be decided in polynomial time, since $D \succeq_i^{+0-} E$ and $D \succeq_i^{+0-} E$ can be decided in polynomial time by Proposition 6. Furthermore, $D \succeq_i^{+C} E$ and $D \succeq_i^{-C} E$ can be decided in polynomial time by Proposition 15.

4. Stability in FEN-Hedonic Games

Now that we have defined our new games and have stated some axiomatic properties, we turn to the complexity of verifying stable outcomes or checking whether there exists one. We have seen that the preference extensions obtained by applying the polarized responsive extension principle to the players' weak rankings with double threshold can lead to incomplete preferences over coalitions. In this section, we consider one possibility to deal with these incomparabilities: We leave them open and consider every possible extension and then study the complexity of related decision problems.

4.1 Possible and Necessary Stability: Properties and Characterizations

We start with formally defining the notions of possible and necessary stability for games with incomplete preference extensions.

Definition 20 (possible and necessary stability) *Let γ be a stability concept for hedonic games, $(A, (\trianglerighteq_1^{+0-}, \dots, \trianglerighteq_n^{+0-}))$ be a FEN-hedonic game, and Γ be a coalition structure. Γ is said to be* possibly γ *if there exists a profile $(\succeq_1, \dots, \succeq_n)$ in $\times_{i=1}^{n} \mathrm{Ext}\left(\succeq_i^{+0-}\right)$ such that Γ satisfies γ in $(A, (\succeq_1, \dots, \succeq_n))$. Γ is said to be* necessarily γ *if for each $(\succeq_1, \dots, \succeq_n)$ in $\times_{i=1}^{n} \mathrm{Ext}\left(\succeq_i^{+0-}\right)$, Γ satisfies γ in $(A, (\succeq_1, \dots, \succeq_n))$.*

Observe first that there always is a necessarily individually rational coalition structure (namely, the coalition structure where every agent is alone). For each extension, there exists a Pareto optimal coalition structure (where different extensions may have different Pareto optimal coalition structures) so that there is always a possibly Pareto optimal coalition structure (see Theorem 30 in Section 4.4 for a formal proof). We can furthermore state the following characterizations for possible/necessary perfectness and individual rationality.

Proposition 21 *Consider a FEN-hedonic game $(A, (\trianglerighteq_1^{+0-}, \dots, \trianglerighteq_n^{+0-}))$ and a coalition structure Γ.*

1. *Γ satisfies necessary perfectness if and only if it satisfies possible perfectness, and both statements are equivalent to the following condition: For each player i, $A_i^+ \subseteq \Gamma(i)$ and $A_i^- \cap \Gamma(i) = \emptyset$, that is, all friends of i are in her coalition, and none of her enemies is.*
2. *Γ is possibly individually rational if and only if for each $i \in A$, $\Gamma(i)$ contains at least a friend of i's or no enemies of i's (i.e., only neutral agents).*
3. *Γ is necessarily individually rational if and only if for each $i \in A$, $\Gamma(i)$ does not contain any enemies of i's.*

Proof. The proofs of these three claims are easy:

1. By definition, a coalition structure is perfect if and only if each player is in one of her favorite coalitions, which in a FEN-hedonic game means that each player is together with all her friends and no enemies.
2. For each $i \in A$, i necessarily prefers $\{i\}$ to $\Gamma(i)$ if and only if $\Gamma(i)$ contains no friend and at least one enemy of i's.
3. For each $i \in A$, i possibly prefers $\{i\}$ to $\Gamma(i)$ if and only if $\Gamma(i)$ contains an enemy of i's.

This completes the proof. ❑

Example 22 *Consider the FEN-hedonic game from Example 9 with three players who can be partitioned in a total of five different coalition structures. Observe that there does not exist a possibly perfect coalition structure.*

While $\{\{1,2,3\}\}$ is possibly Nash stable, there does not exist a necessarily Nash stable coalition structure, as in each of the five different coalition structures, player 1 or player 2, at least possibly, wants to move to another coalition.

Coalition structure $\{\{1,2,3\}\}$ is possibly individually rational, but not necessarily individually rational due to player 2; $\{\{1,2\},\{3\}\}$ is not possibly individually rational; the other three coalition structures are necessarily individually rational.

We furthermore consider the case of a single agent i entering or leaving a coalition. We state the following characterizations, which will be useful for individual stability and contractually individual stability.

Observation 23 *Consider a FEN-hedonic game $(A,(\trianglerighteq_1^{+0-},\dots,\trianglerighteq_n^{+0-}))$.*

1. *For a coalition $C \subseteq A$ and two players, $j \in C$ and $i \notin C$, it holds that $C \succ_j^{+0-} C \cup \{i\}$ if and only if $i \in A_j^-$. Hence, j possibly prefers C to $C \cup \{i\}$ if and only if j necessarily prefers C to $C \cup \{i\}$, which in turn holds if and only if i is an enemy of j's.*
2. *For a coalition $C \subseteq A$ and two players $i,k \in C$, it holds that $C \succ_k^{+0-} C \setminus \{i\}$ if and only if $i \in A_k^+$. Again, k possibly prefers C to $C \setminus \{i\}$ if and only if k necessarily prefers C to $C \setminus \{i\}$, which in turn holds if and only if k considers i as a friend.*

Example 24 *Again, consider the FEN-hedonic game from Example 9 and recall the definitions of individual stability and contractually individual stability.*

For $\{\{1,3\},\{2\}\}$, it holds that player 2 possibly wants to move to $\{1,3\}$, and 1 and 2 do not see 2 as an enemy, so necessary individual stability is not satisfied. Also, since in $\{2\}$ there is no other player who considers 2 a friend, necessary contractually individual stability is not satisfied either. Observe that this coalition structure is, however, possibly individually stable and therefore also possibly contractually individually stable.

Coalition structure $\{\{1\},\{2,3\}\}$ is not possibly individually stable, as player 3 wants to join the coalition $\{1\}$ and player 1 welcomes her. Player 2, however, considers 3 a friend and 1 an enemy. Therefore, as 2 does not want to move and would not welcome 1 to join the coalition $\{2,3\}$, this coalition structure is necessarily contractually individually stable.

To conclude this section, let us now state some easy observations about the connections among the notions of possible and necessary stability defined in Section 2.2. First, observe that if there exists a necessarily strictly popular coalition structure, it is unique, whereas there can be more than one possibly strictly popular coalition structure. Further, if there exists a necessarily strictly popular coalition structure, it is necessarily Pareto optimal. And if there exist possibly strictly popular coalition structures, each of them is possibly Pareto optimal. Finally, if there exists a unique perfect partition, it is always the unique necessarily strictly popular coalition structure.

On the other hand, a necessarily strictly popular coalition structure does not need to be possibly individually rational. Even if the possible core is nonempty, a necessarily strictly popular coalition structure does not need to be possibly core stable. And the same holds for the concepts of Nash stability, individual stability, contractual individual stability, and strict core stability.

4.2 Possible and Necessary Stability: Problem Definitions and Overview of Complexity Results

We are interested in properties of FEN-hedonic games and in characterizations of stability concepts in FEN-hedonic games. For some stability concepts γ, there sometimes (albeit not always) exists a coalition structure satisfying that concept (possibly or necessarily), i.e., some FEN-hedonic games have a (possibly or necessarily) γ-stable coalition structure and some have not. In these nontrivial cases, we ask how hard it is to decide whether for a given FEN-hedonic game a given coalition structure possibly or necessarily satisfies γ, and to decide whether there exists a coalition structure in a given FEN-hedonic game that possibly or necessarily satisfies γ. Similar questions are often analyzed in the context of hedonic games (Woeginger, 2013b; Aziz et al., 2013b; Rey et al., 2016). We now adapt the definition of the verification problem to the notions of possible and necessary verification, and we similarly adapt the definition of the existence problem to possible and necessary existence. Again, let γ be one of the previously defined stability concepts for hedonic games. Possible and necessary verification for γ are defined as follows.

POSSIBLE-γ-VERIFICATION	
Given:	A FEN-hedonic game H and a coalition structure Γ.
Question:	Does Γ possibly satisfy γ in H, that is, does Γ satisfy γ in some profile of preferences resulting from the polarized responsive extension of H?

NECESSARY-γ-VERIFICATION	
Given:	A FEN-hedonic game H and a coalition structure Γ.
Question:	Does Γ necessarily satisfy γ in H, that is, does Γ satisfy γ in all profiles of preferences resulting from the polarized responsive extension of H?

We now define the possible existence and the necessary existence problem. In the former problem, we ask whether there is some coalition structure satisfying the stability concept γ for some preference profile resulting from the polarized responsive extension, while in the latter problem the question is whether there is some coalition structure satisfying γ for all preference profiles resulting from the polarized responsive extension.

POSSIBLE-γ-EXISTENCE	
Given:	A FEN-hedonic game H.
Question:	Does there exist a coalition structure that possibly satisfies γ in H?

NECESSARY-γ-EXISTENCE	
Given:	A FEN-hedonic game H.
Question:	Does there exist a coalition structure that necessarily satisfies γ in H?

Note that the above definition of NECESSARY-γ-EXISTENCE is based on one possible interpretation. Another possible interpretation would change the question of the problem to asking whether for all profiles of preferences resulting from the polarized responsive extension of the game, there exists a coalition structure satisfying the stability concept γ. Observe that the existential and the universal quantifier have been swapped in this alternative interpretation. Consequently, a yes-instance for the latter problem is also a yes-instance for the former. We will illustrate this distinction by Example 25. However, we consider the interpretation underlying NECESSARY-γ-EXISTENCE to be more natural and therefore stick to it here: With the "for all, there exists" definition we would go beyond interpreting FEN-hedonic games as concise descriptions of hedonic games, since it would amount to consider a different hedonic game for each completion. The "there exists, for all" definition, by contrast, leads to the only natural interpretation when one takes FEN-hedonic game to be a concise description of hedonic game.[3]

Example 25 *Consider the following game with three players, $A = \{1,2,3\}$, and with $\trianglerighteq_1^{+0-} = (2 \mid \{3\} \mid \emptyset)$, $\trianglerighteq_2^{+0-} = (1 \mid \{3\} \mid \emptyset)$, and $\trianglerighteq_3^{+0-} = (1 \mid \emptyset \mid 2)$. We obtain the following polarized responsive orders: $\{1,2\} \sim_1 \{1,2,3\} \succ_1 \{1\} \sim_1 \{1,3\}$, $\{1,2\} \sim_2 \{1,2,3\} \succ_2 \{2\} \sim_2 \{2,3\}$, and $\{1,3\} \succ_3 \{3\} \succ_3 \{2,3\}$ and $\{1,3\} \succ_3 \{1,2,3\} \succ_3 \{2,3\}$, while $\{3\}$ and $\{1,2,3\}$ are incomparable for 3. Any coalition structure in which players 1 and 2 are not in the same coalition cannot possibly be Nash stable. On the one hand, $\{\{1,2\},\{3\}\}$ is Nash stable if and only if a preference extension provides $\{3\} \succeq_3 \{1,2,3\}$. On the other hand, $\{\{1,2,3\}\}$ is Nash stable if and only if $\{1,2,3\} \succeq_3 \{3\}$ in a preference extension. Thus, for every preference extension, there certainly exists a Nash stable coalition structure. However, there is no necessarily Nash stable coalition structure.*

Table 3 sums up the computational complexity results for possible and necessary stability verification and existence. In what follows, we will provide our complexity results for the various stability concepts from Section 2.2.

4.3 Perfectness and Individual Rationality

We now give a simple, polynomial characterization of necessary (and, equivalently, possibly) perfectness in FEN-hedonic games. Given a FEN-hedonic game $G = (A, (\trianglerighteq_1^{+0-}, \dots, \trianglerighteq_n^{+0-}))$, the friendship graph $F_G = (A, E)$ is the undirected graph whose set of vertices is A and that contains edge $\{i, j\}$ if and only if $i \in A_j^+$ or $j \in A_i^+$. We write $F_G(i) = \{i\} \cup \{j \mid \{i, j\} \in E\}$. Let F_G^* be the transitive closure of F_G.

[3] We thank an anonymous reviewer for this useful remark.

	VERIFICATION		EXISTENCE	
γ	POSSIBLE	NECESSARY	POSSIBLE	NECESSARY
perfectness	in P (Cor. 28)	in P (Cor. 28)	in P (Cor. 28)	in P (Cor. 28)
ind. rationality	in P (Prop. 29)	in P (Prop. 29)	in P (Prop. 29)	in P (Prop. 29)
Nash stability	in P (Thm. 39)	in P (Thm. 36)	NP-complete (Thm. 43)	NP-complete (Thm. 44)
ind. stability	in P (Thm. 40)	in P (Thm. 37)	in NP	NP-complete (Thm. 48)
contr. ind. stability	in P (Thm. 41)	in P (Thm. 38)	in P (Thm. 45)	in NP
core stability	coNP-complete (Thm. 54)	in coNP (Cor. 53)	in Σ_2^p	in Σ_2^p
str. core stability	coNP-complete (Thm. 54)	in coNP (Cor. 53)	in Σ_2^p	in Σ_2^p
Pareto optimality	coNP-complete (Thm. 35)	coNP-complete (Thm. 35)	in P (Thm. 30)	in Σ_2^p
popularity	coNP-complete (Thm. 61)	coNP-complete (Thm. 61)	in Σ_2^p	in Σ_2^p
str. popularity	coNP-complete (Thm. 61)	coNP-complete (Thm. 61)	coNP-hard, in Σ_2^p (Thm. 62)	coNP-hard, in Σ_2^p (Thm. 62)

Table 3: Overview of complexity results on verification and existence problems for possible and necessary stability notions in FEN-hedonic games

Proposition 26 *There is a necessarily perfect partition (and, equivalently, a possibly perfect partition) for G if for each i, j such that $i \in A_j^-$, F_G contains no path from i to j, and in that case a necessarily perfect partition for G can be computed by Algorithm 1.*

Algorithm 1: Computing a necessarily perfect partition for G according to Proposition 26

```
Let Γ be the set of equivalence classes for F*_G.
if some C ∈ Γ contains a pair (i, j) with i ∈ A⁻_j or j ∈ A⁻_i then
    return failure;
else
    return Γ;
```

Proof. If Algorithm 1 returns Γ, then we check that each agent is in (one of) her best possible coalition(s): For any i, because of the construction of the coalition $\Gamma(i)$ containing i, all friends of i must be in $\Gamma(i)$, and because the failure condition is not satisfied, no enemy of i is in $\Gamma(i)$. Therefore, Γ is necessarily perfect. (Note that there can be more than one perfect partition due to disconnected neutral agents.)

If there are two agents, say i and j, such that $j \in F_G^*(i)$ but $j \notin \Gamma(i)$, then by definition of $F_G^*(i)$ there must be two agents, k and k', such that $k \in F_G^*(i)$, $k' \notin F_G^*(i)$, and F_G contains (k,k'). Then either k or k' is not in her best possible coalition, and Γ is not necessarily perfect. Therefore, if Γ

is a necessarily perfect partition, then (1) for each i we have $\Gamma(i) \supseteq F^*_G(i)$. Now, if the algorithm returns failure, then for any Γ satisfying (1) there must be an agent who is in the same coalition as one of her enemies. Therefore, there exists no necessarily perfect partition. ❑

Example 27 *Let $A = \{1,2,3,4,5,6,7\}$, $A^+_1 = \{2\}$, $A^+_2 = \emptyset$, $A^+_3 = \{1,4\}$, $A^+_4 = \{2\}$, $A^+_5 = \emptyset$, $A^+_6 = \{5\}$, $A^+_7 = \emptyset$, $A^-_1 = \{6\}$, $A^-_2 = \{5\}$, $A^-_3 = A^-_4 = A^-_5 = A^-_7 = \emptyset$, $A^-_6 = \{1,7\}$. We start by constructing $F^*_G(k)$, which contains the equivalence classes $\{1,2,3,4\}$, $\{5,6\}$, and $\{7\}$. None of them contains two agents, i and j, with i being an enemy of j; therefore, $\Gamma = \{\{1,2,3,4\},\{5,6\},\{7\}\}$ is necessarily perfect. Note that $\Gamma = \{\{1,2,3,4,7\},\{5,6\}\}$ is also necessarily perfect (and there is no other necessarily perfect partition). If, on the other hand, we change A^+_2 into $A^+_2 = \{A_5\}$, then $F^*_G(k)$ contains the equivalence classes $\{1,2,3,4,5,6\},\{7\}$, and because 1 is an enemy of 6, there is no necessarily perfect partition.*

Corollary 28 *All variants of verification and existence problems regarding perfectness are in* P.

The same is true for all possible and necessary verification and existence problems with respect to individual rationality.

Proposition 29 *All variants of verification and existence problems regarding individual rationality are in* P.

Proof. Obviously, there always is a necessarily individually rational coalition structure (namely, the coalition structure where every agent is on her own). By the characterization in Proposition 21, possible and necessary individual rationality can be verified easily. ❑

4.4 Pareto Optimality

We now give some results for possible and necessary Pareto optimality and start with the existence problem for possible Pareto optimality.

Theorem 30 *There always exists a possibly Pareto optimal coalition structure in a FEN-hedonic game.*

Proof. We apply serial dictatorship. Let $\mathscr{C}_1$ be the set of all coalition structures most preferred by player 1 (which are all coalition structures in which 1's coalition contains all her friends and none of her enemies); and for each $j = 2,\dots,n$, let $\mathscr{C}_j$ be the set of all coalition structures Γ in $\mathscr{C}_{j-1}$ such that there exists no $\Gamma' \in \mathscr{C}_{j-1}$ with $\Gamma' \succ^{+0-}_j \Gamma$. Let Γ^* be an arbitrary coalition structure in $\mathscr{C}_n$. We claim that Γ^* is possibly Pareto optimal. Assume not: Let Γ' such that $\Gamma'(i) \succeq^{+0-}_i \Gamma^*(i)$ for all i, and $\Gamma'(j) \succ^{+0-}_j \Gamma^*(j)$ for some j. Let j^* be the smallest j with $\Gamma'(j) \succ^{+0-}_j \Gamma^*(j)$. Observe that $\Gamma' \in \mathscr{C}_{j^*-1}$. (If this wasn't the case then Γ^* wouldn't be in $\mathscr{C}_{j^*-1}$ either, because every player weakly prefers Γ' to Γ^*.) Then Γ^* should not have been included in $\mathscr{C}_{j^*}$ since $\Gamma' \in \mathscr{C}_{j^*-1}$ and $\Gamma'(j^*) \succ^{+0-}_{j^*} \Gamma^*(j^*)$: a contradiction. ❑

Consequently, POSSIBLE-PARETO-OPTIMALITY-EXISTENCE is in P. However, the same does not hold for NECESSARY-PARETO-OPTIMALITY-EXISTENCE since there exist both yes- and no-instances for this problem. The following example shows a FEN-hedonic games for which there is no necessarily Pareto optimal coalition structure.

Example 31 *Let $H = (A, \trianglerighteq^{+0-})$ be a FEN-hedonic game with $A = \{1,2,3,f\}$ and let $\trianglerighteq^{+0-}$ consist of the following weak rankings with double threshold: $\trianglerighteq_1^{+0-} = (\ f \mid 3 \mid 2\ )$, $\trianglerighteq_2^{+0-} = (\ f \mid 1 \mid 3\ )$, $\trianglerighteq_3^{+0-} = (\ f \mid 2 \mid 1\ )$, and $\trianglerighteq_f^{+0-} = (\emptyset \mid 1,2,3 \mid \emptyset)$.*

It then holds that there doesn't exist any necessarily Pareto optimal coalition structure, i.e., for every coalition structure Γ there is an extended profile $\succeq = (\succeq_1, \dots, \succeq_n)$ and a coalition structure $\Delta \neq \Gamma$ such that Δ Pareto-dominates Γ in $(A, \succeq)$.

First, every coalition structure, where some of 1, 2, *and* 3 *are in the same coalition, is possibly dominated by $\Delta = \{\{1\},\{2\},\{3\},\{f\}\}$. For example, if $\Gamma_1 = \{\{1,2,f\},\{3\}\}$, then we consider an extended profile with $\{1\} \succ_1 \{1,2,f\}$ and $\{2\} \succ_2 \{1,2,f\}$. These extensions exist since $\{1\}$ and $\{1,2,f\}$ are incomparable with respect to $\succeq_1^{+0-}$, and so they are for $\succeq_2^{+0-}$. However, Δ is also possibly dominated by Γ_1. To see this, we just consider an extended profile with $\{1,2,f\} \succ_1 \{1\}$ and $\{1,2,f\} \succ_2 \{2\}$. Finally, all remaining coalition structures have the form $\Gamma = \{\{i,f\},\{j\},\{k\}\}$ with $\{i,j,k\} = \{1,2,3\}$ and are possibly dominated by $\{\{j,i,f\},\{k\}\}$ where j is the player who sees i as an enemy.*

Now we turn to the verification variants.

Proposition 32 *Γ is not possibly Pareto optimal if and only if there exists a coalition structure Δ such that (1) for all $i \in A$ we have $\Delta(i) \succeq_i^{+\Gamma(i)} \Gamma(i)$, and (2) for some $i \in A$ we have $\Delta(i) \succ_i^{+\Gamma(i)} \Gamma(i)$.*

Proof. From right to left, assume there exists a coalition structure Δ such that (1) and (2) hold. Then, for any $\succeq = (\succeq_1, \dots, \succeq_n) \in \times_{i=1}^{n} \mathrm{Ext}\left(\succeq_i^{+0-}\right)$, by Propositions 19.2(c) and 19.2(d) we have $\Delta(i) \succeq_i \Gamma(i)$ for all $i \in A$, and $\Delta(i) \succ_i \Gamma(i)$ for some $i \in A$, which means that Δ Pareto-dominates Γ with respect to $\succeq$. Therefore, Γ is not possibly Pareto optimal.

From left to right, assume that Γ is not possibly Pareto optimal. Then, for any profile $\succeq \in \times_{i=1}^{n} \mathrm{Ext}\left(\succeq_i^{+0-}\right)$, there exists a coalition structure Δ that Pareto-dominates Γ with respect to $\succeq$. Consider a profile $(\succeq_1, \dots, \succeq_n) \in \times_{i=1}^{n} \mathrm{Ext}\left(\succeq_i^{+\Gamma(i)}\right)$. Such a profile exists because $\succ_i^{+\Gamma(i)}$ is acyclic due to Proposition 16. It then holds for every $i \in A$ that $\succeq_i$ also extends $\succeq_i^{+0-}$. Hence, there also is a coalition structure Δ for this profile that Pareto-dominates Γ, i.e., such that $\Delta(i) \succeq_i \Gamma(i)$ for all $i \in A$, and $\Delta(i) \succ_i \Gamma(i)$ for some $i \in A$. Since $\succeq_i$ extends $\succeq_i^{+\Gamma(i)}$ and with Observation 18 it follows that $\Delta(i) \succeq_i^{+\Gamma(i)} \Gamma(i)$ for every $i \in A$, and $\Delta(i) \succ_i^{+\Gamma(i)} \Gamma(i)$ for some $i \in A$. ❑

Similarly, we get the following result.

Proposition 33 *Γ is not necessarily Pareto optimal if and only if there exists a coalition structure Δ such that (1) for all $i \in A$ we have $\Delta(i) \succeq_i^{-\Gamma(i)} \Gamma(i)$, and (2) for some $i \in A$ we have $\Delta(i) \succ_i^{-\Gamma(i)} \Gamma(i)$.*

Proof. From right to left, assume there exists a coalition structure Δ such that (1) and (2) hold. Then, by Propositions 19.1(c) and 19.1(d), there is a $\succeq = (\succeq_1, \dots, \succeq_n) \in \times_{i=1}^{n} \mathrm{Ext}\left(\succeq_i^{+0-}\right)$ such that we have $\Delta(i) \succeq_i \Gamma(i)$ for all $i \in A$, and $\Delta(i) \succ_i \Gamma(i)$ for some $i \in A$, which means that Δ Pareto-dominates Γ with respect to $\succeq$. Therefore, Γ is not necessarily Pareto optimal.

From left to right, assume that Γ is not necessarily Pareto optimal. Then there exists a profile $(\succeq_1, \dots, \succeq_n) \in \times_{i=1}^{n} \mathrm{Ext}\left(\succeq_i^{+0-}\right)$ and a coalition structure Δ such that $\Delta(i) \succeq_i \Gamma(i)$ for all $i \in A$, and $\Delta(i) \succ_i \Gamma(i)$ for some $i \in A$. By Propositions 19.1(c) and 19.1(d), this means that $\Delta(i) \succeq_i^{-\Gamma(i)} \Gamma(i)$ for all $i \in A$, and $\Delta(i) \succ_i^{-\Gamma(i)} \Gamma(i)$ for some $i \in A$. ❑

By Propositions 32 and 33, we get coNP upper bounds for the two verification variants of Pareto optimality.

Corollary 34 POSSIBLE-PARETO-OPTIMALITY-VERIFICATION *and* NECESSARY-PARETO-OPTIMALITY-VERIFICATION *are in* coNP.

Theorem 35 POSSIBLE-PARETO-OPTIMALITY-VERIFICATION *and* NECESSARY-PARETO-OPTIMALITY-VERIFICATION *are* coNP*-complete.*

Proof. To show coNP-hardness of the first problem, we provide a reduction from the NP-complete problem X3C (which was defined in Section 2.3) to the *complement* of POSSIBLE-PARETO-OPTIMALITY-VERIFICATION. Letting $(B,\mathscr{S})$ with $|B| = 3m$ be a given X3C instance, we construct the following game. The set of players consists of two types of players which we call *element* and *set players*: $A = \{\beta_b \mid b \in B\} \cup \bigcup_{S\in\mathscr{S}} Q_S$. In the following, we describe them with their weak rankings with double threshold. Recall from Section 3.1 that for a set X of players, $X_{\sim_i}$ denotes that player i is indifferent between any pair of players $x,y \in X$. For the sake of readability, we will here and in some later proofs omit the subscript specifying the player i and simply write $X_\sim$ for $X_{\sim_i}$ when i is clear from the context. (Since players always are indifferent with respect to their neutral players, we even omit the subscript $\sim$ and simply write X when X are the neutral players.) We will also drop these subscripts within the weak rankings and write $\rhd$ instead of $\rhd_i$ when i is clear from the context. The set "$\{$other players$\}$" in these weak rankings refers to all players that are not mentioned.

Element players: For each $b \in B$, there is one element player β_b who has all set players (to be defined below) corresponding to sets S containing b as her highest ranked friends, followed by the remaining element players. There are no neutral players, so all other players are in her set of enemies. Formally, for each $b \in B$,

$$\trianglerighteq^{+0-}_{\beta_b} \quad = \quad \left(\bigcup_{\{S \mid b\in S\}} Q_{S\sim} \rhd \{\beta_{b'} \mid b' \neq b\}_\sim \;\middle|\; \emptyset \;\middle|\; \{\text{other players}\}_\sim \right).$$

Set players: For each $S \in \mathscr{S}$, we have $3m-3$ set players in the set $Q_S = \{\zeta_{S,k} \mid 1 \leq k \leq 3m-3\}$. For each fixed $S \in \mathscr{S}$, these players only consider the "next" set player in Q_S to be her friend (except for the last player in each Q_S). The remaining set players in their Q_S-set and the element players corresponding to the elements in the respective S are neutral players while all other players are their enemies. Formally, for each $S \in \mathscr{S}$ and for each k, $1 \leq k \leq 3m-4$,

$$\trianglerighteq^{+0-}_{\zeta_{S,k}} \quad = \quad \big(\{\zeta_{S,k+1} \mid k' \neq k\}_\sim \;\big|\; \{\zeta_{S,k'} \mid k \neq k' \neq k+1\} \cup \{\beta_b \mid b \in S\} \;\big|\; \{\text{other players}\}_\sim\big),$$
$$\trianglerighteq^{+0-}_{\zeta_{S,3m-3}} \quad = \quad \big(\emptyset \;\big|\; \{\zeta_{S,k'} \mid k' \neq k\} \cup \{\beta_b \mid b \in S\} \;\big|\; \{\text{other players}\}_\sim\big).$$

This profile can be determined in polynomial time and is visualized in Figure 5.

For each $S \in \mathscr{S}$, we define $P_S = Q_S \cup \{\beta_b \mid b \in S\}$ and fix $\Gamma = \{\{\beta_b \mid b \in B\}\} \cup \{Q_S \mid S \in \mathscr{S}'\}$ to be our coalition structure. It holds that $(B,\mathscr{S})$ belongs to X3C if and only if Γ is not possibly Pareto optimal.

Only if: Consider a solution $\mathscr{S}'$ for $(B,\mathscr{S})$, assuming there is one. The coalition structure $\Gamma' = \{P_S \mid S \in \mathscr{S}'\} \cup \{Q_S \mid S \notin \mathscr{S}'\}$ necessarily Pareto-dominates Γ: Each player $\zeta_{S,k}$, $S \in \mathscr{S}$, $1 \leq k \leq 3m-3$, is indifferent between Q_S and P_S, as β_b, $b \in S$, is considered as neutral. Furthermore,

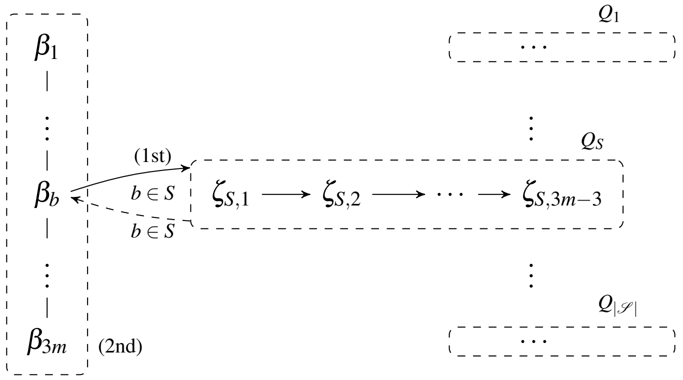


Figure 5: Network of friends for the construction in the proof of Theorem 35

each β_b, $b \in B$, necessarily strictly prefers P_S to $\Gamma(\beta_b)$, since two friends can be mapped to two indifferent friends, and $3m-3$ players can be mapped to higher ranked players, and β_b has got no enemies in either coalition.

If: Assume there exists a coalition structure Γ' that necessarily Pareto-dominates Γ, that is, $\Gamma'(i) \succeq_i \Gamma(i)$ for each player i, and $\Gamma'(j) \succ_j \Gamma(j)$ for at least one player j.

From the point of view of players $\zeta_{S,k}$, $S \in \mathscr{S}$, $1 \leq k \leq 3m-3$, the players in Q_S have to be together in one coalition in Γ' and without any enemies. A player β_b necessarily prefers P_S to $\{\beta_{b'} \mid b' \in B\}$ and the latter possibly to every other coalition containing β_b. Since Γ' necessarily Pareto-dominates Γ, there is a preference extension for which the only possible Γ' assigns each β_b, $b \in B$, to a P_S, which implies that there is an exact cover of B in $\mathscr{S}$.

For the second problem, NECESSARY-PARETO-OPTIMALITY-VERIFICATION, we can show coNP-hardness by slightly changing the construction: Now there are only $3m-2$ players $\zeta_{S,k}$ for each $S \in \mathscr{S}$ and each β_b, $b \in B$, prefers each $\beta_{b'}$, $b \neq b'$ to $\zeta_{S,k}$, $r \in S$, $1 \leq k \leq 3m-2$. Observe that with an analogous argumentation, changing the relations of possible and necessary preferences, $(B, \mathscr{S})$ is a positive instance of X3C if and only if Γ is not necessarily Pareto optimal.

The proof is complete, since POSSIBLE-PARETO-OPTIMALITY-VERIFICATION and NECESSARY-PARETO-OPTIMALITY-VERIFICATION are in coNP by Corollary 34. ❑

4.5 Verification for Nash Stability, Individual Stability, and Contractually Individual Stability

We now turn to the verification problems for Nash, individual, and contractually individual stability. We show that both possible and necessary verification are easy for these three concepts.

Algorithm 2: NECESSARY-NASH-STABILITY-VERIFICATION

Data: A FEN-hedonic game $(A, (\trianglerighteq_1^{+0-}, \dots, \trianglerighteq_n^{+0-}))$ and a coalition structure Γ.
Result: "YES" if Γ is necessarily Nash stable; "NO" otherwise.
1 **for** $i \in A$ **do**
2   **for** $C \in \Gamma \cup \{\emptyset\}$ **do**
3   **if** $\neg \Gamma(i) \succeq_i^{+0-} C \cup \{i\}$ **then**
4   output "NO";
5 output "YES";

Theorem 36 NECESSARY-NASH-STABILITY-VERIFICATION *is in* P.

Proof. Given a FEN-hedonic game $(A, \trianglerighteq^{+0-})$ and a coalition structure Γ, it is possible to determine whether Γ is necessarily Nash stable in polynomial time. This can be done by Algorithm 2.

For a given FEN-hedonic game $(A, (\trianglerighteq_1^{+0-}, \dots, \trianglerighteq_n^{+0-}))$, a coalition structure $\Gamma = \{C_1, \dots, C_k\}$, $1 \le k \le n$, is necessarily Nash stable if for all profiles $(\succeq_1, \dots, \succeq_n) \in \times_{i=1}^n \mathrm{Ext}\left(\succeq_i^{+0-}\right)$, all agents $i \in A$, and all coalitions $C \in \Gamma \cup \{\emptyset\}$ it holds that $\Gamma(i) \succeq_i C \cup \{i\}$. Therefore, we just need to check if we can extend $\succeq^{+0-}$ in such a way that $\Gamma(i) \succeq_i C \cup \{i\}$ does not hold for some $i \in A$ and $C \in \Gamma \cup \{\emptyset\}$. If this is the case, then Γ is not necessarily Nash stable.

Hence, we iterate all $i \in A$ and $C \in \Gamma \cup \{\emptyset\}$. There are four cases possible: (1) $\Gamma(i) \succ_i^{+0-} C \cup \{i\}$, (2) $\Gamma(i) \sim_i^{+0-} C \cup \{i\}$, (3) $\Gamma(i) \prec_i^{+0-} C \cup \{i\}$, or (4) $\Gamma(i)$ and $C \cup \{i\}$ are incomparable. In Cases (1) and (2), Nash stability is clearly not violated. Therefore, the algorithm just continues with the next iteration. If $\Gamma(i) \prec_i^{+0-} C \cup \{i\}$, this is clearly violating Nash stability and "NO" is output. If $\Gamma(i)$ and $C \cup \{i\}$ are incomparable, then it is possible to set $\Gamma(i) \prec_i C \cup \{i\}$ in the extension $\succeq_i$ of $\succeq_i^{+0-}$ such that Nash stability is violated. Accordingly, "NO" is output in this case. If "NO" is not output at any moment, then $\Gamma(i) \succeq_i^{+0-} C \cup \{i\}$ holds for all $i \in A$ and $C \in \Gamma \cup \{\emptyset\}$. Hence, "YES" is output.

The outer for-loop (line 1) runs exactly $|A| = n$ times. The inner for-loop (line 2) runs $|\Gamma \cup \{\emptyset\}| = k + 1 \le n + 1$ times. The relation between $\Gamma(i)$ and $C \cup \{i\}$ (line 3) can be checked in polynomial time by Proposition 6. Therefore, the whole algorithm runs in polynomial time. ❑

Note that Algorithm 2 basically checks whether $\Gamma(i) \succeq_i^{-\Gamma(i)} C \cup \{i\}$ holds for all $i \in A$ and $C \in \Gamma \cup \{\emptyset\}$, which is equivalent to Γ being necessarily Nash stable.

Theorem 37 NECESSARY-INDIVIDUAL-STABILITY-VERIFICATION *is in* P.

Proof. Algorithm 3 solves NECESSARY-INDIVIDUAL-STABILITY-VERIFICATION in polynomial time.

For a given FEN-hedonic game $(A, (\trianglerighteq_1^{+0-}, \dots, \trianglerighteq_n^{+0-}))$, a coalition structure Γ is necessarily individually stable if it holds for all profiles $(\succeq_1, \dots, \succeq_n) \in \times_{i=1}^n \mathrm{Ext}\left(\succeq_i^{+0-}\right)$, all agents $i \in A$, and all coalitions $C \in \Gamma \cup \{\emptyset\}$ that $\Gamma(i) \succeq_i C \cup \{i\}$ or that there is an agent $j \in C$ such that $C \succ_j C \cup \{i\}$. By Observation 23, the last term is equivalent to $i \in A_j^-$.

Therefore, we just need to check if we can extend $\succeq^{+0-}$ in such a way that this condition does not hold: If this is the case, then Γ is not necessarily individually stable. Hence, we iterate all $i \in A$ and $C \in \Gamma \cup \{\emptyset\}$. First, we check whether or not $\Gamma(i) \succeq_i^{+0-} C \cup \{i\}$ is true for the current i and C. If so, this i and C do not contradict individual stability in any extension of $\succeq_i^{+0-}$; and if this is not

Algorithm 3: NECESSARY-INDIVIDUAL-STABILITY-VERIFICATION

Data: A FEN-hedonic game $(A, (\trianglerighteq_1^{+0-}, \dots, \trianglerighteq_n^{+0-}))$ and a coalition structure Γ.
Result: "YES" if Γ is necessarily individually stable; "NO" otherwise.

1 **for** $i \in A$ **do**
2   **for** $C \in \Gamma \cup \{\emptyset\}$ **do**
3   **if** $\neg \Gamma(i) \succeq_i^{+0-} C \cup \{i\}$ **then**
4             $found \longleftarrow$ false;
5   **for** $j \in C$ **do**
6   **if** $i \in A_j^-$ **then**
7                     $found \longleftarrow$ true;
8   **if** $\neg found$ **then**
9   output "NO";
10 output "YES";

the case, there is an extension $\succeq_i$ of $\succeq_i^{+0-}$ such that $\Gamma(i) \prec_i C \cup \{i\}$. Then, in order for this i and C to not contradict individual stability, there has to be an agent $j \in C$ with $i \in A_j^-$. Therefore, we check this condition in lines 5 to 9. If such a $j \in C$ is found, i and C do not contradict individual stability in any extension, and we proceed to the next iteration; otherwise, we output "NO" because we have found i and C witnessing that individual stability does not hold in some extension. If "NO" was not output during any iteration, then the condition stated above has to be true for every profile $(\succeq_1, \dots, \succeq_n) \in \times_{i=1}^n \mathrm{Ext}\left(\succeq_i^{+0-}\right)$ and "YES" is output because Γ is necessarily individually stable.

The outer for-loop (line 1) runs exactly $|A| = n$ times. The second for-loop (line 2) runs $|\Gamma \cup \{\emptyset\}| \leq n+1$ times. The inner for-loop (line 5) runs $|C| \leq n$ times. The relation in line 3 can be checked in polynomial time by Proposition 6. Therefore, the whole algorithm runs in polynomial time. ❑

The same can be shown for the verification problem regarding necessary contractually individual stability.

Theorem 38 NECESSARY-CONTRACTUALLY-INDIVIDUAL-STABILITY-VERIFICATION *is in* P.

Proof. Algorithm 4 solves NECESSARY-CONTRACTUALLY-INDIVIDUAL-STABILITY-VERIFICATION in polynomial time.

For a given FEN-hedonic game $(A, (\trianglerighteq_1^{+0-}, \dots, \trianglerighteq_n^{+0-}))$, a coalition structure Γ is necessarily contractually individually stable if for all profiles $(\succeq_1, \dots, \succeq_n) \in \times_{i=1}^n \mathrm{Ext}\left(\succeq_i^{+0-}\right)$, all agents $i \in A$, and all coalitions $C \in \Gamma \cup \{\emptyset\}$ it holds that $\Gamma(i) \succeq_i C \cup \{i\}$ or there is an agent $j \in C$ with $C \succ_j C \cup \{i\}$ or there is an agent $k \in \Gamma(i)$ with $\Gamma(i) \succ_k \Gamma(i) \setminus \{i\}$. By Observation 23, $C \succ_j C \cup \{i\}$ is equivalent to $i \in A_j^-$ and $\Gamma(i) \succ_k \Gamma(i) \setminus \{i\}$ is equivalent to $i \in A_k^+$.

Thus, Algorithm 4 checks whether for all profiles $(\succeq_1, \dots, \succeq_n) \in \times_{i=1}^n \mathrm{Ext}\left(\succeq_i^{+0-}\right)$ and for every agent $i \in A$, one of the following two conditions holds: (a) there exists an agent $k \in \Gamma(i)$ who sees i as a friend; (b) for all coalitions $C \in \Gamma \cup \{\emptyset\}$, it holds that $\Gamma(i) \succeq_i C \cup \{i\}$ or there exists an agent $j \in C$ who sees i as an enemy. Analogously to the proof of Theorem 37, it is easy to see that Algorithm 4 runs in polynomial time. ❑

Algorithm 4: NECESSARY-CONTRACTUALLY-INDIVIDUAL-STABILITY-VERIFICATION

Data: A FEN-hedonic game $(A, (\trianglerighteq_1^{+0-}, \dots, \trianglerighteq_n^{+0-}))$ and a coalition structure Γ.
Result: "YES" if Γ is necessarily contractually individually stable; "NO" otherwise.

- **for** $i \in A$ **do**
  - *skiprest* $\longleftarrow$ false;
  - **for** $k \in \Gamma(i) \setminus \{i\}$ **do**
    - **if** $i \in A_k^+$ **then**
      - *skiprest* $\longleftarrow$ true;
  - **if** $\neg$*skiprest* **then**
    - **for** $C \in \Gamma \cup \{\emptyset\}$ **do**
      - **if** $\neg \Gamma(i) \succeq_i^{+0-} C \cup \{i\}$ **then**
        - *found* $\longleftarrow$ false;
        - **for** $j \in C$ **do**
          - **if** $i \in A_j^-$ **then**
            - *found* $\longleftarrow$ true;
        - **if** $\neg$*found* **then**
          - output "NO";
- output "YES";

We now turn to the verification problem for *possible* Nash, individual, and contractually individual stability, again showing that these problems can be solved efficiently. We start with possible Nash stability.

Theorem 39 POSSIBLE-NASH-STABILITY-VERIFICATION *is in* P.

Proof. The following algorithm solves POSSIBLE-NASH-STABILITY-VERIFICATION in polynomial time.

Algorithm 5: POSSIBLE-NASH-STABILITY-VERIFICATION

Data: A FEN-hedonic game $(A, (\trianglerighteq_1^{+0-}, \dots, \trianglerighteq_n^{+0-}))$ and a coalition structure Γ.
Result: "YES" if Γ is possibly Nash stable; "NO" otherwise.

- **for** $i \in A$ **do**
  - **for** $C \in \Gamma \cup \{\emptyset\}$ **do**
    - **if** $\Gamma(i) \prec_i^{+0-} C \cup \{i\}$ **then**
      - output "NO";
- output "YES";

A coalition structure Γ is possibly Nash stable if there is an extended profile $(\succeq_1, \dots, \succeq_n) \in \times_{i=1}^{n} \mathrm{Ext}\left(\succeq_i^{+0-}\right)$ such that for all agents $i \in A$ and all coalitions $C \in \Gamma \cup \{\emptyset\}$ we have $\Gamma(i) \succeq_i C \cup \{i\}$.

Hence, Algorithm 5 checks if this condition possibly holds. Again, there are four cases for the relation between $\Gamma(i)$ and $C \cup \{i\}$: We have either (1) $\Gamma(i) \succ_i^{+0-} C \cup \{i\}$, (2) $\Gamma(i) \sim_i^{+0-} C \cup \{i\}$, (3) $\Gamma(i) \prec_i^{+0-} C \cup \{i\}$, or (4) $\Gamma(i)$ and $C \cup \{i\}$ are incomparable.

In Cases (1), (2), and (3), the algorithm acts similarly as Algorithm 2. However, in Case (4), it is possible to set $\Gamma(i) \succeq_i C \cup \{i\}$ in the extension $\succeq_i$ of $\succeq_i^{+0-}$ such that Nash stability is not violated. Accordingly, the algorithm does not output “NO” in this case. Note that setting $\Gamma(i) \succeq_i C \cup \{i\}$ in Case (4) for multiple iterations can never result in a cyclic extension $\succeq_i$ because we just ensure that $\Gamma(i)$ is ranked better than all coalitions $C \in \Gamma \cup \{\emptyset\}$ that are incomparable to $\Gamma(i)$. The transitive closure of the resulting relation is the optimistic extension $\succeq_i^{+\Gamma(i)}$ as introduced in Section 3.4, which is acyclic by Proposition 16. Consequently, the choices made by Algorithm 5—each one individually allowed—to construct an extension can indeed be made simultaneously.

Basically, Algorithm 5 checks whether $\Gamma(i) \succeq_i^{+\Gamma(i)} C \cup \{i\}$ holds for all i and all C, which is equivalent to Γ being possibly Nash stable. Again, it is easy to see that the algorithm runs in polynomial time. ❑

Next, we turn to verifying possible individual stability.

Theorem 40 POSSIBLE-INDIVIDUAL-STABILITY-VERIFICATION *is in* P.

Proof. Algorithm 6 solves the problem in polynomial time.

Algorithm 6: POSSIBLE-INDIVIDUAL-STABILITY-VERIFICATION

Data: A FEN-hedonic game $(A, (\trianglerighteq_1^{+0-}, \dots, \trianglerighteq_n^{+0-}))$ and a coalition structure Γ.
Result: “YES” if Γ is possibly individually stable; “NO” otherwise.

for $i \in A$ **do**
  for $C \in \Gamma \cup \{\emptyset\}$ **do**
    if $\Gamma(i) \prec_i^{+0-} C \cup \{i\}$ **then**
      found $\longleftarrow$ false;
      for $j \in C$ **do**
        if $i \in A_j^-$ **then**
          found $\longleftarrow$ true;
      if $\neg$*found* **then**
        output “NO”;
output “YES” ;

A coalition structure Γ is possibly individually stable if there exists a profile $(\succeq_1, \dots, \succeq_n)$ in $\times_{i=1}^{n} \mathrm{Ext}\left(\succeq_i^{+0-}\right)$ such that it holds for all agents $i \in A$ and all coalitions $C \in \Gamma \cup \{\emptyset\}$ that $\Gamma(i) \succeq_i C \cup \{i\}$ or there exists an agent $j \in C$ who sees i as an enemy. Therefore, Algorithm 6 checks if this condition is true. Again, it is easy to see that the algorithm runs in polynomial time. ❑

Finally, we show that verifying possible contractually individual stability is easy.

Theorem 41 POSSIBLE-CONTRACTUALLY-INDIVIDUAL-STABILITY-VERIFICATION *is in* P.

Algorithm 7: POSSIBLE-CONTRACTUALLY-INDIVIDUAL-STABILITY-VERIFICATION

Data: A FEN-hedonic game $(A, (\trianglerighteq_1^{+0-}, \dots, \trianglerighteq_n^{+0-}))$ and a coalition structure Γ.
Result: "YES" if Γ is possibly contractually individually stable; "NO" otherwise.

for $i \in A$ **do**
    skiprest $\longleftarrow$ false;
    for $k \in \Gamma(i) \setminus \{i\}$ **do**
        if $i \in A_k^+$ **then**
            skiprest $\longleftarrow$ true;
    if $\neg$*skiprest* **then**
        for $C \in \Gamma \cup \{\emptyset\}$ **do**
            if $\Gamma(i) \prec_i^{+0-} C \cup \{i\}$ **then**
                found $\longleftarrow$ false;
                for $j \in C$ **do**
                    if $i \in A_j^-$ **then**
                        found $\longleftarrow$ true;
                if $\neg$*found* **then**
                    output "NO";
output "YES";

Proof. Algorithm 7 solves the problem in polynomial time.

As seen in the proof of Theorem 38, a coalition structure Γ is contractually individually stable for a profile $(\succeq_1, \dots, \succeq_n) \in \times_{i=1}^n \mathrm{Ext}\left(\succeq_i^{+0-}\right)$ exactly if for every agent $i \in A$, one of the following two conditions holds: (a) there exists an agent $k \in \Gamma(i)$ who sees i as a friend; (b) for all coalitions $C \in \Gamma \cup \{\emptyset\}$, it holds that $\Gamma(i) \succeq_i C \cup \{i\}$ or there exists an agent $j \in C$ who sees i as an enemy.

Hence, Algorithm 7 checks if (a) holds or $\succeq^{+0-}$ can be extended in such a way that (b) holds. Again, it is easy to see that the algorithm runs in polynomial time. ❑

We conclude this section with an example showing coalition structures that are possibly or necessarily Nash, individually, or contractually individually stable.

Example 42 *Consider the FEN-hedonic game $(A, (\trianglerighteq_1^{+0-}, \dots, \trianglerighteq_4^{+0-}))$ with $A = \{1,2,3,4\}$ and a coalition structure $\Gamma = \{\{1,2\},\{3,4\}\}$. The weak rankings with double threshold are:*

$$
\begin{aligned}
\trianglerighteq_1^{+0-} &= (2 \rhd_1^+ 3 \mid \emptyset \mid 4),\\
\trianglerighteq_2^{+0-} &= (1 \rhd_2^+ 3 \mid 4 \mid \emptyset),\\
\trianglerighteq_3^{+0-} &= (1 \mid 4 \mid 2),\\
\trianglerighteq_4^{+0-} &= (3 \mid \emptyset \mid 1 \sim_4^- 2).
\end{aligned}
$$

First, we can observe that none of 1, 2, *and* 4 *wants to move to another coalition. However, coalitions $\{3,4\}$ and $\{1,2,3\}$ are incomparable for player* 3 *in the responsive extension. Hence,* 3 *possibly wants to move to coalition $\{1,2\}$. This means that Γ is not necessarily Nash stable. This is* *detected by Algorithm 2 by testing if* $\Gamma(3) = \{3,4\} \succeq_3^{+0-} \{1,2\} \cup \{3\} = C \cup \{3\}$. *Since this is not the case, "NO" is output because player* 3 *possibly prefers* $\{1,2,3\}$ *to his coalition.*

On the other hand, Γ *is possibly Nash stable because no player necessarily wants to move (not even player* 3*). This means that* $\Gamma(i) \prec_i^{+0-} C \cup \{i\}$ *doesn't hold for any* $i \in A$ *and* $C \in \Gamma \cup \{\emptyset\}$, *which is exactly what Algorithm 5 checks.*

Since Γ *is possibly Nash stable, it is a fortiori possibly individually stable and possibly contractually individually stable. Necessary individual stability is not satisfied because* 3 *possibly wants to move to* $\{1,2\}$ *and she is not an enemy of* 1 *or* 2*. Hence, she can possibly deviate without any member of the new coalition rejecting her. Finally,* Γ *is necessarily contractually individually stable. This is so because* 3 *is the only player who possibly wants to move, but* 4 *doesn't want her to leave because* 3 *is a friend of* 4*'s.*

4.6 Existence for Nash Stability, Individual Stability, and Contractually Individual Stability

We now turn to the existence problems for Nash, individual, and contractually individual stability and first show that both possible and necessary existence are NP-complete for Nash stability.

Theorem 43 POSSIBLE-NASH-STABILITY-EXISTENCE *is* NP-*complete.*

Proof. The problem belongs to NP, since we can verify possibly Nash stable coalition structures in polynomial time by Theorem 39. Hence, we can nondeterministically guess a coalition structure and verify it in polynomial time.

NP-hardness can be shown via a polynomial-time many-one reduction from EXACT-COVER-BY-THREE-SETS (X3C), as defined in Section 2.3. Let $(B, \mathscr{S})$ be a given X3C instance, where B has $3m$ elements and $\mathscr{S}$ is a family of 3-element subsets $S \subseteq B$. Without loss of generality, it can be assumed that $m \geq 2$ and each element in B occurs at most three times in a set in $\mathscr{S}$. Given such an X3C instance, we construct the following game.[4] The set A consists of three types of players: $A = \{\alpha_i \mid 1 \leq i \leq 3m-1\} \cup \{\beta_b \mid b \in B\} \cup (\bigcup_{S \in \mathscr{S}} Q_S)$, which we call *chain*, *element*, and *set players* and which we describe below together with their weak rankings with double threshold.

Chain players: For each i, $1 \leq i \leq 3m-1$, we have one chain player α_i. Each of these players considers the "next" chain player to be her only friend, she is indifferent between the remaining chain players, and considers the other players to be enemies. Formally, for each i, $1 \leq i \leq 3m-2$,

$$\begin{aligned} \trianglerighteq_{\alpha_i}^{+0-} &= (\alpha_{i+1} \mid \{\alpha_j \mid i \neq j \neq i+1\} \mid \{\text{other players}\}_\sim), \\ \trianglerighteq_{\alpha_{3m-1}}^{+0-} &= (\ \emptyset \mid \{\alpha_j \mid j \neq 3m-1\} \mid \{\text{other players}\}_\sim). \end{aligned}$$

Element players: For each $b \in B$, there is one element player β_b. Each of these players considers all chain players to be her best ranked friends, followed by all set players (to be defined below) corresponding to sets $S \in \mathscr{S}$ containing b. The least preferred friends are the remaining element players. All other players are considered to be enemies, as there are no neutral players for β_b. Formally, for each $b \in B$,

[4] This construction is inspired by the construction of the proof that it is NP-hard to decide whether there exists a Nash stable coalition structure in an additively separable hedonic game (see Theorem 3 of Sung & Dimitrov, 2010). However, several adjustments are needed to make the construction work to prove Theorem 43.

$$\trianglerighteq^{+0-}_{\beta_b} = \left(\{\alpha_i \mid 1 \le i \le 3m-1\}_\sim \rhd \bigcup_{\{S \mid b \in S\}} Q_{S\sim} \rhd \{\beta_{b'} \mid b' \neq b\}_\sim \,\middle|\, \emptyset \,\middle|\, \{\text{other players}\}_\sim \right).$$

Set players: For each set $S \in \mathscr{S}$, we have one set $Q_S = \{\zeta_{S,k} \mid 1 \le k \le 3m-2\}$ containing $3m-2$ set players. The preferences of the set players in Q_S are similar to the chain players. Each player considers the "next" set player to be her only friend and is indifferent between all remaining set players corresponding to the same set and those element players corresponding to elements in S. All other players are considered to be enemies. Formally, for each $S \in \mathscr{S}$ and for each k, $1 \le k \le 3m-3$,

$$\begin{aligned}
\trianglerighteq^{+0-}_{\zeta_{S,k}} &= \left(\zeta_{S,k+1} \mid \{\zeta_{S,k'} \mid k \neq k' \neq k+1\} \cup \{\beta_b \mid b \in S\} \mid \{\text{other players}\}_\sim\right), \\
\trianglerighteq^{+0-}_{\zeta_{S,3m-2}} &= \left(\emptyset \mid \{\zeta_{S,k'} \mid k' \neq 3m-2\} \cup \{\beta_b \mid b \in S\} \mid \{\text{other players}\}_\sim\right).
\end{aligned}$$

Moreover, for each $S \in \mathscr{S}$, we denote the union of Q_S and all element players belonging to S with $P_S = \{\beta_b \mid b \in S\} \cup Q_S$. This profile can be constructed in polynomial time, since there are $n \le 3m + 3m + 3m \cdot (3m-2) = 9m^2$ players, and each player's preference can be written in time linear in n. Figure 6 gives a visualization of the profile.[5]

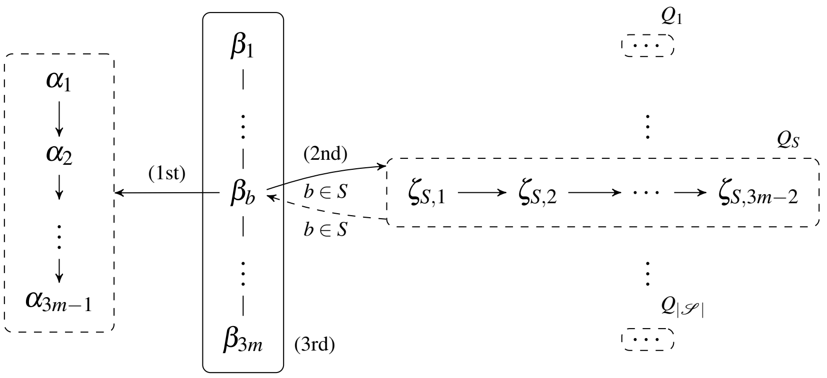


Figure 6: Network of friends for the construction in the proof of Theorem 43

We now show that $(B, \mathscr{S})$ is in X3C if and only if there exists a possibly Nash stable coalition structure in the PR-extension of the constructed game.

Only if: Assume there exists a solution $\mathscr{S}'$ for $(B, \mathscr{S})$. Consider the coalition structure $\Gamma = \{\{\alpha_i \mid 1 \le i \le 3m-1\}\} \cup \{P_S \mid S \in \mathscr{S}'\} \cup \{Q_S \mid S \notin \mathscr{S}'\}$. By a close look at all (possibly empty) coalitions in Γ it can be seen that no α_i, $1 \le i \le 3m-1$, and no $\zeta_{S,k}$, $S \in \mathscr{S}$, $1 \le k \le 3m-2$, wants to move, and each β_b, $b \in B$, possibly does not want to move due to the number of friends; thus Γ is possibly Nash stable.

If: Assume there is a possibly Nash stable coalition structure Γ. We start with the assignment of the chain players α_i to the coalitions in Γ. Let $C = \{\alpha_1, \ldots, \alpha_{3m-1}\}$ be the set containing all the

[5] In Figure 6 and in the upcoming Figures 7 and 5, a solid line/arrow represents a friendship relation (with priorities if required) and a dashed line/arrow stands for a neutral or a friendship relation. Also note that, for the sake of readability, we identify player names with vertex names in illustrations of networks.

chain players. We will argue that C has to be a coalition in Γ: First, Γ cannot contain any coalition that contains a strict subset of $C = \{\alpha_1, \dots, \alpha_{3m-1}\}$ because, as player α_{3m-2}'s only friend is α_{3m-1}, she will always want to move to the coalition α_{3m-1} is contained in. Hence, any coalition structure assigning these two players to different coalitions is not possibly Nash stable. For the same reasons, α_{3m-3} will always follow α_{3m-2}, and so on; thus no coalition structure Δ with $\Delta(\alpha_i) \neq \Delta(\alpha_{i+1})$, $1 \leq i \leq 3m-2$, is possibly Nash stable. Second, as soon as any other player is added to C, player α_{3m-1} necessarily prefers being alone to being in C. Thus Γ does not contain any strict superset of C. Therefore, it holds that $C \in \Gamma$. With an analogous argumentation we can show that for each $S \in \mathscr{S}$ the corresponding set players in Q_S have to be in the same coalition for Γ to be possibly Nash stable. It holds that for each $S \in \mathscr{S}$, each $\zeta_{S,k}$ follows $\zeta_{S,k+1}$ sequentially, $k = 3m-3, \dots, 1$, to a superset D_S of Q_S. D_S cannot contain any set players from other sets, that is, any $\zeta_{S',k'}$ with $S' \neq S$, $1 \leq k' \leq 3m-2$. Furthermore, D_S cannot contain any element player β_b with $b \notin S$, since $\zeta_{S,3m-2}$ is indifferent between everyone but her enemies and will deviate.

This leaves us only the following combinations to consider: For each $S \in \mathscr{S}$, Γ contains the coalition $D_S = Q_S \cup R_S$, where $R_S \subseteq \{\beta_b \mid b \in S\}$. If R_S contains one element player β_b, then there are fewer friends of β_b in coalition D_S than in C. If R_S contains two element players, β_b and $\beta_{b'}$, then both coalitions C and D_S contain the same number of β_b's friends. Since both coalitions contain at most $3m-1$ friends and enemies, and each friend in the first coalition is ranked higher than one in the latter, β_b necessarily prefers $C \cup \{\beta_b\}$ to D_S. Hence, either $D_S = Q_S$ or $D_S = P_S$. If an element player β_b was alone with other element players $\beta_{b'}$, which are in turn not in P_S, she would be with at most $3m-1$ friends, and would rather move to C with the same number of, but higher ranked, friends. This implies that for each $b \in B$, there exists an $S \in \mathscr{S}$ such that $\Gamma(\beta_b) = P_S$, so the coalitions the element players have to be assigned to induce an exact cover of B in $\mathscr{S}$. ❑

With a similar construction, we can show the following.

Theorem 44 NECESSARY-NASH-STABILITY-EXISTENCE *is* NP-*complete.*

Proof. The problem belongs to NP, since it can be verified in polynomial time in the number of players whether a nondeterministically chosen coalition structure is necessarily Nash stable by Theorem 36. NP-hardness can be shown similarly to the proof of Theorem 43. Given an X3C instance $(B, \mathscr{S})$, we again construct a game with three types of players: $A = \{\alpha_i \mid 1 \leq i \leq 3m\} \cup \{\beta_b \mid b \in B\} \cup \{\zeta_{S,k} \mid S \in \mathscr{S}, 1 \leq k \leq 3m-2\}$. In comparison to the proof of Theorem 43, we have an additional chain player, α_{3m}, and we change the order of friends for the element players β_b, $b \in B$.

Chain players: For each i, $1 \leq i \leq 3m$, we have one chain player α_i. Each of these players considers the "next" chain player to be her only friend, she is indifferent between the remaining chain players, and considers all other players to be enemies. Formally, for each i, $1 \leq i \leq 3m-1$,

$$\begin{aligned} \trianglerighteq^{+0-}_{\alpha_i} &= (\alpha_{i+1} \mid \{\alpha_j \mid i \neq j \neq i+1\} \mid \{\text{other players}\}_{\sim}), \\ \trianglerighteq^{+0-}_{\alpha_{3m}} &= (\emptyset \mid \{\alpha_j \mid j \neq 3m\} \mid \{\text{other players}\}_{\sim}). \end{aligned}$$

Element players: For each $b \in B$, there is one element player β_b. Each of these players considers those set players (to be defined below) corresponding to sets $S \in \mathscr{S}$ containing b to be her best friends, followed by all chain players and the remaining element players. There are no neutral players for β_b, so all other players are enemies. Formally, for each $b \in B$,

$$\trianglerighteq^{+0-}_{\beta_b} \quad = \quad \left(\bigcup_{\{S \mid b \in S\}} Q_{S\sim} \rhd \{\alpha_i \mid 1 \leq i \leq 3m-1\}_\sim \rhd \{\beta_{b'} \mid b' \neq b\}_\sim \;\middle|\; \emptyset \;\middle|\; \{\text{other players}\}_\sim \right).$$

Set players: For each set $S \in \mathscr{S}$, we have one set $Q_S = \{\zeta_{S,k} \mid 1 \leq k \leq 3m-2\}$ containing $3m-2$ set players. For a fixed set $S \in \mathscr{S}$, the preferences of the set players in Q_S are similar to those of the chain players. Each player considers the "next" set player to be her only friend and is indifferent between all remaining set players corresponding to the same set and those element players corresponding to the elements in S. All other players are considered to be enemies. Formally, for each $S \in \mathscr{S}$ and for each k, $1 \leq k \leq 3m-3$,

$$\trianglerighteq^{+0-}_{\zeta_{S,k}} \quad = \quad \big(\zeta_{S,k+1} \;\big|\; \{\zeta_{S,k'} \mid k \neq k' \neq k+1\} \cup \{\beta_b \mid b \in S\} \;\big|\; \{\text{other players}\}_\sim\big),$$
$$\trianglerighteq^{+0-}_{\zeta_{S,3m-2}} \quad = \quad \big(\,\emptyset \;\big|\; \{\zeta_{S,k'} \mid k' \neq 3m-2\} \cup \{\beta_b \mid b \in S\} \;\big|\; \{\text{other players}\}_\sim\big).$$

Again, we denote for each $S \in \mathscr{S}$ with $P_S = Q_S \cup \{\beta_b \mid b \in S\}$ the set containing all set players corresponding to S and those element players who are corresponding to the elements in S.

We show that $(B, \mathscr{S})$ is in X3C if and only if there exists a necessarily Nash stable coalition structure in the polarized responsive extension of the constructed game.

Only if: Assume that $\mathscr{S}'$ is a solution for $(B, \mathscr{S})$. Let $C = \{\alpha_1, \dots, \alpha_{3m}\}$ and consider the coalition structure $\Gamma = \{C\} \cup \{P_S \mid S \in \mathscr{S}'\} \cup \{Q_S \mid S \notin \mathscr{S}'\}$. No chain player α_i, $1 \leq i \leq 3m$, and no set player $\zeta_{S,k}$, $S \in \mathscr{S}$, $1 \leq k \leq m$, wants to leave her coalition. Each element player β_b, $b \in B$, now necessarily prefers being in P_S to joining any other existing or the empty coalition. Thus Γ is necessarily Nash stable.

If: Let Γ be a necessarily Nash stable coalition structure. Analogously to the argumentation in the proof of Theorem 43, the coalition $C = \{\alpha_1, \alpha_2, \dots, \alpha_{3m}\}$ has to be in Γ (as argued in the proof of Theorem 43), because of the preferences of the chain players α_i, $1 \leq i \leq 3m$. Furthermore, the coalition $D_S = Q_S \cup R_S$, where $R_S \subseteq \{\beta_b \mid b \in S\}$, has to be in Γ for each $S \in \mathscr{S}$ (as argued in the proof of Theorem 43). The set R_S cannot contain one or two players, since otherwise one element player $\beta_b \in R_S$ would possibly prefer moving to coalition C, which would contradict the assumption of necessary Nash stability of Γ. Consequently, either the coalition $D_S = Q_S$ or the coalition $D_S = P_S$ has to be in Γ for each $S \in \mathscr{S}$. An element player β_b outside of P_S in Γ would also imply a possible deviation to $C \cup \{\beta_b\}$. Thus all element players β_b with $b \in B$ are covered by disjoint sets P_S in Γ; hence, $\{S \mid P_S \in \Gamma\}$ induces a solution to the X3C instance $(B, \mathscr{S})$. ❑

We now turn to individual stability and contractually individual stability. We first show that possible existence is easy for contractually individual stability. Afterwards, we will show that necessary existence is NP-complete for individual stability.

Theorem 45 *There always exists a possibly contractually individually stable coalition structure in a FEN-hedonic game.*

Proof. This follows from Theorem 30 (which says that a possibly Pareto optimal coalition structure always exists in a FEN-hedonic game), together with the fact that Pareto optimality implies contractually individual stability. ❑

We will now show that deciding whether, given a FEN-hedonic game, there exists a necessarily individually stable coalition structure is NP-complete. To this end, Construction 46 is needed, and we briefly explain the ideas behind it. We will provide a polynomial-time many-one reduction from NECESSARY-NASH-STABILITY-EXISTENCE (NNSE, for short), which is NP-complete by Theorem 44. We take a FEN-hedonic game H that is an instance of NNSE and construct another FEN-hedonic game H', which is an instance of NECESSARY-INDIVIDUAL-STABILITY-EXISTENCE, such that there exists a necessarily individually stable coalition structure for H' if and only if there exists a necessarily Nash stable coalition structure for H.

In the upcoming construction, we define so-called *clone players* which have the same preferences as the original players (from H) but unlike the original players are not the enemy of any other player. By this trick we eliminate the possibility that other players can prevent the deviation of a clone player. Furthermore, we include so-called *structure players* to ensure that every necessarily individually stable coalition structure has to satisfy a certain form. Finally, so-called *friend* and *enemy players* help the structure players to fulfill their purpose.

Construction 46 *Let $H = (A, \trianglerighteq^{+0-})$ be a FEN-hedonic game with $A = \{1,\dots,n\}$, $\trianglerighteq^{+0-} = (\trianglerighteq_1^{+0-}, \dots, \trianglerighteq_n^{+0-})$ and $\trianglerighteq_i^{+0-} = (\trianglerighteq_i^{+}|A_i^0|\trianglerighteq_i^{-})$ for every $i \in A$, where $\trianglerighteq_i^{+}$ is the weak order over the set of i's friends A_i^+ and $\trianglerighteq_i^{-}$ is the weak order over the set of i's enemies A_i^-. We now construct a FEN-hedonic game H' in polynomial time. Let $H' = (A', \trianglerighteq^{+0-\prime})$ be a FEN-hedonic game with $A' = A \cup Clone \cup Structure \cup Friend_A \cup Friend_B \cup Enemy$, $Clone = \{c_1,\dots,c_n\}$, $Structure = \{s_1,\dots,s_n\}$, $Friend_A = \{a_1,\dots,a_n\}$, $Friend_B = \{b_1,\dots,b_n\}$, and $Enemy = \{e_1,\dots,e_n\}$. Furthermore, let $\trianglerighteq^{+0-\prime} = (\trianglerighteq_1^{+0-\prime}, \dots, \trianglerighteq_{e_n}^{+0-\prime})$ and, for $1 \le i \le n$, let*

$$
\begin{aligned}
\trianglerighteq_i^{+0-\prime} &= \left(\trianglerighteq_i^{+\prime} \mid A_i^{0\prime} \mid \trianglerighteq_i^{-\prime}\right) &&= \big(&& \big| A' \setminus \{i\} && \big| && \big),\\
\trianglerighteq_{c_i}^{+0-\prime} &= \left(\trianglerighteq_{c_i}^{+\prime} \mid A_{c_i}^{0\prime} \mid \trianglerighteq_{c_i}^{-\prime}\right) &&= \big(\trianglerighteq_i^{+} && \big| A' \setminus (A_i^+ \cup A_i^- \cup \{c_i\}) && \big| \trianglerighteq_i^{-} && \big),\\
\trianglerighteq_{s_i}^{+0-\prime} &= \left(\trianglerighteq_{s_i}^{+\prime} \mid A_{s_i}^{0\prime} \mid \trianglerighteq_{s_i}^{-\prime}\right) &&= \big(i \sim_{s_i} c_i \triangleright_{s_i} a_i \sim_{s_i} b_i && \big| A' \setminus \{i, c_i, s_i, a_i, b_i, e_i\} && \big| e_i && \big),\\
\trianglerighteq_{a_i}^{+0-\prime} &= \left(\trianglerighteq_{a_i}^{+\prime} \mid A_{a_i}^{0\prime} \mid \trianglerighteq_{a_i}^{-\prime}\right) &&= \big(b_i && \big| A' \setminus \{a_i, b_i\} && \big| && \big),\\
\trianglerighteq_{b_i}^{+0-\prime} &= \left(\trianglerighteq_{b_i}^{+\prime} \mid A_{b_i}^{0\prime} \mid \trianglerighteq_{b_i}^{-\prime}\right) &&= \big(e_i && \big| A' \setminus \{b_i, e_i\} && \big| && \big),\\
\trianglerighteq_{e_i}^{+0-\prime} &= \left(\trianglerighteq_{e_i}^{+\prime} \mid A_{e_i}^{0\prime} \mid \trianglerighteq_{e_i}^{-\prime}\right) &&= \big(&& \big| A' \setminus (A \cup \{e_i\}) && \big| 1 \sim_{e_i} \cdots \sim_{e_i} n && \big).
\end{aligned}
$$

This construction can obviously be done in polynomial time.

Example 47 *Consider the FEN-hedonic game $(A, (\trianglerighteq_1^{+0-}, \trianglerighteq_2^{+0-}, \trianglerighteq_3^{+0-}))$ with $A = \{1,2,3\}$ and a coalition structure $\Gamma = \{\{1\},\{2,3\}\}$. The weak rankings with double threshold are:*

$$
\begin{aligned}
\trianglerighteq_1^{+0-} &= (2 \triangleright_1^{+} 3 && \mid \emptyset && \mid \emptyset &&),\\
\trianglerighteq_2^{+0-} &= (3 && \mid 1 && \mid \emptyset &&),\\
\trianglerighteq_3^{+0-} &= (\emptyset && \mid \emptyset && \mid 2 \triangleright_3^{-} 1 &&).
\end{aligned}
$$

Note that Γ is not possibly or necessarily Nash stable because 1 wants to move to $\{2,3\}$. Γ is not possibly or necessarily individually stable because 3 wants to move to the empty coalition $\emptyset$ and there's no player in this empty coalition who could reject 3. Γ is possibly and necessarily contractually individually stable because there are two players who would like to deviate, 1 and 3, but for both there is another player prohibiting the deviation. 1 wants to move to $\{2,3\}$ but 3 doesn't want 1 to join. 3 wants to move to $\emptyset$ but 2 doesn't want her to leave $\{2,3\}$.

The FEN-hedonic game $H' = (A', \trianglerighteq^{+0-\prime})$ as stated in Construction 46 consists of $A' = \{1,2,3\} \cup \{c_1,c_2,c_3\} \cup \{s_1,s_2,s_3\} \cup \{a_1,a_2,a_3\} \cup \{b_1,b_2,b_3\} \cup \{e_1,e_2,e_3\}$, $\trianglerighteq^{+0-\prime} = (\trianglerighteq_1^{+0-\prime}, \ldots, \trianglerighteq_{e_3}^{+0-\prime})$ and for $i \in \{1,2,3\}$ the preferences are

$$\begin{array}{llll}
\trianglerighteq_i^{+0-\prime} = \big(& \big| A' \setminus \{i\} & \big| & \big), \\
\trianglerighteq_{c_1}^{+0-\prime} = \big(2 \rhd_{c_1}^{+} 3 & \big| A' \setminus \{2,3,c_1\}) & \big| & \big), \\
\trianglerighteq_{c_2}^{+0-\prime} = \big(3 & \big| A' \setminus \{3,c_2\}) & \big| & \big), \\
\trianglerighteq_{c_3}^{+0-\prime} = \big(& \big| A' \setminus \{1,2,c_3\}) & \big| 2 \rhd_{c_3}^{-} 1 & \big), \\
\trianglerighteq_{s_i}^{+0-\prime} = \big(i \sim_{s_i} c_i \rhd_{s_i} a_i \sim_{s_i} b_i & \big| A' \setminus \{i,c_i,s_i,a_i,b_i,e_i\} & \big| e_i & \big), \\
\trianglerighteq_{a_i}^{+0-\prime} = \big(b_i & \big| A' \setminus \{a_i,b_i\} & \big| & \big), \\
\trianglerighteq_{b_i}^{+0-\prime} = \big(e_i & \big| A' \setminus \{b_i,e_i\} & \big| & \big), \\
\trianglerighteq_{e_i}^{+0-\prime} = \big(& \big| A' \setminus \{1,2,3,e_i\} & \big| 1 \sim_{e_i} 2 \sim_{e_i} 3 & \big).
\end{array}$$

We will later consider the coalition structure $\Gamma' = \{D_C, E_C \mid C \in \Gamma\}$ with $D_C = \{j, c_j, s_j \mid j \in C\}$ and $E_C = \{a_j, b_j, e_j \mid j \in C\}$. Here, this means $\Gamma' = \{\{1,c_1,s_1\},\{a_1,b_1,e_1\},\{2,c_2,s_2,3,c_3,s_3\},\{a_2,b_2,e_2,a_3,b_3,e_3\}\}$. Due to the construction of H', it holds that Γ not being necessarily Nash stable implies that Γ' is not necessarily individually stable. Since 1 wants to move to $\{2,3\}$ in H, the corresponding clone player c_1 wants to move to $\{2,c_2,s_2,3,c_3,s_3\}$ in H'. And since there is no player who has c_1 as an enemy, no player could prevent the deviation.

Theorem 48 NECESSARY-INDIVIDUAL-STABILITY-EXISTENCE *(*NISE*, for short) is* NP*-complete.*

Proof. To see that NISE is in NP, let the FEN-hedonic game $H = (A, \trianglerighteq^{+0-})$ be a given instance. We nondeterministically guess a coalition structure $\Gamma \in \mathscr{C}_{(A,\trianglerighteq^{+0-})}$ that might be a solution for this instance. Then we check whether Γ indeed is a solution, i.e., whether Γ necessarily satisfies individual stability. This is possible in polynomial time by Theorem 37.

We show NP-hardness of NISE by providing a polynomial-time many-one reduction from NNSE. To do so, we consider the FEN-hedonic games H and H' as defined in Construction 46, where H is considered to be an instance of NNSE and H' an instance of NISE. Obviously, the construction of H' can be done in polynomial time.

We will now show that

$$H \in \mathrm{NNSE} \iff H' \in \mathrm{NISE}.$$

From left to right, assume that $H \in$ NNSE. This means that there exists a coalition structure $\Gamma \in \mathscr{C}_{(A,\trianglerighteq^{+0-})}$ such that for every extended profile $(\succeq_1, \ldots, \succeq_n) \in \times_{i=1}^{n} \mathrm{Ext}\left(\succeq_i^{+0-}\right)$, it holds that $(\forall i \in A)(\forall C \in \Gamma \cup \{\emptyset\})\left[\Gamma(i) \succeq_i C \cup \{i\}\right]$. Since this relation holds for every extended profile, it also has to hold for $(\succeq_1^{+0-}, \ldots, \succeq_n^{+0-})$. Hence, we have

$$(\forall i \in A)(\forall C \in \Gamma \cup \{\emptyset\})\left[\Gamma(i) \succeq_i^{+0-} C \cup \{i\}\right]. \tag{1}$$

We will now show that $H' \in$ NISE, i.e., that there is a coalition structure $\Gamma' \in \mathscr{C}_{(A',\trianglerighteq^{+0-\prime})}$ such that

$$(\forall i \in A')(\forall C' \in \Gamma' \cup \{\emptyset\})\left[\Gamma'(i) \succeq_i^{+0-\prime} C' \cup \{i\} \vee (\exists j \in C')[i \in A_j^{-\prime}]\right]. \tag{2}$$

We consider the coalition structure $\Gamma' = \{D_C, E_C \mid C \in \Gamma\}$ with $D_C = \{j, c_j, s_j \mid j \in C\}$ and $E_C = \{a_j, b_j, e_j \mid j \in C\}$. It then holds for all $i \in A$ that $\Gamma'(i) = \Gamma'(c_i) = \Gamma'(s_i) = \{j, c_j, s_j \mid j \in \Gamma(i)\}$ and $\Gamma'(a_i) = \Gamma'(b_i) = \Gamma'(e_i) = \{a_j, b_j, e_j \mid j \in \Gamma(i)\}$.

We will now show that (2) holds for all players in $A' = A \cup \mathit{Clone} \cup \mathit{Structure} \cup \mathit{Friend}_A \cup \mathit{Friend}_B \cup \mathit{Enemy}$. First, consider the players $i \in A$. It holds that $\Gamma'(i) \succeq_i^{+0-\prime} C' \cup \{i\}$ for all $C' \in \Gamma' \cup \{\emptyset\}$ because i doesn't have any friends or enemies and therefore is indifferent between any two coalitions. Hence, (2) is satisfied for all $i \in A$ and all $C' \in \Gamma' \cup \{\emptyset\}$.

Next, consider the clone players c_i. For all $D_C = \{j, c_j, s_j \mid j \in C\} \in \Gamma'$, it holds that $\Gamma'(c_i) = \{j, c_j, s_j \mid j \in \Gamma(i)\} \succeq_{c_i}^{+0-\prime} \{j, c_j, s_j \mid j \in C\} \cup \{c_i\} = D_C \cup \{c_i\}$ if and only if $\Gamma(i) \cup \{c_i\} \succeq_{c_i}^{+0-\prime} C \cup \{c_i\}$ because c_i is neutral to all other clone players and all structure players. This in turn is equivalent to $\Gamma(i) \setminus \{i\} \cup \{c_i\} \succeq_{c_i}^{+0-\prime} C \cup \{c_i\}$ because c_i is neutral to i. Since c_i has the same friends, order over friends, enemies, and order over enemies as player i has in H, the last preference relation is equivalent to $\Gamma(i) \succeq_i^{+0-} C \cup \{i\}$, which holds by assumption, see Equation (1). Hence, (2) is satisfied for all c_i and $D_C \in \Gamma'$.

Now, consider all $E_C = \{a_j, b_j, e_j \mid j \in C\} \in \Gamma'$. Again, $\Gamma'(c_i) = \{j, c_j, s_j \mid j \in \Gamma(i)\} \succeq_{c_i}^{+0-\prime} \{a_j, b_j, e_j \mid j \in C\} \cup \{c_i\} = E_C \cup \{c_i\}$ is equivalent to $\Gamma(i) \setminus \{i\} \cup \{c_i\} \succeq_{c_i}^{+0-\prime} \{c_i\}$ by removing all neutral players. This is equivalent to $\Gamma(i) \succeq_i^{+0-} \{i\}$, which holds by Equation (1). It is easy to see that the same argumentation is possible for the empty coalition $\emptyset$. Hence, (2) is satisfied for all $c_i \in \mathit{Clone}$ and all $C' \in \Gamma' \cup \{\emptyset\}$.

We now turn to the structure players s_i. $\Gamma'(s_i) = \{j, c_j, s_j \mid j \in \Gamma(i)\}$ contains s_i's two best friends, i and c_i, and no enemy. Every other coalition can only contain at most two other friends of s_i, namely a_i and b_i, which are ranked lower than i and c_i. Hence, s_i prefers $\Gamma'(s_i)$ to every other coalition in $\Gamma' \cup \{\emptyset\}$ and (2) is satisfied for all $s_i \in \mathit{Structure}$.

For all $a_i \in \mathit{Friend}_A$, it holds that $\Gamma'(a_i) = \{a_j, b_j, e_j \mid j \in \Gamma(i)\} \succeq_{a_i}^{+0-\prime} C' \cup \{a_i\}$ for every coalition $C' \in \Gamma' \cup \{\emptyset\}$ because $\Gamma'(a_i)$ contains b_i (a_i's only friend) and no enemies. Therefore, (2) holds for all $a_i \in \mathit{Friend}_A$. Analogously, (2) also holds for all $b_i \in \mathit{Friend}_B$. Finally, consider the enemy players e_i. Since e_i has no friends and $\Gamma'(e_i) = \{a_j, b_j, e_j \mid j \in \Gamma(i)\}$ doesn't contain any enemies of e_i, it holds that $\Gamma'(e_i) \succeq_{e_i}^{+0-\prime} C' \cup \{e_i\}$ for every $C' \in \Gamma' \cup \{\emptyset\}$. So, (2) also holds for all $e_i \in \mathit{Enemy}$.

Thus (2) is satisfied for all players in A' and all $C' \in \Gamma' \cup \{\emptyset\}$, which means that Γ' is necessarily individually stable for H' and $H' \in$ NISE.

From right to left, assume that $H' \in$ NISE. Then, there is a $\Gamma' \in \mathscr{C}_{(A', \trianglerighteq^{+0-\prime})}$ such that (2) holds. Consider such a coalition structure Γ'. We will now show that Γ' necessarily needs to be of the following form because (2) couldn't hold otherwise:

$$\begin{aligned}&\Gamma' = \{D_C \mid C \in \Gamma\} \cup \{E_C \mid C \in \Delta\} \text{ for some partitions } \Gamma \text{ and } \Delta \text{ of } A,\\ &\text{where } D_C = \{j, c_j, s_j \mid j \in C\} \text{ and } E_C = \{a_j, b_j, e_j \mid j \in C\}.\end{aligned}$$

Now consider any $i \in A$. First, note that none of c_i, s_i, a_i, and b_i are the enemy of any other player, which is why the first part of (2) has to hold for them, i.e., $\Gamma'(p) \succeq_p^{+0-\prime} C' \cup \{p\}$ for $p \in \{c_i, s_i, a_i, b_i\}$ and all $C' \in \Gamma' \cup \{\emptyset\}$. Furthermore, for player e_i and coalition $C' = \emptyset$, we have $\Gamma'(e_i) \succeq_{e_i}^{+0-\prime} \{e_i\}$ because there is no player in $\emptyset$ who could see e_i as an enemy.

Since a_i doesn't want to deviate from $\Gamma'(a_i)$, a_i has to be together with b_i because b_i is a_i's only friend and a_i has no enemies. Otherwise, a_i would always prefer the coalition containing b_i. For an analogous reason, b_i has to be together with e_i. Furthermore, i cannot be in the same coalition as e_i because i is an enemy of e_i and e_i would rather be alone otherwise. Hence, we already know that $\{a_i, b_i, e_i\} \subseteq E$ and $\{i\} \subseteq D$ for some $D, E \in \Gamma'$ with $D \neq E$.

There remain ten cases for the allocation of s_i and c_i. By excluding nine of these cases, it will follow that $s_i, c_i \in D$. Recall that $\Gamma'(s_i) \succeq_{s_i}^{+0-\prime} C' \cup \{s_i\}$ holds for all $C' \in \Gamma' \cup \{\emptyset\}$. All of the nine cases presented in the following imply that this is not true for at least one coalition $C' \in \Gamma' \cup \{\emptyset\}$. Hence, they cannot hold. For an overview of the cases, see Table 4.

	$s_i \in E$	$s_i \in D$	$s_i \in F$
$c_i \in E$	E, $D \cup \{s_i\}$ incomparable	D, $E \cup \{s_i\}$ incomparable	F, $E \cup \{s_i\}$ incomparable
$c_i \in D$	$E \prec_{s_i}^{+0-\prime} D \cup \{s_i\}$	holds	F, $E \cup \{s_i\}$ incomparable
$c_i \in F$	E, $D \cup \{s_i\}$ incomparable	D, $E \cup \{s_i\}$ incomparable	F, $E \cup \{s_i\}$ incomparable
$c_i \in G$	—	—	F, $E \cup \{s_i\}$ incomparable

Table 4: Ten cases for the allocation of s_i and c_i and why nine of them cannot hold

- If $s_i, c_i \in E$ (i.e., $\{c_i, s_i, a_i, b_i, e_i\} \subseteq E$ and $\{i\} \subseteq D$), then $\Gamma'(s_i) = E \not\succeq_{s_i}^{+0-\prime} D \cup \{s_i\}$. E and $D \cup \{s_i\}$ are incomparable with respect to $\succeq_{s_i}^{+0-\prime}$ because E contains more friends but also more enemies than $D \cup \{s_i\}$.

- If $s_i \in E$ and $c_i \in D$ (i.e., $\{s_i, a_i, b_i, e_i\} \subseteq E$ and $\{i, c_i\} \subseteq D$), then $\Gamma'(s_i) = E \prec_{s_i}^{+0-\prime} D \cup \{s_i\}$ because $D \cup \{s_i\}$ contains the same number of friends as E, but better friends than E, and no enemies.

- If $s_i \in E$ and $c_i \in F$ for an $F \in \Gamma'$ with $D \neq F \neq E$ (i.e., $\{s_i, a_i, b_i, e_i\} \subseteq E$, $\{i\} \subseteq D$, and $\{c_i\} \subseteq F$), then $\Gamma'(s_i) = E$ and $D \cup \{s_i\}$ are incomparable again because E contains more friends but also more enemies than $D \cup \{s_i\}$. Hence, $\Gamma'(s_i) \not\succeq_{s_i}^{+0-\prime} D \cup \{s_i\}$.

- If $s_i \in F$ for an $F \in \Gamma'$ with $D \neq F \neq E$ (i.e., $\{a_i, b_i, e_i\} \subseteq E$, $\{i\} \subseteq D$, and $\{s_i\} \subseteq F$), then there remain four cases for c_i: $c_i \in E$, $c_i \in D$, $c_i \in F$, or $c_i \in G$ for a $G \in \Gamma'$ with $G \notin \{D, E, F\}$. No matter where c_i is, $\Gamma'(s_i) = F$ and $E \cup \{s_i\}$ are incomparable with respect to $\succeq_{s_i}^{+0-\prime}$ because $E \cup \{s_i\}$ contains more friends but also more enemies than F.

- If $s_i \in D$ and $c_i \in E$ (i.e., $\{c_i, a_i, b_i, e_i\} \subseteq E$ and $\{i, s_i\} \subseteq D$), then $\Gamma'(s_i) = D$ and $E \cup \{s_i\}$ are incomparable for s_i because $E \cup \{s_i\}$ contains more friends but also more enemies than D.

- If $s_i \in D$ and $c_i \in F$ for an $F \in \Gamma'$ with $D \neq F \neq E$ (i.e., $\{a_i, b_i, e_i\} \subseteq E$, $\{i, s_i\} \subseteq D$, and $\{c_i\} \subseteq F$), then $\Gamma'(s_i) = D$ and $E \cup \{s_i\}$ again are incomparable for s_i.

The only remaining case is $s_i, c_i \in D$ (i.e., $\{a_i, b_i, e_i\} \subseteq E$ and $\{i, c_i, s_i\} \subseteq D$). Note that this case indeed fulfills $\Gamma'(s_i) \succeq_{s_i}^{+0-\prime} C' \cup \{s_i\}$ for all $C' \in \Gamma' \cup \{\emptyset\}$. Hence, for every $i \in A$, we have $\{a_i, b_i, e_i\} \subseteq E_i$ and $\{i, c_i, s_i\} \subseteq D_i$ for some $D_i, E_i \in \Gamma'$ with $D_i \neq E_i$. It furthermore holds for any $i, j \in A$ that $E_i \neq D_j$. Otherwise, we had $E_i = D_j \supseteq \{a_i, b_i, e_i, j, c_j, s_j\}$. Since j is an enemy of e_i, e_i would like to deviate to the empty coalition which is a contradiction to the assumption, see Equation (2). It follows that Γ' has the form presented above.

Finally, consider the clone players $c_i \in Clone$. Equation (2) also holds for c_i, i.e., we have

$$(\forall C' \in \Gamma' \cup \{\emptyset\})\left[\Gamma'(c_i) \succeq_{c_i}^{+0-\prime} C' \cup \{c_i\} \vee (\exists x \in C')[c_i \in A_x^{-\prime}]\right].$$

Since c_i is not the enemy of any other player, i.e., $c_i \notin A_x^{-\prime}$ for all $x \in A'$, it follows that

$$(\forall C' \in \Gamma' \cup \{\emptyset\})\left[\Gamma'(c_i) \succeq_{c_i}^{+0-\prime} C' \cup \{c_i\}\right]. \tag{3}$$

Recall that $\Gamma' = \{D_C \mid C \in \Gamma\} \cup \{E_C \mid C \in \Delta\}$ for some partitions Γ and Δ of A with $D_C = \{j, c_j, s_j \mid j \in C\}$ and $E_C = \{a_j, b_j, e_j \mid j \in C\}$. Furthermore, note that $\Gamma'(c_i) = D_{\Gamma(i)}$ and let $D_\emptyset = \emptyset$. Equation (3) in particular holds for all $C' = D_C \in \Gamma' \cup \{\emptyset\}$ with $C \in \Gamma \cup \{\emptyset\}$. Hence, we have

$$(\forall C \in \Gamma \cup \{\emptyset\})\left[D_{\Gamma(i)} \succeq_{c_i}^{+0-\prime} D_C \cup \{c_i\}\right]. \tag{4}$$

Because c_i is neutral to all players $x \in A'$ with $x \notin A, x \neq c_i$, we can remove all these players from (4). With $D_C \cap A = C$, we get $(\forall C \in \Gamma \cup \{\emptyset\})\left[\Gamma(i) \cup \{c_i\} \succeq_{c_i}^{+0-\prime} C \cup \{c_i\}\right]$. c_i is also neutral to i. Hence, we can remove i on the left-hand side and get

$$(\forall C \in \Gamma \cup \{\emptyset\})\left[\Gamma(i) \setminus \{i\} \cup \{c_i\} \succeq_{c_i}^{+0-\prime} C \cup \{c_i\}\right]. \tag{5}$$

Finally, c_i has the same friends, order over friends, enemies, and order over enemies as player i has in H. Therefore, (5) is equivalent to $(\forall C \in \Gamma \cup \{\emptyset\})\left[\Gamma(i) \succeq_i^{+0-} C \cup \{i\}\right]$. Thus the coalition structure Γ is necessarily Nash stable for H and $H \in$ NNSE. ❑

4.7 Core Stability and Strict Core Stability

We now turn to group deviations and start with possible (strict) core stability. We first state some characterizations, which show that possible and necessary verification are in coNP for both core stability and strict core stability.

Proposition 49 *Γ is not possibly core stable if and only if there is a coalition $C \subseteq A, C \neq \emptyset$ such that $\Gamma(i) \prec_i^{+\Gamma(i)} C$ holds for all $i \in C$.*

Proof. From left to right, assume that Γ is not possibly core stable, i.e., that for every profile $(\succeq_1, \dots, \succeq_n) \in \times_{i=1}^n \mathrm{Ext}\left(\succeq_i^{+0-}\right)$ we have a blocking coalition $C \subseteq A, C \neq \emptyset$. Consider some profile $(\succeq_1, \dots, \succeq_n) \in \times_{i=1}^n \mathrm{Ext}\left(\succeq_i^{+\Gamma(i)}\right)$. This extended profile does exist since $\succ_i^{+\Gamma(i)}$ is acyclic for all $i \in A$ by Proposition 16. It then holds for every $i \in A$ that $\succeq_i$ extends $\succeq_i^{+\Gamma(i)}$ and that $\succeq_i^{+\Gamma(i)}$ extends $\succeq_i^{+0-}$ by Proposition 17. Hence, $\succeq_i$ also extends $\succeq_i^{+0-}$. Then, there also is a blocking coalition $C \subseteq A, C \neq \emptyset$ for this profile, i.e., $\Gamma(i) \prec_i C$ for every $i \in C$. Since $\succeq_i$ extends $\succeq_i^{+\Gamma(i)}$ and with Observation 18 it follows that $\Gamma(i) \prec_i^{+\Gamma(i)} C$ for every $i \in C$.

From right to left, assume that there is a coalition $C \subseteq A, C \neq \emptyset$ with $\Gamma(i) \prec_i^{+\Gamma(i)} C$ for every $i \in C$. By Proposition 19.2(d), it follows for each $i \in C$ that all extensions $\succeq_i \in \mathrm{Ext}\left(\succeq_i^{+0-}\right)$ satisfy $C \succ_i \Gamma(i)$. Hence, C blocks Γ for every extended profile and Γ is not possibly core stable. ❑

Proposition 50 *Γ is not necessarily core stable if and only if there is a coalition $C \subseteq A, C \neq \emptyset$ such that $\Gamma(i) \prec_i^{-\Gamma(i)} C$ holds for all $i \in C$.*

Proof. From left to right, assume that Γ is not necessarily core stable. Then there exists a profile $(\succeq_1, \ldots, \succeq_n) \in \times_{i=1}^{n} \mathrm{Ext}\left(\succeq_i^{+0-}\right)$ and a coalition $C \subseteq A, C \neq \emptyset$ such that $\Gamma(i) \prec_i C$ holds for all $i \in C$. With Proposition 19.1(d), it follows that $\Gamma(i) \prec_i^{-\Gamma(i)} C$ for all $i \in C$.

From right to left, assume that there is a coalition $C \subseteq A, C \neq \emptyset$ with $\Gamma(i) \prec_i^{-\Gamma(i)} C$ holds for all $i \in C$. Again, by Proposition 19.1(d), it follows for all $i \in C$ that there is an extension $\succeq_i \in \mathrm{Ext}\left(\succeq_i^{+0-}\right)$ such that $\Gamma(i) \prec_i C$. Hence, Γ is not necessarily core stable. ❑

Proposition 51 *Γ is not possibly strictly core stable if and only if there is a nonempty coalition $C \subseteq A$ such that $\Gamma(i) \preceq_i^{+\Gamma(i)} C$ for each player $i \in C$, and $\Gamma(j) \prec_j^{+\Gamma(j)} C$ for some player $j \in C$.*

Proof. This can be shown similarly to Proposition 49, using Propositions 19.2(c) and 19.2(d). ❑

Proposition 52 *Γ is not necessarily strictly core stable if and only if there is a nonempty coalition $C \subseteq A$ such that $\Gamma(i) \preceq_i^{-\Gamma(i)} C$ for each player $i \in C$, and $\Gamma(j) \prec_j^{-\Gamma(j)} C$ for some player $j \in C$.*

Proof. This can be shown similarly to Proposition 50, using Propositions 19.1(c) and 19.1(d). ❑

Since the relations $\succeq_i^{+C}$ and $\succeq_i^{-C}$ can be decided in polynomial time for any two coalitions, we can choose a coalition structure Δ nondeterministically and verify the characterizations from Propositions 49, 50, 51, and 52 in polynomial time. Therefore, we get the following corollary.

Corollary 53 POSSIBLE-CORE-STABILITY-VERIFICATION, NECESSARY-CORE-STABILITY-VERIFICATION, POSSIBLE-STRICT-CORE-STABILITY-VERIFICATION, *and* NECESSARY-STRICT-CORE-STABILITY-VERIFICATION *are in* coNP.

Theorem 54 *The problems* POSSIBLE-CORE-STABILITY-VERIFICATION *and* POSSIBLE-STRICT-CORE-STABILITY-VERIFICATION *are* coNP-*complete.*

Proof. The coNP upper bounds are already given by Corollary 53. Hardness for coNP of both problems can be shown by means of the reduction from CLIQUE to the complement of the core stability verification problem in the enemy-oriented representation (Sung & Dimitrov, 2007). Note that this representation is a special case of the representation with ordinal preferences and thresholds where there are no neutral players and only indifferences between all friends and between all enemies in a player's preference. Furthermore, note that the enemy-oriented preference extension is contained in $\times_{i=1}^{n} \mathrm{Ext}\left(\succeq_i^{+0-}\right)$. While a "clique" of friends is necessarily preferred by all its members to a coalition containing fewer friends or even more enemies, there does not necessarily exist a blocking coalition if there is no such clique (for example, there is no blocking coalition in the enemy-oriented preference extension). ❑

4.8 Popularity and Strict Popularity

We now consider the different stability problems regarding (strict) popularity. We first state characterizations for possible and necessary (strict) popularity, using the optimistic and pessimistic extensions from Definition 10.

Proposition 55 Γ *is not possibly popular if and only if there is a coalition structure* Δ *such that the number of players* $i \in A$ *with* $\Gamma(i) \succ_i^{+\Gamma(i)} \Delta(i)$ *is smaller than the number of players* $j \in A$ *with* $\Delta(j) \succ_j^{+\Gamma(j)} \Gamma(j)$.

Proof. From left to right, assume that Γ is not possibly popular, i.e., for every $(\succeq_1, \dots, \succeq_n) \in \times_{i=1}^{n} \mathrm{Ext}\left(\succeq_i^{+0-}\right)$ there is a coalition structure $\Delta \neq \Gamma$ such that the number of players $i \in A$ with $\Gamma(i) \succ_i \Delta(i)$ is smaller than the number of players $j \in A$ with $\Delta(j) \succ_j \Gamma(j)$. Consider any profile $(\succeq_1, \dots, \succeq_n) \in \times_{i=1}^{n} \mathrm{Ext}\left(\succeq_i^{+\Gamma(i)}\right)$, which certainly exists due to Proposition 16. Since $(\succeq_1, \dots, \succeq_n)$ $\in \times_{i=1}^{n} \mathrm{Ext}\left(\succeq_i^{+0-}\right)$, the above inequality also holds for this profile.

For all $i \in A$ with $\Gamma(i) \not\succ_i \Delta(i)$, it holds that $\Gamma(i) \not\succ_i^{+\Gamma(i)} \Delta(i)$ because $\succeq_i$ is an extension of $\succeq_i^{+\Gamma(i)}$. Hence, $|\{i \in A | \Gamma(i) \succ_i^{+\Gamma(i)} \Delta(i)\}| \leq |\{i \in A | \Gamma(i) \succ_i \Delta(i)\}|$. Furthermore, consider all $j \in A$ with $\Delta(j) \succ_j \Gamma(j)$. With Observation 18 and because $\succeq_j$ extends $\succeq_j^{+\Gamma(i)}$ it follows that $\Delta(j) \succ_j^{+\Gamma(i)} \Gamma(j)$. Hence, $|\{j \in A | \Delta(j) \succ_j \Gamma(j)\}| \leq |\{j \in A | \Delta(j) \succ_j^{+\Gamma(i)} \Gamma(j)\}|$. Combining all inequalities, we get $|\{i \in A | \Gamma(i) \succ_i^{+\Gamma(i)} \Delta(i)\}| \leq |\{i \in A | \Gamma(i) \succ_i \Delta(i)\}| < |\{j \in A | \Delta(j) \succ_j \Gamma(j)\}| \leq |\{j \in A | \Delta(j) \succ_j^{+\Gamma(j)} \Gamma(j)\}|$.

From right to left, assume that there is a coalition structure $\Delta \neq \Gamma$ such that the number of players $i \in A$ with $\Gamma(i) \succ_i^{+\Gamma(i)} \Delta(i)$ is smaller than the number of players $j \in A$ with $\Delta(j) \succ_j^{+\Gamma(j)} \Gamma(j)$.

Consider all $i \in A$ with $\Gamma(i) \not\succ_i^{+\Gamma(i)} \Delta(i)$. It follows by Observation 18 that $\Delta(i) \succeq_i^{+\Gamma(i)} \Gamma(i)$ which, by Proposition 14, implies $\Delta(i) \succeq_i^{+0-} \Gamma(i)$. Hence, $\Gamma(i) \not\succ_i \Delta(i)$ for any extension $\succeq_i$ of $\succeq_i^{+0-}$. It follows that $|\{i \in A | \Gamma(i) \succ_i \Delta(i)\}| \leq |\{i \in A | \Gamma(i) \succ_i^{+\Gamma(i)} \Delta(i)\}|$ for all $\succeq_i \in \mathrm{Ext}\left(\succeq_i^{+0-}\right)$. Furthermore, consider all $j \in A$ with $\Delta(j) \succ_j^{+\Gamma(j)} \Gamma(j)$. It follows by Proposition 14 that $\Delta(j) \succ_j^{+0-} \Gamma(j)$. Hence, for any extension $\succeq_i$ of $\succeq_i^{+0-}$, it holds that $\Delta(j) \succ_j \Gamma(j)$ and $|\{j \in A | \Delta(j) \succ_j^{+\Gamma(j)} \Gamma(j)\}| \leq |\{j \in A | \Delta(j) \succ_j \Gamma(j)\}|$. Summing up all inequalities, $|\{i \in A | \Gamma(i) \succ_i \Delta(i)\}| \leq |\{i \in A | \Gamma(i) \succ_i^{+\Gamma(i)} \Delta(i)\}| < |\{j \in A | \Delta(j) \succ_j^{+\Gamma(j)} \Gamma(j)\}| \leq |\{j \in A | \Delta(j) \succ_j \Gamma(j)\}|$ holds for all extensions $\succeq_i$ of $\succeq_i^{+0-}$ and Γ is not possibly popular. ❑

Proposition 56 Γ *is not necessarily popular if and only if there is a coalition structure* Δ *such that the number of players* $i \in A$ *with* $\Gamma(i) \succ_i^{-\Gamma(i)} \Delta(i)$ *is smaller than the number of players* $j \in A$ *with* $\Delta(j) \succ_j^{-\Gamma(j)} \Gamma(j)$.

Proof. From left ro right, assume that Γ is not necessarily popular. It then holds that there is a $(\succeq_1, \dots, \succeq_n) \in \times_{i=1}^{n} \mathrm{Ext}\left(\succeq_i^{+0-}\right)$ and a coalition structure $\Delta \neq \Gamma$ such that the number of players $i \in A$ with $\Gamma(i) \succ_i \Delta(i)$ is smaller than the number of players $j \in A$ with $\Delta(j) \succ_j \Gamma(j)$.

By the definition of extensions (Definition 8), it holds for every $i \in A$ with $\Gamma(i) \not\succ_i \Delta(i)$ that $\Gamma(i) \not\succ_i^{+0-} \Delta(i)$. This, by Proposition 14, implies $\Delta(i) \succeq_i^{-\Gamma(i)} \Gamma(i)$. Hence, $\Gamma(i) \not\succ_i^{-\Gamma(i)} \Delta(i)$ for every $i \in A$ with $\Gamma(i) \not\succ_i \Delta(i)$. This means that the number of players $i \in A$ with $\Gamma(i) \succ_i^{-\Gamma(i)} \Delta(i)$ is less than or equal to the number of players $i \in A$ with $\Gamma(i) \succ_i \Delta(i)$. Furthermore, by the definition of extensions, it holds for every $j \in A$ with $\Delta(j) \succ_j \Gamma(j)$ that $\Gamma(j) \not\succeq_j^{+0-} \Delta(j)$. With Proposition 14, this implies $\Gamma(j) \not\succeq_j^{-\Gamma(j)} \Delta(j)$. Since $\succeq_j^{-\Gamma(j)}$ is never undecided concerning $\Gamma(j)$ by Observation 18, we have $\Delta(j) \succ_j^{-\Gamma(j)} \Gamma(j)$. Hence, the number of players $j \in A$ with $\Delta(j) \succ_j^{-\Gamma(j)} \Gamma(j)$ is at least as large as the number of players $j \in A$ with $\Delta(j) \succ_j \Gamma(j)$. Combining all inequalities, we get $|\{i \in A | \Gamma(i) \succ_i^{-\Gamma(i)} \Delta(i)\}| \leq |\{i \in A | \Gamma(i) \succ_i \Delta(i)\}| < |\{j \in A | \Delta(j) \succ_j \Gamma(j)\}| \leq |\{j \in A | \Delta(j) \succ_j^{-\Gamma(j)} \Gamma(j)\}|$.

From right to left, assume that there is a coalition structure $\Delta \neq \Gamma$ such that $|\{i \in A | \Gamma(i) \succ_i^{-\Gamma(i)} \Delta(i)\}| < |\{j \in A | \Delta(j) \succ_j^{-\Gamma(j)} \Gamma(j)\}|$. Consider a profile $(\succeq_1, \dots, \succeq_n) \in \times_{i=1}^{n} \mathrm{Ext}\left(\succeq_i^{-\Gamma(i)}\right)$ of complete extensions, which exists since $\succ_i^{-\Gamma(i)}$ is acyclic (Proposition 16). Note that each $\succeq_i$ is also an extension of $\succeq_i^{+0-}$ due to Proposition 17.

By Observation 18, it holds for every $i \in A$ with $\Gamma(i) \not\succ_i^{-\Gamma(i)} \Delta(i)$ that $\Gamma(i) \preceq_i^{-\Gamma(i)} \Delta(i)$. Hence, it holds for every extension $\succeq_i$ of $\succeq_i^{-\Gamma(i)}$ that $\Gamma(i) \preceq_i \Delta(i)$. This implies $\Gamma(i) \not\succ_i \Delta(i)$. Hence, $|\{i \in A | \Gamma(i) \succ_i \Delta(i)\}| \leq |\{i \in A | \Gamma(i) \succ_i^{-\Gamma(i)} \Delta(i)\}|$. Furthermore, for every $j \in A$ with $\Delta(j) \succ_j^{-\Gamma(j)} \Gamma(j)$ it holds that $\Delta(j) \succ_j \Gamma(j)$ because $\succeq_j$ extends $\succeq_j^{-\Gamma(j)}$. Hence, $|\{j \in A | \Delta(j) \succ_j^{-\Gamma(j)} \Gamma(j)\}| \leq |\{j \in A | \Delta(j) \succ_j \Gamma(j)\}|$. Summing up, it holds for this extended profile that $|\{i \in A | \Gamma(i) \succ_i \Delta(i)\}| \leq |\{i \in A | \Gamma(i) \succ_i^{-\Gamma(i)} \Delta(i)\}| < |\{j \in A | \Delta(j) \succ_j^{-\Gamma(j)} \Gamma(j)\}| \leq |\{j \in A | \Delta(j) \succ_j \Gamma(j)\}|$, which means that Γ is not necessarily popular. ❑

The proofs of the following two propositions are similar to those of Propositions 55 and 56 and are therefore omitted.

Proposition 57 Γ *is not possibly strictly popular if and only if there is a coalition structure* Δ *such that the number of players* $i \in A$ *with* $\Gamma(i) \succ_i^{+\Gamma(i)} \Delta(i)$ *is smaller than or equal to the number of players* $j \in A$ *with* $\Delta(j) \succ_j^{+\Gamma(j)} \Gamma(j)$.

Proposition 58 Γ *is not necessarily popular if and only if there is a coalition structure* Δ *such that the number of players* $i \in A$ *with* $\Gamma(i) \succ_i^{-\Gamma(i)} \Delta(i)$ *is smaller than or equal to the number of players* $j \in A$ *with* $\Delta(j) \succ_j^{-\Gamma(j)} \Gamma(j)$.

Since a coalition structure Δ can be chosen nondeterministically, and it can be verified in polynomial time whether the inequalities in Propositions 55, 56, 57, and 58 hold, we get the following corollary.

Corollary 59 POSSIBLE-POPULARITY-VERIFICATION, NECESSARY-POPULARITY-VERIFICATION, POSSIBLE-STRICT-POPULARITY-VERIFICATION, *and* NECESSARY-STRICT-POPULARITY-VERIFICATION *are in* coNP.

To show NP-hardness of the stability problems regarding (strict) popularity, we can make use of several games constructed from a given X3C instance, which we state in the following construction that is inspired by the proof of Theorem 3 of Sung and Dimitrov (2010) and is also based on ideas in the proof of Theorem 43.

Construction 60 *Let* $(B, \mathscr{S})$ *be an* X3C *instance. All four games have three types of players:* connection players, element players, *and* set players.

1. *Let* $A = \{\alpha_{b,i} \mid b \in B, 1 \leq i \leq 3m+3\} \cup \{\beta_b \mid b \in B\} \cup \bigcup_{S \in \mathscr{S}} Q_S$ *be the set of players and define the weak rankings with double threshold as follows:*

   Connection Players: *For each* $b \in B$*, we have* $3m+3$ *connection players* $\alpha_{b,i}$ *for i,* $1 \leq i \leq 3m+3$*. These players consider all remaining connection players corresponding to the same element* $b \in B$ *to be friends. The first* $3m$ *of these players also consider their corresponding element player* β_b *to be a friend and even rank her on the first position. Formally, for each* $b \in B$*, and for each i,* $2 \leq i \leq 3m+3$*,*

$$
\begin{aligned}
\trianglerighteq^{+0-}_{\alpha_{b,1}} &= \left(\beta_b \sim \{\alpha_{b,j} \mid j \neq 1\}_\sim \;\middle|\; \emptyset \;\middle|\; \{\textit{other players}\}_\sim\right), \\
\trianglerighteq^{+0-}_{\alpha_{b,i}} &= \left(\{\alpha_{b,j} \mid j \neq i\}_\sim \;\middle|\; \{\beta_b\} \;\middle|\; \{\textit{other players}\}_\sim\right).
\end{aligned}
$$

   Element players: *For each* $b \in B$*, there is one element player* β_b*. Each of these players considers all set players corresponding to sets containing b to be her best friends, followed by the remaining element players. The least preferred friends are all connection players corresponding to the specific element b, which are collected in the set* C_b*. As there are no neutral players, all other players are considered to be enemies. Formally, for each* $b \in B$*, define* $C_b = \{\alpha_{b,i} \mid 1 \leq i \leq 3m+3\}$ *and*

$$
\trianglerighteq^{+0-}_{\beta_b} = \left(\bigcup_{\{S \mid b \in S\}} Q_{S\sim} \rhd \{\beta_{b'} \mid b' \neq b\}_\sim \rhd C_{b\sim} \;\middle|\; \emptyset \;\middle|\; \{\textit{other players}\}_\sim \right).
$$

   Set players: *For each* $S \in \mathscr{S}$*, we have* $3m+1$ *set players in the set* $Q_S = \{\zeta_{S,k} \mid 1 \leq k \leq 3m+1\}$*. For each fixed* $S \in \mathscr{S}$*, these players only consider the set players in* Q_S *to be their friends and have all element players corresponding to the elements in S in their set of neutral players. All remaining players are their enemies. Formally, for each* $S \in \mathscr{S}$ *and for each k,* $1 \leq k \leq 3m+1$*,*

$$
\trianglerighteq^{+0-}_{\zeta_{S,k}} = \left(\{\zeta_{S,k'} \mid k' \neq k\}_\sim \;\middle|\; \{\beta_b \mid b \in S\} \;\middle|\; \{\textit{other players}\}_\sim\right).
$$

   The profile is illustrated in Figure 7.

2. *In the second game, we have the same set of players as in the first game (given in Construction 60.1), namely* $A = \{\alpha_{b,i} \mid b \in B, 1 \leq i \leq 3m+3\} \cup \{\beta_b \mid b \in B\} \cup \bigcup_{S \in \mathscr{S}} Q_S$*. The weak rankings with double threshold also are the same, except for the first connection player corresponding to the last element in B:* $\trianglerighteq^{+0-}_{\alpha_{3m,1}} = (\{\alpha_{3m,j} \mid j \neq 1\}_\sim \mid \{\beta_{3m}\} \mid \{\textit{other players}\}_\sim)$*.*

3. *The third game is another modification of the first game (presented in Construction 60.1). Here, for each* $b \in B$*, there are only* $3m+2$ *connection players* $\alpha_{b,i}$*. Thus the set of players is given by* $A = \{\alpha_{b,i} \mid b \in B, 1 \leq i \leq 3m+2\} \cup \{\beta_b \mid b \in B\} \cup \bigcup_{S \in \mathscr{S}} Q_S$*.*

   The preferences of the set players remain the same while those for the element players are changed to $\trianglerighteq^{+0-}_{\beta_b} = \left(C_{b\sim} \rhd \bigcup_{\{S \mid b \in S\}} Q_{S\sim} \rhd \{\beta_{b'} \mid b' \neq b\}_\sim \;\middle|\; \emptyset \;\middle|\; \{\textit{other players}\}_\sim \right)$*.*

4. *The fourth game is a modification of the third game (presented in Construction 60.3). The set of players and their weak rankings with double threshold remain the same, except for the*

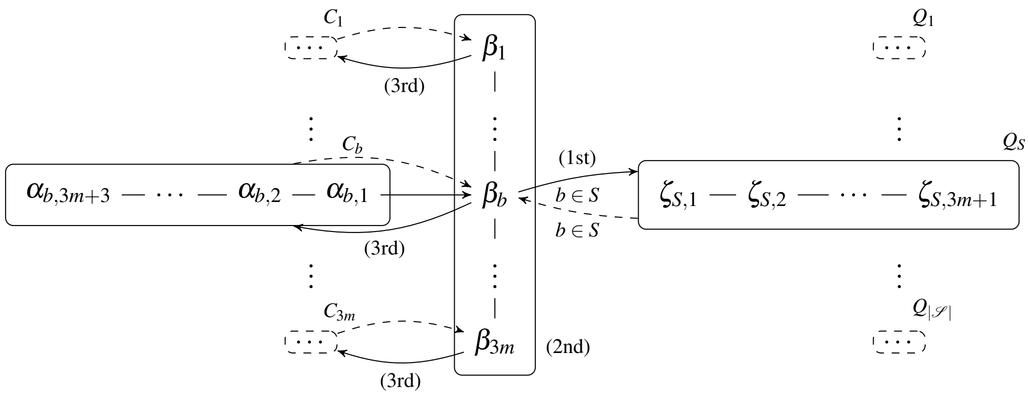


Figure 7: Network of friends from Construction 60.1 that is used in the proof of Theorem 61

preference of the first connection player corresponding to the last element in B, namely $\alpha_{3m,1}$. *In this game, she considers the corresponding element player* β_{3m} *to be a neutral player and not a friend:* $\trianglerighteq^{+0-}_{\alpha_{3m,1}} = (\{\alpha_{3m,j} \mid j \neq 1\}_\sim \mid \{\beta_{3m}\} \mid \{\textit{other players}\}_\sim)$.

We summarize our results for verifying possibly and necessary (strictly) popular coalition structures in the following theorem.

Theorem 61 POSSIBLE-STRICT-POPULARITY-VERIFICATION, POSSIBLE-POPULARITY-VERIFICATION, NECESSARY-STRICT-POPULARITY-VERIFICATION, *and* NECESSARY-POPULARITY-VERIFICATION *are* coNP-*complete.*

Proof. We start with the first part of the theorem, namely the proof of coNP-hardness of POSSIBLE-STRICT-POPULARITY-VERIFICATION. Consider the game from Construction 60.1 and let

$$\Gamma = \{C_b \cup \{\beta_b\} \mid b \in B\} \cup \{Q_S \mid S \in \mathscr{S}\} \tag{6}$$

be the coalition structure of interest. We show that Γ is possibly strictly popular if and only if there is no solution for $(B, \mathscr{S})$.

Only if: Assuming $(B, \mathscr{S})$ has a solution $\mathscr{S}'$, we consider the coalition structure

$$\Gamma' = \{C_b \mid b \in B\} \cup \{P_S \mid S \in \mathscr{S}'\} \cup \{Q_S \mid S \notin \mathscr{S}'\}. \tag{7}$$

There are $3m$ players, namely $\alpha_{b,1}$ for each $b \in B$, who necessarily prefer $\Gamma(\alpha_{b,1})$ to $\Gamma'(\alpha_{b,1})$, and $3m$ players, namely all element players β_b for $b \in B$, who necessarily prefer $\Gamma'(\beta_b)$ to $\Gamma(\beta_b)$. All other players are indifferent between their coalitions in Γ and Γ'. Thus Γ is necessarily prevented from being strictly popular.

If: Assume now that Γ is not possibly strictly popular, that is, for each preference extension, there exists another coalition structure Γ' that beats Γ in pairwise comparison. All players $\alpha_{b,i}$ with $b \in B$ and $1 \leq i \leq 3m+3$, and all players $\zeta_{S,k}$ with $S \in \mathscr{S}$ and $1 \leq k \leq 3m+1$, are in one of their favorite coalitions in Γ; hence, they cannot improve in Γ'. Therefore, there are at most $3m$ players (which have to be element players β_b with $b \in B$) who vote in favor of Γ'.

If, for some $b \in B$, not all players $\alpha_{b,i}$, $1 \leq i \leq 3m+3$, are together, they are all worse off in comparison to Γ. This cannot be counterbalanced by the $3m$ element players; consequently, they have to be in one coalition in Γ'. For the same reason, for each $S \in \mathscr{S}$, the $3m+1$ players in Q_S cannot be separated in Γ'.

If some element player β_b with $b \in B$ wants to improve by adding friends to $C_b \cup \{\beta_b\}$, all $3m+3$ players in C_b will disapprove; hence, this cannot be the case in Γ' either. Thus, for each $b \in B$, the player $\alpha_{b,1}$ is separated from β_b, which sums up in a number of $3m$ players in favor of Γ in comparison to Γ'. This means that, in order for Γ' to be successful, each β_b has to prefer $\Gamma'(\beta_b)$ to $\Gamma(\beta_b)$. It necessarily holds that an element player β_b has the following preferences $Q_S \cup \{\beta_b, \beta_{b'}\} \succ_{\beta_b} Q_S \cup \{\beta_b \mid b \in S\} \succ_{\beta_b} \{\beta_{b'} \mid b \in B\}$, where $\beta_{b'}$ is another element player corresponding to an element from the same $S \in \mathscr{S}$ (i.e., $b \neq b'$ and $b, b' \in S$); it furthermore necessarily holds that $C_b \cup \{\beta_b\} \sim_{\beta_b} Q_S \cup \{\beta_b, \beta_{b'}\}$. However, there exists a preference extension in which this indifference is solved in favor of $\Gamma(\beta_b)$. Thus, for this preference extension, $Q_S \cup \{\beta_b, \beta_{b'}\}$ cannot be in Γ (nor can any coalition even less preferred by β_b be in Γ). Also, there cannot be any enemies of players Q_S in the same coalition, since otherwise there would be at least $3m+1$ more players that disapprove. This leaves only one possibility: There is a coalition structure such that all players β_r are in a coalition P_S with $b \in S$. Consequently, there is an exact cover of B by sets in $\mathscr{S}$. Thus POSSIBLE-STRICT-POPULARITY-VERIFICATION is coNP-hard.

For POSSIBLE-POPULARITY-VERIFICATION, consider the game given in Construction 60.2 and the two coalition structures, Γ and Γ', defined in Equations (6) and (7).

Note that now we need Γ' to strictly defeat Γ, in order to obtain that Γ is not possibly popular. The argumentation is analogous to above, except that now only $3m-1$ players of the form $\alpha_{b,1}$, $b \in B$, dislike being in a different coalition than β_b. Then, for each preference extension, there is such a Γ' if and only if all element players β_b, $b \in B$, can be placed in some P_S, $b \in S$.

Thus it holds that Γ is possibly popular if and only if there is no solution for $(B, \mathscr{S})$.

For the coNP-hardness proof of NECESSARY-STRICT-POPULARITY-VERIFICATION, we consider the game given in Construction 60.3. In this game it holds that possibly $P_S \succ_{\beta_b} C_b \cup \{\beta_b\}$ and necessarily $C_b \cup \{\beta_b\} \succ_{\beta_b} Q_S \cup R_S \succ_{\beta_b} \{\beta_{b'} \mid b \in B\}$, for each $b \in B$, where $R_S \subseteq \{\beta_b \mid b \in S\}$. Similarly to the previous argumentation, it can be shown that the coalition structure Γ defined in Equation (6) is necessarily strictly popular if and only if there is no solution for $(B, \mathscr{S})$.

For the last problem, namely NECESSARY-POPULARITY-VERIFICATION, we can show coNP-hardness via a reduction from X3C given by the game defined in Construction 60.4. With a similar argumentation as above, we have that the coalition structure Γ defined in Equation (6) is necessarily popular if and only if there is no solution for $(B, \mathscr{S})$.

Finally, the coNP upper bounds hold by Corollary 59. ❑

For strict popularity, both existence problems are coNP-hard.

Theorem 62 POSSIBLE-STRICT-POPULARITY-EXISTENCE *and* NECESSARY-STRICT-POPULARITY-EXISTENCE *are* coNP-*hard.*

Proof. To show coNP-hardness of POSSIBLE-STRICT-POPULARITY-EXISTENCE, we consider the game defined in Construction 60.1. We have seen in the proof of Theorem 61 that if there is no solution for the given X3C instance there exists a possibly strictly popular coalition structure (namely Γ defined in Equation (6)). Now we show that, if there is a solution for the given X3C instance, not only is Γ beaten in pairwise comparison, but there is no other strictly popular coalition structure either. Observe that Γ and Γ' (the latter is defined in Equation (7)) tie up in pairwise comparison with the maximal number of positive votes each ($3m$). Thus these two cannot be strictly popular. Any other coalition structure can also only possibly gain at most $3m$ positive votes; hence, there is no coalition structure that beats every other coalition structure in pairwise comparison.

For the coNP-hardness proof of NECESSARY-STRICT-POPULARITY-EXISTENCE, we can use the game defined in Construction 60.3. We have seen that in this game the coalition structure Γ defined in Equation (6) is necessarily strictly popular, so there exists a necessarily strictly popular coalition structure. By analogous arguments as above, while all coalition structures other than Γ and Γ' (where, again, the latter is defined in Equation (7)) are even necessarily worse off, it can be seen that if there is a solution for the original X3C instance, there is no necessarily strictly popular coalition structure at all. ❑

5. Conclusions and Future Work

We have introduced a new representation of preferences in hedonic games using the polarized responsive principle to extend the players' preferences over the other players to preferences over coalitions. Generalizing the responsive extension principle to neutral items in addition to positive and negative items (here called friends and enemies) is novel and original in itself, independently of its use in hedonic games. That is, the polarized extension can be useful more generally in all contexts where some agents may have positive or negative preferences for the presence of some entity. Two important examples are multiwinner elections and fair division.

Regarding multiwinner elections (see, e.g., Faliszewski, Skowron, Slinko, & Talmon, 2017), beyond ranking candidates, it makes sense for voters to specify, for any of the candidates, whether they would prefer to have them in the committee or not; and for some of the candidates voters may not care about whether they are in the committee or not.

Regarding fair division (see, e.g., Brams & Taylor, 1996; Bouveret, Chevaleyre, & Maudet, 2016; Lang & Rothe, 2015), while there is some work on chore division (see, e.g., Aziz, Rauchecker, Schryen, & Walsh, 2017; Bogomolnaia, Moulin, Sandomirskiy, & Yanovskaya, 2016, for recent work), not much is known about settings where an item can be seen as negative for an agent and positive for another one while a third agent does not care about receiving it;[6] still, there are many pratical contexts where this assumption is plausible, such as the allocation of papers to reviewers. If there is no constraint on the allocation, then obviously an item will be assigned to an agent who likes it, provided there is at least one such agent; but if there are constraints (such as balancedness), then it may be the case that an agent gets an item she does not want even though someone expressed a positive preference for it.

We have then looked at several stability concepts in hedonic games with such preferences. The issue of incomparabilities that may remain is tackled by letting these incomparabilities be unresolved and introducing, inspired by the work on necessary and possible winners in voting (Konczak

[6] Recently, Aziz, Caragiannis, Igarashi, and Walsh (2019) studied, from a computational perspective, fair allocation of indivisible goods and chores where an agent may have either a negative or a positive utility for each item. In economics, Bogomolnaia, Moulin, Sandomirskiy, and Yanovskaya (2017) studied fair division of divisible items that they call "mixed manna" (containing both *goods* liked by everyone and *bads* disliked by everyone, but also items that are goods to some agents, yet bads or satiated to other agents).

& Lang, 2005; Xia & Conitzer, 2011),[7] the notions of necessity and possibility for known stability concepts.[8] We have analyzed the computational complexity of the existence and the verification problem of well-known stability concepts for the induced hedonic games. So far, with the help of these solution concepts we can verify whether a coalition structure is a "good" solution, compare two coalition structures, and decide whether there even exists such a coalition structure—sometimes only at great cost in terms of computational complexity, though.

For future work, we propose to consider other solution concepts and to solve the remaining open problems, especially regarding those entries in Table 3 where matching upper and lower bounds on the complexity of problems have not been found yet. One approach to tackle these open problems might be to apply the metaresults from the intriguing work of Peters and Elkind (2015) who establish relations between certain properties of preferences in hedonic games and NP-hardness of certain stability existence problems. Indeed, for our FEN-hedonic games we are able to show, as an easy consequence of the metaresults due to Peters and Elkind (2015), that NASH-STABILITY-EXISTENCE, INDIVIDUAL-STABILITY-EXISTENCE, CORE-STABILITY-EXISTENCE, and STRICT-CORE-STABILITY-EXISTENCE are NP-hard (see the PhD thesis of Rey, 2016, for details). However, this does not say anything about the complexity of the corresponding variants with possible or necessary stability; the open problems in Table 3 remain.

Besides further pursuing the analysis initiated here, for future work we also propose to introduce the notion of *partition correspondences* with the purpose to actually identify "good" coalition structures as an output. In contrast to the original idea of hedonic games where coalitions form in a decentralized manner, a central authority might be used here, in order to decide which coalitions will "best" work together. This might, for example, be the case in a setting where the head of a department has to divide a group of employees into teams. The teams should be stable and/or should have high social welfare (in the sense that the team members are as happy as possible with their group to create a good working atmosphere).

Also, the various notions of altruism in hedonic games may be useful here, such as those introduced by Nguyen et al. (2016) where players care not only about their own preferences but also about their friends' preferences. While they focus on friend-oriented preferences to define altruistic hedonic games of various types, other (compact) representations of hedonic games might be used as well for this purpose, and we propose as a challenging task for future research to study altruism in FEN-hedonic games where the players' preferences on coalitions are based on the polarized responsive extension principle.

Acknowledgments

This work was supported in part by DFG grants RO-1202/{14-1, 14-2, 15-1}, by Agence Nationale de la Recherche under the "programme d'Investissements d'avenir" ANR-19-P3IA-0001 (PRAIRIE), by a DAAD-PPP/PHC PROCOPE project, and by COST Action IC1205 on Computational Social Choice.

[7] Other areas where the concepts of necessary and possible winners from voting has been applied are fair division (Aziz, Walsh, & Xia, 2015; Baumeister, Bouveret, Lang, Nguyen, Nguyen, Rothe, & Saffidine, 2017) and strategy-proofness in judgment aggregation (Baumeister, Erdélyi, Erdélyi, & Rothe, 2015).

[8] A different approach has been taken by Rothe et al. (2018) who consider comparability functions based on Borda-like scoring vectors in order to resolve these incomparabilities (see also Lang et al., 2015).

References

Aziz, H., Brandl, F., Brandt, F., Harrenstein, P., Olsen, M., & Peters, D. (2019). Fractional hedonic games. *ACM Transactions on Economics and Computation*, *7*(2), Article 6, 6:1–6:29.

Aziz, H., Brandt, F., & Harrenstein, P. (2013a). Pareto optimality in coalition formation. *Games and Economic Behavior*, *82*, 562–581.

Aziz, H., Brandt, F., & Seedig, H. (2013b). Computing desirable partitions in additively separable hedonic games. *Artificial Intelligence*, *195*, 316–334.

Aziz, H., Caragiannis, I., Igarashi, A., & Walsh, T. (2019). Fair allocation of indivisible goods and chores. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, pp. 53–59. ijcai.org.

Aziz, H., Gaspers, S., Mackenzie, S., & Walsh, T. (2015). Fair assignment of indivisible objects under ordinal preferences. *Artificial Intelligence*, *227*, 71–92.

Aziz, H., Harrenstein, P., Lang, J., & Wooldridge, M. (2016). Boolean hedonic games. In *Proceedings of the 15th International Conference on Principles of Knowledge Representation and Reasoning*, pp. 166–175. AAAI Press.

Aziz, H., Rauchecker, G., Schryen, G., & Walsh, T. (2017). Algorithms for max-min share fair allocation of indivisible chores. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence*, pp. 335–341. AAAI Press.

Aziz, H., & Savani, R. (2016). Hedonic games. In Brandt, F., Conitzer, V., Endriss, U., Lang, J., & Procaccia, A. (Eds.), *Handbook of Computational Social Choice*, chap. 15, pp. 356–376. Cambridge University Press.

Aziz, H., Walsh, T., & Xia, L. (2015). Possible and necessary allocations via sequential mechanisms. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence*, pp. 468–474. AAAI Press/IJCAI.

Ballester, C. (2004). NP-completeness in hedonic games. *Games and Economic Behavior*, *49*(1), 1–30.

Banerjee, S., Konishi, H., & Sönmez, T. (2001). Core in a simple coalition formation game. *Social Choice and Welfare*, *18*, 135–153.

Baumeister, D., Bouveret, S., Lang, J., Nguyen, N., Nguyen, T., Rothe, J., & Saffidine, A. (2017). Positional scoring-based allocation of indivisible goods. *Journal of Autonomous Agents and Multi-Agent Systems*, *31*(3), 628–655.

Baumeister, D., Erdélyi, G., Erdélyi, O., & Rothe, J. (2015). Complexity of manipulation and bribery in judgment aggregation for uniform premise-based quota rules. *Mathematical Social Sciences*, *76*, 19–30.

Bilò, V., Fanelli, A., Flammini, M., Monaco, G., & Moscardelli, L. (2014). Nash stability in fractional hedonic games. In *Proceedings of the 10th International Workshop on Internet & Network Economics*, pp. 486–491. Springer-Verlag *Lecture Notes in Computer Science #8877*.

Bilò, V., Fanelli, A., Flammini, M., Monaco, G., & Moscardelli, L. (2015). On the price of stability of fractional hedonic games. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems*, pp. 1239–1247. IFAAMAS.

Bogomolnaia, A., & Jackson, M. (2002). The stability of hedonic coalition structures. *Games and Economic Behavior*, *38*(2), 201–230.

Bogomolnaia, A., Moulin, H., Sandomirskiy, F., & Yanovskaya, E. (2016). Dividing goods and bads under additive utilities. Tech. rep. arXiv:1608.01540v1 [cs.GT], ACM Computing Research Repository (CoRR).

Bogomolnaia, A., Moulin, H., Sandomirskiy, F., & Yanovskaya, E. (2017). Competitive division of a mixed manna. *Econometrica*, *85*(6), 1847–1871.

Bossong, U., & Schweigert, D. (2006). Minimal paths on ordered graphs. *Mathematica Slovaca*, *56*(1), 23–31.

Bouveret, S., Chevaleyre, Y., & Maudet, N. (2016). Fair allocation of indivisible goods. In Brandt, F., Conitzer, V., Endriss, U., Lang, J., & Procaccia, A. (Eds.), *Handbook of Computational Social Choice*, chap. 12, pp. 284–310. Cambridge University Press.

Bouveret, S., Endriss, U., & Lang, J. (2010). Fair division under ordinal preferences: Computing envy-free allocations of indivisible goods. In *Proceedings of the 19th European Conference on Artificial Intelligence*, pp. 387–392. IOS Press.

Brams, S., & Sanver, R. (2009). Voting systems that combine approval and preference. In Brams, S., Gehrlein, W., & Roberts, F. (Eds.), *The Mathematics of Preference, Choice, and Order: Essays in Honor of Peter C. Fishburn*, pp. 215–237. Springer.

Brams, S., & Taylor, A. (1996). *Fair Division: From Cake-Cutting to Dispute Resolution*. Cambridge University Press.

Cechlárová, K., & Hajduková, J. (2003). Computational complexity of stable partitions with B-preferences. *International Journal of Game Theory*, *31*(3), 353–364.

Cechlárová, K., & Hajduková, J. (2004). Stable partitions with $\mathscr{W}$-preferences. *Discrete Applied Mathematics*, *138*(3), 333–347.

Cechlárová, K., & Romero-Medina, A. (2001). Stability in coalition formation games. *International Journal of Game Theory*, *29*(4), 487–494.

Darmann, A., Elkind, E., Kurz, S., Lang, J., Schauer, J., & Woeginger, G. (2018). Group activity selection problem with approval preferences. *International Journal of Game Theory*, *47*(3), 767–796.

Delort, C., Spanjaard, O., & Weng, P. (2011). Committee selection with a weight constraint based on a pairwise dominance relation. In *Proceedings of the 2nd International Conference on Algorithmic Decision Theory*, pp. 28–41. Springer-Verlag *Lecture Notes in Artificial Intelligence #6992*.

Dimitrov, D., Borm, P., Hendrickx, R., & Sung, S. (2006). Simple priorities and core stability in hedonic games. *Social Choice and Welfare*, *26*(2), 421–433.

Elkind, E., & Rothe, J. (2015). Cooperative game theory. In Rothe, J. (Ed.), *Economics and Computation. An Introduction to Algorithmic Game Theory, Computational Social Choice, and Fair Division*, chap. 3, pp. 135–196. Springer-Verlag.

Elkind, E., & Wooldridge, M. (2009). Hedonic coalition nets. In *Proceedings of the 8th International Joint Conference on Autonomous Agents and Multiagent Systems*, pp. 417–424. IFAAMAS.

Faliszewski, P., Skowron, P., Slinko, A., & Talmon, N. (2017). Multiwinner voting: A new challenge for social choice theory. In Endriss, U. (Ed.), *Trends in Computational Social Choice*, chap. 2, pp. 27–47. AI Access Foundation.

Garey, M., & Johnson, D. (1979). *Computers and Intractibility: A Guide to the Theory of NP-Completeness*. W. H. Freeman and Company.

Kerkmann, A., & Rothe, J. (2019). Stability in FEN-hedonic games for single-player deviations. In *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems*. IFAAMAS. To appear.

Konczak, K., & Lang, J. (2005). Voting procedures with incomplete preferences. In *Proceedings of the Multidisciplinary IJCAI-05 Workshop on Advances in Preference Handling*, pp. 124–129.

Lang, J., Rey, A., Rothe, J., Schadrack, H., & Schend, L. (2015). Representing and solving hedonic games with ordinal preferences and thresholds. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems*, pp. 1229–1237. IFAAMAS.

Lang, J., & Rothe, J. (2015). Fair division of indivisible goods. In Rothe, J. (Ed.), *Economics and Computation. An Introduction to Algorithmic Game Theory, Computational Social Choice, and Fair Division*, chap. 8, pp. 493–550. Springer-Verlag.

Meyer, A., & Stockmeyer, L. (1972). The equivalence problem for regular expressions with squaring requires exponential space. In *Proceedings of the 13th IEEE Symposium on Switching and Automata Theory*, pp. 125–129. IEEE Computer Society Press.

Nguyen, N., Rey, A., Rey, L., Rothe, J., & Schend, L. (2016). Altruistic hedonic games. In *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems*, pp. 251–259. IFAAMAS. Also presented at the 7th International Workshop on Cooperative Games in Multiagent Systems (CoopMAS 2016), co-located with AAMAS 2016, and at the 6th International Workshop on Computational Social Choice (COMSOC 2016), Toulouse, France, June 2016, both with nonarchival proceedings.

Papadimitriou, C. (1995). *Computational Complexity* (Second edition). Addison-Wesley.

Peters, D. (2016). Complexity of hedonic games with dichotomous preferences. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence*, pp. 579–585. AAAI Press.

Peters, D., & Elkind, E. (2015). Simple causes of complexity in hedonic games. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence*, pp. 617–623. AAAI Press/IJCAI.

Rey, A. (2016). *Beyond Intractability: A Computational Complexity Analysis of Various Types of Influence and Stability in Cooperative Games*. Ph.D. thesis, Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany.

Rey, A., Rothe, J., Schadrack, H., & Schend, L. (2016). Toward the complexity of the existence of wonderfully stable partitions and strictly core stable coalition structures in enemy-oriented hedonic games. *Annals of Mathematics and Artificial Intelligence*, *77*(3), 317–333.

Roth, A. (1985). Common and conflicting interests in two-sided matching markets. *European Economic Review*, *27*(1), 75–96.

Roth, A., & Sotomayor, M. (1992). *Two-sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Cambridge University Press.

Rothe, J. (2005). *Complexity Theory and Cryptology. An Introduction to Cryptocomplexity*. EATCS Texts in Theoretical Computer Science. Springer-Verlag.

Rothe, J., Schadrack, H., & Schend, L. (2018). Borda-induced hedonic games with friends, enemies, and neutral players. *Mathematical Social Sciences*, *96*, 21–36.

Sen, A. (1970). Interpersonal aggregation and partial comparability. *Econometrica*, *38*, 393–409.

Stockmeyer, L. (1976). The polynomial-time hierarchy. *Theoretical Computer Science*, *3*(1), 1–22.

Sung, S., & Dimitrov, D. (2007). On core membership testing for hedonic coalition formation games. *Operations Research Letters*, *35*(2), 155–158.

Sung, S., & Dimitrov, D. (2010). Computational complexity in additive hedonic games. *European Journal of Operational Research*, *203*(3), 635–639.

Woeginger, G. (2013a). Core stability in hedonic coalition formation. In *Proceedings of the 39th Conference on Current Trends in Theory and Practice of Computer Science*, pp. 33–50. Springer-Verlag *Lecture Notes in Computer Science #7741*.

Woeginger, G. (2013b). A hardness result for core stability in additive hedonic games. *Mathematical Social Sciences*, *65*(2), 101–104.

Xia, L., & Conitzer, V. (2011). Determining possible and necessary winners given partial orders. *Journal of Artificial Intelligence Research*, *41*, 25–67.

Chapter 6

Conclusion and Future Work Directions

Based on the current state of research in the field of cooperative game theory, we have proposed new preference formats for hedonic games, established several models of altruism, and studied stability, optimality, and fairness in multiple classes of hedonic games. We will now summarize the contributions of this thesis and highlight some directions for future research.

In Chapter 3, we started our study with topics of altruism. Evolutionary biology has revealed that selfishness is not always a means to success in the real world, but rather friendliness constituted an essential advantage in the evolution of certain species, including humans [74]. Motivated by this fact and with the aim to provide a more realistic model of real world scenarios, we introduced several models of altruism in coalition formation games. In Section 3.1, we presented *altruistic hedonic games* that model agents to behave altruistic towards their friends in a given network. We distinguished between three degrees of altruism and between two ways of aggregating the agents' preferences. We studied the six resulting models with respect to their axiomatic properties, showing that they fulfill some desirable properties while they can represent situations that can not be represented by other preference formats from the literature. We then conducted a computational analysis concerning stability verification and existence, focusing on Nash, individual, contractually individual, core, and strict core stability, individual rationality, and perfectness. For selfish-first altruistic hedonic games, we have settled the complexity of all considered problems. We further initiated the study for altruistic hedonic games where the agents behave according to different degrees of altruism. An important direction for future research is the completion of this study, i.e., the determination of the complexity of all considered verification and existence problems in this case.

In Section 3.2, we continued our study of altruistic hedonic games, focusing on the notions of popularity and strict popularity. We have solved all open cases of popularity and strict popularity verification, showing that verification is coNP-complete for popularity and strict popularity and all considered models of altruism. We even proved the coNP-hardness of strict popularity existence under equal- and altruistic-treatment. Yet, we suspect that these existence problems might be even harder. It is an interesting question for future research whether popularity and strict popularity existence are even Σ_2^p-complete in altruistic hedonic games. An interesting side result of our study is that popularity verification is also coNP-complete for friend-oriented hedonic games.

We have extended our models of altruism in Section 3.3. While altruistic hedonic games model agents to be altruistic to their friends in their current coalitions, we have additionally proposed *altruistic coalition formation games* where agents behave altruistic to all their friends, not only to those in the same coalition. We have seen that this removal of the hedonic restriction brings some axiomatic advantages. Particularly, altruistic coalition formation preferences are unanimous, which is not the case for all altruistic hedonic preferences. Furthermore, altruistic coalition formation preferences fulfill more cases of monotonicity than altruistic hedonic preferences. We have also initiated the study of stability in altruistic coalition formation games. Our results include characterizations of stability in altruistic coalition formation games and computational bounds on the complexity of the associated verification and existence problems.

There are several possible future work directions in the scope of Chapter 3. So far, altruism in coalition formation games was always handled as a static model where agents were only acting according to one selected degree of altruism. We are interested in models where agents individually and dynamically may choose to what degree they wish to act altruistically, which seems to be more realistic: Agents can be expected to act most altruistically when they see that others are suffering, and they are more egoistic if everyone around them is doing well. This can also be observed in reality where solidarity with others increases when social crises occur. We regard modeling such situations as a promising topic for future research. Other research in the scope of altruistic coalition formation could concern the relationship between altruism and fairness: Do altruistic preferences favor the formation of fair outcomes? Furthermore, it could be interesting to apply altruistic coalition formation games to other valuation functions. While we currently use friend-oriented valuations as a basis of our model, one might also consider general additively separable or fractional valuations.

We continued with aspects of fairness in Chapter 4 where we introduced three notions of *local fairness* for hedonic games. We showed that the three notions form a strict hierarchy and related them to other common notions of stability, fairness, and optimality. We intensively studied the three notions of altruism for additively separable hedonic games. Our studies concerning the local fairness notions provide a diverse potential for follow-up research. For instance, it would be interesting to extend the studies concerning the price of local fairness in additively separable hedonic games and find restrictions to the players' preferences such that the price of local fairness is bounded by a nontrivial constant. Another appealing future direction is the investigation of local fairness in other classes of hedonic games such as fractional or modified fractional hedonic games.

Moreover, we provided an elaborate study of *FEN-hedonic games*. In Chapter 5, we introduced the corresponding preference representation that is composed of weak ordinal rankings over the agents which are separated by two thresholds. The new format adds to existing literature that deals with the separation of the agents into friends and enemies (see, e.g., Dimitrov et al. [50], Sung and Dimitrov [137, 136], Rey et al. [122], Ota et al. [111], and Barrot et al. [17]). It is succinct, easy to elicit from the agents, and reasonably expressive. We have examined a variety of stability notions in the context of FEN-hedonic games, distinguishing between possible and necessary satisfaction of these notions. An intriguing topic for future research could be the integration of altruism in FEN-hedonic games. It might be a challenging goal to create a model that uses the rather expressive and at the same time simple representation of weak rankings with double thresholds and lifts these rankings to altruistic preferences over coalitions.

In conclusion, this thesis has made a significant contribution to the field of altruism in coalition formation games, expanded the research on the topic of fairness in hedonic games, and provided a comprehensive study of hedonic games with ordinal preferences and thresholds (FEN-hedonic games). Nevertheless, many exciting questions remain that future research might seek to answer.

Bibliography

[1] J. Alcalde and P. Revilla. “Researching with whom? Stability and manipulation”. In: *Journal of Mathematical Economics* 40.8 (2004), pp. 869–887.

[2] A. Anagnostopoulos, L. Becchetti, B. de Keijzer, and G. Schäfer. “Inefficiency of Games with Social Context”. In: *Proceedings of the 6th International Symposium on Algorithmic Game Theory*. Springer-Verlag *Lecture Notes in Computer Science #8146*, Oct. 2013, pp. 219–230.

[3] K. Apt and G. Schäfer. “Selfishness Level of Strategic Games”. In: *Journal of Artificial Intelligence Research* 49 (2014), pp. 207–240.

[4] S. Arora and B. Barak. *Computational Complexity: A Modern Approach*. Cambridge University Press, 2009.

[5] S. Arora and B. Barak. “More thoughts about P, NP, and all that”. In: *Computational Complexity: A Modern Approach*. Cambridge University Press, 2009. Chap. 2.7, pp. 57–62.

[6] I. Ashlagi, P. Krysta, and M. Tennenholtz. “Social Context Games”. In: *Proceedings of the 4th International Workshop on Internet & Network Economics*. Springer-Verlag *Lecture Notes in Computer Science #5385*, Dec. 2008, pp. 675–683.

[7] H. Aziz and F. Brandl. “Existence of Stability in Hedonic Coalition Formation Games”. In: *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*. IFAAMAS, June 2012, pp. 763–770.

[8] H. Aziz and R. Savani. “Hedonic Games”. In: *Handbook of Computational Social Choice*. Ed. by F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. Procaccia. Cambridge University Press, 2016. Chap. 15, pp. 356–376.

[9] H. Aziz, P. Harrenstein, J. Lang, and M. Wooldridge. “Boolean Hedonic Games”. In: *Proceedings of the 15th International Conference on Principles of Knowledge Representation and Reasoning*. AAAI Press, Apr. 2016, pp. 166–175.

[10] H. Aziz, F. Brandt, and H. Seedig. “Computing Desirable Partitions in Additively Separable Hedonic Games”. In: *Artificial Intelligence* 195 (2013), pp. 316–334.

[11] H. Aziz, F. Brandl, F. Brandt, P. Harrenstein, M. Olsen, and D. Peters. “Fractional Hedonic Games”. In: *ACM Transactions on Economics and Computation* 7.2 (2019), Article 6, 6:1–6:29.

[12] H. Aziz, F. Brandt, and P. Harrenstein. “Pareto Optimality in Coalition Formation”. In: *Games and Economic Behavior* 82 (2013), pp. 562–581.

[13] J. Balcázar, J. Díaz, and J. Gabarró. *Structural Complexity II*. EATCS Monographs on Theoretical Computer Science. Springer-Verlag, 1990.

[14] C. Ballester. “NP-Completeness in Hedonic Games”. In: *Games and Economic Behavior* 49.1 (2004), pp. 1–30.

[15] S. Banerjee, H. Konishi, and T. Sönmez. “Core in a Simple Coalition Formation Game”. In: *Social Choice and Welfare* 18.1 (2001), pp. 135–153.

[16] N. Barrot and M. Yokoo. “Stable and Envy-free Partitions in Hedonic Games”. In: *Proceedings of the 28th International Joint Conference on Artificial Intelligence*. ijcai.org, July 2019, pp. 67–73.

[17] N. Barrot, K. Ota, Y. Sakurai, and M. Yokoo. “Unknown Agents in Friends Oriented Hedonic Games: Stability and Complexity”. In: *Proceedings of the 33rd AAAI Conference on Artificial Intelligence*. AAAI Press, 2019, pp. 1756–1763.

[18] E. Bell. “The iterated exponential integers”. In: *Annals of Mathematics* (1938), pp. 539–557.

[19] V. Bilò, A. Fanelli, M. Flammini, G. Monaco, and L. Moscardelli. “Nash Stable Outcomes in Fractional Hedonic Games: Existence, Efficiency and Computation”. In: *Journal of Artificial Intelligence Research* 62 (2018), pp. 315–371.

[20] V. Bilò, A. Celi, M. Flammini, and V. Gallotti. “Social Context Congestion Games”. In: *Theoretical Computer Science* 514 (2013), pp. 21–35.

[21] A. Bogomolnaia and M. Jackson. “The Stability of Hedonic Coalition Structures”. In: *Games and Economic Behavior* 38.2 (2002), pp. 201–230.

[22] É. Borel. “La théorie du jeu et les équations intégrales à noyau symétrique gauche”. In: *Comptes rendus hebdomadaires des séances de l’Académie des sciences*. Vol. 173. 1304–1308, 1921.

[23] S. Bouveret and M. Lemaître. “Characterizing Conflicts in Fair Division of Indivisible Goods Using a Scale of Criteria”. In: *Journal of Autonomous Agents and Multi-Agent Systems* 30.2 (2016), pp. 259–290.

[24] S. Bouveret, Y. Chevaleyre, and N. Maudet. “Fair Allocation of Indivisible Goods”. In: *Handbook of Computational Social Choice*. Ed. by F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. Procaccia. Cambridge University Press, 2016. Chap. 12, pp. 284–310.

[25] F. Brandl, F. Brandt, and M. Strobel. “Fractional Hedonic Games: Individual and Group Stability”. In: *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems*. IFAAMAS, May 2015, pp. 1219–1227.

[26] F. Brandt and M. Bullinger. “Finding and Recognizing Popular Coalition Structures”. In: *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems*. IFAAMAS, 2020, 195–203.

[27] F. Brandt, M. Bullinger, and A. Wilczynski. “Reaching Individually Stable Coalition Structures in Hedonic Games”. In: *Proceedings of the 35th AAAI Conference on Artificial Intelligence*. AAAI Press, Feb. 2021, pp. 5211–5218.

[28] F. Brandt, M. Bullinger, and L. Tappe. "Single-Agent Dynamics in Additively Separable Hedonic Games". In: *Proceedings of the 36th AAAI Conference on Artificial Intelligence*. To appear. 2022.

[29] S. Brânzei and K. Larson. "Coalitional Affinity Games and the Stability Gap". In: *Proceedings of the 21st International Joint Conference on Artificial Intelligence*. AAAI Press/IJCAI, 2009, pp. 79–84.

[30] S. Brânzei and K. Larson. "Social Distance Games". In: *Proceedings of the 22nd International Joint Conference on Artificial Intelligence*. AAAI Press/IJCAI, 2011, pp. 91–96.

[31] R. Bredereck, E. Elkind, and A. Igarashi. "Hedonic Diversity Games". In: *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems*. IFAAMAS, May 2019, 565–573.

[32] E. Budish. "The Combinatorial Assignment Problem: Approximate Competitive Equilibrium from Equal Incomes". In: *Journal of Political Economy* 119.6 (2011), pp. 1061–1103.

[33] M. Bullinger. "Pareto-Optimality in Cardinal Hedonic Games". In: *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems*. 2020, pp. 213–221.

[34] M. Bullinger and S. Kober. "Loyalty in Cardinal Hedonic Games". In: *Proceedings of the 30th International Joint Conference on Artificial Intelligence*. ijcai.org, July 2021, pp. 66–72.

[35] N. Burani and W. Zwicker. "Coalition Formation Games with Separable Preferences". In: *Mathematical Social Sciences* 45.1 (2003), pp. 27–52.

[36] J. Carlson, A. Jaffe, and A. Wiles, eds. *The Millennium Prize Problems*. American Mathematical Society, Providence, RI, 2006.

[37] R. Carosi, G. Monaco, and L. Moscardelli. "Local Core Stability in Simple Symmetric Fractional Hedonic Games". In: *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems*. IFAAMAS, May 2019, pp. 574–582.

[38] K. Cechlárová and J. Hajduková. "Computational Complexity of Stable Partitions with B-preferences". In: *International Journal of Game Theory* 31.3 (2003), pp. 353–364.

[39] K. Cechlárová and J. Hajduková. "Stable Partitions with $\mathscr{W}$-Preferences". In: *Discrete Applied Mathematics* 138.3 (2004), pp. 333–347.

[40] K. Cechlárová and A. Romero-Medina. "Stability in Coalition Formation Games". In: *International Journal of Game Theory* 29.4 (2001), pp. 487–494.

[41] G. Chalkiadakis, E. Elkind, and M. Wooldridge. *Computational Aspects of Cooperative Game Theory*. Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan and Claypool Publishers, 2011.

[42] P. Chen, B. de Keijzer, D. Kempe, and G. Schäfer. “Altruism and Its Impact on the Price of Anarchy”. In: *ACM Transactions on Economics and Computation* 2.4 (2014), Article 17, 17:1–17:45.

[43] A. Cobham. “The Intrinsic Computational Difficulty of Functions”. In: *Proceedings of the 1964 International Congress for Logic Methodology and Philosophy of Science*. North Holland, 1964, pp. 24–30.

[44] S. Cook. “The Complexity of Theorem-Proving Procedures”. In: *Proceedings of the 3rd ACM Symposium on Theory of Computing*. ACM Press, 1971, pp. 151–158.

[45] S. Cook. “The P versus NP Problem”. In: *The Millennium Prize Problems*. Ed. by J. Carlson, A. Jaffe, and A. Wiles. American Mathematical Society, Providence, RI, 2006, pp. 87–104.

[46] S. Cramer. “Minimum-Based Altruism in Coalition Formation Games”. Bachelor’s thesis. Heinrich-Heine-Universität Düsseldorf, Jan. 2021.

[47] A. Darmann, E. Elkind, S. Kurz, J. Lang, J. Schauer, and G. Woeginger. “Group Activity Selection Problem with Approval Preferences”. In: *International Journal of Game Theory* 47.3 (2018), pp. 767–796.

[48] D. Dimitrov and S. Sung. “On Top Responsiveness and Strict Core Stability”. In: *Journal of Mathematical Economics* 43.2 (2007), pp. 130–134.

[49] D. Dimitrov and S. Sung. “Top Responsiveness and Nash Stability in Coalition Formation Games”. In: *Kybernetika* 42.4 (2006), pp. 453–460.

[50] D. Dimitrov, P. Borm, R. Hendrickx, and S. Sung. “Simple Priorities and Core Stability in Hedonic Games”. In: *Social Choice and Welfare* 26.2 (2006), pp. 421–433.

[51] R. Downey and M. Fellows. *Fundamentals of Parameterized Complexity*. Springer-Verlag, 2013.

[52] R. Downey and M. Fellows. *Parameterized Complexity*. Springer-Verlag, 1999.

[53] J. Drèze and J. Greenberg. “Hedonic Coalitions: Optimality and Stability”. In: *Econometrica* 48.4 (1980), pp. 987–1003.

[54] J. Edmonds. “Paths, Trees and Flowers”. In: *Canadian Journal of Mathematics* 17 (1965), pp. 449–467.

[55] E. Elkind and J. Rothe. “Cooperative Game Theory”. In: *Economics and Computation. An Introduction to Algorithmic Game Theory, Computational Social Choice, and Fair Division*. Ed. by J. Rothe. Springer Texts in Business and Economics. Springer-Verlag, 2015. Chap. 3, pp. 135–193.

[56] E. Elkind and M. Wooldridge. “Hedonic Coalition Nets”. In: *Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems*. IFAAMAS, May 2009, pp. 417–424.

[57] E. Elkind, T. Rahwen, and N. Jennings. “Computational Coalition Formation”. In: *Multiagent Systems*. Ed. by G. Weiss. 2nd ed. MIT Press, 2013. Chap. 8, pp. 329–380.

[58] E. Elkind, A. Fanelli, and M. Flammini. "Price of Pareto Optimality in Hedonic Games". In: *Proceedings of the 30th AAAI Conference on Artificial Intelligence*. AAAI Press, Feb. 2016, pp. 475–481.

[59] M. Flammini, G. Monaco, and Q. Zhang. "Strategyproof Mechanisms for Additively Separable Hedonic Games and Fractional Hedonic Games". In: *Proceedings of the 15th Workshop on Approximation and Online Algorithms*. Springer, 2017, pp. 301–316.

[60] J. Flum and M. Grohe. *Parameterized Complexity Theory*. EATCS Texts in Theoretical Computer Science. Springer-Verlag, 2006.

[61] D. Foley. "Resource Allocation and the Public Sector". In: *Yale Economic Essays* 7.1 (1967), pp. 45–98.

[62] M. Gairing and R. Savani. "Computing Stable Outcomes in Hedonic Games". In: *Proceedings of the 3rd International Symposium on Algorithmic Game Theory*. Springer-Verlag *Lecture Notes in Computer Science #6386*, Oct. 2010, pp. 174–185.

[63] M. Gairing and R. Savani. "Computing stable outcomes in hedonic games with voting-based deviations". In: *Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems*. IFAAMAS, May 2011, pp. 559–566.

[64] D. Gale and L. Shapley. "College Admissions and the Stability of Marriage". In: *The American Mathematical Monthly* 69.1 (1962), pp. 9–15.

[65] P. Gärdenfors. "Match making: Assignments based on bilateral preferences". In: *Behavioral Science* 20.3 (1975), pp. 166–173.

[66] M. Garey and D. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman and Company, 1979.

[67] J. Gill. "Computational Complexity of Probabilistic Turing Machines". In: *SIAM Journal on Computing* 6.4 (1977), pp. 675–695.

[68] T. Gonzalez. "Clustering to Minimize the Maximum Intercluster Distance". In: *Theoretical Computer Science* 38 (1985), pp. 293–306.

[69] L. Gourvès, J. Monnot, and L. Tlilane. "Near Fairness in Matroids". In: *Proceedings of the 21st European Conference on Artificial Intelligence*. IOS Press, Aug. 2014, pp. 393–398.

[70] F. Gurski, I. Rothe, J. Rothe, and E. Wanke. *Exakte Algorithmen für schwere Graphenprobleme*. eXamen.Press. Springer-Verlag, 2010.

[71] D. Gusfield and R. Irving. *The Stable Marriage Problem: Structure and Algorithms*. MIT press, 1989.

[72] J. Hajduková. "Coalition Formation Games: A Survey". In: *International Game Theory Review* 8.4 (2006), pp. 613–641.

[73] T. Hanaka and M. Lampis. "Hedonic Games and Treewidth Revisited". In: *ACM Computing Research Repository* (Feb. 2022). arXiv: `2202.06925`.

[74] B. Hare and V. Woods. *Survival of the Friendliest: Understanding Our Origins and Rediscovering Our Common Humanity*. New York: Penguin Random House, 2020.

[75] M. Hoefer and A. Skopalik. “Altruism in Atomic Congestion Games”. In: *ACM Transactions on Economics and Computation* 1.4 (2013), 21:1–21:21.

[76] M. Hoefer, M. Penn, M. Polukarova, A. Skopalik, and B. Vöcking. “Considerate Equilibrium”. In: *Proceedings of the 22nd International Joint Conference on Artificial Intelligence*. AAAI Press/IJCAI, July 2011, pp. 234–239.

[77] M. Hoefer, D. Vaz, and L. Wagner. “Dynamics in Matching and Coalition Formation Games with Structural Constraints”. In: *Artificial Intelligence* 262 (2018), pp. 222–247.

[78] A. Igarashi and E. Elkind. “Hedonic Games with Graph-Restricted Communication”. In: *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems*. IFAAMAS, May 2016, 242–250.

[79] A. Igarashi, K. Ota, Y. Sakurai, and M. Yokoo. “Robustness against Agent Failure in Hedonic Games”. In: *Proceedings of the 28th International Joint Conference on Artificial Intelligence*. ijcai.org, July 2019, pp. 364–370.

[80] C. Kaklamanis, P. Kanellopoulos, and K. Papaioannou. “The Price of Stability of Simple Symmetric Fractional Hedonic Games”. In: *Proceedings of the 9th International Symposium on Algorithmic Game Theory*. Ed. by M. Gairing and R. Savani. Springer-Verlag, 2016, pp. 220–232.

[81] M. Karakaya. “Hedonic Coalition Formation Games: A New Stability Notion”. In: *Mathematical Social Sciences* 61.3 (2011), pp. 157–165.

[82] R. Karp. “Reducibility among Combinatorial Problems”. In: *Complexity of Computer Computations*. Ed. by R. Miller, J. Thatcher, and J. Bohlinger. Springer, 1972, pp. 85–103.

[83] A. Kerkmann. “Stability in FEN-Hedonic Games”. Master’s thesis. Heinrich-Heine-Universität Düsseldorf, Mar. 2019.

[84] A. Kerkmann and J. Rothe. “Altruism in Coalition Formation Games”. In: *Proceedings of the 29th International Joint Conference on Artificial Intelligence*. ijcai.org, July 2020, pp. 347–353.

[85] A. Kerkmann and J. Rothe. “Altruism in Coalition Formation Games”. In: *Nonarchival website proceedings of the 16th International Symposium on Artificial Intelligence and Mathematics*. Jan. 2020. URL: `http://isaim2020.cs.ou.edu/papers/ISAIM2020_Kerkmann_Rothe.pdf`.

[86] A. Kerkmann and J. Rothe. “Four Faces of Altruistic Hedonic Games”. In: *Nonarchival website proceedings of the 8th International Workshop on Computational Social Choice*. June 2021. URL: `https://drive.google.com/file/d/1VlOMwobBVyEiJekDR5pQ1LB69_CyWg1j/view`.

[87] A. Kerkmann and J. Rothe. “Popularity and Strict Popularity in Altruistic Hedonic Games and Minimum-Based Altruistic Hedonic Games (Extended Abstract)”. In: *Proceedings of the 21th International Conference on Autonomous Agents and Multiagent Systems*. To appear. IFAAMAS, May 2022.

[88] A. Kerkmann and J. Rothe. “Popularity and Strict Popularity in Average-Based and Minimum-Based Altruistic Hedonic Games”. Submitted to the *47th International Symposium on Mathematical Foundations of Computer Science*. 2022.

[89] A. Kerkmann and J. Rothe. “Stability in FEN-Hedonic Games for Single-Player Deviations”. In: *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems*. IFAAMAS, May 2019, pp. 891–899.

[90] A. Kerkmann, S. Cramer, and J. Rothe. “Altruism in Coalition Formation Games”. Submitted to the *Annals of Mathematics and Artificial Intelligence*. 2022.

[91] A. Kerkmann, N. Nguyen, A. Rey, L. Rey, J. Rothe, L. Schend, and A. Wiechers. “Altruistic Hedonic Games”. Submitted to the *Journal of Artificial Intelligence Research*. 2022.

[92] A. Kerkmann, J. Lang, A. Rey, J. Rothe, H. Schadrack, and L. Schend. “Hedonic Games with Ordinal Preferences and Thresholds”. In: *Journal of Artificial Intelligence Research* 67 (2020), pp. 705–756.

[93] A. Kerkmann, N. Nguyen, and J. Rothe. “Local Fairness in Hedonic Games via Individual Threshold Coalitions”. In: *Theoretical Computer Science* 877 (2021), pp. 1–17.

[94] B. Klaus, D. Manlove, and F. Rossi. “Matching under Preferences”. In: *Handbook of Computational Social Choice*. Ed. by F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. Procaccia. Cambridge University Press, 2016. Chap. 14, pp. 333–355.

[95] J. Lang and J. Rothe. “Fair Division of Indivisible Goods”. In: *Economics and Computation. An Introduction to Algorithmic Game Theory, Computational Social Choice, and Fair Division*. Ed. by J. Rothe. Springer Texts in Business and Economics. Springer-Verlag, 2015. Chap. 8, pp. 493–550.

[96] J. Lang, A. Rey, J. Rothe, H. Schadrack, and L. Schend. “Representing and Solving Hedonic Games with Ordinal Preferences and Thresholds”. In: *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems*. IFAAMAS, May 2015, pp. 1229–1237.

[97] L. Levin. “Universal Sorting Problems”. In: *Problemy Peredaci Informacii* 9 (1973). In Russian. English translation by B. Trakhtenbrot [140], pp. 115–116.

[98] D. Manlove. *Algorithmics of Matching Under Preferences*. Vol. 2. Series on Theoretical Computer Science. World Scientific Publishing, 2013.

[99] A. Meyer and L. Stockmeyer. “The Equivalence Problem for Regular Expressions with Squaring Requires Exponential Space”. In: *Proceedings of the 13th IEEE Symposium on Switching and Automata Theory*. IEEE Computer Society Press, 1972, pp. 125–129.

[100] L. Miles. “A Simulator for Hedonic Games”. In: *ACM Computing Research Repository* (July 2017). arXiv: `1706.08501`.

[101] G. Monaco, L. Moscardelli, and Y. Velaj. “On the Performance of Stable Outcomes in Modified Fractional Hedonic Games with Egalitarian Social Welfare”. In: *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems*. IFAAMAS, May 2019, pp. 873–881.

[102] G. Monaco, L. Moscardelli, and Y. Velaj. “Stable Outcomes in Modified Fractional Hedonic Games”. In: *Autonomous Agents and Multi-Agent Systems* 34.1 (2019), 4:1–4:29.

[103] J. von Neumann. “Zur Theorie der Gesellschaftsspiele”. In: *Mathematische Annalen* 100.1 (1928), pp. 295–320.

[104] J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 1944.

[105] N. Nguyen and J. Rothe. “Local Fairness in Hedonic Games via Individual Threshold Coalitions”. In: *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems*. IFAAMAS, May 2016, pp. 232–241.

[106] N. Nguyen and J. Rothe. “Local Fairness in Hedonic Games via Individual Threshold Coalitions”. In: *Nonarchival website proceedings of the 7th International Workshop on Cooperative Games in Multiagent Systems*. May 2016. URL: `https://mimuw.edu.pl/~tpm/CoopMas16/wp-content/uploads/2016/05/CoopMAS-2016-Proceedings.pdf`.

[107] N. Nguyen, A. Rey, L. Rey, J. Rothe, and L. Schend. “Altruistic Hedonic Games”. In: *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems*. IFAAMAS, May 2016, pp. 251–259.

[108] N. Nguyen, A. Rey, L. Rey, J. Rothe, and L. Schend. “Altruistic Hedonic Games”. In: *Nonarchival website proceedings of the 7th International Workshop on Cooperative Games in Multiagent Systems*. May 2016. URL: `https://mimuw.edu.pl/~tpm/CoopMas16/wp-content/uploads/2016/05/CoopMAS-2016-Proceedings.pdf`.

[109] N. Nguyen, A. Rey, L. Rey, J. Rothe, and L. Schend. “Altruistic Hedonic Games”. In: *Nonarchival website proceedings of the 6th International Workshop on Computational Social Choice*. June 2016. URL: `http://research.illc.uva.nl/COMSOC/proceedings/comsoc-2016/NguyenReyEtAlCOMSOC2016.pdf`.

[110] M. Olsen. “On Defining and Computing Communities”. In: *Proceedings of the 18th Computing: The Australasian Theory Symposium*. Vol. 128. 2012, pp. 97–102.

[111] K. Ota, N. Barrot, A. Ismaili, Y. Sakurai, and M. Yokoo. “Core Stability in Hedonic Games among Friends and Enemies: Impact of Neutrals”. In: *Proceedings of the 26th International Joint Conference on Artificial Intelligence*. AAAI Press, 2017, pp. 359–365.

[112] C. Papadimitriou. *Computational Complexity*. 2nd ed. Addison-Wesley, 1995.

[113] B. Peleg and P. Sudhölter. *Introduction to the Theory of Cooperative Games*. 2nd ed. Springer-Verlag, 2007.

[114] D. Peters. "Complexity of Hedonic Games with Dichotomous Preferences". In: *Proceedings of the 30th AAAI Conference on Artificial Intelligence*. AAAI Press, Feb. 2016, pp. 579–585.

[115] D. Peters. "Graphical Hedonic Games of Bounded Treewidth". In: *Proceedings of the 30th AAAI Conference on Artificial Intelligence*. AAAI Press, Feb. 2016, pp. 586–593.

[116] D. Peters. "Precise Complexity of the Core in Dichotomous and Additive Hedonic Games". In: *Proceedings of the 5th International Conference on Algorithmic Decision Theory*. Springer-Verlag, 2017, pp. 214–227.

[117] D. Peters and E. Elkind. "Simple causes of complexity in hedonic games". In: *Proceedings of the 24th International Joint Conference on Artificial Intelligence*. AAAI Press/IJCAI, July 2015, pp. 617–623.

[118] W. Poundstone. *Prisoner's Dilemma: John von Neumann, Game Theory, and the Puzzle of the Bomb*. Anchor, 1993.

[119] M. Rahn and G. Schäfer. "Bounding the Inefficiency of Altruism through Social Contribution Games". In: *Proceedings of the 9th International Workshop on Internet & Network Economics*. Springer-Verlag *Lecture Notes in Computer Science #8289*, 2013, pp. 391–404.

[120] G. von Randow. *Das Ziegenproblem: Denken in Wahrscheinlichkeiten*. Rowohlt, 2004.

[121] A. Rey and L. Rey. "FEN-Hedonic Games with Distance-Based Preferences". In: *ACM Computing Research Repository* (Jan. 2022). arXiv: `2201.13158`.

[122] A. Rey, J. Rothe, H. Schadrack, and L. Schend. "Toward the Complexity of the Existence of Wonderfully Stable Partitions and Strictly Core Stable Coalition Structures in Enemy-Oriented Hedonic Games". In: *Annals of Mathematics and Artificial Intelligence* 77.3 (2016), pp. 317–333.

[123] G. Rota. "The number of partitions of a set". In: *The American Mathematical Monthly* 71.5 (1964), pp. 498–504.

[124] A. Roth and M. Sotomayor. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Econometric Society Monographs. Cambridge University Press, 1990.

[125] J. Rothe. *Complexity Theory and Cryptology. An Introduction to Cryptocomplexity*. EATCS Texts in Theoretical Computer Science. Springer-Verlag, 2005.

[126] J. Rothe, ed. *Economics and Computation. An Introduction to Algorithmic Game Theory, Computational Social Choice, and Fair Division*. Springer Texts in Business and Economics. Springer-Verlag, 2015.

[127] J. Rothe. “Inside NP”. In: *Complexity Theory and Cryptology. An Introduction to Cryptocomplexity*. Ed. by J. Rothe. EATCS Texts in Theoretical Computer Science. Springer-Verlag, 2005. Chap. 3.6, pp. 106–114.

[128] J. Rothe. *Komplexitätstheorie und Kryptologie. Eine Einführung in Kryptokomplexität*. eXamen.Press. Springer-Verlag, 2008.

[129] J. Rothe. “Thou Shalt Love Thy Neighbor as Thyself When Thou Playest: Altruism in Game Theory”. In: *Proceedings of the 35th AAAI Conference on Artificial Intelligence*. AAAI Press, Feb. 2021, pp. 15070–15077.

[130] J. Schlueter and J. Goldsmith. “Super Altruistic Hedonic Games”. In: *Proceedings of the 33rd International Florida Artificial Intelligence Research Society Conference*. AAAI Press, May 2020, pp. 160–165.

[131] S. Selvin. “A Problem in Probability”. In: *The American Statistician* 29.1 (1975), pp. 67–71.

[132] S. Selvin. “On the Monty Hall Problem”. In: *The American Statistician* 29.3 (1975), p. 134.

[133] Y. Shoham and K. Leyton-Brown. *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge University Press, 2009.

[134] L. Stockmeyer. “The Polynomial-Time Hierarchy”. In: *Theoretical Computer Science* 3.1 (1977), pp. 1–22.

[135] W. Suksompong. “Individual and group stability in neutral restrictions of hedonic games”. In: *Mathematical Social Sciences* 78 (Nov. 2015), pp. 1–5.

[136] S. Sung and D. Dimitrov. “Computational Complexity in Additive Hedonic Games”. In: *European Journal of Operational Research* 203.3 (2010), pp. 635–639.

[137] S. Sung and D. Dimitrov. “On Core Membership Testing for Hedonic Coalition Formation Games”. In: *Operations Research Letters* 35.2 (2007), pp. 155–158.

[138] S. Sung and D. Dimitrov. “On Myopic Stability Concepts for Hedonic Games”. In: *Theory and Decision* 62.1 (2007), pp. 31–45.

[139] K. Suzuki and S. Sung. *Hedonic Coalition Formation in Conservative Societies*. Tech. rep. SSRN-1700921. Social Science Research Network, Nov. 2010.

[140] B. Trakhtenbrot. “A Survey of Russian Approaches to Perebor (Brute-Force Search) Algorithms”. In: *Annals of the History of Computing* 6.4 (1984), pp. 384–400.

[141] A. Turing. “On Computable Numbers, with an Application to the Entscheidungsproblem”. In: *Proceedings of the London Mathematical Society* 42 (1936), pp. 230–265.

[142] A. Turing. “On Computable Numbers, with an Application to the Entscheidungsproblem. A correction”. In: *Proceedings of the London Mathematical Society* 43 (1937), pp. 544–546.

[143] S. Ueda. “Envy Based Fairness in Hedonic Games”. In: *Proceedings of the Genetic and Evolutionary Computation Conference Companion*. GECCO ’18. Kyoto, Japan: ACM, 2018, pp. 1852–1853. ISBN: 9781450357647.

[144] D. West. *Introduction to Graph Theory*. 2nd ed. Pearson Education, 2001.

[145] A. Wiechers and J. Rothe. “Stability in Minimization-Based Altruistic Hedonic Games”. In: *Proceedings of the 9th European Starting AI Researchers’ Symposium*. Vol. 2655. CEUR-WS.org, Aug. 2020, paper 3.

[146] G. Woeginger. “A Hardness Result for Core Stability in Additive Hedonic Games”. In: *Mathematical Social Sciences* 65.2 (2013), pp. 101–104.

[147] G. Woeginger. “Core Stability in Hedonic Coalition Formation”. In: *Proceedings of the 39th International Conference on Current Trends in Theory and Practice of Computer Science*. Springer-Verlag *Lecture Notes in Computer Science #7741*, Jan. 2013, pp. 33–50.

[148] M. Wright and Y. Vorobeychik. “Mechanism Design for Team Formation”. In: *Proceedings of the 29th AAAI Conference on Artificial Intelligence*. AAAI Press, Jan. 2015, pp. 1050–1056.

List of Figures

List of Tables

Index

Eidesstattliche Erklärung

entsprechend §5 der Promotionsordnung vom 15.06.2018

Ich versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist.

Des Weiteren erkläre ich, dass ich eine Dissertation in der vorliegenden oder in ähnlicher Form noch bei keiner anderen Institution eingereicht habe.

Teile dieser Dissertation wurden bereits in den folgenden Schriften veröffentlicht, zur Publikation angenommen oder zur Begutachtung eingereicht:

- Kerkmann et al. [91] mit den vorläufigen Versionen [107, 108, 109, 145, 129, 86];
- Kerkmann and Rothe [88] mit der vorläufigen Version [87];
- Kerkmann et al. [90] mit den vorläufigen Versionen [84, 85, 46, 86];
- Kerkmann et al. [93] mit den vorläufigen Versionen [105, 106]
- Kerkmann et al. [92] mit den vorläufigen Versionen [96, 89, 83].

Meine Anteile an diesen Schriften werden auf den Seiten 36, 78, 102, 132 und 152 erläutert.

______________________________ ______________________________

Ort, Datum Anna Maria Kerkmann