Altruistic Hedonic Games

Anna Maria Kerkmann   ANNA.KERKMANN@HHU.DE
Nhan-Tam Nguyen   NHAN-TAM.NGUYEN@HHU.DE
Anja Rey   ANJA.REY@HHU.DE
Lisa Rey   LISA.REY@HHU.DE
Jörg Rothe   ROTHE@HHU.DE
Lena Schend   LENA.SCHEND@GOOGLEMAIL.COM
Alessandra Wiechers   ALESSANDRA.WIECHERS@HHU.DE
Heinrich-Heine-Universität Düsseldorf,
Universitätsstraße 1, 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, 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
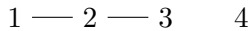
$$1 — 2 — 3 \qquad 4$$

Figure 1: Example of a network of friends

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.

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.

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.[1] 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, as we will see, 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 with respect to axiomatic properties such as anonymity, monotonicity, and friend

---

1. 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."

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* (Coop-MAS'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 6.2, Theorems 5.6, 5.7, 5.8, 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*[2]) 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 2) is guaranteed.

Hoefer, Penn, Polukarov, 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.

---

2. A fundamental property of congestion games is that they always have a *Nash equilibrium in pure strategies* (Rosenthal, 1973, see also, e.g., Ashlagi et al., 2008), 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 altruism in *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),[3] 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 (min-based EQ AHGs) are equivalent to their loyal variant of symmetric friend-oriented hedonic games. Their results for this model imply that, given some min-based EQ AHG, a designated agent, and some threshold, deciding whether there exists a coalition for which this agent's utility reaches the threshold is NP-complete (Bullinger & Kober, 2021, Theorem 2). Furthermore, given some min-based EQ AHG where the underlying network of friends is connected and regular, the coalition structure consisting of the grand coalition is strictly core stable (Bullinger & Kober, 2021, Proposition 10). Yet, they leave the following problems open: the complexity of computing core stable coalition structures for min-based EQ AHGs, and the question of whether there might exist min-based EQ AHGs without core stable coalition structures.

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

---

3. (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.

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.

3. Preliminaries

A *hedonic game* is given by a pair $(N, \succeq)$, where $N = \{1, \ldots, n\}$ is a set of players and $\succeq = (\succeq_1, \ldots, \succeq_n)$ 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 $\Gamma = \{C_1, \ldots, C_k\}$ of the players into k coalitions $C_1, \ldots, C_k \subseteq N$ (i.e., $\bigcup_{r=1}^{k} C_r = N$ and $C_r \cap C_s = \emptyset$ for all distinct $r, s \in \{1, \ldots, k\}$). 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

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

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

$$v_i(A) = n|A \cap F_i| - |A \cap E_i|. \quad (2)$$

$$\text{perfect} \longrightarrow \text{SCS} \longrightarrow \text{CS} \longrightarrow \text{IR}$$
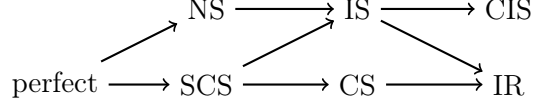
Figure 2: Relations between the stability notions defined in Section 3.2. In this figure, there is a directed path from notion A to notion B if and only if A implies B.

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 $i \in C$ 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.
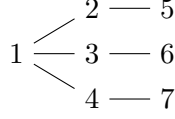
$$2 \relbar\joinrel\relbar 5$$

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

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:

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

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

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

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

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

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

$$u_i^{avgSF}(A) = M \cdot v_i(A) + \text{avg}_i^F(A) \text{ and} \qquad (7)$$
$$u_i^{minSF}(A) = M \cdot v_i(A) + \min_i^F(A) \qquad (8)$$

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

$$u_i^{avgEQ}(A) = \text{avg}_i^{F+}(A) \text{ and} \qquad (9)$$
$$u_i^{minEQ}(A) = \min_i^{F+}(A), \qquad (10)$$

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

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

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

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. *$\text{avg}_i^F(A) > \text{avg}_i^F(B)$ implies $A \succ_i^{avgAL} B$, and*

4. *$\text{min}_i^F(A) > \text{min}_i^F(B)$ 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 $\text{avg}_i^F(A) \geq \text{avg}_i^F(B)$. For $\text{avg}_i^F(A) < \text{avg}_i^F(B)$, it holds if and only if $M > \frac{\text{avg}_i^F(B) - \text{avg}_i^F(A)}{v_i(A) - v_i(B)}$. 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 $M > \frac{v_i(B) - v_i(A)}{\text{avg}_i^F(A) - \text{avg}_i^F(B)}$. 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 $\text{avg}_i^F(A) > \text{avg}_i^F(B)$ 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

$$\text{avg}_i^F(A) - \text{avg}_i^F(B) = \frac{\alpha}{|A \cap F_i|} - \frac{\beta}{|B \cap F_i|} = \frac{\alpha|B \cap F_i| - \beta|A \cap F_i|}{|A \cap F_i||B \cap F_i|} \geq \frac{1}{|A \cap F_i||B \cap F_i|} \geq \frac{1}{n^2}.$$
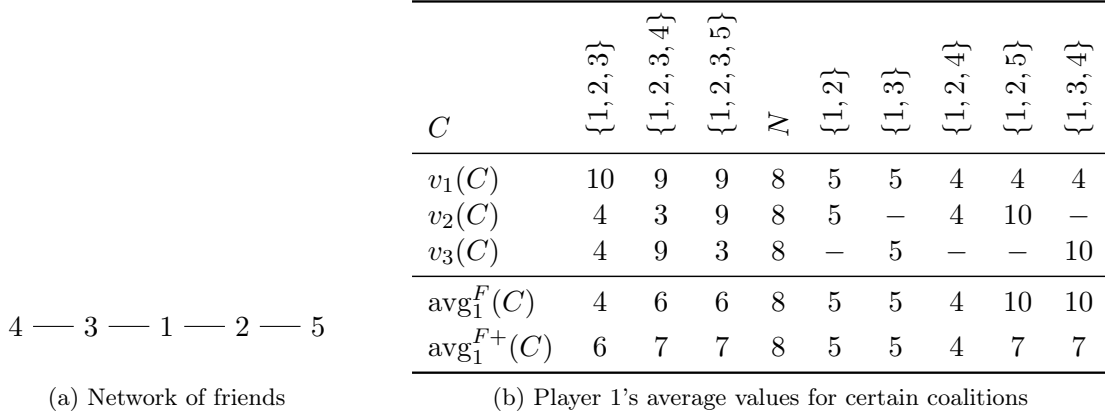
Overall, $M \geq n^4$ suffices for statement (3). □

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

$$4 - 3 - 1 - 2 - 5$$

(a) Network of friends   (b) Player 1's average values for certain coalitions

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

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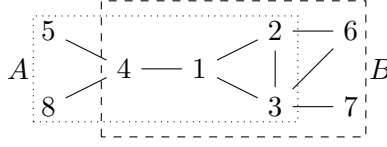
Figure 5: Network of friends with coalitions A and B in Example 4.4

	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

Example 4.4. *Let $N = \{1, \ldots, 8\}$ 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.*

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

Our altruistic models are not fully expressive because players are indifferent between friends and enemies, respectively. Also, for coalitions that only consist of enemies, all our altruistic preference extensions correlate with the original definition of friend-oriented preferences. Yet, AHGs can express games that cannot be expressed by friend-oriented hedonic games and vice versa. To wit, consider the friend-oriented hedonic game that is induced by the network of friends in Figure 4a. If we consider the friend-oriented preferences over size-two coalitions in this game, we see that, for example, agent 1 prefers $\{1, 2\}$ to $\{1, 4\}$. Similar observations can be made for the other agents and all size-two coalitions. Now, assuming that the same preferences can be represented by an AHG (i.e., assuming that this friend-oriented hedonic game is an AHG), 1's preference of $\{1, 2\}$ over $\{1, 4\}$ implies that 2 is a friend of 1's in the network that underlies this AHG, whereas 4 is an enemy of 1's. By consulting all preferences over size-two coalitions, we can then deduce that the AHG with these preferences needs to have the exactly same network of friends as the friend-oriented hedonic game (namely the network in Figure 4a). Yet, we have seen in Example 4.3 that this network of friends leads to different preferences under the friend-oriented and under the altruistic extensions. This is a contradiction to the assumption that the friend-oriented hedonic game that is induced

143

by the network of friends in Figure 4a is also an AHG. Analogously, any AHG induced by the network of friends in Figure 4a is not friend-oriented. Hence, we have seen that friend-oriented hedonic games are neither a subclass of AHGs nor vice versa.

Next, we show that AHGs are 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 models 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 models of altruism. We will now show that there are AHGs that cannot be expressed by any of the above mentioned classes. While similar examples exist for all altruistic models, we now focus on the avg-based EQ extension:

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 $u_1^{avgEQ}(\{1, 2, 3\}) = \frac{(3-1)+3\cdot2}{2} = 4$. 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, fractional hedonic games, hedonic games with $\mathcal{B}$- or $\mathcal{W}$-preferences, or FEN-hedonic games. However, all these classes 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 altruistic hedonic preferences because of indifference between friends and enemies, respectively. (Comparing three coalitions of size two under altruistic hedonic preferences, there will always be an indifference.)

Overall, neither are altruistic hedonic preferences more expressive than any of the other considered models nor the other way around.

5.2 Properties of Preference Extensions

Aiming at determining the axiomatic properties of our altruistic preference extensions, we now give a selection of properties of preference extensions that deduce preferences over coalitions from the friends-and-enemies encoding. These properties are inspired by properties from various related fields such as social choice theory and resource allocation, which also are concerned with preferences.

Let $N = \{1, \ldots, n\}$ 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: Given any coalition $A \in \mathcal{N}^i$ and a friend $f \in F_i \setminus A$, the coalition $A \cup \{f\}$ is acceptable for i.

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).[4] 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).[5] 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.

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 $A \succeq^F_j B$, then $A \succ'_i B$, and (2) if $A \sim_i B$, $j \in A \cap B$, and $A \succeq^F_j B$, 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 $F_1, \ldots, F_n$ 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.

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.

We start by showing some basic properties.

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.*

	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

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. ❑

We continue with the following lemma that will be very useful for the subsequent analysis.

Lemma 5.3. *Under 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

147

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., $v_i(C) > 0$, $\mathrm{avg}_i^F(C) > 0$, and $\min_i^F(C) > 0$, which implies a positive utility.

The second statement follows directly from the first one. ❑

Theorem 5.4. *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: Let $i \in N$, $A \in \mathcal{N}^i$, and $f \in F_i \setminus A$. Since i has at least one friend in $A \cup \{f\}$, it follows by Lemma 5.3.1 that i has a positive utility for $A \cup \{f\}$. Hence, $A \cup \{f\}$ is acceptable for i.

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

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

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.[6] Before turning to Theorem 5.6, we state the following interesting lemma.

Lemma 5.5. *Any avg-based or min-based altruistic preference is locally friend-dependent if and only if it is friend-dependent.*

Proof. We consider some player $i \in N$ and her altruistic preference $\succeq_i$ that was obtained under any of the three degrees of avg-based or min-based altruism. By definition, local friend-dependence implies friend-dependence. So, it remains to show that every pair (A, B) of coalitions that is friend-dependent under $\succeq_i$ is also locally friend-dependent under $\succeq_i$. 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, $\succeq_i$ is locally friend-dependent if and only if $\succeq_i$ is friend-dependent. ❑

By means of Lemma 5.5, we now show the following results.

Theorem 5.6. *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. 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. Due to Lemma 5.5, this also implies that $\succeq_i^\star$ is locally friend-dependent in exactly these cases. 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 $u_i^{minEQ}(C)$ 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.

---

6. 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).

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 $\mathrm{avg}_i^F(\{i, f, e_1\})$ and $\mathrm{avg}_i^F(\{i, f, e_2\})$. If $\mathrm{avg}_i^F(\{i, f, e_1\}) = \mathrm{avg}_i^F(\{i, f, e_2\})$, we change F_f such that $\mathrm{avg}_i^F(\{i, f, e_1\}) \neq \mathrm{avg}_i^F(\{i, f, e_2\})$, 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 $f_1, \ldots, f_{n-1} \in N \setminus \{i\}$, 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 5.7. *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 $\succeq_a^{avgSF'}$ 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 $\succeq_i^{avgSF'}$ 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 $v_j'(B) = v_j(B) + n + 1$.

150

Furthermore,

$$\mathrm{avg}_i^{F'}(A) = \sum_{a \in A \cap F_i'} \frac{v_a'(A)}{|A \cap F_i'|} = \sum_{a \in (A \cap F_i) \cup \{j\}} \frac{v_a'(A)}{|(A \cap F_i) \cup \{j\}|}$$

$$= \sum_{a \in A \cap F_i} \frac{v_a(A)}{|A \cap F_i| + 1} + \frac{v_j'(A)}{|A \cap F_i| + 1}$$

$$= \frac{|A \cap F_i|}{|A \cap F_i| + 1} \cdot \mathrm{avg}_i^F(A) + \frac{v_j(A) + n + 1}{|A \cap F_i| + 1} \quad \text{and} \tag{13}$$

$$\mathrm{avg}_i^{F'}(B) = \frac{|B \cap F_i|}{|B \cap F_i| + 1} \cdot \mathrm{avg}_i^F(B) + \frac{v_j(B) + n + 1}{|B \cap F_i| + 1}. \tag{14}$$

If $A \succ_i^{avgSF} B$ then either (i) $v_i(A) = v_i(B)$ and $\mathrm{avg}_i^F(A) > \mathrm{avg}_i^F(B)$, 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 $|A \cap F_i| = |B \cap F_i|$, $\mathrm{avg}_i^F(A) > \mathrm{avg}_i^F(B)$, and $v_j(A) \geq v_j(B)$ to (13) and (14), we get $\mathrm{avg}_i^{F'}(A) > \mathrm{avg}_i^{F'}(B)$. Then, $v_i'(A) = v_i'(B)$ and $\mathrm{avg}_i^{F'}(A) > \mathrm{avg}_i^{F'}(B)$ implies $A \succ_i^{avgSF'} B$. In case (ii), $v_i(A) > v_i(B)$ implies $v_i'(A) > v_i'(B)$. Hence, $A \succ_i^{avgSF'} B$.

If $A \sim_i^{avgSF} B$ then $v_i(A) = v_i(B)$ and $\mathrm{avg}_i^F(A) = \mathrm{avg}_i^F(B)$. $v_i(A) = v_i(B)$ implies $v_i'(A) = v_i'(B)$ and $|A \cap F_i| = |B \cap F_i|$. Applying $|A \cap F_i| = |B \cap F_i|$, $\mathrm{avg}_i^F(A) = \mathrm{avg}_i^F(B)$, and $v_j(A) \geq v_j(B)$ to (13) and (14), we get $\mathrm{avg}_i^{F'}(A) \geq \mathrm{avg}_i^{F'}(B)$. Hence, $v_i'(A) = v_i'(B)$ and $\mathrm{avg}_i^{F'}(A) \geq \mathrm{avg}_i^{F'}(B)$ implies $A \succeq_i^{avgSF'} B$.

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 $A \succ_i^{avgSF'} B$.

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, $A \succ_i^{avgSF'} B$.

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 5.8. *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 $v_1(A) = v_1(B) = 11$, $v_2(A) = v_2(B) = -3$, $v_3(A) = v_4(A) = 11$, 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,
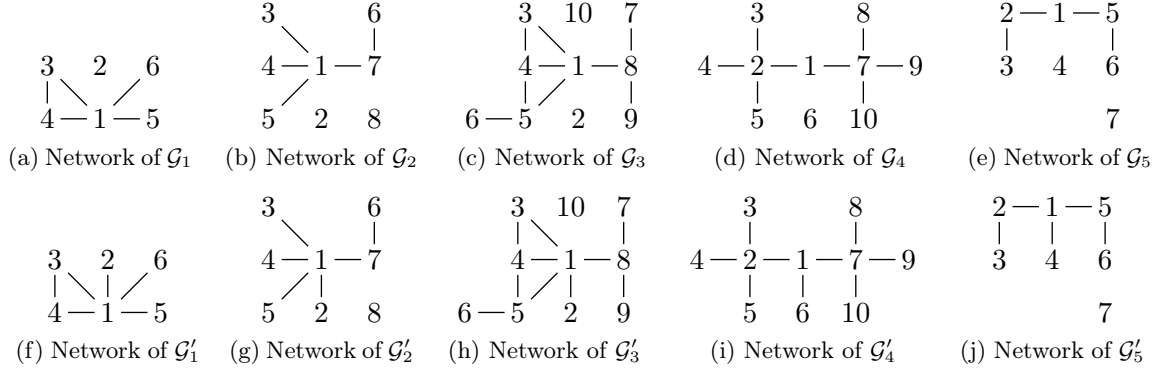
(a) Network of $\mathcal{G}_1$ (b) Network of $\mathcal{G}_2$ (c) Network of $\mathcal{G}_3$ (d) Network of $\mathcal{G}_4$ (e) Network of $\mathcal{G}_5$

(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 5.8

	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 $B = \{1,\ldots,6\}$: $A \succ_1^{avgAL} B$, $2 \in A \cap B$, $v_2(A) \geq v_2(B)$ 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 5.8

$A \sim_1^{minSF} B$, $A \sim_1^{minEQ} B$, 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. □

152

Note that the above results of Theorems 5.7 and 5.8 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. Note that whenever the verification problem is in P, the existence problem is in NP, as we can then guess and verify a stable coalition structure in nondeterministic polynomial time. Our complexity results for these problems are summarized in Table 4.

Starting our analysis with individual rationality, note that the first statement of Lemma 5.3 provides the following characterization.

Corollary 6.1. *For any AHG $(N, \succeq)$, a coalition structure $\Gamma \in \mathcal{C}_N$ 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 6.1, individual rationality verification is in P and existence is trivial because $\Gamma = \{\{1\}, \ldots, \{n\}\}$ is always individually rational.

The next useful lemma only considers SF preferences.

	VERIFICATION		EXISTENCE	
	general	SF	general	SF
ind. rationality	in P	in P	YES	YES
Nash stability	in P	in P	YES	YES
ind. stability	in P	in P	YES	YES
contr. ind. 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.

Lemma 6.2. *Given avg-based or min-based SF preferences and coalitions $C, D \in \mathcal{N}^i$, where D is a clique of size k (in the underlying network of friends), 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.*

Proof. Let $C, D \in \mathcal{N}^i$ such that D is a clique of size k. Then i and all her friends in D each have $k-1$ friends in D, i.e., $v_i(D) = n(k-1)$, $\text{avg}_i^F(D) = n(k-1)$, and $\min_i^F(D) = n(k-1)$.

First, assume that i prefers C to D. Then, either **(a)** $v_i(C) > v_i(D)$ or **(b)** $v_i(C) = v_i(D)$ and $\text{avg}_i^F(C) > \text{avg}_i^F(D)$ (or $\min_i^F(C) > \min_i^F(D)$). If **(a)** holds, then i obviously has at least k friends in C. If **(b)** holds, then $v_i(C) = v_i(D) = n(k-1)$. Therefore, C consists of i and $k-1$ of i's friends, and no further players. However, this is a contradiction because, for this coalition, it is impossible that $\text{avg}_i^F(C) > \text{avg}_i^F(D) = n(k-1)$ (or $\min_i^F(C) > \min_i^F(D) = n(k-1)$). Hence, case **(a)** has to be true.

Second, assume that i is indifferent between C and D. Then, again, $v_i(C) = n(k-1)$ implies that C consists of i and $k-1$ of i's friends. Since $\text{avg}_i^F(C) = n(k-1)$ (or $\min_i^F(C) = n(k-1)$), it is further implied that all of i's friends in C have $k-1$ friends. Hence, 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. First, as for any reasonable class of hedonic games, we can verify these concepts in polynomial time.

Proposition 6.3. *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 6.4. *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 5.3.1 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 5.3.1. 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 138 or Aziz & Savani, 2016). ❑

We now turn to core stability. Theorem 6.5 is inspired by a result of Dimitrov et al. (2006).

Theorem 6.5. *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 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

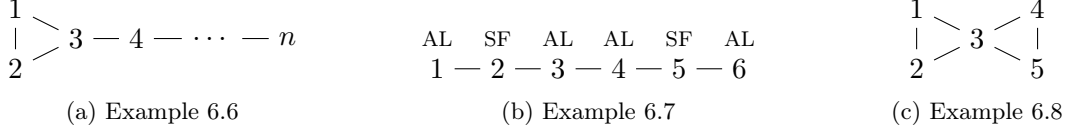(a) Example 6.6   (b) Example 6.7   (c) Example 6.8

Figure 7: Three networks of friends for Examples 6.6, 6.7, and 6.8

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 6.5—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 the other degrees of altruism. Indeed, the following example shows that this coalition structure may not be core stable in these cases.

Example 6.6. *Consider an AHG $(N, \succeq)$ with $n \geq 8$ agents and the network of friends in Figure 7a. Furthermore, consider the coalition structure $\Gamma = \{N\}$ consisting of the grand coalition, i.e., of the only connected component in this network of friends.*

It holds that $C = \{1, 2, 3\}$ blocks Γ under avg-based AL preferences: $\text{avg}_i^F(C) = 2n$ for $i \in \{1, 2, 3\}$ but $\text{avg}_1^F(N) = \frac{v_2(C) + v_3(C)}{2} = \frac{(n+3) + (2n+4)}{2} = 1.5n + 3.5$, $\text{avg}_2^F(N) = 1.5n + 3.5$, and $\text{avg}_3^F(N) = \frac{v_1(C) + v_2(C)}{2} = \frac{(n+3) + (n+3)}{2} = n + 3$. Hence, 1, 2, and 3 prefer C to N for $n \geq 8$. Similar calculations show that C blocks Γ under avg-based and min-based EQ and under min-based AL preferences.

Note that for general AHGs with mixed preferences even a short path of only six agents can have a blocking coalition.

Example 6.7. *Consider an AHG $(N, \succeq)$ with six agents and the network of friends in Figure 7b. 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. It holds that $C = \{1, 2, 3\}$ blocks Γ: 1 prefers C to N because $\text{avg}_1^F(C) = \min_1^F(C) = v_2(C) = 2n > 2n - 3 = v_2(N) = \text{avg}_1^F(N) = \min_1^F(N)$; 2 prefers C to N because $v_2(C) = 2n > 2n - 3 = v_2(N)$; and 3 prefers C to N because $\text{avg}_3^F(C) = \min_3^F(C) = v_2(C) = 2n > 2n - 3 = \text{avg}_3^F(N) = \min_3^F(N)$.

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 138 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 7c. 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* $B = \{3, 4, 5\}$, *and*

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 $|S_i| = 3$ 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

$$N = \{\beta_b \mid b \in B\} \cup \{\zeta_S \mid S \in \mathscr{S}\} \cup \{\alpha_{S,1}, \alpha_{S,2}, \alpha_{S,3} \mid S \in \mathscr{S}\} \cup \{\delta_{S,1}, \ldots, \delta_{S,4k-3} \mid S \in \mathscr{S}\}.$$
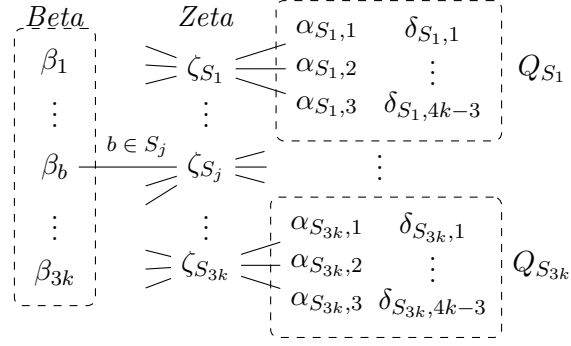
Figure 8: Network of friends in the proof of Theorem 6.9

Define the sets

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

The network of friends is given in Figure 8, 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 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, which in turn is true 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 6.2 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 $\#_\beta = 3k - 1$: In this case, each of the β-players in C has $3k - 2$ β-friends in C and additionally needs at least 2 ζ-friends in C. It follows that there are at least $2(3k - 1) = 6k - 2$ edges between the β- and ζ-players in C. Then there are at least $2k$ ζ-players in C. Thus $\#_\beta + \#_\zeta \geq (3k - 1) + (2k) = 5k - 1$, which is a contradiction (for $k > 2$) to $\#_\beta + \#_\zeta \leq 4k + 1$. 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 one 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 $\operatorname{avg}_{\zeta_S}^F(C) \geq \operatorname{avg}_{\zeta_S}^F(\Gamma(\zeta_S))$. In C, ζ_S has three β-friends who all have the same valuation: $\operatorname{avg}_{\zeta_S}^F(C) = n \cdot (3k) - k$. In $\Gamma(\zeta_S)$, ζ_S has three α-friends who all have the same valuation: $\operatorname{avg}_{\zeta_S}^F(\Gamma(\zeta_S)) = n \cdot (4k)$. Hence, $\operatorname{avg}_{\zeta_S}^F(C) < \operatorname{avg}_{\zeta_S}^F(\Gamma(\zeta_S))$, 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.   *Only if:*   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. *If:*   This implication 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 5.3.2, 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:

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

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

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

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, local friend dependence is a crucial property that distinguishes altruistic preferences from previous models, e.g., from friend-oriented preferences (Dimitrov et al., 2006). 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 types of 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. In particular, we consider it an important task to determine the complexity of finding or deciding the existence of (strictly) core stable coalition structures in equal-treatment and altruistic-treatment altruistic hedonic games. Also, there remain some open questions concerning the perfectness verification and existence problems in equal-treatment and altruistic-treatment altruistic hedonic games.

Comparing our altruistic extensions to the underlying friend-oriented preferences, we can see that our complexity results for selfish-first altruistic hedonic games mirror the known results for friend-oriented hedonic games (Aziz et al., 2013b; Bogomolnaia & Jackson, 2002; Dimitrov et al., 2006; Chen, Csáj, Roy, & Simola, 2022). So, the introduction of selfish-first altruism has no impact on the complexity of the considered stability problems. Yet, the introduction of equal-treatment or altruistic-treatment altruism might cause an increase in the complexity of the (strict) core stability existence problem. While it is easy to com-

pute strictly core stable coalition structures in friend-oriented hedonic games (cf. Dimitrov et al. (2006) or Theorem 6.5), this might not be the case for equal-treatment and altruistic-treatment altruistic hedonic games. As mentioned earlier, we consider determining the precise complexity of these problems an important open problem.

Another future direction might be the consideration of further altruistic models. It might be useful to extend our altruistic models 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

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.

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, J., Csáj, G., Roy, S., & Simola, S. (2022). Cores in friend-oriented hedonic games: Verification is surprisingly harder than searching. Tech. rep. arXiv:2203.09655v1 [cs.GT], ACM Computing Research Repository (CoRR).

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.

Hoefer, M., Penn, M., Polukarov, 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.

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.

Rosenthal, R. (1973). A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory*, *2*, 65–67.

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.