Unknown Agents in Friends Oriented Hedonic Games: Stability and Complexity

Nathanaël Barrot,[1,2] Kazunori Ota,[2] Yuko Sakurai,[3] Makoto Yokoo[2,1]
[1]RIKEN, Center for Advanced Intelligence Project (AIP), [2]Kyushu University,
[3]National Institute of Advanced Industrial Science and Technology (AIST)
nathanaelbarrot@gmail.com, ota@agent.inf.kyushu-u.ac.jp
yuko.sakurai@aist.go.jp, yokoo@inf.kyushu-u.ac.jp

Abstract

We study hedonic games under friends appreciation, where each agent considers other agents *friends*, *enemies*, or *unknown agents*. Although existing work assumed that unknown agents have no impact on an agent's preference, it may be that her preference depends on the number of unknown agents in her coalition. We extend the existing preference, *friends appreciation*, by proposing two alternative attitudes toward unknown agents, *extraversion* and *introversion*, depending on whether unknown agents have a slightly positive or negative impact on preference. When each agent prefers coalitions with more unknown agents, we show that both core stable outcomes and individually stable outcomes may not exist. We also prove that deciding the existence of the core and the existence of an individual stable coalition structure are respectively NP^{NP}-complete and NP-complete.

Introduction

In many real-life examples, ranging from sports clubs to political parties, individuals (*agents*) carry out activities in groups (*coalitions*) to achieve common goals. Coalition formation with hedonic preferences, or *hedonic games*, is concerned with settings where each agent's payoff depends only on the coalition that she joins. A natural question in hedonic games is the existence of stable partitions of the set of agents (*coalition structures*). Various stability concepts have been studied; among the most prominent are *core* and *individual stability* (see handbook Aziz and Savani (2016)).

In hedonic games, each agent's preference is a weak order over the exponentially many coalitions that she can join. Therefore, several compact preference representations have been proposed in the literature. Dimitrov et al. (2006) proposed a notably simple preference representation, where each agent partitions the set of agents into *friends* and *enemies*. They studied a preference extension called *friends appreciation* where each agent prefers coalitions with more friends and, in case of a tie, with fewer enemies. Recently Ohta et al. (2017) slightly extended this model by introducing agents who do not impact preferences (*neutral agents*).

We propose to extend this analysis by generalizing friends appreciation. Indeed it is quite natural to assume that only a subset of agents significatively affects our preferences, i.e. our friends and our enemies. Agents that we don't know enough to identify as friends or enemies, i.e. *unknown agents*, have a minor impact on our preferences. However, we argue that the attitude toward unknown agents is not restricted to *neutrality*. In personality theory, the broadly recognized five-factor model presents *extraversion* as one of the five core personality traits (see introduction (McCrae and John 1992) and handbook (John and Srivastava 1999)). Roughly speaking, individuals with high extraversion tend to be sociable and to enjoy meeting new people, while those with low extraversion tend to be reserved and to avoid large social gatherings. In our setting, neutrality toward unknown agents would lie in the middle of this continuous scale. In this paper, we consider two alternative preferences which take into account this personality trait: *friends appreciation with extraverted agents* and *friends appreciation with introverted agents*, modelling agents who prefer coalitions with respectively more or fewer unknown agents.

From a theoretical viewpoint, analyzing how unknown agents affect stability concepts more precisely is meaningful. Indeed, under friends appreciation without unknown agents, an individually stable coalition structure always exists (Dimitrov and Sung 2004), while when considering unknown agents with no impact on preferences, the question has still to be addressed. In both models, there always exists a core stable coalition structure, which can be found in polynomial time. Since in the more general model of additively separable preferences a core/individually stable outcome may not exist, we hope to clarify a boundary case where core/individual stability is guaranteed.

Our results show that non-neutrality toward unknown agents affects the existence of the core and of an individually stable coalition structure: they are not guaranteed to exist in the presence of extraverted agents. We also study individual stability for friends appreciation under neutrality, showing that an individually stable coalition structure may not exist. For both stability notions under friends appreciation with extraverted agents, we investigate the complexity of verifying whether a given coalition structure is stable (VERIF) and of deciding the existence of a stable coalition structure (EXIST). We show that deciding whether core stable and individually stable outcomes exist are respectively NP^{NP}-complete and NP-complete. The results contrast strongly in the presence of introverted agents: both core stable outcomes and individually stable outcomes always exist and can be computed in polynomial time.

Related work In hedonic games, initiated by Banerjee, Konishi, and Sönmez (2001) and Bogomolnaia and Jackson (2002), a fundamental question is to identify conditions which guarantee the existence of stable coalition structures. Aziz and Brandl (2012) clarified the relations between well-studied stability concepts. In fractional hedonic games, Aziz, Brandt, and Harrenstein (2014) described sufficient conditions for the existence of the core, Brandl, Brandt, and Strobel (2015) showed that an individually stable outcome may not exist, and Monaco, Moscardelli, and Velaj (2018) studied Nash and core stability in a slightly modified setting.

The idea of neutral agents was introduced by Lang et al. (2015), using the generalized Bossong-Schweigert extension principle. They characterized coalition structures that are necessarily/possibly stable. Furthermore, Peters (2016) proposed a graphical representation of hedonic games, where an agent's preference only depends on her neighbors.

Regarding computational complexity, Ballester (2004) showed that the existence problem is NP-complete for core and individual stability under individually rational coalition lists. In additively separable hedonic games, Woeginger (2013) showed that existence of the core is $\mathrm{NP}^{\mathrm{NP}}$-complete, result later extended to symmetric preferences in Peters (2017). Peters and Elkind (2015) identified sufficient conditions on expressivity for proving the NP-hardness of existence problems, which apply to various hedonic games such as hedonic coalition nets (Elkind and Wooldridge 2009). Bilò et al. (2018) showed NP-hardness of computing a best Nash stable outcome in fractional hedonic games.

Outline Preliminaries introduce the model and relevant notions. In the next section, with extraverted agents, we show that the core may be empty and we examine the complexity of deciding its existence. The following section focuses on individual stability; we provide counter-examples and, with extraverted agents, we study the complexity of the existence problem. Finally, we discuss stability in presence of introverted agents and stability when preferences are symmetric.

Preliminaries

Let $N = \{1, \ldots, n\}$ denote the *set of agents*. A *coalition* $C \subseteq N$ is a subset of agents. A *coalition structure* π is a partition of N. Let $\pi(i)$ denote the coalition to which agent i belongs in π. Let C^N denote the set of all coalition structures. A *hedonic game* (N, P) is defined by set of agents N and a *preference profile* $P = (\succsim_i)_{i \in N}$. For every agent i, her *preference* $\succsim_i$ is based on the coalitions to which she belongs; let $\succ_i$ and $\sim_i$ respectively denote the strict preference and the indifference relation derived from $\succsim_i$.

Additively separable hedonic games form a natural class of hedonic game where each agent has a value for any other agent and the utility that an agent derives from a coalition is the sum of the values of its members.

Definition 1 (Additively Separable)**.** A hedonic game $(N, \succsim)$ is *additively separable* if for each agent $i \in N$ there exists a utility function $v_i : N \mapsto \mathbb{R}$ such that $v_i(i) = 0$ and for any two coalitions $S, T \subseteq N$ such that $i \in S, T$,

$$S \succsim_i T \Leftrightarrow \sum_{j \in S} v_i(j) \geq \sum_{j \in T} v_i(j).$$

Hedonic games under friends appreciation and its extensions proposed in this paper are additively separable. Moreover, an additively separable hedonic game is *symmetric* if any two agents associate the same value to each other.

Definition 2 (Symmetry)**.** An additively separable hedonic game satisfies *symmetry* if for all $i, j \in N$, $v_i(j) = v_j(i)$.

We focus on two prominent stability concepts, *core* and *individual stability*. These stability concepts are among the least restrictive -after *individual rationality*- which concern respectively coalition and individual deviations. Individual rationality is a minimal requirement which guarantees that each player weakly prefers her coalition over being alone.

Definition 3 (Individual Rationality)**.** A coalition structure $\pi \in C^N$ is individually rational if there exists no agent that prefers to deviate alone, i.e., for all $i \in N$, $\pi(i) \succsim_i \{i\}$.

Definition 4 (Core Stability)**.** A coalition structure $\pi \in C^N$ admits a blocking coalition $X \subseteq N$ ($X \neq \emptyset$) if for all $i \in X$, $X \succ_i \pi(i)$ holds. The core is the set of coalition structures that do not admit any blocking coalition.

Definition 5 (Individual Stability)**.** A coalition structure $\pi \in C^N$ is individually stable if there exists no agent $i \in N$ and coalition $C \in \pi \cup \{\emptyset\}$ such that $C \cup \{i\} \succ_i \pi(i)$, and for all $j \in C$, $C \cup \{i\} \succsim_j C$.

Intuitively, a coalition structure is core stable if no group of agents benefits from forming a deviating coalition, and it is individually stable if no individual agent benefits from joining an existing coalition without harming any agent in this coalition. Furthermore, a coalition C is *acceptable* to agent i if and only if $C \succsim_i \{i\}$ holds. Thus, if $\pi(i)$ is unacceptable to agent i, π cannot be a member of the core or individually stable.

We consider a simple and compact preference called *friends appreciation*, first proposed by Dimitrov et al. (2006) and extended with neutral agents by Ohta et al. (2017).[1] For each agent i, the set of other agents $N \setminus \{i\}$ is partitioned into friends, enemies, and unknown agents, and we denote the respective sets of agents by F_i, E_i, and $\bot_i$. Under *friends appreciation*, when comparing two coalitions, an agent first compares the number of her friends in each coalition, and then the number of her enemies. Unknown agents have no impact on preferences, they are considered neutral agents. For two coalitions S and T such that agent $i \in S, T$, agent i prefers the coalition with more friends, and in case of a tie, she prefers the one with fewer enemies: $S \succ_i T \Leftrightarrow$ (1) $|S \cap F_i| > |T \cap F_i|$, or (2) $|S \cap F_i| = |T \cap F_i|$ and $|S \cap E_i| < |T \cap E_i|$.

Notice that $S \sim_i T$ holds iff $|S \cap F_i| = |T \cap F_i|$ and $|S \cap E_i| = |T \cap E_i|$. The set of preference profiles under

[1] Both papers also study an alternative preference called *enemies aversion*, where each agent prefers coalitions with fewer enemies and in case of a tie, with more friends.

friends appreciation is denoted by $\mathcal{P}^F$. A preference in $\mathcal{P}^F$ is additively separable on weights n for a friend, 0 for an unknown agent, and -1 for an enemy.

We propose two alternative preferences, *friends appreciation with extraverted agents* and *friends appreciation with introverted agents*, illustrated in Example 1, by allowing agents to take into account in their preferences the number of unknown agents. An *extraverted* agent considers that unknown agents have a positive impact, whereas an *introverted* agent considers that unknown agents have a negative impact. Under friends appreciation with extraverted agents, in the case of a tie for both the number of friends and enemies, agent i prefers the coalition with more unknown agents, i.e., for two coalitions, S and T such that $i \in S, T$: $S \succ_i T \Leftrightarrow$ (1) $|S \cap F_i| > |T \cap F_i|$, or (2) $|S \cap F_i| = |T \cap F_i|$ and $|S \cap E_i| < |T \cap E_i|$, or (3) $|S \cap F_i| = |T \cap F_i|$ and $|S \cap E_i| = |T \cap E_i|$ and $|S \cap \perp_i| > |T \cap \perp_i|$. Under friends appreciation with introverted agents, agent i prefers the coalition with fewer unknown agents.

For both preferences, $S \sim_i T$ holds iff $|S \cap F_i| = |T \cap F_i|$, $|S \cap E_i| = |T \cap E_i|$, and $|S \cap \perp_i| = |T \cap \perp_i|$. The set of preference profiles under friends appreciation with extraverted (resp. introverted) agents is denoted by $\mathcal{P}^{F+}$ (resp. $\mathcal{P}^{F-}$). A preference in $\mathcal{P}^{F+}$ (resp. $\mathcal{P}^{F-}$) is additively separable with weights n for a friend, $1/n$ for an unknown agent, and -1 for an enemy (resp. n, $-1/n$, and -1).

Definition 6 (HG/F, HG/F+, and HG/F-)**.** An *HG/F* (resp. *HG/F+*, *HG/F-*) is a hedonic game $(N, (\succsim_i)_{i \in N})$ such that each $\succsim_i$ lies in $\mathcal{P}^F$ (resp. $\mathcal{P}^{F+}$, $\mathcal{P}^{F-}$).

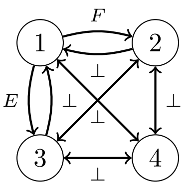
The preferences $\mathcal{P}^F$, $\mathcal{P}^{F+}$, and $\mathcal{P}^{F-}$ are strictly more expressive than the original model, i.e. $\mathcal{P}^F$ with only friends and enemies ($\mathcal{P}_o^F$), and moreover, their pairwise intersections correspond exactly to $\mathcal{P}_o^F$:
1) Consider an agent whose preference lies in $\mathcal{P}^F$. When she has an unknown agent, her preference is not representable by $\mathcal{P}_o^F$, $\mathcal{P}^{F+}$, or $\mathcal{P}^{F-}$. Indeed, these preferences do not allow for similarity between coalitions of different sizes.
2) Now consider an agent whose preference lies in $\mathcal{P}^{F+}$. If she has an unknown agent and another relation (friend or enemy), then her preference is not representable by $\mathcal{P}_o^F$, $\mathcal{P}^F$, or $\mathcal{P}^{F-}$. Indeed an unknown agent and a friend both have a positive impact but not an interchangeable one.
3) Finally consider an agent whose preference lies in $\mathcal{P}^{F-}$. If she has an unknown agent and an enemy, then her preference is not representable by $\mathcal{P}_o^F$, $\mathcal{P}^F$, or $\mathcal{P}^{F+}$. Indeed an unknown agent and an enemy have both a negative impact on preference but not interchangeable. Notice however that if this agent has only unknown agents and friends, then her preference is representable by $\mathcal{P}_o^F$ and thus by $\mathcal{P}^F$, or $\mathcal{P}^{F+}$.

In other words, $\mathcal{P}^F$, $\mathcal{P}^{F+}$, and $\mathcal{P}^{F-}$ generalise the original friends appreciation $\mathcal{P}_o^F$ in three distinct directions.

Furthermore HG/F, HG/F+, and HG/F- are related to *anonymous* hedonic games, which are games where agents' preferences depend only on the size of the coalitions they belong to. Indeed, these games are anonymous *within* the sets of friends, enemies, and unknown agents, e.g., two friends of an agent have interchangeable impacts on her preference.

Hedonic games HG/F, HG/F+, or HG/F- can be represented by a labeled directed graph, $G_{EF\perp} = (N, A_E \cup A_F \cup A_\perp)$, where each vertex corresponds to an agent, and arc (i, j) in set A_E (resp. A_F, $A_\perp$) represents that agent i considers agent j an enemy (resp. a friend, an unknown agent).

Example 1. Consider four agents $\{1, 2, 3, 4\}$ and the game which relations are described in the graph below:



Agent 1 has one friend, one enemy, and one neutral. When the game is HG/F, HG/F+, or HG/F-, it leads respectively to preference a, b, or c (restricted to acceptable coalitions):

a: $\{1,2,4\} \sim_1 \{1,2\} \succ_1 \{1,2,3,4\} \sim_1 \{1,2,3\} \succ_1 \{1,4\} \sim_1 \{1\}$
b: $\{1,2,4\} \succ_1 \{1,2\} \succ_1 \{1,2,3,4\} \succ_1 \{1,2,3\} \succ_1 \{1,4\} \succ_1 \{1\}$
c: $\{1,2\} \succ_1 \{1,2,4\} \succ_1 \{1,2,3\} \succ_1 \{1,2,3,4\} \succ_1 \{1\}$

In the context of hedonic games, two well-studied decision problems are Existence and Verification. Given a stability concept and a hedonic game, the former decides whether there exists a stable coalition structure, and the latter verifies whether a given coalition structure is stable. We define HG/F+/IS/EXIST and HG/F+/C/EXIST as the existence problems related to individual and core stability of an HG/F+. Similarly, we define HG/F+/IS/VERIF and HG/F+/C/VERIF as the verification problems.
We assume that the reader is familiar with concepts from complexity theory, particularly with time complexity classes NP, NP^{NP}, and their complements (Garey and Johnson 2002). In our complexity proofs, we utilize the NP-complete problem MAXCLIQUE and the coNP^{NP}-complete problem MINMAXCLIQUE (Ko and Lin 1995), defined below. First, a *clique* is a graph $\mathcal{G} = (\mathcal{V}, \mathcal{A})$ such that for any pair of nodes x, y in $\mathcal{V}$, edge (x, y) belongs to $\mathcal{A}$.

Definition 7 (Problem MAXCLIQUE)**.** Consider a graph $\mathcal{G} = (\mathcal{V}, \mathcal{A})$ and a threshold $k \in \mathbb{N}$. Does there exist a subset of k vertices $\mathcal{W} \subseteq \mathcal{V}$ such that subgraph $\mathcal{G}[\mathcal{W}]$ is a clique?

Definition 8 (Problem MINMAXCLIQUE)**.** Consider a graph $\mathcal{G} = (\mathcal{V}, \mathcal{A})$, two sets I, J that partition $\mathcal{V}$ into $\{\mathcal{V}_{i,j} \mid i \in I, j \in J\}$, and a threshold $k \in \mathbb{N}$. For every function $t : I \to J$, does the subgraph $\mathcal{G}[\cup_{i \in I} \mathcal{V}_{i,t(i)}]$ contain a clique of size k?

In other words, $\mathcal{V}$ is partitioned into $|I| \cdot |J|$ subsets $\mathcal{V}_{i,j}$. Then, for each function $t : I \to J$, we consider a MAXCLIQUE problem for the subgraph induced by $\mathcal{G}[\cup_{i \in I} \mathcal{V}_{i,t(i)}]$. Notice that the proof in Ko and Lin (1995) showing coNP^{NP}-completeness of MINMAXCLIQUE, holds even when $J = \{0, 1\}$ and $|\mathcal{V}_{i,0}| = |\mathcal{V}_{i,1}|$ for all $i \in I$.

Core Stability with Extraverted Agents

First, we discuss the existence of a core stable coalition structure under friends appreciation with extraverted agents, as well as the complexity of the existence problem.

The core may be empty

When unknown agents have no impact on preferences, Ohta et al. (2017) showed that a core stable coalition structure always exists and that it can be computed in polynomial time as the strongly connected components of graph G_F. These results also hold in the original friends and enemies model with no unknown agents (Dimitrov et al. 2006). However, when agents prefer coalitions with more unknown agents, the existence of the core is no longer guaranteed.

Theorem 1. *In an HG/F+, the core may be empty.*

To prove this theorem, we utilize the following example:

Example 2. Consider six agents $\{1, 2, 3, 4, 5, 5'\}$ with the following preferences. For $i \in \{1, 2, 3\}$, agent i considers agent $i + 1$ a friend, whereas $i + 1$ regards i as unknown. Agents 4, 5, and $5'$ view each other as unknown agents. Agents 5 and $5'$ see 1 as a friend, but 1 considers them as unknown agents. Other relations are enemy relations. Preferences are illustrated with graph $G_{F\perp}$ in Figure 1.

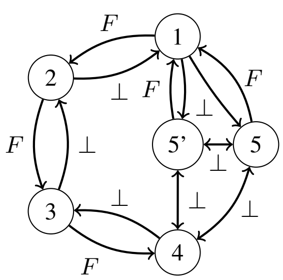


Figure 1: Example of an HG/F+ with an empty core.

Proof. By contradiction, assume that a core stable outcome π exists in Example 2.
First, assume that agents 5 and $5'$ belong to different coalitions in π. If agent 1 does not belong to $\pi(2)$, then coalition $\{1, 5, 5'\}$ is a deviation; thus $\pi(1) = \pi(2)$. If 5 belongs to $\pi(1)(= \pi(2))$, then 3 belongs to $\pi(1)$ (otherwise 2 deviates alone), which implies that 4 also belongs to $\pi(1)$ (otherwise 3 deviates alone). However, 4 has enemies but no friend in this coalition, thus 4 deviates alone. Therefore 5 does not belong to $\pi(1)$, and by symmetry between 5 and $5'$, $\{5, 5'\} \cap \pi(1) = \emptyset$. Now, if agent 4 does not belong to $\pi(3)$, coalition $\{4, 5, 5'\}$ is a deviation, but when 4 belongs to $\pi(3)$, coalition $\{5, 5'\}$ is a deviation. Therefore, 5 and $5'$ belong to the same coalition. Since they have similar preferences, we consider them as one agent 5-$5'$ in the following.
Now, assume that there exists a coalition with three agents or more from $\{1, 2, 3, 4, 5\text{-}5'\}$. This coalition must include an agent with no friend and at least one enemy, who then prefers to deviate in a singleton coalition. Therefore, each coalition consists of at most two agents, which implies that at least one agent is in a singleton coalition. However, if agent 1 (resp. 2, 3, 4) is in a singleton coalition, then coalition $\{1, 5\text{-}5'\}$ (resp. $\{1, 2\}$, $\{2, 3\}$, $\{3, 4\}$) is a deviation, since agent 5-$5'$ (resp. 1, 2, 3) gains one friend. Similarly, if agent 5-$5'$ is in a singleton coalition, then $\{4, 5\text{-}5'\}$ is a deviation since 4 gains one unknown agent. Thus, the core is empty. □

Even if the impact of unknown agents on preferences is negligible compared to the impact of friends or enemies, Theorem 1 shows that they can greatly affects core stability. An interesting question is then how difficult it is to decide if a given HG/F+ has an empty core.

Computational Complexity

In this subsection, we study the complexity of the existence of a core stable outcome under friends appreciation with extraverted agents. First, notice that an HG/F+ where there exists no friend relation is equivalent to a hedonic game under *enemies aversion* (HG/E, see Footnote 1) where unknown arcs from the original graph become friend arcs in the new one. Since under enemies aversion, verifying that a given coalition structure is in the core is coNP-complete (Sung and Dimitrov 2007), it extends to our setting:

Theorem 2. *Problem* HG/F+/C/VERIF *is coNP-complete.*

This result implies that the corresponding existence problem is in NP^{NP}. Moreover, we show the following:

Theorem 3. *Problem* HG/F+/C/EXIST *is* $\textit{NP}^{\textit{NP}}$*-complete.*

The main argument of the proof resembles to the one of Theorem 4 in (Ohta et al. 2017) concerning HG/E, however we have to adapt it to our setting. Indeed, while in an HG/E, coalitions in a core stable partition are necessarily cliques (in the graph $G_{F\perp}$), this does not hold in an HG/F+. We adapt it by introducing cliques where only single nodes were necessary in their proof and by utilizing the following remarks.

Remark 1. Consider an HG/F+ where there exists a clique of friends, K, which agents have no friends outside of K. Then a coalition structure that divides the agents of K into different coalitions is not core stable (since K is a deviation).

Remark 2. Consider an HG/F+ composed of b cliques of friends $\{K^1, \dots, K^b\}$ and assume that 1) agents have no friends outside of their clique, 2) agents in the same clique have the same set of unknown agents, which is a union of cliques in $\{K^1, \dots, K^b\}$, and 3) the game is symmetric. Then, a core stable coalition structure exists.

Proof of Theorem 3. First, by Theorem 2, HG/F+/C/EXIST is in NP^{NP}. We prove NP^{NP}-hardness by reducing the coNP^{NP}-complete MINMAXCLIQUE to the complement of HG/F+/C/EXIST. Let an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{A})$, a set I partitioning set $\mathcal{V}$ into $\{\mathcal{V}_{i,0}, \mathcal{V}_{i,1} \mid i \in I\}$, and an integer $k \in \mathbb{N}$, define a restricted instance of MINMAXCLIQUE, where $\forall i \in I, |\mathcal{V}_{i,0}| = |\mathcal{V}_{i,1}|$. We construct the corresponding instance of coHG/F+/C/EXIST as follows, illustrated by a partial representation of graph $G_{F\perp}$ in Figure 2:
(1) For each x in $\mathcal{V}$, we create *vertex-clique* K^x that contains k' mutual friends (k' is specified at the end of the proof), and then $V = \bigcup_{i \in I} V_{i,0} \cup V_{i,1}$, where $V_{i,j} = \bigcup_{x \in \mathcal{V}_{i,j}} K^x$. For each edge (x, y) in graph $\mathcal{G} = (\mathcal{V}, \mathcal{A})$, we introduce mutually unknown arcs between each $x' \in K^x$ and $y' \in K^y$.
(2) We introduce a *generalization of Example 2* where agents 5 and $5'$ are replaced by a clique of $k''-1$ mutually unknown agents, K^μ. We set k'' such that a clique of size k in the original graph $\mathcal{G}$ induces a clique of size k'' in graph $G_{F\perp}$.
(3) We introduce a *fulcrum-clique* K^φ of 2 mutual friends. Agents in K^φ consider agents in K^μ and in V as unknown agents. Agents in K^μ view agents in K^φ as friends.

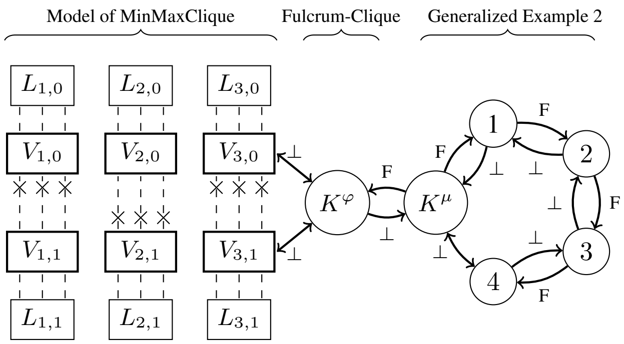


Figure 2: Corresponding instance of coHG/F+/C/EXIST.

(4) Between each pair $V_{i,0}$ and $V_{i,1}$, we introduce $|\mathcal{V}_{i,0}|$ *inhibitors* (marked by $\times$ in Figure 2). Each K^x in $V_{i,0}$ is paired to one K^y in $V_{i,1}$ through an inhibitor which makes exactly one of them available for a core stable coalition with K^φ.
(5) We connect to each $V_{i,j}$ a *logic game* $L_{i,j}$ which contains a blocking coalition iff condition 'all agents in $V_{i,j}$ are inhibited, or none is inhibited' is not satisfied.
(6) All others relations are enemies relations.

Before the main argument, notice that all the vertex-cliques $K^x, x \in \mathcal{V}$, the fulcrum-clique K^φ, the inhibitors, and the cliques from the logic games (described below) satisfy Remark 1. Therefore we do not discuss the stability of coalition structures that divide the agents of those cliques.

Main argument
First, observe that no core stable coalition structure contains a coalition with agents from the left and the right of K^φ (w.r.t. Figure 2). Thus, in a core stable coalition, K^φ is either grouped with K^μ of size $k''-1$ or with a clique which size is at least k'' in V. If clique K^φ goes to the left, the game from the generalization of Example 2 is isolated and thus has an empty core. However, if clique K^φ goes to the right, the core of game $K^\varphi \cup$ Example 2 is non-empty: $\{\{K^\varphi, K^\mu\}, \{1,2\}, \{3,4\}\}$ is core stable.

Assume that there exists a function $t^* : I \to \{0,1\}$ such that subgraph $\mathcal{G}[\cup_{i\in I}\mathcal{V}_{i,t^*(i)}]$ contains no clique of size k. We claim that a core stable coalition structure exists. For each $i \in I$, we set the inhibitors to $V_{i,1-t^*(i)}$. Thus no logic game generates a blocking coalition and only agents in $\cup_{i\in I}V_{i,t^*(i)}$ are available to K^φ. Thus, if we group cliques K^φ and K^μ together, K^φ has no interest to deviate left, which implies that the partial coalition structure $\{\{K^\varphi, K^\mu\}, \{1,2\}, \{3,4\}\}$ admits no blocking coalition (in the whole game). Furthermore, the subgraph composed of vertex-cliques K^x in $\cup_{i\in I}V_{i,t^*(i)}$ represents a game that satisfies Remark 2, and thus, a core stable partition exists.

Conversely, assume that for every function $t : I \to \{0,1\}$, subgraph $\mathcal{G}[\cup_{i\in I}\mathcal{V}_{i,t(i)}]$ contains a clique of size k. By contradiction, assume that a core stable coalition structure π exists. Then there exists function t^π such that, for each $i \in I$, all inhibitors between $V_{i,0}$ and $V_{i,1}$ are set to $1-t^\pi(i)$, otherwise at least one logic game is not core stable. Thus K^φ is grouped with a clique which size is at least k'' in $G_{F\perp}[\cup_{i\in I}V_{i,t^\pi(i)}]$, which exists for any function $t : I \to \{0,1\}$ by assumption. Thus, the game from Example 2 is isolated and contains a blocking coalition.

Inhibitors and logic games (sketch)
Inhibitors and logic games help us model the MINMAXCLIQUE problem; the former make vertex-cliques (un)available for a deviation with K^φ, and the later impose that the set of available vertex-cliques models a function $t : I \to \{0,1\}$. Due to space limit, we only sketch them.

An *inhibitor* is a clique that contains k^* mutual friends (k^* is specified below) and is mutually unknown with one vertex-clique K^x in $V_{i,0}$ and with one vertex-clique K^y in $V_{i,1}$. We enforce that an inhibitor either joins K^x or K^y in a core stable outcome by properly fixing the size k^*.

Each *logic game* $L_{i,j}$ relies on a combination of gadget games, presented in Figure 3, that model logical gates with the understanding that an available agent amounts to Boolean True. All K^* in Figure 3 are cliques of friends. In gate NOT, we assume that K^x and K^y have identical sizes; in gate OR, we assume that K^{x_1} and K^{x_2} have size $s \geq 2$, and that K^α and K^y have size $s-1$; in gate DUPLIC, we assume that the size of K^x is $s \geq 3$, the size of each K^{β_1} and K^{β_2} is $s-2$, and the size of each K^{y_1} and K^{y_2} is $s-1$.

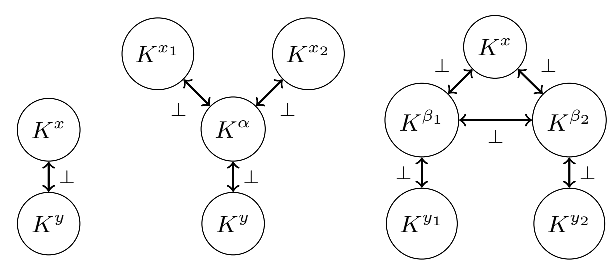


Figure 3: Gates NOT, OR and DUPLIC (In:K^x, Out:K^y).

By combining these logic gates and taking the vertex-cliques $K^x \in V_{i,j}$ as input, each $L_{i,j}$ is constructed to obtain formula $\left(\bigwedge_{K^x\in V_{i,j}} K^x\right) \bigvee \left(\bigwedge_{K^x\in V_{i,j}} \neg K^x\right)$, i.e. "all vertex-cliques $K^x \in V_{i,j}$ are available or none is". The output of each $L_{i,j}$ is then connected to a specific instance of Example 2, such that the core of the game $L_{i,j} \cup$ Example 2 exists if and only if the formula is valid.
Finally, we briefly explain the values of k', k'', and k^*. The construction of the logic games leads us to set the size of each vertex-clique K^x to 7, i.e., $k' = 7$. Furthermore, each vertex-clique K^x is associated with a separating-clique of size k' (not described here), which separates it from the logic game. It implies that a clique of size k in the original MinMaxClique instance leads to a core stable coalition with $(2 \times k') \times k$ unknown agents for K^φ. Thus, we set $k'' = 14 \times k$. Lastly, excluding inhibitors, the maximal number of unknown agents for agent x in $V_{i,j}$ is $7 + 14 \times (|\mathcal{V}| - 1) + 2$, that is, 7 unknown agents from x's separating-clique, $14 \times (|\mathcal{V}| - 1)$ unknown agents from the $(|\mathcal{V}| - 1)$ other vertex-cliques and their separating-cliques, and 2 unknown agents from the fulcrum-clique. So we set $k^* = 14 \times (|\mathcal{V}| - 1) + 10$. $\square$

Hence, although HG/F+ are notably simple (additive on three values), they encompass the highest computational complexity reachable by any class of hedonic game (with polynomially computable preference) w.r.t. core stability.

Individual Stability

Now we investigate the existence of an individually stable coalition structure first under friends appreciation and then under friends appreciation with extraverted agents.

Friends Appreciation under Neutrality

Dimitrov and Sung (2004) proved that, in an HG/F with no unknown agents, individually stable outcomes always exist. When adding unknown agents with no impact on preferences, although Ohta et al. (2017) argued that the core always exists, they did not address individual stability. Theorem 4 proves that individual stability is not guaranteed.

Theorem 4. *In an HG/F, an individually stable coalition structure may not exist.*

To prove this theorem, we utilize the following example.

Example 3. Consider 12 agents $\{0, \dots, 11\}$, divided into four groups: $\{0, 1, 2\}$, $C_0 = \{3, 4, 5\}$, $C_1 = \{6, 7, 8\}$, and $C_2 = \{9, 10, 11\}$, and the following preferences, illustrated in Figure 4, a partial representation of graph $G_{EF\perp}$.

In this example and the following proof, when we write [3], we mean $\pmod 3$. First, for $i \in \{0, 1, 2\}$, preferences of agents in group C_i are defined as follows:

- agents $3(i+1)+1$ and $3(i+1)+2$ consider agent $3(i+1)$ a friend, and other relations within C_i are unknown relations;
- all agents in C_i consider agents i and $i+1$ [3] unknowns, agent $i+2$ [3] an enemy, and agents from $C_j, j \neq i$, enemies.

For $i \in \{0, 1, 2\}$, agent i's preferences are such that:

- agent i considers agent $i+1$ [3] as friend, but agent $i+1$ [3] regards agent i as an unknown agent;
- for C_i, agent i considers agent $3(i+1)$ as an enemy and agents $3(i+1)+1$ and $3(i+1)+2$ as friends;
- for $C_{i+2[3]}$, agent i regards agent $3((i+2\,[3])+1)$ as an unknown agent and agents $3((i+2\,[3])+1)+1$ and $3((i+2\,[3])+1)+2$ as friends;
- Finally, agent i is mutually enemy with agents in $C_{i+1[3]}$.

Proof. Assume that an individually stable coalition structure π exists in Example 3. We first focus on C_0. Notice that agent 3 has no friend, which implies that $\pi(3)$ does not contain any enemy of 3; otherwise 3 deviates alone. Notice also that agents 3, 4, and 5 have identical preferences toward agents outside of C_0, and that 4 and 5 share a unique friend, that is agent 3. This implies that agents in C_0 are in the same coalition, which does not contain any enemy of 3.

By symmetry, for $i \in \{0, 1, 2\}$, agents in C_i belong to the same coalition, which does not contain any enemy of agent $3(i+1)$. Furthermore, for $i \in \{0, 1, 2\}$, agent i belongs to the coalition of C_i or of $C_{i+2[3]}$, where i has at least two friends. Thus, for $i \in \{0, 1, 2\}$, C_i is either alone, grouped with agent i, with agent $i+1$ [3], or with both.

Then, there are two cases to consider: (1) agents 0, 1, and 2 join three different coalitions, or (2) two agents from

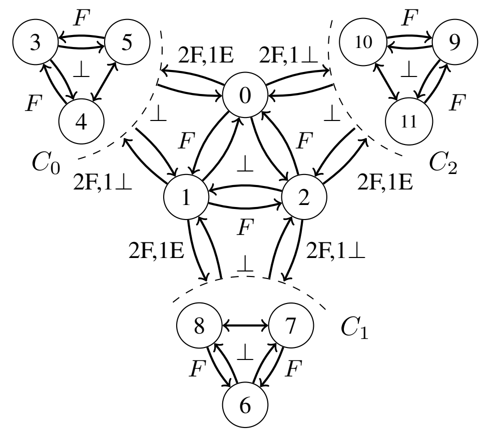


Figure 4: An HG/F with no individually stable outcome.

$\{0, 1, 2\}$ join the same coalition.

Case (1): If agent 0 joins C_0, then 0 deviates toward coalition $\{2\} \cup C_2$, where she has two friends and no enemy. However if 0 joins C_2, then 0 deviates toward coalition $\{1\} \cup C_0$, where she has three friends.

Case (2): Without loss of generality, assume that agents 0 and 1 join the same coalition, i.e., $\pi(0) = \pi(1) = \{0, 1\} \cup C_0$. If agent 2 joins C_2, then 2 deviates toward C_1, where she has no enemy. However, if agent 2 joins C_1, then 1 deviates toward coalition $\{2\} \cup C_1$, where she has three friends.

As a result, no individually stable outcome exists. ☐

Friends Appreciation with Extraverted Agents

Back to our model in the presence of extraverted agents, we prove now that an individually stable coalition structure may not exist and that deciding its existence is NP-complete.

Theorem 5. *In an $HG/F+$, an individually stable coalition structure may not exist.*

Proof. The proof is based on the same example as in the proof of Theorem 1. Assume that an individual stable coalition structure π exists in Example 2. Furthermore, assume that agents 5 and $5'$ belong to different coalitions.

Notice first that $\pi(4) \cap \{1, 2\} = \emptyset$, since otherwise agent 4 deviates in a singleton coalition. It implies that agent 3 does not belong to $\pi(1)$, since otherwise 3 deviates in a singleton coalition. Then, assume that agent 1 does not belong to $\pi(2)$. If $\pi(1) \cap \{5, 5'\} = \emptyset$, then 5 deviates toward $\{1\}$. However, if $\pi(1) = \{1, 5\}$ (resp. $\{1, 5'\}$), then $5'$ (resp. 5) deviates toward $\{1, 5\}$ (resp. $\{1, 5'\}$).

Thus, 1 belongs to $\pi(2)$, which implies that $\pi(1) \cap \{5, 5'\} = \emptyset$, since otherwise agent 2 deviates in a singleton coalition. Then, 3 does not belong to $\pi(5)$ or $\pi(5')$, since otherwise 5 or $5'$ deviates in singleton coalition. In other words, agents 5 and $5'$ are either both in singleton coalitions or just one of them is and the other belongs to $\pi(4)$. If 5 and $5'$ are both in singleton coalitions, then 5 deviates toward $\{5'\}$. Thus, $\pi(4) \cap \{5, 5'\} \neq \emptyset$, but then 5 (resp. $5'$) deviates toward $\{4, 5'\}$ (resp. $\{4, 5\}$). Therefore, 5 and $5'$ belong to the same coalition and a similar argument as in Theorem 1 applies, showing that no individually stable outcome exists. □

Similarly to Theorem 1, this result shows that adding unknown agents in the presence of extraverts impacts individual stability. Hence, we turn to the complexity of deciding the existence of an individually stable outcome. Notice first that the verification problem for individual stability is trivially polynomial for any class of hedonic game, and thus, the existence problem is in NP. Moreover, we show:

Theorem 6. *Problem* HG/F+/IS/EXIST *is NP-complete.*

Proof. To prove NP-hardness, we reduce the NP-complete problem MAXCLIQUE to problem HG/F+/IS/EXIST.

Let graph $\mathcal{G} = (\mathcal{V}, \mathcal{A})$ and threshold $k \in \mathbb{N}$ define an instance of MAXCLIQUE. We construct the corresponding instance of HG/F+/IS/EXIST with n vertex-agents in $V \equiv \mathcal{V}$ (modeling graph $\mathcal{G}$), n cliques of unknown agents $(K_i)_{i \in V}$, each of size k, and three agents $1', 2', 3'$; therefore the set of agents is $N = V \cup (K_i)_{i \in V} \cup \{1', 2', 3'\}$. The preferences are define as follows, illustrated in Figure 5.
In set V, for each edge (i, j) in graph $\mathcal{G} = (\mathcal{V}, \mathcal{A})$, we construct unknown arcs (i, j) and (j, i). Moreover, all agents in V consider agent $3'$ unknown, whereas agent $3'$ considers each of them friends. For $i \in V$, all the k agents in K_i are mutually unknown toward each other, mutually unknown with agent $i \in V$, and they consider agent $1'$ a friend, whereas agent $1'$ considers them unknown agents. Furthermore, agent $1'$ (resp. $2'$) considers agent $2'$ (resp. $3'$) as friend, whereas agent $2'$ (resp. $3'$) considers agent $1'$ (resp. $2'$) as unknown. Finally, all other arcs are enemies; in particular agents from different cliques in $(K_i)_{i \in V}$ are enemies.
First, notice that the structure of the whole game is based

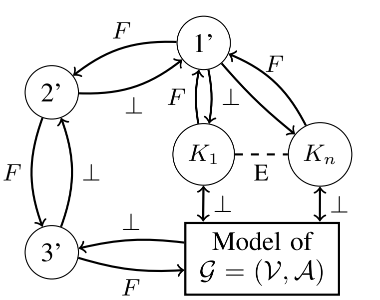


Figure 5: Instance of HG/F+/IS/EXIST from MAXCLIQUE.

on Example 2, with cliques $(K_i)_{i \in N}$ in place of agents 5 and $5'$, and agents in V (model of $\mathcal{G}$) embedded in place of agent 4. Thus, with a similar argument, we see that if there is no clique of size k in V, then no individually stable coalition structure exists. However, when a clique C of size k exists, the following coalition structure is individually stable: $\{C \cup \{3'\}, \{1', 2'\}, (K_i \cup \{i\})_{i \notin C}, (K_i)_{i \in C}\}$. Indeed, no agent $i \in V$ in coalition $C \cup \{3'\}$ has an incentive to deviate toward K_i since the number of unknown agents is the same. Thus agent $2'$ cannot join $3'$ and the deviation cycle in Example 2 is interrupted. □

Hence HG/F+ contain some of the most difficult problems in hedonic games related to individual stability.

Discussion

Friends appreciation with introverted agents When all agents are introverted, i.e. they value unknown agents negatively, enemies and unknown agents have similar (but not equivalent) impact on preferences. Thus, we obtain similar results as under friends appreciation with no unknown agents, that is, core and individually stable coalition structures always exist. We prove this result by showing that the strict core always exists, with the same argument developed under friends appreciation with no unknown agents in (Dimitrov et al. 2006). By definition, a strict core stable coalition structure is also core and individually stable.

Theorem 7. *Under friends appreciation with introverted agents (1) a strict core outcome always exists and (2) it can be computed in polynomial time as the strongly connected components of graph G_F.*

Symmetric hedonic games Previous sections focused on non-existence results for core/individually stable outcomes. An interesting question is then what restrictions on preference can guarantee the existence of stable outcomes. We address it for a natural restriction, namely *symmetry*.

Proposition 1. *In symmetric HG/F+, the core always exists. Moreover, in both symmetric HG/F and HG/F+, an individually stable outcome always exists.*

Proof. First, an HG/F+ with symmetric preferences satisfies the top-coalition property which is a sufficient condition for the existence of the core (Banerjee, Konishi, and Sönmez 2001). Moreover, Bogomolnaia and Jackson (2002) proved that an individually stable coalition structure always exists in a symmetric additively separable hedonic game. □

Conclusion

We studied the impact of different attitudes toward unknown agents on stability in friends oriented hedonic games. We proved that three distinct extensions of the existing preference lead to a diverse stability and complexity landscape. With extraverted agents, we provided counterexamples showing that both core and individually stable coalition structures may not exist, whereas the strict core is guaranteed in the presence of introverted agents. Then we proved that deciding the existence of such outcomes is $\mathrm{NP}^{\mathrm{NP}}$-complete for core stability and NP-complete for individual stability. We also proved that an individually stable coalition structure may not exist under friends appreciation with neutrality. An open question is to prove the complexity of deciding the existence of individual stable outcomes under friends appreciation with neutrality.

Acknowledgments

This work was partially supported by JSPS KAKENHI Grant Number JP17H00761, JP18H03299 and JST Strategic International Collaborative Research Program, SICORP. We thank Anisse Ismaili for figures from Ohta et al. (2017).

References

Aziz, H., and Brandl, F. 2012. Existence of stability in hedonic coalition formation games. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2012)*, 763–770.

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

Aziz, H.; Brandt, F.; and Harrenstein, P. 2014. Fractional hedonic games. In *Proceedings of the 13th International Conference on Autonomous Agents and Multi-agent Systems (AAMAS-2014)*, 5–12.

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

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

Bilò, V.; Fanelli, A.; Flammini, M.; Monaco, G.; and Moscardelli, L. 2018. Nash stable outcomes in fractional hedonic games: Existence, efficiency and computation. *Journal of Artificial Intelligence Research* 62:315–371.

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

Brandl, F.; Brandt, F.; and Strobel, M. 2015. Fractional hedonic games: Individual and group stability. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2015)*, 1219–1227.

Dimitrov, D., and Sung, S. C. 2004. Enemies and friends in hedonic games: individual deviations, stability and manipulation.

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

Elkind, E., and Wooldridge, M. 2009. Hedonic coalition nets. In *Proceedings of the 8th International Conference on Autonomous Agents and Multi-agent Systems (AAMAS-2009)*, 417–424.

Garey, M. R., and Johnson, D. S. 2002. *Computers and intractability*, volume 29. wh freeman New York.

John, O. P., and Srivastava, S. 1999. The big five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and research* 2(1999):102–138.

Ko, K.-I., and Lin, C.-L. 1995. On the complexity of min-max optimization problems and their approximation. *Nonconvex optimization and its applications* 4:219–240.

Lang, J.; Rey, A.; Rothe, J.; Schadrack, H.; and Schend, L. 2015. Representing and solving hedonic games with ordinal preferences and thresholds. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2015)*, 1229–1237.

McCrae, R. R., and John, O. P. 1992. An introduction to the five-factor model and its applications. *Journal of personality* 60(2):175–215.

Monaco, G.; Moscardelli, L.; and Velaj, Y. 2018. Stable outcomes in modified fractional hedonic games. *arXiv preprint arXiv:1804.05040*.

Ohta, K.; Barrot, N.; Ismaili, A.; Sakurai, Y.; and 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 (IJCAI-2017)*, 359–365.

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

Peters, D. 2016. Graphical hedonic games of bounded treewidth. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI-2016)*, 586–593.

Peters, D. 2017. Precise complexity of the core in dichotomous and additive hedonic games. In *International Conference on Algorithmic DecisionTheory*, 214–227. Springer.

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

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