Non-obvious Manipulability in Hedonic Games with Friends Appreciation Preferences

Michele Flammini[1,2], Maria Fomenko[1,*] and Giovanna Varricchio[2]

[1]Gran Sasso Science Institute, L'Aquila, Italy
[2]University of Calabria, Rende, Italy

Abstract

We study non-obvious manipulability (NOM), a relaxed form of strategyproofness, in the context of Hedonic Games (HGs) with Friends Appreciation (FA) preferences. In HGs, the aim is to partition agents into coalitions according to their preferences, which solely depend on the coalition they are assigned to. Under FA preferences, agents consider any other agent either a friend or an enemy, preferring coalitions with more friends and, in case of ties, the ones with fewer enemies. Prior research established that computing a welfare maximizing (optimum) partition for FA preferences is not strategyproof, and the best-known approximation to the optimum subject to strategyproofness is linear in the number of agents. In this work, we explore NOM to improve approximation results. We first prove the existence of a NOM mechanism that always outputs the optimum; however, we also demonstrate that the computation of an optimal partition is NP-hard. In turn, we also propose a NOM mechanism guaranteeing a $(4 + o(1))$-approximation in polynomial time. Finally, we briefly discuss NOM in the case of Enemies Aversion (EA) preferences, the counterpart of FA, where agents give priority to coalitions with fewer enemies and show that no mechanism computing the optimum can be NOM.

Keywords

Hedonic Games, Strategyproofness, Non-obvious Manipulability

1. Introduction

Hedonic Games (HGs) [1] are a game-theoretic model describing the coalition formation of selfish agents and have been extensively studied in the literature (e.g.,[2, 3, 4, 5, 6, 7, 8]). In such games, the objective is to partition a set of agents into disjoint coalitions, with each agent's satisfaction determined solely by the members of her coalition. Different HG classes capture various social preferences, such as *additively separable* HGs (ASHGs) [9] or HGs with *friends appreciation* (FA) preferences [10]. A recent stream of research is focusing on designing *strategyproof* (SP) mechanisms [11] which can prevent agents from manipulating the outcome by misrepresenting their preferences. Unfortunately, combining strategyproofness with good social welfare – defined as the sum of the agents' utilities in the outcome – is challenging: even in the simple case of FA preferences the best-known SP mechanism guarantees an approximation of the optimal social welfare linear in the number of agents [12]. Strategyproofness has also been studied in several game-theoretic settings and turned out to be often incompatible with other desirable properties or even impossible to achieve [13, 14, 15, 16]. Moreover, according to the definition of strategyproofness, to successfully manipulate, an agent has to possess the knowledge of others' strategies and deeply understand the underlying mechanics of the game; otherwise, she might end up with an outcome that is worse than the one she attempted to avoid. However, the ability of a cognitively limited agent to satisfy this requirement seems unrealistic, leading to the notion of *non-obviously manipulable* (NOM) mechanisms, which are unlikely to be manipulated in practice [17].

Our Contribution. We initiated the study of NOM within the context of HGs, focusing specifically on FA preferences. Our contribution is threefold: i) we show that a NOM mechanism that computes a social welfare-maximizing partition always exists (Theorem 1); ii) we prove that the underlying

ICTCS 2025: Italian Conference on Theoretical Computer Science, September 10–12, 2025, Pescara, Italy
*Corresponding author.
michele.flammini@gssi.it (M. Flammini); maria.fomenko@gssi.it (M. Fomenko); giovanna.varricchio@unical.it (G. Varricchio)

optimization problem is NP-hard (Theorem 2); iii) we design a polynomial-time NOM mechanism achieving a $(4+o(1))$-approximation (Theorem 3). Finally, we complement these results demonstrating that optimality and NOM may be incompatible in HGs; we show this under the simple and natural class with *enemies aversion* (EA) preferences, the counterpart to FA, where agents prioritize minimizing enemies. Further details can be found in the conference version of this work [18].

Related Work. Alongside the extensive research on HGs [19, 20, 8, 21, 5, 4, 6, 22], the classes with FA and EA preferences have received extensive attention [23, 24, 25, 26, 27]. In terms of strategyproofness, part of the literature has focused on SP mechanisms that ensure some form of stability [10, 23, 28, 29]. More recently, attention has shifted toward approximating maximum social welfare [11, 30, 15, 12]; however, the strategyproofness requirement may have a negative impact on the quality of the outcome. For instance, in the FA setting, the best-known SP mechanism achieves only a linear approximation in the number of agents [12]. In turn, in the case of EA, the best-known polynomial algorithm achieves a linear approximation in the number of agents, while a constant approximation ratio is possible when time complexity is not a concern [12]. For ASHGs, a superclass of FA and EA preferences, it has been shown that no SP mechanism can guarantee a bounded approximation ratio [15].

In contrast to strategyproofness, recently, non-obvious manipulability has been introduced [17]. This notion turned out to be a relaxation good enough to circumvent the inherent impossibility results of strategyprofness in several game-theoretic settings [31, 32, 33]. Since in HGs strategyproof mechanisms fail to approximate the maximum social welfare within a constant ratio or are computationally inefficient, this provides us with additional motivation to study NOM mechanisms in this setting.

2. Preliminaries

In the classical framework of HGs, we are given a set of n agents, denoted by $\mathcal{N} = \{1, \ldots, n\}$, and aim to create a disjoint partition $\pi = \{C_1, \ldots, C_m\}$ such that $\cup_{h=1}^{m} C_h = \mathcal{N}$ and $C_h \cap C_k = \emptyset$ for $h \neq k$. Such a partition is also called an *outcome* or a *coalition structure*. We denote by Π the set of all possible outcomes of the game, i.e., all possible partitions, and by $\pi(i)$ the coalition that agent i belongs to in a given outcome $\pi \in \Pi$. The main characteristic of HGs is that agents, when evaluating an outcome, consider only the coalition they belong to, and not how the other agents aggregate. A simple yet interesting scenario for HGs is when each agent i partitions the others into a set of friends F_i and a set of enemies E_i, with $F_i \cup E_i = \mathcal{N} \setminus \{i\}$ and $F_i \cap E_i = \emptyset$. Here, the agents' preferences depend solely on how many friends and enemies are in their own coalition. Specifically, in this work we focus on *friends appreciation* (FA) preferences, where agents give priority to the number of friends in their coalition (the higher the better) and in case of ties prefer coalitions with fewer enemies. Games with FA preferences are a proper subclass of ASHGs, where each agent i has a value $v_i(j)$ for every other agent j and her utility for being in a given coalition $C \in \mathcal{N}_i$ is $u_i(C) = \sum_{j \in C \setminus \{i\}} v_i(j)$. To comply with the FA preferences, v_i can be defined as $v_i(j) = 1,$ if $j \in F_i$, and $v_i(j) = -\frac{1}{n},$ if $j \in E_i$.Since the utility of an agent depends only on the coalition she belongs to, we might write $u_i(\pi)$ to denote $u_i(\pi(i))$. An FA instance is given by $\mathcal{I} = (\mathcal{N}, \{v_i\}_{i \in \mathcal{N}})$.

One of the main challenges in HGs is to find a partition that maximizes the overall happiness of the agents measured by the *social welfare* (SW). Specifically, in an HG instance $\mathcal{I}$ the *utilitarian* social welfare of a partition π is given by $\mathsf{SW}(\pi) = \sum_{i \in \mathcal{N}} u_i(\pi)$. We call *social optimum*, or simply *optimum*, any outcome OPT in $\arg\max_{\pi \in \Pi} \mathsf{SW}(\pi)$ and denote by opt the value $\mathsf{SW}(\mathrm{OPT})$.

A very convenient representation of an FA instance $\mathcal{I}$ is by a directed unweighted graph where the agents are the vertices. With E_i being $\mathcal{N} \setminus \{F_i \cup \{i\}\}$, it is sufficient to represent only friendship relationships through edges: if for $i \neq j$ $\{i, j\}$ is an edge of this graph, it means $j \in F_i$; otherwise, $j \in E_i$. We call this graph the *friendship graph* of $\mathcal{I}$ and denote it by $G^f = (\mathcal{N}, F)$, where $F = \{\{i, j\} \mid j \in F_i\}$.

Strategyproofness and Non-obvious Manipulability. The sets F_i and E_i might be private information of the agent i; therefore, to compute the outcome, we need to receive this information from the agents. Let us denote by $\mathbf{d} = (d_1, \ldots, d_n)$ the agents' *declarations* vector, where d_i contains the information related to agent i. We assume direct revelation, and hence $d_i(j) \in \{1, -\frac{1}{n}\}$ represents the value i declared for an agent j. We denote by $\mathcal{D}$ the space of feasible declarations $\mathbf{d}$. For our convenience, we denote by $\mathbf{d}_{-i}$ the agents' declarations except the one of i, by $\mathcal{D}_{-i}$ the set of all feasible $\mathbf{d}_{-i}$, and by $\mathcal{D}_i$ the feasible declarations for i.

In this setting, the natural challenge is to design algorithms, a.k.a. *mechanisms*, inducing truthful behavior of the agents. We shall denote by $\mathcal{M}$ a mechanism and by $\mathcal{M}(\mathbf{d})$ the output of the mechanism – a partition upon the declaration $\mathbf{d}$ of the agents. The agents might be strategic, which means that an agent i could declare $d_i \neq t_i$, where $t_i \in \mathcal{D}_i$ is the real information of agent i, also called her *real type*. For this reason, the design of mechanisms preventing manipulations is fundamental. The most desirable characteristic for such kind of mechanisms is *strategyproofness*.

Definition 1 (Strategyproofness and Manipulability)**.** A mechanism $\mathcal{M}$ is said to be *strategyproof* (SP) if, for each $i \in \mathcal{N}$ having real type t_i, and any declaration of the other agents $\mathbf{d}_{-i}$, $u_i(\mathcal{M}(t_i, \mathbf{d}_{-i})) \geq u_i(\mathcal{M}(d_i, \mathbf{d}_{-i}))$ holds true for any possible false declaration $d_i \neq t_i$ of agent i.

In turn, a mechanism is said to be *manipulable* if there exists an agent i, a real type t_i and declarations $\mathbf{d}_{-i}$ and $d_i \neq t_i$ such that this condition does not hold. Then, such d_i is called a *manipulation*.

Since SP mechanisms may be quite inefficient w.r.t. the truthful opt, we aim to understand if mechanisms satisfying milder conditions lead to more efficient outcomes. Considering that i might be unaware of which are the declarations $\mathbf{d}_{-i}$ of the other agents, she could not be able to determine a manipulation without knowing $\mathbf{d}_{-i}$. Thus, we next consider a relaxation of SP where an agent i decides to misreport her true values only if it is clearly profitable for her. Given a mechanism $\mathcal{M}$, let us denote by $\Pi_i(d_i, \mathcal{M}) = \{\mathcal{M}(d_i, \mathbf{d}_{-i}) \mid \mathbf{d}_{-i} \in \mathcal{D}_{-i}\}$, the space of possible outcomes of $\mathcal{M}$ given the declaration d_i of i. Notice the space $\Pi_i(d_i, \mathcal{M})$ is finite.

Definition 2 (Non-obvious Manipulability)**.** A mechanism $\mathcal{M}$ is said to be *non-obviously manipulable* (NOM) if for every $i \in \mathcal{N}$, real type t_i, and any other declaration d_i the following hold true:

Condition 1 (best case): $\max_{\pi \in \Pi_i(t_i, \mathcal{M})} u_i(\pi) \geq \max_{\pi \in \Pi_i(d_i, \mathcal{M})} u_i(\pi)$

Condition 2 (worst case): $\min_{\pi \in \Pi_i(t_i, \mathcal{M})} u_i(\pi) \geq \min_{\pi \in \Pi_i(d_i, \mathcal{M})} u_i(\pi)$

If there exist i, t_i, and d_i such that Condition 1 or 2 is violated, then, $\mathcal{M}$ is *obviously manipulable* and d_i is an *obvious manipulation*.

Observation 1. *Strategyproofness implies non-obvious manipulability.*

In what follows, we always denote by t_i the real type of i and by $e_i = |E_i|$ and $f_i = |F_i|$, where E_i and F_i are the truthful set of friends and enemies of i, respectively.

3. An optimal and NOM mechanism

In [12], it has been shown that no strategyproof mechanism can have an approximation better than 2. In contrast, we next show there is a way to simultaneously guarantee optimality and NOM.

Theorem 1. *There exists a mechanism $\mathcal{M}_{OPT}$ that is optimal and NOM.*

To show the theorem, we at first need to understand which are the worst/best outcomes for i in the space of possible outcomes of the optimal mechanism when i reports t_i. We will then compare their utility for i with the worst/best outcomes for any other feasible d_i. Also, since in the worst/best case instances there might be more than one optimum, we have to define a tie-breaking rule. One of the possible ways to do it while maintaining non-obvious manipulability is to choose the optimal partition minimizing the number of coalitions.

When i truthfully reports t_i, in the best case, i ends up in the coalition $C = \{i\} \cup F_i$, which happens when $\mathbf{d}_{-i}$ is so that C is a bidirectional clique in the friendship graph G^f and all other nodes are

(a) Octopus graph.

(b) Generalized octopus graph.

Figure 1: Octopus graph structures having center i. Undirected edges represent bidirectional edges in G^f.

isolated. This coalition is of maximum utility for i, and therefore the best case cannot be improved by any misreport $d_i \neq t_i$. Understanding the worst case upon any possible declaration, instead, is less trivial. When truthfully reporting, it is achieved when $\mathbf{d}_{-i}$ induces in G^f a specific graph structure:

Definition 3 (Octopus Graph)**.** Given an agent i and $H \subseteq \mathcal{N} \setminus \{i\}$, $G^f = (\mathcal{N}, F)$ is an i-centered *octopus graph* with the head H if i) H is a bidirectional clique in G^f; ii) for each $j \in H$, $\{j, i\} \in F$; iii) for each $j \in \mathcal{N} \setminus i$ and $k \in \mathcal{N} \setminus (\{i\} \cup H)$, none of $\{j, k\}$, $\{k, j\}$, $\{k, i\}$ belongs to F while $\{i, j\}$ may belong to F. See Figure 1a for an example.

Then, if H consists of E_i and $\max\{\lceil \frac{N}{2} \rceil - |E_i|, 0\}$ i's friends, then, regardless of i's preferences, in the optimum i ends up in the coalition $H \cup \{i\}$, which is in fact the worst case by truthfully reporting. To prove it is not possible to improve the worst case reporting $d_i \neq t_i$, we use another graph structure, the *generalized octopus graph*, where agents from $\mathcal{N} \setminus (\{i\} \cup H)$ may form arbitrary cliques (see Figure 1b). Then, Condition 2 is not violated as we show that a) for generalized octopus graphs, the space of possible optimal outcomes coincides with the corresponding space for any friendship graph and b) there is no generalized octopus graph where in the optimal outcome i ends up in the coalition worse than $H \cup \{i\}$.

4. Computing the optimum is NP-hard

In this section, we show that, unfortunately, computing an optimum partition is NP-hard.

Theorem 2. *For* FA *preferences, computing the optimum is NP-hard.*

Theorem 2 is proven with a reduction from the following 3-Partition problem:
Input: A ground set $\{x_1, x_2, \ldots, x_{3m}\}$ of $3m$ elements such that for some $T > 0$: (i) $\sum_{h=1}^{3m} x_h = mT$; (ii) for each $h \in [3m]$, $x_h \in \mathbb{N}$; (iii) for each $h \in [3m]$, $\frac{T}{4} < x_h < \frac{T}{2}$.
Question: Does there exist a partition of the ground set into m disjoint subsets $S_1, \ldots, S_m$ such that, for every $k \in [m]$, $S_k = \{s_k^1, s_k^2, s_k^3\}$ and $s_k^1 + s_k^2 + s_k^3 = T$?

Let us note that in the standard formulation of 3-Partition, condition (iii) is usually not required, however, the problem remains strongly NP-hard even under such a condition [34]. Moreover, condition (iii) also implies that for any $S \subseteq \{x_1, x_2, \ldots, x_{3m}\}$ if $\sum_{x \in S} x = T$, then $|S| = 3$. Consequently, any partition into subsets, each having sum T, is a partition into triples.

Given a 3-Partition instance, we construct the graph G^f representing an FA instance as follows:
Element-cliques: Each of these cliques represents a specific element in the ground set of the 3-Partition instance: For every $h \in [3m]$, we create a bidirectional clique K^h of size x_h.
Set-cliques: We create m bidirectional cliques $K_X^1, \ldots, K_X^m$ each one of size $X = 4m^2T$. The choice of X is made in such a way that we can use the cliques $K_X^1, \ldots, K_X^m$ to interpret a coalition in an optimum partition, for the FA instance, as a triple in a partition for the 3-Partition instance.
Connections between cliques: We add x_h bidirectional edges between K^h and each K_X^k in such a way that there is exactly one bidirectional edge between each vertex of K^h and some node in K_X^k. Since $|K^h| = x_h < X$, this is always possible.

Notice that the number of agents is $n = \sum_{h=1}^{3m} x_h + mX = mT + 4m^3T$; thus, with 3-Partition being strongly NP-hard the correctness of the reduction proves the NP-hardness of our problem.

5. An approximation mechanism

For the sake of achieving NOM in polynomial time, in this section, we present a $(4+o(1))$-approximation mechanism. We recall that in [12] it was shown that creating a coalition for each weakly connected component of G^f is SP and guarantees an n-approximation to the optimum. This is so far the best approximation achieved by an SP mechanism. The bad performances of this mechanism can be attributed to the fact that when G^f is weakly connected but really sparse (e.g., a directed path), it would be convenient to split the unique weakly connected component of G^f into smaller coalitions.

To circumvent this, our mechanism partitions the agents into two sets, P_1 and P_2, of size $\lceil \frac{n}{2} \rceil$ and $\lfloor \frac{n}{2} \rfloor$, respectively. It then updates P_1 and P_2, through the subroutine ImproveSW more formally described in the full paper. ImproveSW repeatedly tries to improve $\mathsf{SW}(\{P_1, P_2\})$ by swapping two agents, that is, simultaneously moving $i \in P_1$ to P_2 and $j \in P_2$ to P_1, or moving an agent from the largest to the smallest coalition (in case the two sets have the same size the algorithm will never perform a move). ImproveSW terminates when no swap or move can increase the SW. The mechanism then computes the weakly connected components in P_1 and P_2 which will be the coalitions of the returned partition.

To show the mechanism is NOM, the initialization of $\{P_1, P_2\}$ will be crucial. The mechanism will create the initial $\{P_1, P_2\}$ by greedily adding agents to the set P_1 in the following way: At first, it inserts an agent $i \in \mathcal{N}$ with highest $\delta(i)$, then, iteratively proceeds by including an agent $j \in N(P_1) \setminus P_1$ with highest $\delta(j)$ – ties broken arbitrarily. This process continues until P_1 contains exactly $\lceil \frac{n}{2} \rceil$ agents. If at some point $N(P_1) \setminus P_1 = \emptyset$, the mechanism selects a new agent $i \in \mathcal{N} \setminus P_1$ with highest $\delta(i)$, and proceeds as before. We call this partition a *greedy 2-partition* of $\mathcal{N}$. In summary:

Mechanism $\mathcal{M}_1$. Given $\mathcal{N}$ and the declarations $\mathbf{d}$, it creates a greedy 2-partition $\{P_1, P_2\}$. Then, while possible, it updates the partition using ImproveSW: $\{P_1, P_2\} \leftarrow \text{ImproveSW}(P_1, P_2)$. Finally, it computes $C_1, \ldots, C_m$, the weakly connected components in P_1 and P_2, and returns $\pi = \{C_1, \ldots, C_m\}$.

Theorem 3. *For* FA *instances, Mechanism $\mathcal{M}_2$ is NOM and guarantees a $(4 + o(1))$-approximation of the optimum in polynomial time.*

We note that our approach is similar to the 4-approximating local search algorithm for the Max-Cut problem in directed and unweighted graphs. However, due to the presence of weights $\{-\frac{1}{n}, 1\}$, our approximation factor slightly deteriorates.

6. Discussion

In this paper, we investigated NOM in HGs with FA preferences, aiming at designing mechanisms optimizing the social welfare while preventing manipulation. Despite proving that computing a welfare-maximizing partition is NP-hard, we showed that a NOM mechanism having a constant approximation always exists. Moreover, if time complexity is not a concern, there exists a NOM and optimal mechanism as well. Interestingly enough, we were also able to show that it is not always the case that optimality is compatible with NOM. In particular, an optimal outcome cannot be NOM when agents have Enemies Aversion (EA) preferences, the natural counterpart of FA preferences, where agents give priority to coalitions with fewer enemies, and when the number of enemies is the same, they prefer coalitions with a higher number of friends, i.e., $v_i(j) \in \{1, -n\}$, for $i \neq j$. Independent of interest, our approximation algorithm represents the first deterministic constant-factor approximation for FA preferences; this is an interesting contrast to EA preferences for which it is known to be hard to approximate the optimum within a factor of $O(n^{1-\epsilon})$ [12]. We refer the interested reader to the full paper for further details [18].

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

[1] J. H. Dreze, J. Greenberg, Hedonic coalitions: Optimality and stability, Econometrica: Journal of the Econometric Society (1980) 987–1003.

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

[3] H. Aziz, F. Brandt, H. G. Seedig, Computing desirable partitions in additively separable hedonic games, Artificial Intelligence 195 (2013) 316–334.

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

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

[6] E. Elkind, M. J. Wooldridge, Hedonic coalition nets., in: AAMAS (1), Citeseer, 2009, pp. 417–424.

[7] E. Elkind, A. Fanelli, M. Flammini, Price of pareto optimality in hedonic games, Artificial Intelligence 288 (2020) 103357.

[8] M. Gairing, R. Savani, Computing stable outcomes in hedonic games, in: International Symposium on Algorithmic Game Theory, Springer, 2010, pp. 174–185.

[9] J. Hajduková, On coalition formation games, IM Preprints series A (2004).

[10] D. Dimitrov, S. C. Sung, Enemies and friends in hedonic games: Individual deviations, stability and manipulation, SSRN Electronic Journal (2004). doi:10.2139/ssrn.639483.

[11] G. Varricchio, On approximate strategyproof mechanisms for hedonic games and the group activity selection problem, in: Proceedings of IPS, volume 3585 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2023.

[12] M. Flammini, B. Kodric, G. Varricchio, Strategyproof mechanisms for friends and enemies games, Artif. Intell. 302 (2022) 103610.

[13] S. Özyurt, M. R. Sanver, A general impossibility result on strategy-proof social choice hyperfunctions, Games and economic behavior 66 (2009) 880–892.

[14] G. Amanatidis, G. Birmpas, G. Christodoulou, E. Markakis, Truthful allocation mechanisms without payments: Characterization and implications on fairness, in: Proceedings of the 2017 ACM Conference on Economics and Computation, 2017, pp. 545–562.

[15] M. Flammini, B. Kodric, G. Monaco, Q. Zhang, Strategyproof mechanisms for additively separable and fractional hedonic games, Journal of Artificial Intelligence Research 70 (2021) 1253–1279.

[16] F. Brandl, F. Brandt, M. Eberl, C. Geist, Proving the incompatibility of efficiency and strategyproofness via smt solving, Journal of the ACM (JACM) 65 (2018) 1–28.

[17] P. Troyan, T. Morrill, Obvious manipulations, Journal of Economic Theory 185 (2020) 104970.

[18] M. Flammini, M. Fomenko, G. Varricchio, Non-obvious manipulability in hedonic games with friends appreciation preferences, in: Proceedings of the 24th International Conference on Autonomous Agents and Multiagent Systems, 2025.

[19] F. Bloch, E. Diamantoudi, Noncooperative formation of coalitions in hedonic games, International Journal of Game Theory 40 (2011) 263–280.

[20] M. Feldman, L. Lewin-Eytan, J. S. Naor, Hedonic clustering games, ACM Trans. Parallel Comput. 2 (2015) 4:1–4:48.

[21] M. Flammini, B. Kodric, M. Olsen, G. Varricchio, Distance hedonic games, in: SOFSEM 2021: Theory and Practice of Computer Science - 47th International Conference on Current Trends in Theory and Practice of Computer Science, SOFSEM 2021, Bolzano-Bozen, Italy, January 25-29, 2021, Proceedings, volume 12607, Springer, 2021, pp. 159–174.

[22] A. Igarashi, E. Elkind, Hedonic games with graph-restricted communication, in: Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems, International Foundation for Autonomous Agents and Multiagent Systems, 2016, pp. 242–250.

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

[24] J. Rothe, H. Schadrack, L. Schend, Borda-induced hedonic games with friends, enemies, and neutral

players, Mathematical Social Sciences 96 (2018) 21–36.

[25] A. M. Kerkmann, N.-T. Nguyen, A. Rey, L. Rey, J. Rothe, L. Schend, A. Wiechers, Altruistic hedonic games, Journal of Artificial Intelligence Research 75 (2022) 129–169.

[26] N. Barrot, K. Ota, Y. Sakurai, M. Yokoo, Unknown agents in friends oriented hedonic games: Stability and complexity, in: Proceedings of the AAAI Conference on Artificial Intelligence, volume 33, 2019, pp. 1756–1763.

[27] J. Chen, G. Csáji, S. Roy, S. Simola, Hedonic games with friends, enemies, and neutrals: Resolving open questions and fine-grained complexity, in: Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2023, London, United Kingdom, 29 May 2023 - 2 June 2023, ACM, 2023, pp. 251–259.

[28] B.-E. Klaus, F. Klijn, S. Özbilen, Core stability and strategy-proofness in hedonic coalition formation problems with friend-oriented preferences, Available at SSRN 4677609 (2023).

[29] C. Rodríguez-Álvarez, Strategy-proof coalition formation, International Journal of Game Theory 38 (2009) 431–452.

[30] M. Flammini, G. Varricchio, Approximate strategyproof mechanisms for the additively separable group activity selection problem, in: Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI 2022, Vienna, Austria, 23-29 July 2022, 2022, pp. 300–306.

[31] H. Aziz, A. Lam, Obvious manipulability of voting rules, in: Algorithmic Decision Theory - 7th International Conference, ADT 2021, Toulouse, France, November 3-5, 2021, Proceedings, volume 13023 of *Lecture Notes in Computer Science*, Springer, 2021, pp. 179–193.

[32] P. Troyan, (non-)obvious manipulability of rank-minimizing mechanisms, Journal of Mathematical Economics (2024) 103015.

[33] A. Psomas, P. Verma, Fair and efficient allocations without obvious manipulations, in: Advances in Neural Information Processing Systems 35: Annual Conference on Neural Information Processing Systems 2022, NeurIPS 2022, New Orleans, LA, USA, November 28 - December 9, 2022, 2022.

[34] M. R. Garey, D. S. Johnson, Computers and intractability, volume 174, freeman San Francisco, 1979.