Pareto Optimality in Coalition Formation

Haris Aziz Felix Brandt Paul Harrenstein
Department of Informatics
Technische Universität München
85748 Garching bei München, Germany
{aziz,brandtf,harrenst}@in.tum.de

Abstract

A minimal requirement on allocative efficiency in the social sciences is Pareto optimality. In this paper, we exploit a strong structural connection between Pareto optimal and perfect partitions that has various algorithmic consequences for coalition formation. In particular, we show that computing and verifying Pareto optimal partitions in general hedonic games and B-hedonic games is intractable while both problems are tractable for roommate games and W-hedonic games. The latter two positive results are obtained by reductions to maximum weight matching and clique packing, respectively.

1 Introduction

Topics concerning coalitions and coalition formation have come under increasing scrutiny of computer scientists. The reason for this may be obvious. For the proper operation of distributed and multiagent systems, cooperation may be required. At the same time, collaboration in very large groups may also lead to unnecessary overhead, which may even exceed the positive effects of cooperation. To model such situations formally, concepts from the social and economic sciences have proved to be very helpful and thus provide the mathematical basis for a better understanding of the issues involved.

Coalition formation games, which were first formalized by Drèze and Greenberg [1980], model coalition formation in settings in which utility is *non-transferable*. In many such situations it is natural to assume that a player's appreciation of a coalition structure only depends on the coalition he is a member of and not on how the remaining players are grouped. Initiated by Banerjee *et al.* [2001] and Bogomolnaia and Jackson [2002], much of the work on coalition formation now concentrates on these so-called *hedonic games*. In this paper, we focus on Pareto optimality and individual rationality in this rich class of coalition formation games.

The main question in coalition formation games is which coalitions one may reasonably expect to form. To get a proper formal grasp of this issue, a number of stability concepts have been proposed for hedonic games—such as the core or Nash stability—and much research concentrates on conditions for existence, the structure, and computation of stable and efficient partitions. *Pareto optimality*—which holds if no coalition structure is strictly better for some player without being strictly worse for another—and *individual rationality*—which holds if every player is satisfied in the sense that no player would rather be on his own—are commonly considered minimal requirements for any reasonable partition.

Another reason to investigate Pareto optimal partitions algorithmically is that, in contrast to other stability concepts like the core, they are guaranteed to exist. This even holds if we additionally require individual rationality. Moreover, even though the *Gale-Shapley algorithm* returns a core stable matching for marriage games, it is already NP-hard to check whether the core is empty in various classes and representations of hedonic games, such as roommate games [Ronn, 1990], general hedonic games [Ballester, 2004], and games with $\mathscr{B}$- and $\mathscr{W}$-preferences [Cechlárová and Hajduková, 2004a,b]. Interestingly, when the status-quo partition cannot be changed without the mutual consent of all players, Pareto optimality defines stability [Morrill, 2010].

In this paper, we investigate both the problem of finding a Pareto optimal and individually rational partition and the problem of deciding whether a partition is Pareto optimal. In particular, our results concern *general hedonic games*, *B-hedonic* and *W-hedonic games* (two classes of games in which each player's preferences over coalitions are based on his most preferred and least preferred player in his coalition, respectively), and *roommate* games.

Many of our results, both positive and negative, rely on the concept of *perfection* and how it relates to Pareto optimality. A *perfect* partition is one that is most desirable for every player. We find (*a*) that under extremely mild conditions, NP-hardness of finding a perfect partition implies NP-hardness of finding a Pareto optimal partition (Lemma 1), and (*b*) that under stronger but equally well-specified circumstances, feasibility of finding a perfect partition implies feasibility of finding a Pareto optimal partition (Lemma 2). The latter we show via a Turing reduction to the problem of computing a perfect partition. At the heart of this algorithm, which we refer to as the *Preference Refinement Algorithm (PRA)*, lies a fundamental insight of how perfection and Pareto optimality are related. It turns out that a partition is Pareto optimal for a particular preference profile if and only if the partition is perfect for another but related one (Theorem 1). In this way PRA is also applicable to any other discrete allocation setting.

For general allocation problems, *serial dictatorship*—which chooses subsequently the most preferred allocation for a player given a fixed ranking of all players—is well-established as a procedure for finding Pareto optimal solutions [see, e.g., Abdulkadiroğlu and Sönmez, 1998]. However, it is only guaranteed to do so, if the players' preferences over outcomes are strict, which is not feasible in many compact representations. Moreover, when applied to coalition formation games, there may be Pareto optimal partitions that serial dictatorship is unable to find, which may have serious repercussions if also other considerations, like fairness, are taken into account. By contrast, PRA handles weak preferences well and and is complete in the sense that it may return any Pareto optimal partition, provided that the subroutine that calculates perfect partitions can compute any perfect partition (Theorem 2).

2 Preliminaries

In this section, we review the terminology and notation used in this paper.

Hedonic games Let N be a set of n players. A *coalition* is any non-empty subset of N. By $\mathscr{N}_i$ we denote the set of coalitions player i belongs to, i.e., $\mathscr{N}_i = \{S \subseteq N : i \in S\}$. A *coalition structure*, or simply a *partition*, is a partition π of the players N into coalitions, where $\pi(i)$ is the coalition player i belongs to.

A *hedonic game* is a pair (N, R), where $R = (R_1, \ldots, R_n)$ is a *preference profile* specifying the preferences of each player i as a binary, complete, reflexive, and transitive *preference relation* R_i over $\mathscr{N}_i$. If R_i is also anti-symmetric we say that i's preferences are *strict*. We adopt the conventions of social choice theory by writing $S\ P_i\ T$ if $S\ R_i\ T$ but not $T\ R_i\ S$—i.e., if i *strictly prefers* S to T—and $S\ I_i\ T$ if both $S\ R_i\ T$ and $T\ R_i\ S$—i.e., if i is *indifferent* between S and T.

For a player i, a coalition S in $\mathscr{N}_i$ is *acceptable* if for i being in S is at least preferable as being alone—i.e., if $S\ R_i\ \{i\}$—and *unacceptable* otherwise.

In a similar fashion, for X a subset of $\mathscr{N}_i$, a coalition S in X is said to be *most preferred in X by i* if $S\ R_i\ T$ for all T in X and *least preferred in X by i* if $T\ R_i\ S$ for all $T \in X$. In case $X = \mathscr{N}_i$ we generally omit the reference to X. The sets of most and least preferred coalitions in X by i, we denote by $\max_{R_i}(X)$ and $\min_{R_i}(X)$, respectively.

In hedonic games players are only interested in the coalition they are in. Accordingly, preferences over coalitions naturally extend to preferences over partitions and we write $\pi\ R_i\ \pi'$ if $\pi(i)\ R_i\ \pi'(i)$. We also say that partition π is *acceptable* or *unacceptable* to a player i according to whether $\pi(i)$ is acceptable or unacceptable to i, respectively. Moreover, π is *individually rational* if π is acceptable to all players. A partition π is *Pareto optimal in R* if there is no partition π' with $\pi'\ R_j\ \pi$ for all players j and $\pi'\ P_i\ \pi$ for at least one player i. Partition π is, moreover, said to be *weakly Pareto optimal in R_i* if there is no π' with $\pi'\ P_i\ \pi$ for all players i.

Classes of hedonic games The number of potential coalitions grows exponentially in the number of players. In this sense, hedonic games are relatively large objects and for algorithmic purposes it is often useful to look at classes of games that allow for concise representations.

For *general hedonic games*, we will assume that each player expresses his preferences only over his acceptable coalitions. This representation is alternatively known as *Representation by Individually Rational Lists of Coalitions* [Ballester, 2004].

We now describe classes of hedonic games in which the players' preferences over coalitions are induced by their preferences over the other players. For R_i such preferences of player i over players, we say that a player j is *acceptable* to i if $j\ R_i\ i$ and *unacceptable* otherwise. Any coalition containing an unacceptable player is unacceptable to player i.

Roommate games. The class of *roommate games*, which are well-known from the literature on matching theory, can be defined as those hedonic games in which only coalitions of size one or two are acceptable.

B-hedonic and W-hedonic games. For a subset J of players, we denote by $\max_{R_i}(J)$ and $\min_{R_i}(J)$ the sets of the most and least preferred players in J by i, respectively. We will assume that $\max_{R_i}(\emptyset) = \min_{R_i}(\emptyset) = \{i\}$. In a *B-hedonic game* the preferences R_i of a player i over players extend to preferences over coalitions in such a way that, for all coalitions S and T in $\mathscr{N}_i$, we have $S\ R_i\ T$ if and only if $\max_{R_i}(S \setminus \{i\})\ R_i \max_{R_i}(T \setminus \{i\})$ or some j in T is unacceptable to i. Analogously, in a *W-hedonic game* (N, R), we have $S\ R_i\ T$ if and only if $\min_{R_i}(S \setminus \{i\})\ R_i \min_{R_i}(T \setminus \{i\})$ or some j in T is unacceptable to i.[1]

3 Perfection and Pareto Optimality

Pareto optimality constitutes a rather minimal efficiency requirement on partitions. A much stronger property is that of *perfection*. We say that a partition π is *perfect* if $\pi(i)$ is a most preferred coalition for all players i. Thus, every perfect partition is Pareto optimal but not necessarily the other way round. Perfect partitions are obviously very desirable, but, in contrast to Pareto optimal ones, they are not guaranteed to exist. Still, a strong structural connection exists between the two concepts, which, in the next section, we exploit in our algorithm for finding Pareto optimal partitions.

The problem of finding a perfect partition (PP) we formally specify as follows: given a preference profile R, find a perfect partition for R and if no perfect partition exists in R, output "none".

We will later see that the complexity of PP depends on the specific class of hedonic games that is being considered. By contrast, the related problem of *checking* whether a partition is perfect is an almost trivial problem for virtually all reasonable classes of games. If perfect partitions exist, they clearly coincide with the Pareto optimal ones. Hence, an oracle to compute a Pareto optimal partition can be used to solve PP.

[1] W-hedonic games are equivalent to hedonic games with $\mathscr{W}$-preferences if individually rational outcomes are assumed. Unlike hedonic games with $\mathscr{B}$-preferences, B-hedonic games are defined in analogy to W-hedonic games and the preferences are not based on coalition sizes [cf. Cechlárová and Hajduková, 2004a].

If this Pareto optimal partition is perfect we are done, if it is not, no perfect partitions exist. Thus, we obtain the following lemma, which we will invoke in our hardness proofs for computing Pareto optimal partitions.

Lemma 1 *For every class of hedonic games for which checking whether a given partition is perfect can be solved in polynomial time, NP-hardness of* PP *implies NP-hardness of computing a Pareto optimal partition.*

It might be less obvious that a procedure solving PP can also be deployed as an oracle for an algorithm to compute Pareto optimal partitions. To do so, we first give a characterization of Pareto optimal partitions in terms of perfect partitions, which forms the mathematical heart of the Preference Refinement Algorithm to be presented in the next section.

This characterization depends on the concept of a coarsening of a preference profile and the lattices these coarsenings define. To make things precise, we say that a preference profile $R = (R_1, \ldots, R_n)$ is a *coarsening of* or *coarsens* another preference profile $R' = (R'_1, \ldots, R'_n)$ whenever for every player i we have $R'_i \subseteq R_i$. In that case we also say that R' *refines* R and write $R \leq R'$. Moreover, we write $R{<}R'$ if $R{\leq}R'$ but not $R'{\leq}R$. Thus, if R' refines R, i.e., if $R \leq R'$, then for each i and all coalitions S and T we have that $S\,R'_i\,T$ implies $S\,R_i\,T$, but not necessarily the other way round. Intuitively, a player i may be indifferent in R between coalitions over which i entertains strict preferences in R'. It is worth observing that, if a partition is perfect in some preference profile R, then it is also perfect in any coarsening of R. The same holds for Pareto optimal partitions.

For preference profiles R and R' with $R \leq R'$, let $[R, R']$ denote the set $\{R'' : R \leq R'' \leq R'\}$, i.e., $[R, R']$ is the set of all coarsenings of R' that are not coarser than R. Then, $([R, R'], \leq)$ is a complete lattice with R and R' as bottom and top element, respectively. We say that R *covers* R' if R is a minimal refinement of R', i.e., if $R' < R$ and there is no R'' such that $R' < R'' < R$. R *strongly covers* R' if among all preference profiles that cover R', R is one that, for all players, allows for a maximal number of most preferred alternatives, i.e., $\max_{R''_i}(\mathscr{N}_i) \subseteq \max_{R_i}(\mathscr{N}_i)$ for all players i and each R'' that covers R'. We are now in a position to prove the following theorem, which characterizes Pareto optimal partitions given a preference profile R as those that are perfect in particular coarsenings R' of R. These R' are such that no perfect partitions exist in any preference profile that strongly covers R'.

Theorem 1 *Let $(N, R^\top)$ and $(N, R^\perp)$ be hedonic games and π a partition such that $R^\perp \leq R^\top$ and π is a perfect partition in $R^\perp$. Then, π is Pareto optimal in $R^\top$ if and only if there is some $R \in [R^\perp, R^\top]$ such that (i) π is a perfect partition in R and (ii) there is no perfect partition for any $R' \in [R^\perp, R^\top]$ that strongly covers R.*

Proof: For the if-direction, assume there is some $R \in [R^\perp, R^\top]$ such that π is perfect in R and there is no perfect partition in any $R' \in [R^\perp, R^\top]$ that strongly covers R. (Observe that this implies that, for all i, R_i and $R_i^\top$ coincide on coalitions less preferred by i than $\pi(i)$.) For contradiction, also assume π is not Pareto optimal in $R^\top$. Then, there is some π' such that $\pi'\,R_j^\top\,\pi$ for all j and $\pi'\,P_i^\top\,\pi$ for some i. By $R \leq R^\top$ and transitivity of preferences, π' is a perfect partition in R as well. Let π'' be such that $\pi''(i) \in \min_{R_i^\top}(\max_{R_i}(\mathscr{N}_i))$ and define $R'_i = R_i \setminus \{(X, Y) : \pi''(i)\,R_i^\top\,X \text{ and } Y\,P_i^\top\,\pi''(i)\}$. Thus, $\pi''(i)$ is one of i's least preferred coalitions according to $R_i^\top$ among i's most preferred coalitions in R_i. Intuitively, R'_i is exactly like R_i be it that i strictly prefers Y to X in R'_i if $X \in \min_{R_i^\top}(\max_{R_i}(\mathscr{N}_i))$ and $Y\,P_i^\top\,X$. Observe that $R' = (R_1, \ldots, R_{i-1}, R'_i, R_{i+1}, \ldots, R_n)$ is in $[R^\perp, R^\top]$ and covers R. By choice of π'', R' even strongly covers R. Moreover, as $\pi'\,P_i^\top\,\pi$ and, therefore, $\pi' \notin \min_{R_i^\top}(\max_{R_i}(\mathscr{N}_i))$, π' is still a perfect partition in R', a contradiction.

For the only-if direction assume that π is Pareto optimal in $R^\top$. Let R be the finest coarsening of $R^\top$ in which π is perfect. Observe that $R = (R_1, \ldots, R_n)$ can be defined such that $R_i = R_i^\top \cup \{(X, Y) : X\,R_i^\top\,\pi \text{ and } Y\,R_i^\top\,\pi\}$ for all i. Also observe that $R^\perp \leq R$. If $R = R^\top$, we are done immediately. Otherwise, consider an arbitrary $R' \in [R^\perp, R^\top]$ that strongly covers R and assume for contradiction that there is some perfect partition π' in R'. Then, in particular, $\pi'\,R'_k\,\pi$ for all k. Since R' covers R, there is exactly one i with $R'_i \neq R_i$, whereas $R'_j = R_j$ for all $j \neq i$. As π is perfect in R, we also have $\pi\,R'_j\,\pi'$ for all $j \neq i$. With R' being a finer coarsening of $R^\top$ than R, however, π is not perfect in R'. Hence, it is not the case that $\pi\,R'_i\,\pi'$ and, therefore, $\pi'\,P'_i\,\pi$. We may now conclude that π is not Pareto optimal in R'. Since, $R' \leq R^\top$, moreover, π not Pareto optimal in $R^\top$ either, a contradiction. □

4 The Preference Refinement Algorithm

In this section, we present the *Preference Refinement Algorithm (PRA)*, a general algorithm to compute Pareto optimal and individually rational partitions. The algorithm invokes an oracle solving PP and is based on the formal connection between Pareto optimality and perfection made explicit in Theorem 1.

The idea underlying the algorithm is as follows. To calculate a Pareto optimal and individually rational partition for a hedonic game (N, R), first find that coarsening R' of R in which each player is indifferent among all his acceptable coalitions and his preferences among unacceptable coalitions are as in R. In this coarsening, a perfect and individually rational partition—which we also refer to as the *coarsest acceptable coarsening*—is guaranteed to exist. From there on, start moving up in the lattice $([R', R], \leq)$ to strongly covering preference profiles for which a perfect partition exists, until you reach a preference profile for which this is no longer possible. By calculating a perfect partition for this last preference profile, in virtue of Theorem 1, you find a Pareto optimal partition for R. A formal specification of PRA is given in Algorithm 1. It is worth mentioning that Algorithm 1 is an anytime algorithm that can return an intermediate result when stopped prematurely.

Theorem 2 *For any hedonic game (N, R),*

(i) PRA returns an individually rational and Pareto optimal partition.

Algorithm 1 Preference Refinement Algorithm (PRA)

Input: Hedonic game (N, R)
Output: Pareto optimal and individually rational partition

1 $Q_i \leftarrow R_i \cup \{(X, Y) : X\, R_i\, \{i\}$ and $Y\, R_i\, \{i\}\}$, for each $i \in N$
2 $Q \leftarrow (Q_1, \ldots, Q_n)$
3 $J \leftarrow N$
4 **while** $J \neq \emptyset$ **do**
5 $\quad i \in J$
6 $\quad S \in \min_{R_i}(\max_{Q_i}(\mathcal{N}_i))$
7 $\quad Q'_i \leftarrow Q_i \setminus \{(X, Y) : S\, R_i\, X$ and $Y\, P_i\, S\}$
8 $\quad Q' \leftarrow (Q_1, \ldots, Q_{i-1}, Q'_i, Q_{i+1}, \ldots, Q_n)$
9 $\quad$ **if** PP$(N, Q') \neq$ none **then**
10 $\qquad Q \leftarrow Q'$
11 $\quad$ **else**
12 $\qquad J \leftarrow J \setminus \{i\}$
13 $\quad$ **end if**
14 **end while**
15 **return** PP(N, Q)

(*ii*) *For every individually rational and Pareto optimal partition π', there is an execution of PRA that returns a partition π such that $\pi\, I_i\, \pi'$ for all i in N.*

Proof: For (*i*), we prove that during the running of PRA, for each assignment of Q, there exists a perfect partition π for that assignment. This claim certainly holds for the first assignment of Q which is the coarsest acceptable coarsening of R. Furthermore, Q is only refined via the strong covering relation (Steps 6 through 7), if there exists a perfect partition for a strong covering of Q. Let Q^* be the final assignment of Q. Then, we argue that the partition π returned by PRA is Pareto optimal and individually rational. By Theorem 1, if π were not Pareto optimal, there would exist a strong covering of Q^* for which a perfect partition still exists and Q^* would not be the final assignment of Q. Since, each player at least gets one of his acceptable coalitions, π is also individually rational.

For (*ii*), first observe that, by Theorem 1, for each Pareto optimal and individually rational partition π for a preference profile R there is some coarsening R^* of R where π is perfect and no perfect partitions exist for any strong covering of R^*. By individual rationality of π, it follows that R^* is a refinement of the initial assignment of Q. An appropriate number of strong coverings of the initial assignment of Q with respect to each player results in a final assignment Q^* of Q to R^*. The perfect partition for Q^* that is returned by PRA is then such that $\pi\, I_i\, \pi'$ for all i in N. □

Note that for each player's preferences over coalitions induces equivalence classes in which a player is indifferent between coalitions in the same equivalence class. We specify the conditions under which PRA runs in polynomial time.

Lemma 2 *Let (N, R) be a hedonic game such that for each player the number of equivalence classes of acceptable outcomes is polynomial in the input, the coarsest acceptable coarsening of R as well as the strong coverings of each of its refinements can be computed in polynomial time, and* PP *can be solved in polynomial time for all coarsenings of R. Then, PRA runs in polynomial time.*

Proof: Under the given conditions, we prove that PRA runs in polynomial time. In each iteration of the while-loop, either the preference profile Q is strongly covered (Step 10) or a player i which cannot be further improved is removed from J (Step 12). Both of these steps take polynomial time due to the conditions specified. Since each player has a polynomial number of acceptable equivalence classes in R_i, there can only be a polynomial number of reassignments of Q and therefore the while-loop iterates a polynomial number of times. As the crucial subroutine PP (Step 9) takes polynomial time, PRA runs in polynomial time. □

PRA applies not only to general hedonic games but to many natural classes of hedonic games in which equivalence classes (of possibly exponentially many coalitions) for each player are implicitly defined.[2]

Note that PRA as it is presented does not leverage the potential benefit of preferences being strict because when preferences are coarsened, the strictness of the preferences is lost and PP becomes NP-hard (see Theorem 3). *Serial dictatorship* is a well-studied mechanism in resource allocation, in which an arbitrary player is chosen as the 'dictator' who is then given his most favored allocation and the process is repeated until all players or resources have been dealt with. In the context of coalition formation, *serial dictatorship* is well-defined only if in every iteration, the dictator has a unique most preferred coalition.

Proposition 1 *For general hedonic games, W-hedonic games, and roommate games, a Pareto optimal partition can be computed in polynomial time when preferences are strict.*

Proposition 1 follows from the application of serial dictatorship to hedonic games with strict preferences over the coalitions. If the preferences over coalitions are not strict, then the decision to assign one of the favorite coalitions to the dictator may be sub-optimal. Serial dictatorship does not work for hedonic games in which preferences over coalitions are not strict, not even for B-hedonic games with strict preferences over players. Observe that PRA can be tweaked so as to obtain an individually rational version of the serial dictatorship algorithm, which also achieves the positive results of Proposition 1. Abdulkadiroğlu and Sönmez [1998] showed that in the case of strict preferences and house allocation settings, every Pareto optimal allocation can be achieved by serial dictatorship. In the case of coalition formation, however, it is easy to construct a four-player hedonic game with strict preferences for which there is a Pareto optimal partition that serial dictatorship cannot return.

5 Computational results

In this section, we consider the problem of Verification (verifying whether a given partition is Pareto optimal) and Computation (computing a Pareto optimal partition).

[2] For example, in W-hedonic games, $\max_{R_i}(N)$ specifies the set of favorite players of player i but can also implicitly represent all those coalitions S such that the least preferred player in S is also a favorite player for i.

5.1 General hedonic games

As shown in Proposition 1, Pareto optimal partitions can be found efficiently for general hedonic games with strict preferences. If preferences are not strict, the problem becomes NP-hard. We can prove the following statement by utilizing Lemma 1 and showing that PP is NP-hard by a reduction from ExactCoverBy3Sets (X3C).

Theorem 3 *For a general hedonic game, computing a Pareto optimal partition is NP-hard even when each player has a maximum of four acceptable coalitions and the maximum size of each coalition is three.*

Interestingly, *verifying* Pareto optimality is coNP-complete even for strict preferences.

Theorem 4 *For any general hedonic game, verifying whether a partition π is Pareto optimal and whether π is weakly Pareto optimal is coNP-complete even when preferences are strict and π consists of the grand coalition of all players.*

5.2 Roommate games

For roommate games, we observe that PP is equivalent to solving a perfect matching of the graph in which two vertices (players) are connected if and only if they consider each other as a favorite player. Therefore, we obtain the following as a corollary of Lemma 2.

Theorem 5 *For roommate games, an individually rational and Pareto optimal coalition can be computed in polynomial time.*

We found that in the case of general hedonic games, verifying Pareto optimality can be significantly harder than computing a Pareto optimal partition when preferences are strict. Abraham and Manlove [2004] and Morrill [2010] showed that there are efficient algorithms to verify whether a partition is Pareto optimal for roommate games with strict preferences. The more general case of non-strict preferences is left open.[3] We answer this problem in the next theorem.

Theorem 6 *For roommate games, it can be checked in polynomial time whether a partition is Pareto optimal.*

Proof sketch: We reduce the problem to computing a maximum weight matching of a graph.

For roommate game (N, R), let π be the partition which we want to check for Pareto optimality. Since π contains coalitions of size one or two, we can construct an undirected graph $G = (V, E)$ where $V = N \cup (N \times \{0\})$, $E = V \times V \setminus (\{\{i, j\} : \pi(i)\, P_i\, \{i\}\} \cup \{\{i, (i, 0)\} : \pi(i)\, P_i\, \{i\}\}$. For graph (V, E), consider the matching $M = \{S \in \pi : |S| = 2\} \cup \{\{i, (i, 0)\} : \{i\} \in \pi\}$.

We now define a weight function such that for all $i \in V$, $w_i : E \to \mathbb{R}^+$ where w_i is defined inductively in the following way: $w_{(i,0)}(e) = 0$ for all e such that $(i, 0) \in e \in E$ and $i \in N$; $w_i(\pi(i)) = n$ if $\pi(i) \neq \{i\}$ and $\pi(i) = \{i, j\}$; $w_i(\{i, (i, 0)\}) = n$ if $\pi(i) = \{i\}$; $w_i(S) = -n$ if $i \notin S$; $w_i(T) = w_i(S) + 1/n$ if there is a coalition T such that $i \in T$, $T\, P_i\, S$, and there exists no coalition T' such that $T\, P_i\, T'\, P_i\, S$; and $w_i(T) = w_i(S)$ if $S\, R_i\, \pi(i)$ and T is coalition such that $T\, I_i\, S$.

Define a weight function $w' : E \to \mathbb{R}^+$ such that for any $S = \{i, j\} \in E$, $w'(S) = w_i(S) + w_j(S)$. For $E'' \subseteq E$, denote by $w'(E'')$, the value $\sum_{e \in E''} w'(e)$. We can then prove that π is Pareto optimal if and only if π is the maximum weight matching of $G^{w'}$, the graph G, weighted by weight function w'. The complete proof is omitted due to space limitations. Since we have a linear-time reduction to maximum weight matching [Gabow and Tarjan, 1991], the complexity of the algorithm is $O(n^3)$. □

Note that Theorem 6 allows us to find a Pareto optimal Pareto improvement for any given partition if the partition is not Pareto optimal.

5.3 W-hedonic games

We now turn to Pareto optimality in W-hedonic games.

Theorem 7 *For W-hedonic games, a partition that is both individually rational and Pareto optimal can be computed in polynomial time.*

Proof sketch: The statement follows from Lemma 2 and the fact that PP can be solved in polynomial time for W-hedonic games. The latter is proved by a polynomial-time reduction of PP to a polynomial-time solvable problem called *clique packing*.

We first introduce the more general notion of graph packing. Let $\mathscr{F}$ be a set of undirected graphs. An $\mathscr{F}$-packing of a graph G is a subgraph H such that each component of H is (isomorphic to) a member of $\mathscr{F}$. The size of $\mathscr{F}$-packing H is $|V(H)|$. We will informally say that vertex i is *matched* by $\mathscr{F}$-packing H if i is in a connected component in H. Then, a maximum $\mathscr{F}$-packing of a graph G is one that matches the maximum number of vertices. It is easy to see that computing a maximum $\{K_2\}$-packing of a graph is equivalent to maximum cardinality matching. Hell and Kirkpatrick [1984] and Cornuéjols *et al.* [1982] independently proved that there is a polynomial-time algorithm to compute a maximum $\{K_2, \ldots, K_n\}$-packing of a graph. Cornuéjols *et al.* [1982] note that finding a $\{K_2, \ldots, K_n\}$-packing can be reduced to finding a $\{K_2, K_3\}$-packing.

We are now in a position to reduce PP for W-hedonic games to computing a maximum $\{K_2, K_3\}$-packing. For a W-hedonic game (N, R), construct a graph $G = (N \cup (N \times \{0, 1\}), E)$ such that $\{(i, 0), (i, 1)\} \in E$ for all $i \in N$; $\{i, j\} \in E$ if and only if $i \in \max_{R_j}(N)$ and $j \in \max_{R_i}(N)$ for $i, j \in N$ such that $i \neq j$; and $\{i, (i, 0)\}, \{i, (i, 1)\} \in E$ if and only if $i \in \max_{R_i}(N)$ for all $i \in N$. Let H be a maximum $\{K_2, K_3\}$-packing of G.

It can then be proved that there exists a perfect partition of N according to R if and only if $|V(H)| = 3|N|$. We omit the technical details due to space restrictions.

Since PP for W-hedonic games reduces to checking whether graph G can be packed perfectly by elements in $\mathscr{F} = \{K_2, K_3\}$, we have a polynomial-time algorithm to solve

[3] In fact, Abraham and Manlove [2004] state that 'the case where preference lists [...] may include ties merits further investigation.'

PP for W-hedonic games. Denote by $CC(H)$ the set of connected components of graph H. If $|V(H)| = 3|N|$ and a perfect partition does exist, then $\{V(S) \cap N : S \in CC(H)\} \setminus \emptyset$ is a perfect partition. □

Similarly, the following is evident from the arguments in the proof of Theorem 7.

Theorem 8 *For W-hedonic games, it can be checked in polynomial time whether a given partition is Pareto optimal or weakly Pareto optimal.*

Our positive results for W-hedonic games also apply to hedonic games with $\mathscr{W}$-preferences.

5.4 B-hedonic games

We saw that for W-hedonic games, a Pareto optimal partition can be computed efficiently, even in the presence of unacceptable players. In the absence of unacceptable players, computing a Pareto optimal and individually rational partition is trivial in B-hedonic games, as the partition consisting of the grand coalition is a solution. Interestingly, if preferences do allow for unacceptable players, the same problem becomes NP-hard.

Theorem 9 *For B-hedonic games, computing a Pareto optimal partition is NP-hard.*

Proof sketch: It can be checked in polynomial time whether a partition is perfect in a B-hedonic game. Hence, by Lemma 1, it suffices to show that PP is NP-hard. We do so by a reduction from SAT. Let $\varphi = X_1 \wedge \cdots \wedge X_k$ a Boolean formula in conjunctive normal form in which the Boolean variables $p_1, \ldots, p_m$ occur. Now define the B-hedonic game (N, R), where $N = \{X_1, \ldots, X_k\} \cup \{p_1, \neg p_1, \ldots, p_m, \neg p_m\} \cup \{0, 1\}$ and the preferences for each literal p or $\neg p$, and each clause $X = (x_1 \vee \cdots \vee x_\ell)$ are denoted by lists of equivalence classes of equally preferred players in decreasing order of preference, as follows,

$$\begin{aligned}
p &: \{0,1\}, N \setminus \{0,1,\neg p\}, \{\neg p\} \\
\neg p &: \{0,1\}, N \setminus \{0,1,p\}, \{p\} \\
X &: \{x_1, \ldots, x_\ell\}, N \setminus \{0, x_1, \ldots, x_\ell\}, \{0\} \\
0 &: N \setminus \{0,1\}, \{0\}, \{1\} \\
1 &: N \setminus \{0,1\}, \{1\}, \{0\}
\end{aligned}$$

We prove that φ is satisfiable if and only if a perfect (and individually rational) partition for (N, R) exists. The proof details are omitted due to space limitations. □

By using similar techniques, the following can be proved.

Theorem 10 *For B-hedonic games, verifying whether a partition is weakly Pareto optimal is coNP-complete.*

6 Conclusions

Pareto optimality and individual rationality are important requirements for desirable partitions in coalition formation. In this paper, we examined computational and structural issues related to Pareto optimality in various classes of hedonic games (see Table 1). We saw that unacceptability and ties are a major source of intractability when computing Pareto optimal outcomes. In some cases, checking whether a given partition is Pareto optimal can be significantly harder than finding one. We expect Theorem 10 to also hold for Pareto optimality instead of weak Pareto optimality.

It should be noted that most of our insights gained into Pareto optimality and the resulting algorithmic techniques—especially those presented in Section 3 and Section 4—do not only apply to coalition formation but to any discrete allocation setting.

Game	VERIFICATION	COMPUTATION
General	coNP-C (Th. 4)	NP-hard (Th. 3)
General (strict)	coNP-C (Th. 4)	in P (Prop. 1)
Roommate	in P (Th. 6)	in P (Th. 5)
B-hedonic	coNP-C (Th. 10, weak PO)	NP-hard (Th. 9)
W-hedonic	in P (Th. 8)	in P (Th. 7)

Table 1: Complexity of Pareto optimality in hedonic games: positive results hold for both Pareto optimality and individual rationality.

References

A. Abdulkadiroğlu and T. Sönmez. Random serial dictatorship and the core from random endowments in house allocation problems. *Econometrica*, 66(3):689–702, 1998.

D. J. Abraham and D. F. Manlove. Pareto optimality in the roommates problem. Technical Report TR-2004-182, University of Glasgow, Department of Computing Science, 2004.

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

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

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

K. Cechlárová and J. Hajduková. Stability of partitions under $\mathscr{BW}$-preferences and $\mathscr{WB}$-preferences. *International Journal of Information Technology and Decision Making*, 3(4):605–618, 2004.

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

G. Cornuéjols, D. Hartvigsen, and W. Pulleyblank. Packing subgraphs in a graph. *Operations Research Letters*, 1(4):139–143, 1982.

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

H. N. Gabow and R. Tarjan. Faster scaling algorithms for general graph matching problems. *Journal of the ACM*, 38(4):815–853, October 1991.

P. Hell and D. G. Kirkpatrick. Packings by cliques and by finite families of graphs. *Discrete Mathematics*, 49(1):45–59, 1984.

T. Morrill. The roommates problem revisited. *Journal of Economic Theory*, 145(5):1739–1756, 2010.

E. Ronn. NP-complete stable matching problems. *Journal of Algorithms*, 11:285–304, 1990.