The Price of Neutrality for the Ranked Pairs Method

Markus Brill
Institüt für Informatik
Technische Universität München
85748 Garching, Germany

Felix Fischer
Statistical Laboratory
University of Cambridge
Cambridge CB3 0WB, UK

Abstract

The complexity of the winner determination problem has been studied for almost all common voting rules. A notable exception, possibly caused by some confusion regarding its exact definition, is the method of ranked pairs. The original version of the method, due to Tideman, yields a social preference function that is irresolute and neutral. A variant introduced subsequently uses an exogenously given tie-breaking rule and therefore fails neutrality. The latter variant is the one most commonly studied in the area of computational social choice, and it is easy to see that its winner determination problem is computationally tractable. We show that by contrast, computing the set of winners selected by Tideman's original ranked pairs method is NP-complete, thus revealing a trade-off between tractability and neutrality. In addition, several known results concerning the hardness of manipulation and the complexity of computing possible and necessary winners are shown to follow as corollaries from our findings.

1 Introduction

The fundamental problem of social choice theory can be concisely described as follows: given a number of individuals, or *voters*, each having a preference ordering over a set of *alternatives*, how can we aggregate these preferences into a collective, or *social*, preference ordering that is in some sense faithful to the individual preferences? By a preference ordering we here understand a (transitive) ranking of all alternatives, and a function aggregating individual preferences orderings into social preference orderings is called a *social preference function* (SPF).[1]

A natural idea to construct an SPF is by letting an alternative a be socially preferred to another alternative b if and only if a majority of voters prefers a to b. However, it was observed as early as the 18th century that this approach might lead to paradoxical situations: the collective preference relation may be cyclic even when all individual preferences are transitive (de Condorcet 1785).

To remedy this situation, a large number of SPFs have been suggested, together with a variety of criteria that a reasonable SPF should satisfy (see Arrow, Sen, and Suzumura 2002 for an overview). *Neutrality* and *anonymity*, for instance, are basic fairness criteria which require, loosely speaking, that all alternatives and all voters are treated equally. Another criterion we will be interested in is the *computational effort* required to evaluate an SPF. Computational tractability of the winner determination problem is obviously a significant property of any SPF: the inability to efficiently compute social preferences would render the method virtually useless, at least for large problem instances that do not exhibit additional structure. As a consequence, computational aspects of preference aggregation have received tremendous interest in recent years (see, e.g., Faliszewski et al. 2009; Conitzer 2010).

In this paper, we study the computational complexity of the *ranked pairs* method (Tideman 1987). To the best of our knowledge, this question has not been considered before, which is particularly surprising given the extensive literature that is concerned with computational aspects of ranked pairs.[2] A possible reason for this gap might be the confusion of two variants of the method, only one of which satisfies neutrality. In Section 2, we address this confusion and describe both variants as well as a closely related SPF known as *Kemeny's rule*. After introducing the necessary notation in Section 3, we show in Section 4 that deciding whether a given alternative is a ranked pairs winner for the neutral variant is NP-complete. Section 5 shows that a number of known results follow as corollaries from our results, and Section 6 discusses variants of the ranked pairs method that are not anonymous. Finally, Section 7 concludes with a complexity-theoretic comparison of the ranked pairs method and Kemeny's rule.

2 Kemeny's Rule and Two Variants of the Ranked Pairs Method

In this section we address the difference between two variants of the ranked pairs method that are commonly studied

[1] In contrast to a *social welfare function* as studied by Arrow (1951), an SPF can output *multiple* social preference orderings with the interpretation that all those rankings are tied for winner. The rationale behind this is to allow for a symmetric outcome when individual preferences are symmetric, like in the case of two individuals with diametrically opposed preferences.

[2] Typical problems include the hardness of manipulation (Betzler, Hemmann, and Niedermeier 2009; Xia and Conitzer 2011) and the complexity of computing possible and necessary winners (Xia et al. 2009; Obraztsova and Elkind 2011).

in the literature. Both variants are anonymous, i.e., treat all voters equally. Non-anonymous variants of the ranked pairs method have been suggested by Tideman (1987) and Zavist and Tideman (1989), and will be discussed in Section 6.

It will be instructive to first consider Kemeny's rule (Kemeny 1959; Young 1995). The latter, which is also known as the *Kemeny-Young method*, chooses the ranking or rankings with maximal *Kemeny score*, where the Kemeny score of a ranking measures the extent to which it agrees with individual preferences: for every pair (a, b) of alternatives, each voter i contributes one point to the Kemeny score of a ranking if and only if i ranks a and b in the same order as the ranking does. As the number of possible rankings grows rapidly with the number of alternatives (for n alternatives there are $n!$ rankings), computing all Kemeny scores is computationally demanding. Indeed, it was shown by Bartholdi, III, Tovey, and Trick (1989) that finding the so-called *Kemeny rankings*, i.e., the rankings with a maximal Kemeny score, is NP-hard.[3] This is commonly seen as strong evidence that no efficient algorithm exists for this problem.

The method of ranked pairs, introduced by Tideman (1987) and extended by Zavist and Tideman (1989), can be viewed as a heuristic to find a ranking with high—if not maximal—Kemeny score. This is not to say that the main motivation for ranked pairs was the approximation of Kemeny's rule. Actually, the ranked pairs method is interesting in its own right and satisfies a number of desirable properties, some of which are not satisfied by Kemeny's rule (see Tideman 1987; Lamboray 2009; Parkes and Xia 2012).

The easiest way to describe the ranked pairs method is to formulate it as a procedure. The procedure first defines a priority ordering over the set of all (ordered) *pairs* (a, b) of alternatives by giving priority to pairs (a, b) with a larger number of voters preferring a to b. Then, it constructs a ranking of the alternatives by starting with the empty ranking and iteratively considering pairs in order of priority. When a pair (a, b) is considered, the ranking is extended by fixing that a precedes b—unless fixing this pairwise comparison would create a cycle together with the previously fixed pairs, in which case the pair (a, b) is discarded. This procedure is guaranteed to terminate with a complete ranking of all alternatives.

What is missing from the above description is a *tie-breaking rule* for cases where two or more pairwise comparisons have the same support from the voters. This turns out to be a rather intricate issue. In principle, it is possible to employ an arbitrary tie-breaking rule. However, each fixed tie-breaking rule biases the method in favor of some alternative and thereby destroys neutrality.[4] In order to repair this flaw, Tideman (1987) originally defined the ranked pairs method to return the set of all those rankings that result from the above procedure for *some* tie-breaking rule.[5] We will henceforth denote this variant by RP.

In a subsequent paper, Zavist and Tideman (1989) showed that a tie-breaking rule is in fact necessary in order to achieve the property of *independence of clones*, which was the main motivation for introducing the ranked pairs method. While Zavist and Tideman (1989) proposed a way to define a tie-breaking rule based on the preferences of a distinguished voter (see Section 6 for details), the variant that is most commonly studied in the literature considers the tie-breaking rule to be exogenously given and fixed for all profiles. This variant of ranked pairs will be denoted by RPT. Whereas RP may output a *set* of rankings, with the interpretation that all the rankings in the set are tied for winner, RPT always outputs a single ranking. In social choice terminology, RP is an *irresolute* SPF, and RPT is a *resolute* one. It is straightforward to see that RP is neutral, i.e., treats all alternatives equally, and that RPT is not. An easy example for the latter statement is the case of two alternatives and two voters who each prefer a different alternative.

Rather than completely ranking all alternatives, it is often sufficient to identify the socially "best" alternatives. This is the purpose of a *social choice function* (SCF). An SCF has the same input as an SPF, but returns alternatives instead of rankings. Each SPF gives rise to a corresponding SCF that returns the top elements of the rankings instead of the rankings themselves, and we will frequently switch between these two settings. Interestingly, deciding whether a given ranking is chosen by an SPF can be considerably easier than deciding whether a given alternative is chosen by the corresponding SCF (see Table 1).

From a computational perspective, RPT is easy: constructing the ranking for a given tie-breaking rule takes time polynomial in the size of the input (see Proposition 1). For RP, however, the picture is different: as the number of tie-breaking rules is exponential, executing the iterative procedure for every single tie-breaking rule is infeasible. Of course, this does not preclude the existence a clever algorithm that efficiently computes the set of all alternatives that are the top element of some chosen ranking.[6] Our main result states that such an algorithm does not exist unless P equals NP.

3 Preliminaries

For a finite set X, let $\mathcal{L}(X)$ denote the set of all *rankings* of X, where a ranking of X is a complete, transitive, and asymmetric relation on X. The *top element* of a ranking $L \in \mathcal{L}(X)$, denoted by $top(L)$, is the unique element $x \in X$ such that $x \mathrel{L} y$ for all $y \in X \setminus \{x\}$.

[3] Later, Hemaspaandra, Spakowski, and Vogel (2005) strenghtened this result by showing that identifying Kemeny rankings is complete for $\Theta_2^p = \mathrm{P}_{||}^{\mathrm{NP}}$, the class of problems solvable via parallel access to NP.

[4] Neutrality can be maintained if the tie-breaking rule varies with the individual preferences (see Section 6).

[5] This definition, sometimes called *parallel universes tie-breaking (PUT)*, can also be used to "neutralize" other voting rules that involve tie-breaking (Conitzer, Rognlie, and Xia 2009). PUT can be interpreted as a *possible winner* notion: if the ranked pairs method is used with an unknown tie-breaking rule, the PUT version of ranked pairs selects exactly those alternatives that have a chance to be chosen in the actual election.

[6] As the number of chosen rankings might be exponential (for completely tied instances such as R^* in the proof of Proposition 4, *all* rankings are chosen), it immediately follows that computing all of them requires exponential time in the worst case.

Let $N = \{1, \dots, n\}$ be a set of voters with preferences over a finite set A of alternatives. The preferences of voter $i \in N$ are represented by a ranking R_i on A. The interpretation of $(a, b) \in R_i$, usually denoted by $a \; R_i \; b$, is that voter i strictly prefers a to b. A *preference profile* is an ordered list containing a ranking for each voter.

A *social choice function* (SCF) f associates with every preference profile R a non-empty set $f(R) \subseteq A$ of alternatives. A *social preference function* (SPF) f associates with every preference profile R a non-empty set $f(R) \subseteq \mathcal{L}(A)$ of rankings of A.

An SCF or SPF is *neutral* if permuting the alternatives in the individual rankings also permutes the set of chosen alternatives, or the set of chosen rankings, in the exact same way. Formally, f is neutral if $f(\pi(R)) = \pi(f(R))$ for all preference profiles R and all permutations π of A. An SCF or SPF is *anonymous* if the set of chosen alternatives, or the set of chosen rankings, does not change when the voters are permuted.

For a given preference profile $R = (R_1, \dots, R_n)$ and two distinct alternatives $a, b \in A$, we denote by $n_R(a, b)$ the difference between the number of voters who prefer a to b and the number of voters who prefer b to a, i.e.,

$$n_R(a, b) = |\{i \in N : a \; R_i \; b\}| - |\{i \in N : b \; R_i \; a\}|.$$

The resolute variant of the ranked pairs method takes as input a preference profile R and a *tie-breaking rule* $\tau \in \mathcal{L}(A \times A)$. A ranking $\succ_\tau^R$ of $A \times A$ is constructed by ordering all pairs in accordance with $n_R(\cdot, \cdot)$, using τ to break ties: $(a, b) \succ_\tau^R (c, d)$ if and only if $n_R(a, b) > n_R(c, d)$ or $(n_R(a, b) = n_R(c, d)$ and $(a, b) \; \tau \; (c, d))$.

The relation L_τ^R on A is constructed by iteratively considering the pair ranked highest by $\succ_\tau^R$ among all pairs that have not been considered so far. The pair is then added to the relation L_τ^R *unless* this addition would create an L_τ^R-cycle with the pairs that have been added before.[7] After all pairs in $A \times A$ have been considered, L_τ^R is a ranking of A. The resolute variant of ranked pairs, interpreted as an SCF, returns the top element of L_τ^R.

Definition 1. $\mathit{RPT}(R, \tau) = \{\mathit{top}(L_\tau^R)\}$.

The outcome of RPT depends on the choice of τ, and RPT is not neutral. Tideman (1987) therefore defined an irresolute and neutral variant that chooses all alternatives that are at the top of L_τ^R for *some* tie-breaking rule τ.

Definition 2. $\mathit{RP}(R) = \{a \in A :$ *there exists* $\tau \in \mathcal{L}(A \times A)$ *such that* $a = \mathit{top}(L_\tau^R)\}$.

The alternatives in $\mathsf{RP}(R)$ are called *ranked pairs winners for* R. In the SPF setting, RP returns the rankings $\{L_\tau^R : \tau \in \mathcal{L}(A \times A)\}$, which are henceforth called *ranked pairs rankings for* R.

We will work with an alternative characterization of ranked pairs rankings that was introduced by Zavist and Tideman (1989). Given a preference profile R, a ranking L of A, and two alternatives a and b, we say that *a attains b through L* if there exists a sequence of distinct alternatives $a_1, a_2, \dots, a_\ell$, where $\ell \geq 2$, such that $a_1 = a$, $a_\ell = b$, $a_i \; L \; a_{i+1}$, and $n_R(a_i, a_{i+1}) \geq n_R(b, a)$ for all i with $1 \leq i < \ell$. In this case, we will say that a attains b *via* $(a_1, a_2, \dots, a_\ell)$. A ranking L is called a *stack* if for any pair of alternatives a and b it holds that $a \; L \; b$ implies that a attains b through L.

Lemma 1 (Zavist and Tideman 1989)**.** *A ranking of A is a ranked pairs ranking if and only if it is a stack.*

It follows that an alternative is a ranked pairs winner if and only if it is the top element of a stack.

4 Results

We are now ready to study the computational complexity of RP. We first consider the SPF setting and observe that finding and checking ranked pairs rankings is easy. This also provides an efficient way to find *some* ranked pairs winner, i.e., some alternative that is chosen in the SCF setting. The problem of deciding whether a *particular* alternative is a ranked pairs winner, on the other hand, turns out to be NP-complete. We then demonstrate that for the ranked pairs method neutrality and tractability are incompatible in a more general sense. Finally, we extend the hardness result to a variant of the winner determination problem that asks for unique winners. Some easy proofs are omitted due to space constraints and can be found in the full version of the paper.

4.1 Ranked Pairs Rankings

It can easily be seen that an arbitrary ranked pairs ranking can be found efficiently.

Proposition 1. *Finding a ranked pairs ranking is in P.*

Deciding whether a given ranking is a ranked pairs ranking is also feasible in polynomial time, by checking whether the given ranking is a stack.

Proposition 2. *Deciding whether a given ranking is a ranked pairs ranking is in P.*

It is worth noting that Proposition 2 can also be shown directly, without referring to stacks. For a given ranking L, define a tie-breaking rule τ_L such that $(a, b) \; \tau_L \; (c, d)$ for all $(a, b) \in L$ and $(c, d) \notin L$. It can be shown that L is a ranked pairs ranking if and only if $L = L_{\tau_L}^R$. The advantage of this alternative proof is that for each "yes" instance it constructs a witnessing tie-breaking rule.

4.2 Ranked Pairs Winners

We now consider the SCF setting. As every ranked pairs ranking yields a ranked pairs winner, Proposition 1 immediately implies that an arbitrary element of $\mathsf{RP}(R)$ can be found efficiently.

Proposition 3. *Finding a ranked pairs winner is in P.*

Deciding whether a given alternative is a ranked pairs winner, on the other hand, turns out to be NP-complete.

Theorem 1. *Deciding whether a given alternative is a ranked pairs winner is NP-complete.*

[7] A P-cycle of a relation $P \subseteq A \times A$ is a set of pairs $\{(a_1, a_2), (a_2, a_3), \dots, (a_{\ell-1}, a_\ell)\}$ with $a_\ell = a_1$ and $a_i \; P \; a_{i+1}$ for all i with $1 \leq i < \ell$. Pairs of the form (a, a) are considered cycles of length 1, and are therefore never added to L_τ^R.

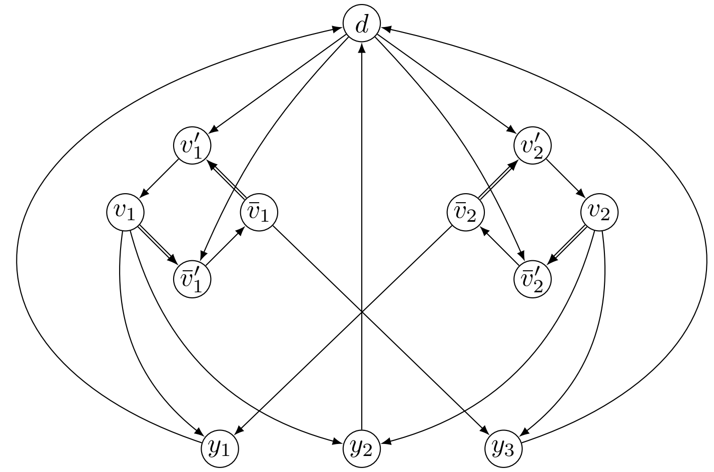

Figure 1: Graphical representation of $n_{R_\varphi}(\cdot,\cdot)$ for the Boolean formula $\varphi = \{v_1, \bar{v}_2\} \wedge \{v_1, v_2\} \wedge \{\bar{v}_1, v_2\}$. The relation $\succ^2$ is represented by arrows, and $\succ^4$ is represented by double-shafted arrows. For all pairs (a, b) that are not connected by an arrow, we have $n(a, b) = n(b, a) = 0$.

Membership in NP follows from Proposition 2. For hardness, we give a reduction from the NP-complete Boolean satisfiability problem (SAT, see, e.g., Papadimitriou 1994). An instance of SAT consists of a Boolean formula $\varphi = C_1 \wedge \cdots \wedge C_k$ in conjunctive normal form over a finite set $V = \{v_1, \ldots, v_m\}$ of variables. Denote by $X = \{v_1, \bar{v}_1, \ldots, v_m, \bar{v}_m\}$ the set of all *literals*, where a literal is either a variable or its negation. Each *clause* C_j is a set of literals. An *assignment* $\alpha \subseteq X$ is a subset of the literals with the interpretation that all literals in α are set to "true." Assignment α is *valid* if $\ell \in \alpha$ implies $\overline{\ell} \notin \alpha$ for all $\ell \in X$, and α *satisfies* clause C_j if $C_j \cap \alpha \neq \emptyset$. A valid assignment that satisfies all clauses of φ will be called a *satisfying assignment* for φ, and a formula that has a satisfying assignment will be called *satisfiable*.

For a particular Boolean formula $\varphi = C_1 \wedge \cdots \wedge C_k$ over a set $V = \{v_1, \ldots, v_m\}$ of variables, we will construct a preference profile R_φ over a set A_φ of alternatives such that a particular alternative $d \in A_\varphi$ is a ranked pairs winner for R_φ if and only if φ is satisfiable.

Let us first define the set A_φ of alternatives. For each variable $v_i \in V$, $1 \leq i \leq m$, there are four alternatives v_i, $\bar{v}_i$, v_i', and $\bar{v}_i'$. For each clause C_j, $1 \leq j \leq k$, there is one alternative y_j. Finally, there is one alternative d for which we want to decide membership in $\mathsf{RP}(R_\varphi)$.

Instead of constructing R_φ explicitly, we will specify a number $n(a, b)$ for each pair $(a, b) \in A_\varphi \times A_\varphi$. Debord (1987) has shown that there exists a preference profile R such that $n_R(a, b) = n(a, b)$ for all a, b, as long as $n(a, b) = -n(b, a)$ for all a, b and all the numbers $n(a, b)$ have the same parity.[8] In order to conveniently define $n(\cdot,\cdot)$, the following notation will be useful: for a natural number w, $a \succ^w b$ denotes setting $n(a, b) = w$ and $n(b, a) = -w$.

For each variable $v_i \in V$, $1 \leq i \leq m$, let $v_i \succ^4 \bar{v}_i' \succ^2 \bar{v}_i \succ^4 v_i' \succ^2 v_i$. For each clause C_j, $1 \leq j \leq k$, let $v_i \succ^2 y_j$ if variable $v_i \in V$ appears in clause C_j as a positive literal, and $\bar{v}_i \succ^2 y_j$ if variable v_i appears in clause c_j as a negative literal. Finally let $y_j \succ^2 d$ for $1 \leq j \leq k$ and $d \succ^2 v_i'$ and $d \succ^2 \bar{v}_i'$ for $1 \leq i \leq m$. For all pairs (a, b) for which $n(a, b)$ has not been specified so far, let $n(a, b) = n(b, a) = 0$. An example is shown in Figure 1.

As $n(a, b) \in \{-4, -2, 0, 2, 4\}$ for all $a, b \in A_\varphi$, Debord's (1987) theorem guarantees the existence of a preference profile R_φ with $n_{R_\varphi}(a, b) = n(a, b)$ for all $a, b \in A_\varphi$, and such a profile can in fact be constructed efficiently.

The following two lemmata show that alternative d is a ranked pairs winner for R_φ if and only if the formula φ is satisfiable.

Lemma 2. *If $d \in \mathsf{RP}(R_\varphi)$, then φ is satisfiable.*

Proof. Assume that d is a ranked pairs winner for R_φ and let L be a stack with $top(L) = d$. Consider an arbitrary j with $1 \leq j \leq k$. As L is a stack and $d \; L \; y_j$, d attains y_j through L, i.e., there exists a sequence $P_j = (a_1, a_2, \ldots, a_\ell)$ with $a_1 = d$ and $a_\ell = y_j$ such that $a_i \; L \; a_{i+1}$ and $n(a_i, a_{i+1}) \geq 2$ for all i with $1 \leq i < \ell$. If d attains y_j via several sequences, fix one of them arbitrarily.

The definition of $n(\cdot,\cdot)$ implies that

$$P_j = (d, \overline{\ell}', \overline{\ell}, \ell', \ell, y_j) \quad \text{or}$$
$$P_j = (d, \ell', \ell, y_j),$$

where ℓ is some literal. The former is in fact not possible because $\overline{\ell}'$ does not attain ℓ through L. Therefore, each P_j is of the form $P_j = (d, \ell', \ell, y_j)$ for some $\ell \in X$.

Now define assignment α as the set of all literals that are contained in one of the sequences $P_j, 1 \leq j \leq k$, i.e., $\alpha = X \cap (\bigcup_{j=1}^{k} P_j)$. We claim that α is a satisfying assignment for φ.

In order to show that α is valid, suppose there exists a literal $\ell \in X$ such that both ℓ and $\overline{\ell}$ are contained in α.

[8] See also Le Breton (2005).

This implies that there exist i and j such that d attains y_i via $P_i = (d, \ell', \ell, y_i)$ and d attains y_j via $P_i = (d, \overline{\ell}', \overline{\ell}, y_j)$. In particular, $\ell' \; L \; \ell$ and $\overline{\ell}' \; L \; \overline{\ell}$. As L is transitive and asymmetric, it follows that either $\ell' \; L \; \overline{\ell}$ or $\overline{\ell}' \; L \; \ell$. However, neither does ℓ' attain $\overline{\ell}$ through L, nor does $\overline{\ell}'$ attain ℓ through L, a contradiction.

In order to see that α satisfies φ, consider an arbitrary clause C_j. As d attains y_j via $P_j = (d, \ell', \ell, y_j)$ and $n(y_j, d) = 2$, we have that $n(\ell, y_j) \geq 2$. By definition of $n(\cdot,\cdot)$, this implies that $\ell \in C_j$. □

Lemma 3. *If φ is satisfiable, then $d \in \mathsf{RP}(R_\varphi)$.*

Proof. Assume that φ is satisfiable and let α be a satisfying assignment. Let $V_i = \{v_i, \bar{v}_i, v_i', \bar{v}_i'\}$, $1 \leq i \leq m$, and $Y = \{y_1, y_2, \ldots, y_k\}$. We define a ranking L of A_φ as follows, using $B \; L \; C$ as shorthand for $b \; L \; c$ for all $b \in B$ and $c \in C$.

- For all $1 \leq i \leq m$, let $d \; L \; V_i$ and $V_i \; L \; Y$.
- For all $1 \leq i < j \leq m$, let $V_i \; L \; V_j$.
- For the definition of L within V_i, we distinguish two cases. If $v_i \in \alpha$, i.e., if v_i is set to "true" under α, let $\overline{v}_i \; L \; v_i' \; L \; v_i \; L \; \overline{v}_i'$. If, on the other hand, $v_i \notin \alpha$, let $v_i \; L \; \overline{v}_i' \; L \; \overline{v}_i \; L \; v_i'$.
- Within Y, define L arbitrarily.

We now prove that L is a stack. For each pair (a, b) with $a \; L \; b$, we need to verify that a attains b through L. If $n(b, a) \leq 0$, it is easily seen that a attains b through L. We can therefore assume that $n(b, a) > 0$. By definition of L and $n(\cdot,\cdot)$, a particular such pair (a, b) satisfies either

$$a = d \text{ and } b \in Y \text{, or}$$
$$a, b \in V_i \text{ for some } i \text{ with } 1 \leq i \leq m.$$

First consider a pair of the former type, i.e., $(a, b) = (d, y_j)$ for some j with $1 \leq j \leq k$. As α satisfies C_j, there exists $\ell \in C_j$ with $\ell \in \alpha$. Consider the sequence $P_j = (d, \ell', \ell, y_j)$. As $n(y_j, d) = 2$ and $d \succ^2 \ell' \succ^2 \ell \succ^2 y_j$, d attains y_j via P_j.

Now consider a pair of the latter type, i.e., $a, b \in V_i$ for some i with $1 \leq i \leq m$. Assume that $v_i \in \alpha$ and, therefore, $\overline{v}_i \; L \; v_i' \; L \; v_i \; L \; \overline{v}_i'$. The only non-trivial case is the pair $(\overline{v}_i, \overline{v}_i')$ with $\overline{v}_i \; L \; \overline{v}_i'$ and $n(\overline{v}_i', \overline{v}_i) = 2$. But $\overline{v}_i$ attains $\overline{v}_i'$ via $(\overline{v}_i, v_i', v_i, \overline{v}_i')$ because $\overline{v}_i \succ^4 v_i' \succ^2 v_i \succ^4 \overline{v}_i'$. The case $v_i \notin \alpha$ is analogous.

We have shown that L is a stack. Lemma 1 now implies that $d \in \mathsf{RP}(R_\varphi)$, which completes the proof. □

Combining Lemma 2 and Lemma 3, and observing that both A_φ and R_φ can be constructed efficiently, completes the proof of Theorem 1.

4.3 Neutrality versus Tractability

In a sense, RP and RPT are very different variants of the ranked pairs method: whereas RPT uses only a single tie-breaking rule, the definition of RP ranges over the whole set $\mathcal{L}(A \times A)$. It is a natural question whether there exists a set $\mathcal{T} \subseteq \mathcal{L}(A \times A)$ of tie-breaking rules such that the corresponding variant is both neutral and tractable.

Formally, for a preference profile R and a set $\mathcal{T}$ of tie-breaking rules, define

$$\mathsf{RP}_{\mathcal{T}}(R) = \{a \in A : \exists \tau \in \mathcal{T} \text{ such that } a = top(L_\tau^R)\}.$$

Thus, in particular, $\mathsf{RP}(\cdot) = \mathsf{RP}_{\mathcal{L}(A \times A)}(\cdot)$ and $\mathsf{RPT}(\cdot, \tau) = \mathsf{RP}_{\{\tau\}}(\cdot)$. While $\mathsf{RP}_{\mathcal{T}}$ is anonymous for all $\mathcal{T} \subseteq \mathcal{L}(\mathcal{A} \times \mathcal{A})$, other properties of $\mathsf{RP}_{\mathcal{T}}$ obviously depend on $\mathcal{T}$. We say that $\mathcal{T}$ is *neutral* if $\mathsf{RP}_{\mathcal{T}}$ is neutral. Furthermore, we call $\mathcal{T}$ *intractable* if deciding membership in $\mathsf{RP}_{\mathcal{T}}(R)$ is NP-complete, and *tractable* if the problem is in P. Our previous results imply that $\mathcal{L}(A \times A)$ is neutral but intractable, and that $\{\tau\}$ is tractable but not neutral for any $\tau \in \mathcal{L}(A \times A)$.

It turns out that RP is the only variant in this framework that is neutral.

Proposition 4. *If $\mathcal{T}$ is neutral, then $\mathsf{RP}_{\mathcal{T}} = \mathsf{RP}$.*

Proof. Let $\mathcal{T}$ be neutral. We have to show that $\mathsf{RP}_{\mathcal{T}}(R) = \mathsf{RP}(R)$ for all preference profiles R. The inclusion from left to right follows from the definition of $\mathsf{RP}_{\mathcal{T}}$. For the inclusion from right to left, we need the following observation.

Define R^* as the preference profile that contains each element of $\mathcal{L}(A)$ exactly once. Since $n_{R^*}(a, b) = 0$ for all $a, b \in A$ and R^* is completely symmetric, neutrality and anonymity of $\mathsf{RP}_{\mathcal{T}}$ imply that $\mathsf{RP}_{\mathcal{T}}(R^*) = \mathcal{L}(A)$. It follows that for every $L \in \mathcal{L}(A)$ there exists a tie-breaking rule $\tau_L \in \mathcal{T}$ such that $L = L_{\tau_L}^{R^*}$.

Now let R be an arbitrary preference profile and $L \in \mathsf{RP}(R)$. Let furthermore $\tau_L \in \mathcal{T}$ be a tie-breaking rule such that $L = L_{\tau_L}^{R^*}$. It is now easily verified that $L_{\tau_L}^{R} = L$, which implies that $L \in \mathsf{RP}_{\mathcal{T}}(R)$. □

By combining Proposition 4 and Theorem 1, it immediately follows that neutrality and tractability are incompatible even within this generalized class of ranked pairs methods.

Corollary 1. *If $\mathcal{T}$ is neutral, then it is intractable.*

4.4 Uniqueness of Winners

An interesting variant of the winner determination problem concerns the question whether a given alternative is the *unique* winner for a given preference profile. Despite its similarity to the original winner determination problem, this problem is sometimes considerably easier.

For RP, the picture is different: verifying unique winners is not feasible in polynomial time, unless P equals coNP.

Theorem 2. *Deciding whether a given alternative is the unique ranked pairs winner is coNP-complete.*

Proof. Membership in coNP follows from the observation that for every "no" instance there is a stack whose top element is different from the alternative in question.

For hardness, we modify the construction from Section 4.2 to obtain a reduction from the problem UNSAT, which asks whether a given Boolean formula is *not* satisfiable. For a Boolean formula φ, define $A_\varphi' = A_\varphi \cup \{d^*\}$, where d^* is a new alternative and A_φ is defined as in Section 4.2. R_φ' is defined such that $d \succ^2 d^*$ and $d^* \succ^4 a$ for all $a \in A_\varphi \setminus \{d\}$. Within A_φ, R'_φ coincides with R_φ. We show that $\mathsf{RP}(R'_\varphi) = \{d^*\}$ if and only if φ is unsatisfiable.

For the direction from left to right, assume for contradiction that $\mathsf{RP}(R'_\varphi) = \{d^*\}$ and φ is satisfiable. Consider a satisfying assignment α and let L be the ranking of A_φ defined in the proof of Lemma 3. Define the ranking L' of A'_φ by

$$L' = L \cup \{(d, d^*)\} \cup \{(d^*, a) : a \in A_\varphi \setminus \{d\}\}.$$

That is, L' extends L by inserting the new alternative d^* in the second position. As in the proof of Lemma 3, it can be shown that L' is a stack. It follows that $top(L') = d \in \mathsf{RP}(R'_\varphi)$, contradicting the assumption that $\mathsf{RP}(R'_\varphi) = \{d^*\}$.

For the direction from right to left, assume for contradiction that φ is unsatisfiable and $\mathsf{RP}(R'_\varphi) \neq \{d^*\}$. Then there exists a tie-breaking rule τ such that $top(L_\tau^{R'_\varphi}) = a \neq d^*$. From the definition of R'_φ it follows that $a = d$, as $d^* \succ^4 b$ for all $b \in A_\varphi \setminus \{d\}$ and there are no $\succ^4$-cycles. By the same argument as in the proof of Lemma 2, it can be shown that φ is satisfiable, contradicting our assumption. □

5 Alternative Proofs for Known Results

In this section we briefly review results from the literature that follow from our findings. All results concern the neutral variant RP. We refer to the respective papers for formal definitions of the computational problems.

An alternative a is a *possible winner* for a partially specified preference profile R if there exists a completion R' of R such that a is a winner for R'. It is a *necessary winner* if it is a winner for every completion of R. Both the possible and the necessary winner problem have a variant that requires an alternative to be the *unique* winner for the completion.

Corollary 2 (Xia and Conitzer 2011)**.** *Computing possible (unique) ranked pairs winners for partially specified preference profiles is NP-complete.*

Proof. Membership in NP is straightforward. Hardness follows from the fact that the possible (unique) winner problem is equivalent to the (unique) winner determination problem in the special case when the preference profile is completely specified. □

Corollary 3 (Xia and Conitzer 2011)**.** *The following hold:*

- *Computing necessary unique ranked pairs winners for partially specified preference profiles is coNP-complete.*
- *Computing necessary ranked pairs winners for partially specified preference profiles is coNP-hard.*[9]

Proof. The "unique" variant is in coNP because for every "no" instance there is a completion and a tie-breaking rule that produces a different winner. Hardness of both variants follows from the fact that the necessary (unique) winner problem is equivalent to the (unique) winner determination problem in the special case when the preference profile is completely specified. □

The *unweighted coalitional manipulation* (UCM) problem asks whether it is possible for a group of voters to cast their votes in a way that achieves a given outcome.

Corollary 4 (Xia et al. 2009)**.** *The UCM problem under ranked pairs is NP-complete for any number of manipulators.*[10]

Proof. Membership in NP is again straightforward. Hardness follows from the fact that the UCM problem with zero manipulators is equivalent to the unique winner determination problem. □

6 Non-Anonymous Variants

As mentioned in Section 2, Tideman (1987) and Zavist and Tideman (1989) suggested ways to use the preferences of a distinguished voter, say, a chairperson, to render the ranked pairs method resolute. There are essentially two ways to achieve this, which differ in the point in time when ties are broken. For the sake of simplicity, we only consider the SCF setting in this section.

The *a priori* variant uses the preferences of the chairperson in order to construct a tie-breaking rule $\tau \in \mathcal{L}(A \times A)$, which is then used to compute $\mathsf{RPT}(\cdot, \tau)$. The *a posteriori* variant first computes $\mathsf{RP}(\cdot)$ and then chooses the alternative from this set that is most preferred by the chairperson. Both variants are neutral: if the alternatives are permuted in each ranking, including the ranking of the chairperson, the tie-breaking rule and thus the chosen alternative will change accordingly.

Whereas the *a priori* variant is a special case of RPT and therefore efficiently computable, the *a posteriori* variant is intractable by the results in Section 4. It follows that neutrality and tractability can be reconciled at the expense of anonymity. By moving to *non-deterministic* SCFs, one can even regain anonymity: choosing the chairperson for the *a priori* variant uniformly at random results in a procedure that is neutral, anonymous, and tractable, for appropriate generalizations of anonymity and neutrality to the case of non-deterministic SCFs. The winner determination problem for the *a posteriori* variant remains intractable when the chairperson is chosen randomly.

7 Conclusion

We have studied the complexity of the ranked pairs method. While *some* ranked pairs winner is easy to find, deciding whether a given alternative is a winner turns out to be NP-complete. If one is interested in ranked pairs rankings, both problems are computationally easy. Table 1 summarizes these results and contrasts them with the corresponding results for Kemeny's rule. Interestingly, all four problems are computationally harder for the latter under plausible complexity-theoretic assumptions.

[9] Xia and Conitzer (2011) state that the "non-unique" variant is coNP-complete as well. However, their argument for membership in coNP is incorrect since it assumes that winner determination is in P.

[10] The proof of Theorem 4.1 by Xia et al. (2009) actually works for both RP and RPT (Xia, personal communication, March 29, 2012).

	Ranked pairs	Kemeny's rule
find ranking	in P	NP-hard[a]
find winner	in P	NP-hard[a]
is ranking	in P	coNP-complete[b]
is winner	NP-complete	Θ_2^p-complete[b]

[a] Bartholdi, III, Tovey, and Trick (1989)

[b] Hemaspaandra, Spakowski, and Vogel (2005)

Table 1: Computational aspects of the ranked pairs method and Kemeny's rule

From a practical point of view, the ranked pairs method is easier than Kemeny's rule as well. The reason is that the *expected* number of ties among two or more pairs is rather small. This is particularly true when the number of voters is large compared to the number of alternatives, which is the case in many realistic settings. It is therefore to be expected that ranked pairs winners are easy to compute *on average* for most reasonable distributions of individual preferences.

Our results reveal a trade-off between neutrality and tractability in the context of the ranked pairs method: while the efficiently computable variant RPT fails neutrality, the neutral variant RP is intractable. In fact, this tension cannot be resolved even when moving to an arbitrary set of fixed tie-breaking rules. A very similar trade-off can be observed for the *single transferable vote* rule (Conitzer, Rognlie, and Xia 2009; Wichmann 2004).

We have finally discussed variants of the ranked pairs method that achieve neutrality at the expense of anonymity, by using individual preferences to break ties. The tractability of those variants depends on the point in time ties are broken.

Acknowledgements

This material is based on work supported by the Deutsche Forschungsgemeinschaft under grants BR 2312/10-1 and FI 1664/1-1. We thank three anonymous reviewers for valuable comments and suggestions.

References

Arrow, K. J.; Sen, A. K.; and Suzumura, K., eds. 2002. *Handbook of Social Choice and Welfare*, volume 1. North-Holland.

Arrow, K. J. 1951. *Social Choice and Individual Values*. New Haven: Cowles Foundation, 1st edition. 2nd edition 1963.

Bartholdi, III, J.; Tovey, C. A.; and Trick, M. A. 1989. Voting schemes for which it can be difficult to tell who won the election. *Social Choice and Welfare* 6(3):157–165.

Betzler, N.; Hemmann, S.; and Niedermeier, R. 2009. A multivariate complexity analysis of determining possible winners given incomplete votes. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI)*, 53–58.

Conitzer, V.; Rognlie, M.; and Xia, L. 2009. Preference functions that score rankings and maximum likelihood estimation. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI)*, 109–115.

Conitzer, V. 2010. Making decisions based on the preferences of multiple agents. *Communications of the ACM* 53(3):84–94.

de Condorcet, M. 1785. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Imprimerie Royale. Facsimile published in 1972 by Chelsea Publishing Company, New York.

Debord, B. 1987. Caractérisation des matrices des préférences nettes et méthodes d'agrégation associées. *Mathématiques et sciences humaines* 97:5–17.

Faliszewski, P.; Hemaspaandra, E.; Hemaspaandra, L.; and Rothe, J. 2009. A richer understanding of the complexity of election systems. In Ravi, S., and Shukla, S., eds., *Fundamental Problems in Computing: Essays in Honor of Professor Daniel J. Rosenkrantz*. Springer-Verlag.

Hemaspaandra, E.; Spakowski, H.; and Vogel, J. 2005. The complexity of Kemeny elections. *Theoretical Computer Science* 349(3):382–391.

Kemeny, J. G. 1959. Mathematics without numbers. *Daedalus* 88:577–591.

Lamboray, C. 2009. A prudent characterization of the ranked pairs rule. *Social Choice and Welfare* 32(1):129–155.

Le Breton, M. 2005. On the uniqueness of equilibrium in symmetric two-player zero-sum games with integer payoffs. *Économie publique* 17(2):187–195.

Obraztsova, S., and Elkind, E. 2011. On the complexity of voting manipulation under randomized tie-breaking. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI)*, 319–324. AAAI Press.

Papadimitriou, C. H. 1994. *Computational Complexity*. Addison-Wesley.

Parkes, D., and Xia, L. 2012. A complexity-of-strategic-behavior comparison between Schulze's rule and ranked pairs. In *Proceedings of the 26th AAAI Conference on Artificial Intelligence (AAAI)*. AAAI Press. Forthcoming.

Tideman, T. N. 1987. Independence of clones as a criterion for voting rules. *Social Choice and Welfare* 4(3):185–206.

Wichmann, B. A. 2004. Tie breaking in STV. *Voting Matters* 19:1–5.

Xia, L., and Conitzer, V. 2011. Determining possible and necessary winners under common voting rules given partial orders. *Journal of Artificial Intelligence Research* 41:25–67.

Xia, L.; Zuckerman, M.; Procaccia, A. D.; Conitzer, V.; and Rosenschein, J. S. 2009. Complexity of unweighted coalitional manipulation under some common voting rules. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI)*, 348–353. AAAI Press.

Young, H. P. 1995. Optimal voting rules. *Journal of Economic Perspectives* 9(1):51–64.

Zavist, T. M., and Tideman, T. N. 1989. Complete independence of clones in the ranked pairs rule. *Social Choice and Welfare* 6(2):167–173.