Nonmanipulable Selections from a Tournament

Alon Altman
Stanford University
epsalon@stanford.edu

Ariel D. Procaccia
Microsoft Israel R&D Center
arielpro@gmail.com

Moshe Tennenholtz
Microsoft Israel R&D Center
and Technion, IIT
moshet@microsoft.com

Abstract

A tournament is a binary dominance relation on a set of alternatives. Tournaments arise in many contexts that are relevant to AI, most notably in voting (as a method to aggregate the preferences of agents). There are many works that deal with choice rules that select a desirable alternative from a tournament, but very few of them deal directly with incentive issues, despite the fact that game-theoretic considerations are crucial with respect to systems populated by selfish agents.

We deal with the problem of the manipulation of choice rules by considering two types of manipulation. We say that a choice rule is *monotonic* if an alternative cannot get itself selected by losing on purpose, and *pairwise nonmanipulable* if a pair of alternatives cannot make one of them the winner by reversing the outcome of the match between them. Our main result is a combinatorial construction of a choice rule that is monotonic, pairwise nonmanipulable, and onto the set of alternatives, for any number of alternatives besides three.

1 Introduction

A *tournament* is a binary dominance relation on a set of alternatives A, i.e., for every two alternatives $x, y \in A$ either x dominates y or y dominates x, but not both. Tournaments are of course ubiquitous in sports but arise in a wide variety of settings, many of them highly relevant to AI and, in particular, to multiagent systems. For example, tournaments can be used to model a defeat relation between different arguments in argumentation theory [Dung, 1995; Dunne, 2007], and have also been used in multi-criteria decision making (see, e.g., [Bouyssou *et al.*, 2006]).

In addition, tournaments are often used to model the majority dominance relation in an election. In this context, a voter has a ranking over the set of alternatives; alternative x is said to dominate y if the majority of voters prefer x to y. As early as the eighteenth century, the Marquis de Condorcet noticed that the majority relation may contain cycles [de Condorcet, 1785]. Almost two centuries later, Condorcet's observation was generalized by McGarvey [1953], who established that *any* tournament can be obtained as the dominance relation of an election. The connection between social choice theory and multiagent systems is by now very well-established (see, e.g., [Conitzer *et al.*, 2007; Procaccia and Rosenschein, 2007; Hemaspaandra *et al.*, 2007] and the many references therein).

A recurring theme in the literature is how to select the "most desirable" element given a tournament. This desirable element is then deemed the winner of the sports competition, the winner of the election, etc. Due to the fact that these winner determination problems are so central, they have become an important subject of study in multiagent systems and computational social choice (see, e.g., [Brandt *et al.*, 2008] and the references therein).

In settings where multiple heterogeneous, selfish agents must agree on a common choice, incentive issues naturally become prevalent. Therefore, it is not surprising that recent years have seen a fast growing interest in game theory and mechanism design within the AI community. Game-theoretic considerations are very common, for instance, in computational social choice, but are now also being investigated in other areas that are relevant to the study of tournaments, such as argumentation theory [Procaccia and Rosenschein, 2005; Rahwan and Larson, 2008].

Despite the discussion above, the direct study of incentives in the context of selecting a winner from a tournament has so far been very limited. This is our focus in this paper.

Our Results. We consider choice rules that select an alternative given a tournament, that is, functions from the set of all tournaments over A to A. The players in our game are the alternatives. In our model, an alternative cannot announce that it dominates an alternative that it loses to. However, an alternative can lose on purpose to another alternative. In other words, if $x \in A$ dominates $y \in A$, x can cheat by reversing the outcome of the match between itself and y, but y cannot reverse the outcome of this match unless x agrees. Specifically, we consider two basic types of manipulation:

1. An alternative loses a match, with the purpose making itself the winner.
2. Two alternatives reverse the result of their match in order to make one of them the winner.

Of course, the question of whether manipulation is possible depends on the choice rule used to select a winner. A choice

Dagstuhl Seminar Proceedings 10101
Computational Foundations of Social Choice
http://drops.dagstuhl.de/opus/volltexte/2010/2560

rule that is immune to the former type of manipulation is said to be *monotonic*, whereas a rule that is immune to the latter is called *pairwise nonmanipulable* (PNM).

In addition to nonmanipulability, we wish to consider choice rules that satisfy minimal notions of fairness. One notion is known as *Condorcet-consistency*: the rule must select an alternative that dominates every other alternative, if such an alternative exists in the given tournament. *Non-imposition* is a strictly weaker notion of fairness: the rule must be onto the set of alternatives, that is, any alternative can be selected.

Our first result is an impossibility result: we prove that there are no choice rules that are both PNM and Condorcet-consistent. We therefore relax our fairness criterion by requiring non-imposition. Even then, we establish that when $|A| = 3$ there are no choice rules that are both PNM and non-imposing. However, we complement this impossibility by a surprising possibility theorem, which is also the main result (conceptually and technically) of the paper: for any $|A| \neq 3$, there is a choice rule that is PNM, non-imposing, and also monotonic. Thus, if the number of alternatives is not three, there exists a non-imposing choice rule that is immune to both types of manipulation discussed above.

Related Work. There is a rich literature on axiomatic characterizations of tournament choice rules and choice sets (the latter select a *set* of desirable alternatives). Over the years, many tournament solutions have been suggested, and their properties have been studied in detail. The goal is usually to design tournament solutions that satisfy a set of axioms (as we do here, with respect to our manipulation-related axioms). A comprehensive review of this literature is given in the book by Laslier [1997].

As mentioned above, a tournament choice rule can be interpreted as a voting rule, which maps a preference profile (a vector of rankings of the alternatives) to a winning alternative. In this context there are many impossibility theorems in the social choice literature regarding rules that are immune to manipulation, albeit when the players are the voters. In particular, the Gibbard-Satterthwaite Theorem [Gibbard, 1973; Satterthwaite, 1975] asserts that if there are at least three alternatives, any nonmanipulable voting rule must be a dictatorship, that is, there is one voter that decides the outcome of the election. However, we discuss an inherently different setting where the potential manipulators are the alternatives themselves.

A rare example of a paper that directly studies manipulation of tournament choice rules is the work of Dutta et al. [2002]. They examined a specific class of choice rules, in the context of voting. Specifically, Dutta et al. investigated a setting where alternatives can decide whether or not to enter the election; they characterize the set of alternatives that can be outcomes of the election in equilibrium.

Finally, there are some mathematical, if not conceptual, connections between our results and the work of Altman and Tennenholtz [2008]. In that work, selection functions based on aribtrary input from the agents were discussed, while in our work an agent's influence is limited to manipulating games that involve it. That said, in the case of three agents the mathematical representations of both results converge.

Structure of the paper. In Section 2, we introduce existing concepts and definitions. Our contribution appears in Sections 3 and 4: Section 3 deals with monotonic choice rules, while Section 4 deals with pairwise nonmanipulable choice rules and contains our main results. We then discuss our results in Section 5.

2 Preliminaries

Let A be a set of alternatives. A *tournament* T over A is a complete asymmetric binary relation over A. In other words, for every two distinct alternatives $x, y \in A$, exactly one of the following holds: xTy (read: x *dominates* y), or yTx. We denote the set of tournaments over A by $\mathcal{T}(A)$.

A common visual way to represent tournaments is via graphs. A tournament $T \in \mathcal{T}(A)$ corresponds to a directed graph $G = (V, E)$, where $V = A$, and the directed edge from x to y is in E if and only if xTy. In other words, G is an orientation of the complete graph on A.

A second (uncommon) way to represent a tournament, which will prove very helpful in the sequel, is via a string of bits. Let $|A| = n$; a tournament T over A has $\binom{n}{2}$ edges, and each edge has two possible directions. Therefore, we can represent T by a string of $\binom{n}{2}$ bits. In order for this representation to be meaningful, we must specify the correspondence between edges and bits. For example, if $A = \{a, b, c\}$, we have three edges; in this case, we can specify that the leftmost, middle, and rightmost bits correspond to the edges (a, b), (a, c), and (b, c), respectively. Now, for every bit, we have to specify which edge direction corresponds to 0; we can do this, e.g., by fixing some tournament as the all-zeros string (000 in our example). This gives us a unique one-to-one and onto correspondence between tournaments and strings of bits. We use the three representations of tournaments (as a binary relation, a graph, and a string of bits) interchangeably.

Given the representation of tournaments as strings of bits, we can talk about the *Hamming distance* between $T \in \mathcal{T}(A)$ and $T' \in \mathcal{T}(A)$, that is, the number of edges that must be flipped in T in order to obtain T'. For two tournaments T and T' with Hamming distance one, we say that T and T' are (x, y)-*adjacent* if and only if the two tournaments disagree exactly on the edge (x, y), that is, $xTy \Leftrightarrow yT'x$, and for all edges $(z, w) \neq (x, y)$, $zTw \Leftrightarrow zT'w$.

Let $C \subseteq A$, and $T \in \mathcal{T}(A)$. We say that C is a *component* of T if for all $x, y \in C$ and $z \in A \setminus C$, $xTz \Leftrightarrow yTz$. Informally, an alternative outside the component either dominates *all* the alternatives in the component or is dominated by *all* the alternatives in the component.

We are interested in choice rules $r : \mathcal{T}(A) \to A$ that select a winning alternative given a tournament. We would like these rules to satisfy one of two basic desiderata. Alternative x is a *Condorcet winner* in $T \in \mathcal{T}(A)$ if xTy for all $y \in A \setminus \{x\}$; note that most tournaments do not have a Condorcet winner. A choice rule r is *Condorcet-consistent* if r always selects a Condorcet winner in a given tournament, if one exists. A choice rule r is *non-imposing* if it is onto A, i.e., for every $x \in A$ there exists $T \in \mathcal{T}(A)$ such that $r(T) = x$.

Clearly Condorcet-consistency implies non-imposition. Indeed, given any alternative $x \in A$, consider a tournament T where x is a Condorcet winner; then if r is Condorcet-consistent, $r(T) = x$.

3 Monotonic Choice Rules

We begin our analysis of incentives in tournaments with a crucial observation. Given a tournament T and $x, y \in A$ such that xTy, it is plausible to assume that x can lose on purpose to y. On the other hand, if yTx, x cannot unilaterally reverse this situation. In other words, we focus on manipulations where an alternative reverses an edge that is outgoing from itself in the tournament, but assume an alternative cannot reverse an edge that is incoming to itself.

In this section we briefly discuss a setting where an alternative only cares about whether it is selected by the choice rule r. In other words, a successful manipulation is one where an alternative is not elected by a rule, but is elected after reversing an outgoing edge. We wish to examine choice rules that are immune to this type of manipulation.

Definition 3.1. A choice rule $r : \mathcal{T}(A) \to A$ is *monotonic* if and only if for all $T \in \mathcal{T}(A)$, for all $x \in A$ such that $r(T) \neq x$, and for all $y \in A$ such that xTy, if T and T' are (x, y)-adjacent then $r(T') \neq x$.

This definition was formulated thus to make the connection to manipulation obvious. A more intuitive interpretation of monotonicity might be the following: r is monotonic if for all T where $r(T) = x$, and for all $y \in A$ such that yTx, if T and T' are (x, y)-adjacent then $r(T') = x$. However, it is easy to see that the two definitions are equivalent, hence we stand by the former definition.

The definition of monotonicity exists in the literature, albeit in the context of choice *sets* (which select a set of winning alternatives) rather than choice *rules* (see, e.g., [Laslier, 1997, page 38]). Nevertheless, to the best of our knowledge, it has never been interpreted in the context of manipulation by an alternative, but rather as an axiom of "social justice".

We also remark that the definition of monotonicity directly implies that an alternative cannot make itself a winner by reversing multiple edges instead of just a single edge.

Monotonicity is quite easy to achieve; below we give two examples of monotonic choice rules. A first example is known as the *Copeland* rule: simply select an alternative with maximum outdegree in the tournament, i.e., an alternative that dominates a maximum number of other alternatives. Since there may be multiple such alternatives, we need some tie-breaking rule, so choose the alternative with lexicographically smallest name. If an alternative loses on purpose it only decreases its outdegree, hence this choice rule is monotonic.

Another example of a monotonic choice rule is any voting tree where each alternative appears in the leaves at most once. Voting trees are sequential procedures for choosing from a tournament. A voting tree is given by a binary tree whose leaves are labeled by alternatives. When the tree is applied to a tournament T, in every stage two sibling leaves x and y compete according to T; the father of the two leaves is labeled by the winner (that is, by x if xTy and by y if yTx), and the two leaves are pruned. The label of the root of the tree is the selected alternative. If an alternative reverses an edge against itself, it can only lose in a competition that it could have won. However, if each alternative appears only once in the leaves, an alternative that is eliminated at some stage cannot be ultimately selected, hence the tree is monotonic as a choice rule. Further, notice that the monotonic rules that we mentioned are also Condorcet-consistent.

So, monotonicity is an easy desirable property to satisfy. We will return to monotonicity in the sequel, when we will require it in conjunction with other, less easily satisfied, properties.

4 Pairwise Nonmanipulable Choice Rules

We now turn to a second, natural type of manipulation. Consider a pair of alternatives with shared interests or goals. Given that neither of the alternatives is elected, the pair may conspire to make one of them a winner by flipping the edge between them. A pairwise manipulation is then given by a pair $\{x, y\}$ and a tournament T such that $r(T) \notin \{x, y\}$, but $r(T') \in \{x, y\}$, where T and T' are (x, y)-adjacent. Formally:

Definition 4.1. A choice rule $r : \mathcal{T}(A) \to A$ is *pairwise nonmanipulable* (PNM) if and only if for all $T \in \mathcal{T}(A)$ and all $x, y \in A$, if T and T' are (x, y)-adjacent then $r(T) \in \{x, y\} \Leftrightarrow r(T') \in \{x, y\}$.

PNM, which is introduced here for the first time, is somewhat reminiscent of a choice set property called *independence of non-winners*, or INW for short (see, e.g, [Laslier, 1997, page 38]). Under INW, a reversal of an edge that is incident upon a non-winning alternative cannot change the outcome. In the context of choice *rules* (which select a single winner), any INW rule must be constant. PNM is much weaker, as the outcome can change, but not in a way that benefits the manipulating pair.

Note that PNM is incomparable with monotonicity. Indeed, monotonicity clearly does not imply PNM. In the other direction, we construct a simple function that is PNM and not monotonic. Let $A = \{a, b\}$. We have that aTb and $bT'a$ in the two possible tournaments T and T'. Set $r(T) = b$ and $r(T') = a$; then r is vacuously PNM, but is not monotonic, since a can gain by moving from T to T', and b can gain by moving from T' to T. It is also easy to construct an example for more than two alternatives.

Our first result is an impossibility: we show that PNM and Condorcet-consistency are mutually exclusive.

Theorem 4.2. *Let A such that $|A| \geq 3$. Any PNM choice rule $r : \mathcal{T}(A) \to A$ is not Condorcet-consistent.*

Proof. Let $A = \{a, b, c, \ldots, \}$, and assume for contradiction that r is both PNM and Condorcet-consistent; we define a tournament $T \in \mathcal{T}(A)$ as follows. We let $C = \{a, b, c\}$ be a component in T, where CTx for all $x \in A \setminus C$. Moreover, inside C we have that aTb, bTc, and cTa. See Figure 1 for an illustration of T.

Now, there must be at least two alternatives (there might be three) from C that are different from $r(T)$; without loss of generality $r(T) \notin \{a, b\}$. Let T' be an adjacent tournament such that $bT'a$. Crucially, b is a Condorcet winner in T'.

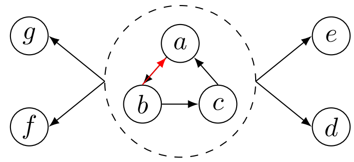

Figure 1: The tournament T constructed in the proof of Theorem 4.2, for $A = \{a, b, c, d, e, f, g\}$. The subset of alternatives $\{a, b, c\}$ is a component of T; these alternatives dominate all the alternatives outside the component.

Since r is Condorcet-consistent, it must hold that $r(T') = b$, in contradiction to PNM. $\square$

Theorem 4.2 can be seen as a strong impossibility result. Its implications are especially striking in the context of sports, where it common practice to use Condorcet-consistent solutions, e.g., select the Copeland winner (defined in Section 3). Hence, the theorem implies that sports tournaments are prone to simple manipulation by pairs of alternatives. Put another way, any "reasonable" selection of a winner from a sports tournament is susceptible to match fixing!

However, in social choice theory Condorcet-consistency is far from being universally accepted since, especially in voting-related interpretations of tournaments, it precludes other very basic properties (e.g., *participation*: voters might be hurt by participating in the election). It is therefore quite natural to ask whether we can obtain PNM by relaxing the notion of Condorcet-consistency. Indeed, in the sequel we ask whether there is a choice rule that is both PNM and non-imposing. Notice that if there is one alternative or two alternatives, there clearly exists such a rule. However, it turns out that there is no such rule when there are exactly three alternatives.

Theorem 4.3. *Let A such that $|A| = 3$. Any PNM choice rule $r : \mathcal{T}(A) \to A$ is not non-imposing.*

Proof. Let $A = \{a, b, c\}$, and assume for contradiction that r is PNM and non-imposing. We represent tournaments over A by a string of three bits, where the leftmost, middle, and rightmost bits represent the direction of the edges (a, b), (a, c), and (b, c), respectively. Fix an arbitrary tournament as 000.

Assume without loss of generality that $r(000) = a$. By PNM $r(001) = a$, otherwise the coalition $\{b, c\}$ gains from flipping the edge (b, c). We also have from PNM that $r(100) \in \{a, b\}$, since otherwise the coalition a, b can gain by switching from 100 to 000; finally, $r(010) \in \{a, c\}$, otherwise $\{a, c\}$ can gain by switching from 010 to 000. We differentiate two cases:

Case 1: $r(111) \in \{b, c\}$. By symmetry we can assume without loss of generality that $r(111) = b$. By PNM, $r(101) = b$, $r(110) \in \{b, c\}$, $r(011) \in \{a, b\}$. By non-imposition, we must have a tournament where c wins under r. By the above, we only have two options, 010 and 110. By PNM, since the two tournaments are (a, b)-adjacent, if the winner in one of these two tournaments is c, the winner in the other is c as well, hence it holds that $r(010) = c$, $r(110) = c$. From $r(110) = c$ and PNM it follows that $r(100) \in \{a, c\}$. We already know that $r(100) \in \{a, b\}$, hence $r(100) = a$. Now by PNM $r(101) = a$, in contradiction to our previous conclusion that $r(101) = b$.

Case 2: $r(111) = a$. By PNM we conclude that $r(110) = a$, $r(101) \in \{a, c\}$, and $r(011) \in \{a, b\}$. By non-imposition, we must have a tournament where b wins. We have only two options, 100 and 011. If $r(100) = b$, by PNM $r(110) = b$, in contradiction to $r(110) = a$. Symmetrically, if $r(011) = b$, we must have $r(001) = b$, in contradiction to $r(001) = a$. $\square$

One might expect the above impossibility result to also hold when $|A| > 3$. In other words, given Theorem 4.3, the intuition is that any choice rule $r : \mathcal{T}(A) \to A$ when $|A| > 3$ cannot be both PNM and non-imposing. Surprisingly, this turns out to be false. In fact, we establish that for any number of alternatives except three there is a choice rule that satisfies both properties. Moreover, the choice rule we construct is also monotonic. So, if one relaxes the requirement of Condorcet-consistency and only asks for non-imposition, and the number of alternatives is not three, then there are rules that are immune to both types of manipulation discussed above. This is, conceptually and technically, the main result of this paper.

Theorem 4.4. *Let A such that $|A| \neq 3$. There exists a choice rule $r : \mathcal{T}(A) \to A$ that is monotonic, PNM, and non-imposing.*

Before turning to the proof of the theorem, let us give some general intuitions. Our construction is inductive; we design a function that switches between the outcomes in $\{a, b, c\}$ based only on the direction of the edges (a, b) and (a, c). In order to achieve non-imposition, one of the four configurations of the two edges gives us outcomes in $A \setminus \{a, b, c\}$; in order to decide between alternatives in this set, we use the function for $|A| - 3$ alternatives whose existence is guaranteed by the induction assumption. This technique does not work for three alternatives as the set $A \setminus \{a, b, c\}$ is empty. The main obstacle is using the same technique to obtain a monotonic, PNM, non-imposing rule for six alternatives (despite the impossibility result for three); we are able to do this using a tailor-made construction for choosing between three alternatives when six alternatives are available.

Proof (sketch) of Theorem 4.4. We first prove the theorem for any $|A| = n \equiv 1 \pmod 3$. The proof is by induction. For $n = 1$, the claim holds trivially.

Now, let $n > 1$ such that $n \equiv 1 \pmod 3$, and denote $A = \{a, b, c, \ldots\}$. Let $A' = A \setminus \{a, b, c\}$. By the induction assumption, since $|A'| = n - 3 \equiv 1 \pmod 3$, there exists a function $r' : \mathcal{T}(A') \to A'$ that is monotonic, PNM, and onto A'. Given a tournament T, let $T \downarrow_{A'}$ be the restriction of T to A'. The construction of r is given by

$$r(T) = \begin{cases} a & aTb \text{ and } aTc \\ b & bTa \text{ and } aTc \\ c & aTb \text{ and } cTa \\ r'(T \downarrow_{A'}) & bTa \text{ and } cTa \end{cases} \tag{1}$$

In other words, the transitions between a, b, c, and A' depend only on the direction of the edges (a, b) and (a, c).

It is quite straightforward to see that r is onto A. Indeed, it is possible to obtain a, b and c as values of r by configuring the edges (a, b) and (a, c). Now, by the non-imposition of r', for any $x \in A'$, there is $T' \in \mathcal{T}(A')$ such that $r'(T') = x$. We then have that $r(T) = x$ for any tournament T where bTa, cTa, and $T \downarrow_{A'} = T'$.

We claim that r is monotonic. Let $T, T' \in \mathcal{T}(A)$ be two adjacent tournaments such that $r(T) \neq a$, $r(T') = a$. Then, by the construction of r, either bTa and $aT'b$, or cTa and $aT'c$, hence monotonicity is not violated. Similarly, let T and T' be two adjacent tournaments where $r(T) \neq b$ and $r(T') = b$, then either aTb and $bT'a$, or cTa and $aT'c$. In both cases, monotonicity is not violated. A symmetric argument holds for adjacent tournaments such that $r(T) \neq c$ and $r(T') = c$. Finally, let $x \in A'$, and consider two (x, y)-adjacent tournaments T and T'. Assume for contradiction that $r(T) \neq x$, $r(T') = x$. Since $x \notin \{a, b, c\}$ but the outcome changes as a result of flipping the edge (x, y), it follows from the construction of r that $y \in A'$ and the outcomes $r(T)$ and $r(T')$ are both determined by r'. Therefore, we have obtained a contradiction to the monotonicity of r'.

It is left to verify that r is PNM. This is easy to verify with respect to the transitions between a, b, and c. Consider two adjacent tournaments T and T' such that $r(T) = b$, $r(T') = x \in A'$. The two tournaments must differ in the edge (a, c). PNM is preserved since $b \notin \{a, c\}$ and $x \notin \{a, c\}$. A symmetric argument holds when $r(T) = c$, $r(T') \in A'$. Finally, if $r(T) = x \in A'$, $r(T') = x' \in A'$, then PNM holds by the PNM of r' using arguments that are similar to the ones given for monotonicity.

Let us now prove the theorem for any $n \equiv 2 \pmod 3$. The proof is almost identical, except for a small difference in the base of the induction. Indeed, we must prove that if $A = \{a, b\}$, there is a function $r : \mathcal{T}(A) \to A$ that is monotonic, PNM, and non-imposing. Let $r(T) = a$ if aTb and $r(T) = b$ if bTa. This function trivially satisfies all three properties.

It remains to prove the theorem for $n \equiv 0 \pmod 3$; this part of the proof is the most involved one. The trouble is that, by Theorem 4.3, there is no PNM and non-imposing choice rule when $n = 3$. Hence, the base of our induction is $n = 6$. In other words, we can use the induction as before given that we can explicitly construct a monotonic, PNM, and non-imposing choice rule $r : \mathcal{T}(A) \to A$ when $A = \{a, b, c, d, e, f\}$; this is the task we presently turn to.

Define r as in Equation (1), with one exception: if bTa and cTa then $r(T) = r'(T)$, where $r' : \mathcal{T}(A) \to \{d, e, f\}$ is defined in the sequel. Crucially, r' takes into account the entire tournament and not its restriction to $\{d, e, f\}$, but only returns outcomes in $\{d, e, f\}$. This is required in order to avoid the implications of Theorem 4.3.

We construct r' as follows. First, the value of r' depends only on the direction of the edges (a, d), (a, e), and (b, c). This choice of (b, c) as one of the edges is crucial, as it greatly increases our degrees of freedom in the construction of r'.

Let us represent the direction of the three critical edges by a string of three bits, where the left bit represents the direction of (a, d), the middle bit represents (a, e), and the right bit represents (b, c). Fix arbitrary directions as 000. In other words, although the input of r' has $\binom{6}{2} = 15$ edges (and so,

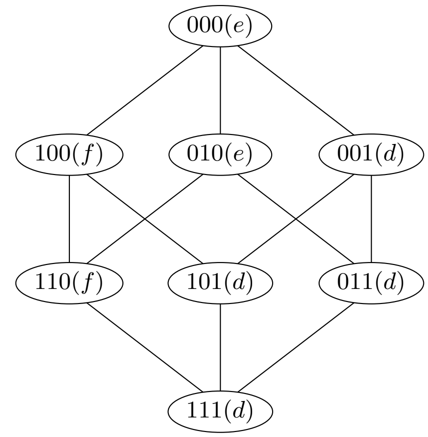

Figure 2: A visual representation of the construction of r' in the proof of Theorem 4.4. A vertex contains a string of three bits that represents the direction of the edges (a, d), (a, e), and (b, c). The letter in parentheses is the value of r' given the corresponding tournament. Every two adjacent tournaments are connected by an edge.

strictly speaking, should be represented by a string of 15 bits), r' disregards all but three, so for simplicity we can represent the input of r' as a string of three bits.

We define: $r'(000) = e$, $r'(100) = f$, $r'(010) = e$, $r'(001) = d$, $r'(110) = f$, $r'(101) = d$, $r'(011) = d$, $r'(111) = d$. See Figure 2 for an illustration. Clearly r' is non-imposing. Moreover, by checking all pairs of adjacent tournaments, one can verify that r' is also monotonic and PNM. Now, the fact that r itself is monotonic, PNM, and non-imposing is implied by essentially the same arguments as before (with minor changes that are required due to the fact that r' depends on (a, d), (a, e) and (b, c) rather than (d, e), (d, f) and (e, f)). □

Theorem 4.4 asserts that when the number of alternatives is not three, there is a choice rule that is monotonic, PNM, and non-imposing. In fact, by symmetry, the theorem implies that there are multiple rules that satisfy the three desiderata. These rules may differ in the additional properties that they satisfy. It is therefore interesting to ask whether it is possible to obtain a list of all choice rules satisfying the foregoing three properties, or, in other words, whether it is possible to obtain a list of all monotonic, PNM, and non-imposing choice rules, for a given number of alternatives.

The answer to this question is obviously positive when the number of alternatives is one or two, and when it is three there are no such choice rules (by Theorem 4.3). What about four alternatives? Notice that in this case, each tournament has $\binom{4}{2} = 6$ edges, therefore there are 64 possible tournaments, and the number of possible choice rules is already $4^{64} \approx 10^{39}$. It is therefore impossible (with current technology) to check all possible choice rules. However, observe that, using PNM, fixing the values of specific tournaments uniquely fixes the values of adjacent tournaments. Using this observation, we can reduce the number of choice rules to be checked to 2048 rules (we omit the nontrivial details due to lack of space). By enumerating these choice rules, we have found that when there are four alternatives, the number of monotonic, PNM, and non-imposing choice rules is 146. It seems possible to apply the same ideas to enumerate all monotonic, PNM, and non-imposing choice rules when there are five alternatives. When $|A| > 5$, new combinatorial techniques are required.

5 Discussion

We have succeeded in obtaining a surprising positive result (Theorem 4.4) with respect to two types of manipulations: an individual alternative trying to make itself a winner (monotonicity), and a pair of alternatives conspiring to make one of them a winner (PNM—pairwise nonmanipulability).

The reader might be concerned, however, that the voting rule constructed in Theorem 4.4 does not satisfy additional social choice desiderata, and hence might be "unreasonable" as a choice rule. Indeed, now that we know that monotonic, PNM, non-imposing rules exist, it is worthwhile to explore the existence of rules that satisfy additional properties. Note that, given a specific additional property, this can be done directly by enumeration with respect to a small number of alternatives.

Another (related) concern might be regarding different types of manipulations. For instance, one can think of a type of pairwise "external" manipulation where x loses on purpose to y in order to get z to win. More generally, it is possible to think of a scenario where each alternative has a ranking of the alternatives, and is trying to obtain a more preferred outcome according to its ranking. Intuitively, immunity to these types of manipulations is a very strong property, hence it is not surprising that, under both definitions, any nonmanipulable choice rule must be constant (the proof is by induction on the Hamming distance from some initial tournament, and is omitted due to lack of space).

On the other hand, the definition of PNM can be strengthened by requiring immunity to manipulation by triplets. In other words, it is possible to look at coalitions of size three that reverse the edges between them in order to get one of them elected. Is there a choice rule that is non-imposing and immune to manipulations by triplets? If there exists such a choice rule, is there one that is immune to similar manipulations by larger coalitions? We leave these questions as intriguing directions for future research.

Acknowledgments

The authors thank Felix Fischer for very helpful comments on a draft of this paper.

References

[Altman and Tennenholtz, 2008] A. Altman and M. Tennenholtz. Strategyproof deterministic lotteries under broadcast communication. In *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 1549–1552, 2008.

[Bouyssou *et al.*, 2006] D. Bouyssou, T. Marchant, M. Pirlot, A. Tsoukiàs, and P. Vincke. *Evaluation and Decision Models: Stepping Stones for the Analyst*. Springer-Verlag, 2006.

[Brandt *et al.*, 2008] F. Brandt, F. Fischer, P. Harrenstein, and M. Mair. A computational analysis of the Tournament Equilibrium Set. In *Proceedings of the 23rd AAAI Conference on Artificial Intelligence (AAAI)*, pages 38–43, 2008.

[Conitzer *et al.*, 2007] V. Conitzer, T. Sandholm, and J. Lang. When are elections with few candidates hard to manipulate? *Journal of the ACM*, 54(3):1–33, 2007.

[de Condorcet, 1785] Marquis de Condorcet. Essai sur l'application de l'analyse à la probabilité de décisions rendues à la pluralité de voix. Imprimerie Royal, 1785. Facsimile published in 1972 by Chelsea Publishing Company, New York.

[Dung, 1995] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77:321–357, 1995.

[Dunne, 2007] P. E. Dunne. Computational properties of argumentation systems satisfying graph-theoretic constraints. *Artificial Intelligence*, 171(10–15):701–729, 2007.

[Dutta *et al.*, 2002] B. Dutta, M. O. Jackson, and M. Le Breton. Voting by successive elimination and strategic candidacy. *Journal of Economic Theory*, 103:190–218, 2002.

[Gibbard, 1973] A. Gibbard. Manipulation of voting schemes. *Econometrica*, 41:587–602, 1973.

[Hemaspaandra *et al.*, 2007] E. Hemaspaandra, L. A. Hemaspaandra, and J. Rothe. Anyone but him: The complexity of precluding an alternative. *Artificial Intelligence*, 171(5–6):255–285, 2007.

[Laslier, 1997] J.-F. Laslier. *Tournament Solutions and Majority Voting*. Springer, 1997.

[McGarvey, 1953] D. C. McGarvey. A theorem on the construction of voting paradoxes. *Econometrica*, 21:608–610, 1953.

[Procaccia and Rosenschein, 2005] A. D. Procaccia and J. S. Rosenschein. Extensive-form argumentation games. In *Proceedings of the 3rd European Workshop on Multi-Agent Systems (EUMAS)*, pages 312–322, 2005.

[Procaccia and Rosenschein, 2007] A. D. Procaccia and J. S. Rosenschein. Junta distributions and the average-case complexity of manipulating elections. *Journal of Artificial Intelligence Research*, 28:157–181, 2007.

[Rahwan and Larson, 2008] I. Rahwan and K. Larson. Mechanism design for abstract argumentation. In *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 1031–1038, 2008.

[Satterthwaite, 1975] M. Satterthwaite. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10:187–217, 1975.