Voting Almost Maximizes Social Welfare Despite Limited Communication

Ioannis Caragiannis
RACTI & Department of Computer
Engineering and Informatics
University of Patras, Greece
caragian@ceid.upatras.gr

Ariel D. Procaccia
School of Engineering and Applied Sciences
Harvard University, USA
arielpro@seas.harvard.edu

Abstract

In cooperative multiagent systems an alternative that maximizes the *social welfare*—the sum of utilities—can only be selected if each agent reports its full utility function. This may be infeasible in environments where communication is restricted. Employing a voting rule to choose an alternative greatly reduces the communication burden, but leads to a possible gap between the social welfare of the optimal alternative and the social welfare of the one that is ultimately elected. Procaccia and Rosenschein (2006) have introduced the concept of *distortion* to quantify this gap.

In this paper, we present the notion of *embeddings into voting rules*: functions that receive an agent's utility function and return the agent's vote. We establish that very low distortion can be obtained using randomized embeddings, especially when the number of agents is large compared to the number of alternatives. We investigate our ideas in the context of three prominent voting rules with low communication costs: Plurality, Approval, and Veto. Our results arguably provide a compelling reason for employing voting in cooperative multiagent systems.

Introduction

A major challenge that arises in the design and implementation of multiagent systems is the aggregation of the preferences of the agents. Voting theory provides a neat solution by giving extremely well-studied methods of preference aggregation. In recent years the theoretical aspects of computational voting have been enthusiastically investigated, especially within the AI community.

While the appeal of voting in the context of heterogeneous, competitive multiagent systems is apparent, some multiagent systems are *centrally designed* and *fully cooperative* (e.g., systems for planning and scheduling, recommender systems, collaborative filtering, and so on). We believe that, to date, the benefit of employing voting in such domains was unclear. Indeed, agents are normally assumed to compute a utility for every possible alternative. If the agents are cooperative then they can simply communicate their utilities for the different alternatives, and subsequently select an alternative that maximizes the *social welfare*, i.e., the sum of utilities.

However, accurately conveying an agent's utility function for each alternative may be very costly in terms of communication. This could prove to be a serious obstacle in domains where communication is restricted. Communication may be limited by the physical properties of the system (e.g., slow or error-prone transmitters) or the representation of the full utility functions may require a huge amount of information. Moreover, "reducing the communication requirements lessens the burden placed on the agents ... [when they] need to invest effort (such as computation or information gathering) to determine their preferences" (Conitzer & Sandholm 2005, page 78). Blumrosen et al. (2007) outline additional persuasive reasons why communication should be restricted in multiagent settings. Fortunately, some prominent voting rules—functions that select an alternative given the preferences of the agents—impose a very small communication burden (Conitzer & Sandholm 2005), and are moreover resistant to errors in communication (Procaccia, Rosenschein, & Kaminka 2007).

For example, consider the paradigmatic cooperative multiagent system domain: scanning an area on Mars with multiple rovers (which are known to have limited communication capabilities). Suppose the rovers must select or update their joint plan (this may happen very often), and there are one million alternatives. Moreover, suppose each rover computes a utility for each alternative on a scale of one to one million (this is, in fact, a very coarse scale). A rover would need to communicate $10^6 \cdot \log(10^6) \approx 20 \cdot 10^6$ bits in order to report its utility function. In contrast, under the *Plurality* voting rule, where each agent votes for a single alternative and the alternative with most votes wins, a rover must transmit only twenty bits.

In this paper we shall argue that, in cooperative multiagent systems, exact maximization of the social welfare can be replaced by very simple voting rules (given an extra ingredient that we present below). The benefit is a huge reduction in the communication burden, whereas the cost, a deterioration in the social welfare of the outcome, will be shown to be almost negligible in some settings. This arguably provides a pivotal reason for employing voting in cooperative multiagent systems, and in AI in general.

Our approach. The degree to which the social welfare of the outcome can decrease when voting is used is captured by the notion of *distortion*, introduced by Procaccia and Rosenschein (2006). They focus on voting rules that receive as input a ranking of the alternatives, and, crucially, assume that each agent reports a ranking such that the alternative that is ranked in the k'th place has the k'th highest utility. Under this assumption, they define the distortion of a voting rule to be the worst-case ratio between the maximum social welfare over all the alternatives, and the social welfare of the winner of the election; the worst-case is taken over all the possible utility functions of the agents. After proving some impossibility results, Procaccia and Rosenschein further restrict the structure of the utility functions. Even under this additional (very strong) assumption, they show that the distortion of most prominent voting rules is linear in the number of alternatives. The approach of Procaccia and Rosenschein is *descriptive*: they propose to use the notion of distortion as a criterion in the comparison of different voting rules.

Our main conceptual contribution is the consideration of *embeddings into voting rules*. An embedding is a set of instructions that informs each agent how to vote, based only on the agent's own utility function, that is, without any communication or coordination between different agents. More accurately, an embedding into a specific voting rule is a function from utility functions to votes that are valid under the voting rule. For instance, consider the simple Plurality rule described above. Given a utility function, an embedding into Plurality returns the alternative that the agent votes for. Procaccia and Rosenschein implicitly use one specific embedding, but many different embeddings exist. In this sense, our approach is *algorithmic*: we wish to *design* embeddings in a way that minimizes the distortion.

We redefine the notion of distortion to take embeddings into account. The *distortion of an embedding into a voting rule* is still the worst-case ratio between the maximum social welfare and the social welfare of the winner, but now the winner depends both on the voting rule and on the embedding, that is, on the way the utilities of the agents are translated into votes. The worst-case is taken over all possible utilities; we do not make any assumption regarding the utilities, except that they are normalized.

We take the idea of embeddings into voting rules one step further by allowing *randomized embeddings*. A randomized embedding randomly chooses the agent's vote, according to some probability distribution. The distortion is defined similarly, by taking into account the *expected* social welfare of the winner of the election. As we shall see, randomization gives us great power and flexibility, and ultimately provides us with the tools to design truly low-distortion embeddings.

We wish to design low distortion embeddings into voting rules with low communication complexity. Indeed, given that each of our cooperative agents votes according to the instructions provided by the embedding (in a fully *decentralized* way), then an alternative with social welfare close to optimal may be elected in the face of restricted communication. We find the existence of low distortion embeddings rather striking, as the social welfare is a *centralized* concept.

Our results. We study the distortion of embeddings into three voting rules: Plurality, Approval (each agent approves a subset of alternatives), and Veto (each agent gives a "negative point" to one alternative). Plurality and Veto have the smallest communication burden among all prominent voting rules: only $\log m$ bits per agent, where m is the number of alternatives. Approval requires more communication, m bits per agent, but still less than other prominent voting rules.

We first deal with the Plurality rule. We show that any deterministic embedding into Plurality has distortion $\Omega(m^2)$, and also provide a matching upper bound. Our main result deals with randomized embeddings into Plurality: we show that the naïve embedding into Plurality, which selects an alternative with probability proportional to its utility, yields constant distortion when $n = \Omega(m \ln m)$, where n is the number of agents, and has extremely low distortion, specifically $1 + o(1)$, for larger values of n.

Next we investigate the Approval rule. We give a lower bound of $\Omega(m)$ for deterministic embeddings, and also present a matching upper bound. Our randomized upper bounds for Approval follow directly from the upper bounds for Plurality, since any embedding into Plurality is also an embedding into Approval.

Finally, we consider the Veto rule. We show that any deterministic embedding into Veto has infinite distortion, and the same is true for randomized embeddings if $n < m - 1$. We further show that low-distortion embeddings into Veto can be obtained, albeit using a large number of agents.

Embeddings into Plurality

We denote by $N = \{1, \ldots, n\}$ the set of *agents*, and by A, $|A| = m$, the set of *alternatives*.

We assume that the agents have normalized cardinal utilities over A. Specifically, let $\mathcal{U} = \mathcal{U}(A)$ be the set of utility functions u over A such that for each $x \in A$, $u(x) \geq 0$, and $\sum_{x \in A} u(x) = 1$. Each agent i has a utility function $u \in \mathcal{U}$. A *utility profile* is a vector of utility functions

$$\mathbf{u} = \langle u_1, \ldots, u_n \rangle \in \mathcal{U}^n \ .$$

The *social welfare* of an alternative $x \in A$ with respect to $\mathbf{u} \in \mathcal{U}^n$, denoted $\mathrm{sw}(x, \mathbf{u})$, is the sum of the utilities of x for all agents:

$$\mathrm{sw}(x, \mathbf{u}) = \sum_{i \in N} u_i(x) \ .$$

In our formal presentation, a voting rule is defined as a function that selects a *set* of alternatives rather than a single alternative. Such a function is formally known as a *voting correspondence*, hence the term *voting rule* is slightly abused. We must deal with sets of winners since our rules are based on notions of score, and there might be a tie with respect to the maximum score.

Under the *Plurality* rule, each agent casts its vote in favor of a single alternative. The set of winners is the set of alternatives with a maximum number of votes.

A *deterministic embedding into Plurality* is a function $f : \mathcal{U} \to A$. Informally, given an agent $i \in N$ with a utility function $u \in \mathcal{U}$, $f(u)$ is the alternative which agent i votes for under the embedding f. Given a utility profile $\mathbf{u} \in \mathcal{U}^n$ and an embedding f, denote the (Plurality) *score* of $x \in A$ by

$$\mathrm{sc}(x, f, \mathbf{u}) = |\{i \in N \ : \ f(u_i) = x\}| \ ,$$

and denote the set of *winners* by

$$\text{win}(f, \mathbf{u}) = \text{argmax}_{x\in A}\text{sc}(x, f, \mathbf{u}) \ .$$

Note that the argmax function returns a *set* of maximal alternatives.

A *randomized embedding* randomly selects one of the alternatives, that is, it is a function $f : \mathcal{U} \to \Delta(A)$, where $\Delta(A)$ is the space of probability distributions over A. Put another way, given $u \in \mathcal{U}$, $f(u)$ is a random variable that takes the value $x \in A$ with probability $p(x)$, i.e., $\sum_{x\in A} p(x) = 1$. With respect to a randomized embedding f, $\text{sc}(x, f, \mathbf{u})$ is a random variable that takes values in $\{1, \ldots, n\}$, and $\text{win}(f, \mathbf{u})$ is a random variable that takes values in 2^A, the powerset of A. Less formally, given a randomized embedding f, a utility profile $\mathbf{u}$, and $S \subseteq A$, we have some probability (possibly zero) of S being the set of winners when f is applied to $\mathbf{u}$.

As a measure of the quality of an embedding, we use the notion of distortion, introduced by Procaccia and Rosenschein (2006), but adapt it to apply to general embeddings.

Definition 1 (Distortion)**.**

1. Let $f : \mathcal{U} \to A$ be a deterministic embedding, $\mathbf{u} \in \mathcal{U}^n$. The *distortion of f at* $\mathbf{u}$ is

$$\text{dist}(f, \mathbf{u}) = \frac{\max_{y\in A} \text{sw}(y, \mathbf{u})}{\min_{x\in\text{win}(f,\mathbf{u})} \text{sw}(x, \mathbf{u})} \ .$$

2. Let $f : \mathcal{U} \to \Delta(A)$ be a randomized embedding, $\mathbf{u} \in \mathcal{U}^n$. The *distortion of f at* $\mathbf{u}$ is

$$\text{dist}(f, \mathbf{u}) = \frac{\max_{y\in A} \text{sw}(y, \mathbf{u})}{\mathbb{E}\left[\min_{x\in\text{win}(f,\mathbf{u})} \text{sw}(x, \mathbf{u})\right]} \ .$$

3. Let f be a deterministic or randomized embedding. The *distortion of f* is

$$\text{dist}(f) = \max_{\mathbf{u}\in\mathcal{U}^n} \text{dist}(f, \mathbf{u}) \ .$$

Let us give an intuitive interpretation of this important definition. The distortion of a deterministic embedding is the worst-case ratio between the social welfare of the most popular alternative, and the social welfare of the least popular winner, where the worst-case is with respect to all possible utility profiles. In other words, we are interested in the question: how small can the social welfare of one of the winners be, when compared to the alternative with maximum social welfare?

Our focus on the social welfare of the "worst" winner is appropriate since the analysis is worst-case. Alternatively, it is possible to think of voting rules that elect only one of the alternatives with maximum score, but in the worst-case the most unpopular one is elected, that is, in the worst-case ties are broken in favor of alternatives with lower social welfare; this is the interpretation of Procaccia and Rosenschein (2006).

The definition of distortion with respect to randomized embeddings is slightly more subtle. Here there is no definite winner. However, given a utility profile $\mathbf{u} \in \mathcal{U}$, we can talk about the expected minimum social welfare among the winners, since the set of winners is simply a random variable that takes values in 2^A, hence $\min_{x\in\text{win}(f,\mathbf{u})} \text{sw}(x, \mathbf{u})$ is a random variable that takes values in the interval $[0, n]$ and its expectation is well-defined. The rest of the definition is identical to the deterministic case.

Deterministic Embeddings

Procaccia and Rosenschein (2006) consider a specific, naïve deterministic embedding into Plurality. Their embedding simply maps a utility function $u \in \mathcal{U}$ to an alternative with maximum utility, that is, $f(u) \in \text{argmax}_{x\in A} u(x)$. They show that its distortion is $m - 1$ under a very restricted definition of distortion (called *misrepresentation*) that assumes a specific structure of utility functions.

It is easy to see that the distortion of this naïve embedding, according to Definition 1, is at most m^2. Indeed, let $\mathbf{u} \in \mathcal{U}$, and let $x \in \text{win}(f, \mathbf{u})$, where f is the naïve embedding. By the Pigeonhole Principle, it must hold that $\text{sc}(x, f, \mathbf{u}) \geq n/m$. Now, for each agent $i \in N$ such that $f(u_i) = x$, it must hold that $u_i(x) \geq 1/m$, since x has maximum utility and there must exist an alternative with utility $1/m$ (again, by the Pigeonhole principle). We deduce that $\text{sw}(x, \mathbf{u}) \geq n/m^2$. On the other hand, for any $y \in A$, $\text{sw}(y, \mathbf{u}) \leq n$. Therefore,

$$\text{dist}(f) = \max_{\mathbf{u}\in\mathcal{U}^n} \frac{\max_{y\in A} \text{sw}(y, \mathbf{u})}{\min_{x\in\text{win}(f,\mathbf{u})} \text{sw}(x, \mathbf{u})} \leq \frac{n}{n/m^2} = m^2 \ .$$

We wish to ask whether there is a clever deterministic embedding into Plurality that beats the m^2 upper bound given by the naïve one. Our first theorem answers this question in the negative.

Theorem 2. *Let $|A| = m \geq 3$, $|N| = n \geq \left\lceil \frac{m+1}{2} \right\rceil$, and let $f : \mathcal{U} \to A$ be a deterministic embedding into Plurality. Then dist$(f) = \Omega(m^2)$.*

Proof. Let f be a deterministic embedding into Plurality. For every pair of distinct alternatives $x, y \in A$, let $u^{xy} \in \mathcal{U}$ such that $u^{xy}(x) = 1/2$, $u^{xy}(y) = 1/2$, and $u(z) = 0$ for every $z \in A \setminus \{x, y\}$. We claim that we can assume that $f(u^{xy}) \in \{x, y\}$, since otherwise the distortion is infinite. Indeed, if $f(u^{xy}) = z \notin \{x, y\}$, then consider a utility profiles $\mathbf{u}$ where $u_i \equiv u^{xy}$ for all $i \in N$. Then $\text{win}(f, \mathbf{u}) = \{z\}$, but $\text{sw}(z, \mathbf{u}) = 0$, whereas, say, $\text{sw}(x, \mathbf{u}) = n/2 > 0$.

Let T be a tournament on A, that is, a complete asymmetric binary relation (see, e.g., (Moon 1968)). For every two alternatives $x, y \in A$, we have that xTy (read: x dominates y) if $f(u^{xy}) = x$, and yTx if $f(u^{xy}) = y$. By our claim above, T is well-defined.

Since the number of pairs of alternatives is $\binom{m}{2} = \frac{m(m-1)}{2}$, by the Pigeonhole Principle there must be an alternative that is dominated by at least $\frac{m-1}{2}$ other alternatives; without loss of generality this alternative is $a \in A$. Let A' be a subset of alternatives of size $\left\lceil \frac{m-1}{2} \right\rceil$ such that for all $x \in A'$, xTa. Further, let $A'' = A \setminus (A' \cup \{a\})$ and notice that $|A''| = \left\lfloor \frac{m-1}{2} \right\rfloor$. Define a utility function $u^* \in \mathcal{U}$ as follows: for every $x \in A''$, $u^*(x) = 1/|A''|$; for every $x \in A \setminus A''$, $u^*(x) = 0$. Then without loss of generality

$f(u^*) = b$, with $b \in A''$, otherwise the distortion is infinite by the same reasoning as above.

Now, we have $A' + 1$ blocks of agents, each of size either $\lceil n/(|A'|+1) \rceil$ or $\lfloor n/(|A'|+1) \rfloor$. All the agents in the first block, which is at least as large as any other, have the utility function u^* (therefore they vote for b). For each $x \in A'$, there is a block of agents with the utility function u^{ax} (hence they vote for x). Given this utility profile $\mathbf{u}$, b must be among the winners, that is, $b \in \text{win}(f, \mathbf{u})$. We have that

$$\text{sw}(b, \mathbf{u}) \leq \left\lceil \frac{n}{\lceil \frac{m-1}{2} \rceil + 1} \right\rceil \cdot \frac{1}{\lfloor \frac{m-1}{2} \rfloor} \leq \frac{8n}{m^2} \ ,$$

whereas

$$\text{sw}(a, \mathbf{u}) \geq \left(n - \left\lceil \frac{n}{\lceil \frac{m-1}{2} \rceil + 1} \right\rceil \right) \cdot \frac{1}{2} \geq \frac{n}{6} \ .$$

The distortion is at least as large as the ratio between the maximum social welfare and the social welfare of a winner with respect to the specific utility profile $\mathbf{u}$, that is,

$$\text{dist}(f) \geq \text{dist}(f, \mathbf{u}) \geq \frac{\text{sw}(a, \mathbf{u}^*)}{\text{sw}(b, \mathbf{u}^*)} = \Omega(m^2) \ .$$

□

Randomized Embeddings

Theorem 2 implies that the distortion of any deterministic embedding into Plurality is quite high. Can we do better using randomized embeddings? In general, the answer is definitely positive. However, we start our investigation of randomized embeddings into Plurality with a negative result that holds when the number of agents is very small.

Theorem 3. *Let $|N| = \{1, 2\}$ and $|A| = m$. Then any randomized embedding $f : \mathcal{U} \to \Delta(A)$ into Plurality has distortion $\Omega(m)$.*

The idea of the proof (which appears in the appendix) is the design of a utility profile where agent 1 has equal utility for all alternatives, whereas agent 2 has utility 1 for one special alternative. Although the special alternative has high social welfare, it is the unique winner with small probability.

We presently turn to our presentation of low-distortion embeddings. It turns out that when the number of agents is at least as large as the number of alternatives, a huge reduction in the distortion can be achieved using randomized embeddings. If the number of agents is significantly larger, the distortion can be very close to one. Indeed, consider the following embedding.

Embedding 1 (Naïve randomized embedding into Plurality)**.** Given a utility function $u \in \mathcal{U}$, select alternative $x \in A$ with probability $u(x)$.

The following powerful theorem is our main result.

Theorem 4. *Let $|N| = n \geq m = |A|$, and denote Embedding 1 by f. Then:*

1. *dist*$(f) = \mathcal{O}\left(m^{2m/n}\right)$.
2. *dist*$(f) = \mathcal{O}\left(\sqrt{m}\right)$.
3. *Let $n \geq 3$ and $\epsilon(n, m) = 4\sqrt{\frac{m \ln n}{n}}$. If $\epsilon(n, m) < 1$, then dist$(f) \leq \frac{1}{1-\epsilon(n,m)}$.*

All three bounds on the distortion are required, since each has values of n and m where its guarantees are stronger than the others. Asymptotically, the most powerful bound is the one given in Part 1: it guarantees that the distortion of Embedding 1 is already constant when $n = \Theta(m \ln m)$, that is, when the number of agents is slightly larger than the number of alternatives. Part 1 yields a very weak result for the case $n = m$; in this case, we get by Part 2 that the distortion is $\mathcal{O}\left(\sqrt{m}\right)$. Finally, for large values of n we do not find it sufficient to show that the distortion is *constant*, we want to establish that it is almost one. This does not follow from Part 1 due to the constant hidden in the $\mathcal{O}$ notation. However, from Part 3 we get that for, e.g., $n \geq m^2$, the distortion is $1 + o(1)$.

We presently prove Part 1 of the theorem, which is the easiest one. Part 2 is similar but slightly more involved. The proof of Part 3 is quite different and significantly more complicated. We relegate the proofs of Parts 2 and 3 to the appendix. Before proving Part 1, we require two results regarding the sums of random variables (many more such results are required for Part 3).

Lemma 5. *Let $X_1, ..., X_n$ be independent heterogeneous Bernoulli trials. Denote by μ the expectation of the random variable $X = \sum_i X_i$. Then:*

1. *(Jogdeo & Samuels 1968)* $\Pr[X < \lfloor \mu \rfloor] < 1/2$.
2. *(Chernoff 1952)* $\Pr\left[X \geq (1+\delta)\mu\right] \leq \left(\frac{e}{1+\delta}\right)^{(1+\delta)\mu}$.

Proof of Part 1 of Theorem 4. Let $\mathbf{u} \in \mathcal{U}^n$ be a utility profile. First notice that the expected Plurality score of $x \in A$ under the embedding f is $\text{sw}(x, \mathbf{u})$. Let $x^* \in \text{argmax}_{x \in A} \text{sw}(x, \mathbf{u})$ be an alternative with maximum social welfare. We have that $\sum_{x \in A} \text{sw}(x, \mathbf{u}) = n$; by the assumption that $n \geq m$, it follows that $\text{sw}(x^*, \mathbf{u}) \geq n/m \geq 1$. By Part 1 of Lemma 5, with probability at least $1/2$ it holds that $\text{sc}(x^*, f, \mathbf{u}) \geq \lfloor \text{sw}(x^*, \mathbf{u}) \rfloor$.

Consider some alternative $x \in A$ such that

$$\text{sw}(x, \mathbf{u}) < \frac{\text{sw}(x^*, \mathbf{u})}{2e(4m)^{2m/n}} \ . \qquad (1)$$

We apply the upper tail Chernoff bound (Part 2 of Lemma 5) to the random variable $\text{sc}(x, f, \mathbf{u})$ with expectation $\mu = \text{sw}(x, \mathbf{u})$ using $(1+\delta)\mu = \lfloor \text{sw}(x^*, \mathbf{u}) \rfloor$. By (1) and since $\lfloor \text{sw}(x^*, \mathbf{u}) \rfloor > \text{sw}(x^*, \mathbf{u})/2$, we also have $1 + \delta > e(4m)^{2m/n}$. Therefore,

$$\begin{aligned}
\Pr[\text{sc}(x, f, \mathbf{u}) \geq \lfloor sw(x^*, \mathbf{u}) \rfloor] &\leq \left(\frac{1}{(4m)^{2m/n}}\right)^{\lfloor \text{sw}(x^*, \mathbf{u}) \rfloor} \\
&\leq \left(\frac{1}{(4m)^{2m/n}}\right)^{\text{sw}(x^*, \mathbf{u})/2} \\
&\leq \frac{1}{4m} \ ,
\end{aligned}$$

where the last inequality follows since $\text{sw}(x^*, \mathbf{u}) \geq n/m$.

By the union bound, the probability that either $\mathrm{sc}(x^*, f, \mathbf{u}) < \lfloor \mathrm{sw}(x^*, \mathbf{u}) \rfloor$, or some alternative x that satisfies (1) has $\mathrm{sc}(x, f, \mathbf{u}) \geq \lfloor \mathrm{sw}(x^*, \mathbf{u}) \rfloor$, is at most $3/4$. Therefore, with probability $1/4$ all the winners have social welfare at least $\mathrm{sw}(x^*, \mathbf{u})/(2e(4m)^{2m/n})$. Hence

$$\mathrm{dist}(f, \mathbf{u}) \leq \frac{\mathrm{sw}(x^*, \mathbf{u})}{\frac{1}{4} \cdot \frac{\mathrm{sw}(x^*, \mathbf{u})}{2e(4m)^{2m/n}}} = 8e \cdot (4m)^{2m/n} \ .$$

Since $n \geq m$, we have that $4^{2m/n} \leq 16$. It follows that the distortion of f is as announced. □

Our final result regarding embeddings into Plurality asserts that the upper bound of $\mathcal{O}\left(\sqrt{m}\right)$ for the case of $n = m$, which follows from Part 2 of Theorem 4, is almost tight. This case is especially interesting since for slightly larger values of n the distortion is constant.

Theorem 6. *Let $|N| = n = m = |A|$, and denote Embedding 1 by f. Then dist$(f) = \Omega\left(\sqrt{\frac{m}{\ln m}}\right)$.*

The proof of the theorem appears in the appendix. The main idea is the construction of a utility profile with "heavy" (high social welfare) alternatives and "light" (low social welfare) alternatives; the heavy alternatives have social welfare and score of exactly two, whereas the light alternatives have low social welfare and expected score. However, since there are many light alternatives, with high probability at least one such alternative has a score of two.

Embeddings into Approval

Under the Approval rule, each agent approves a subset of the alternatives. Each approved alternative receives one point. The set of winners includes the alternatives with most points, summed over all the agents.

We must reformulate some of our definitions in order to apply our notions to Approval voting. A deterministic embedding into Approval is a function $f : \mathcal{U} \rightarrow 2^A$, where 2^A is the powerset of alternatives. In words, an agent with a utility function u approves each of the alternatives in $f(u)$. The (Approval) score of an alternative is redefined to be

$$\mathrm{sc}(x, f, \mathbf{u}) = |\{i \in N \ : \ x \in f(u_i)\}| \ .$$

A randomized embedding is a function $f : \mathcal{U} \rightarrow \Delta(2^A)$. The rest of the definitions (in particular, the definition of distortion) are the same as before.

Deterministic Embeddings

In Section we have seen that no deterministic embedding into Plurality can achieve distortion better than $\Omega(m^2)$ (Theorem 2). As it turns out, better results can be achieved with respect to Approval. Indeed, consider the following Embedding.

Embedding 2 (Deterministic embedding into Approval)**.** Given a utility function u, approve the subset of alternatives $x \in A$ such that $u(x) \geq 1/m$.

The following straightforward result (whose proof appears in the appendix) establishes that the distortion of this embedding is $\mathcal{O}(m)$.

Theorem 7. *Let $|N| = n$, $|A| = m$, and denote Embedding 2 by f. Then dist$(f) \leq 2m - 1$.*

Unfortunately, it is impossible to design low-distortion deterministic embeddings into Approval. In fact, the following theorem asserts that the simple Embedding 2 is asymptotically optimal.

Theorem 8. *Let $|N| = n \geq 2$, $|A| = m \geq 3$, and let $f : \mathcal{U} \rightarrow 2^A$ be a deterministic embedding into Approval. Then dist$(f) \geq (m-1)/2$.*

The proof of the theorem is given in the appendix of the paper.

Randomized Embeddings

In the context of deterministic embeddings, we have seen that there is a gap between the distortion of embeddings into Approval and embeddings into Plurality. It turns out that there is also a huge gap with respect to randomized embeddings, when the number of agents is very small.

Indeed, consider the randomized embedding f into Approval that, with probability 1/2, approves an alternative with maximum utility, and with probability 1/2 approves all the alternatives. Further, assume that $N = \{1, 2\}$, and let $\mathbf{u} \in \mathcal{U}^2$. Without loss of generality there exists $x^* \in A$ such that $u_1(x^*) \geq u_i(x)$ for all $x \in A$ and all $i \in N$. Then clearly, for every $x \in A$, $\mathrm{sw}(x, \mathbf{u}) \leq 2 \cdot \mathrm{sw}(x^*, \mathbf{u})$. Moreover, it holds that $\mathrm{win}(f, \mathbf{u}) = \{x^*\}$ with probability at least 1/4. Hence, the distortion of this embedding is at most eight, i.e., constant. This reasoning can easily be extended to obtain constant distortion with respect to any constant n. Compare this result with Theorem 3.

However, as before, we are mostly interested in the case of a large number of agents. Crucially, every embedding into Plurality can also be seen as an embedding into Approval, where for every utility function exactly one alternative is approved. Hence, the powerful positive result regarding Embedding 1, namely Theorem 4, also holds with respect to Approval. It remains open whether there is a gap in the distortion of randomized embeddings into Plurality and randomized embeddings into Approval when $n \geq m$. Interestingly enough, the lower bound of $\Omega(\sqrt{m/\ln m})$ for $n = m$ (Theorem 6) also holds with respect to some natural embeddings into Approval, which may approve multiple alternatives.

Embeddings into Veto

Under the Veto rule, each agent vetoes a single (presumably least preferred) alternative. The set of winners includes all the alternatives that are vetoed the least number of times. Equivalently, each agent awards one point to all the alternatives except one, and the alternatives with most points are the winners.

Let us quickly reformulate our definitions for Veto. A deterministic embedding into Veto is a function $f : \mathcal{U} \rightarrow A$, as with Plurality; however, now $f(u)$ is interpreted as the alternative vetoed by an agent with utility function u. The (Veto) score of an alternative is

$$\mathrm{sc}(x, f, \mathbf{u}) = |\{i \in N \ : \ x \neq f(u_i)\}| \ ,$$

that is, the number of agents that do not veto x. A randomized embedding is, once again, a function $f : \mathcal{U} \to \Delta(A)$.

Deterministic Embeddings

The Plurality and Veto rules are closely related in the sense that agents must award an equal number of points to almost all the alternatives, and therefore cannot make a distinction in their votes between very desirable and very undesirable alternatives. However, this turns out to be a more acute problem under Veto, since agents cannot even single out one good alternative. The following easy result formalizes this intuition.

Theorem 9. *Let $|N| = n \geq 1$ and $|A| = m \geq 3$, and let $f : \mathcal{U} \to A$ be a deterministic embedding into Veto. Then dist$(f) = \infty$.*

Proof. Let $A = \{a, b, c, \ldots\}$. Let $\mathbf{u}$ be a utility profile such that for all $i \in N$, $u_i(a) = 1$, and $u_i(x) = 0$ for all $x \in A \setminus \{a\}$. Let f be an embedding into Veto; we have that either $f(u) \neq b$ or $f(u) \neq c$, hence either $b \in \text{win}(f, \mathbf{u})$ or $c \in \text{win}(f, \mathbf{u})$. Since $\text{sw}(a, \mathbf{u}) = n$, $\text{sw}(b, \mathbf{u}) = \text{sw}(c, \mathbf{u}) = 0$, it follows that the distortion of f is infinite. □

Randomized Embeddings

First, observe that if $n < m - 1$, then any randomized embedding into Veto has infinite distortion, for reasons similar to the deterministic case (Theorem 9). Indeed, consider n agents with utility 1 for alternative $x^* \in A$ and 0 for the remaining $m-1$ alternatives. At least one of these $m-1$ alternatives is not vetoed, and hence it is included in the set of winners. However, for larger values of n it is possible to obtain positive results; we consider the following (interesting) embedding.

Embedding 3 (Randomized embedding into Veto)**.** Given a utility function $u \in \mathcal{U}$, select an alternative $x \in A$ with probability $u(x)$; denote the selected alternative by x^*. Now, the vetoed alternative $f(u)$ is selected uniformly at random from $A \setminus \{x^*\}$ (that is, each alternative in $A \setminus \{x^*\}$ is selected with probability $1/(m-1)$).

We have the following upper bound on the distortion of Embedding 3.

Theorem 10. *Let $|N| = n \geq m = |A|$, $n \geq 3$, and denote Embedding 3 by f. Furthermore, let*

$$\epsilon(n, m) = 2m(m-1)\sqrt{\frac{\ln n}{n}} \ .$$

If $\epsilon(n, m) < 1$, then dist$(f) \leq \frac{1}{1-\epsilon(n,m)}$.

The proof of the theorem, which appears in the appendix, is similar to Part 3 of Theorem 4. As corollaries, we have that if $n/\ln n \geq 16m^2(m-1)^2$, then the distortion of Embedding 3 is at most two. In addition, for instances with $\epsilon(n, m) = o(1)$, the distortion is at most $1 + o(1)$.

When n is not much larger than m, we can show an exponential lower bound on the distortion of Embedding 3 by exploiting the relation to the well-known coupon collector problem; we omit the details. More generally we have the following theorem, whose proof appears in the appendix.

Theorem 11. *Let $n = |N| \geq |A| = m$, and let $f : \mathcal{U} \to \Delta(A)$ be a randomized embedding into Veto. Then dist$(f) = \Omega\left(m/\sqrt{n}\right)$.*

This statement provides a necessary condition $n = \Omega(m^2)$ in order to obtain constant distortion into Veto, whereas Embedding 1 yields a constant upper bound when $n = \Theta(m \ln m)$ with respect to Plurality and Approval. Hence, randomized embeddings into Veto are provably less efficient even when the number of agents is larger than the number of alternatives.

Discussion

Our notion of embeddings into voting rules is extremely decentralized, that is, the agents cast their votes independently according to the embedding. On the other extreme, if full coordination is allowed, the distortion would always be one, as the agents would be able to find out which alternative maximizes social welfare and coordinate their votes in a way that this alternative is elected (assuming the voting rule is onto the set of alternatives). It would be interesting to investigate a notion of embedding that allows for partial communication between the agents.

Our strongest positive results hold in settings where the number of agents is larger than the number of alternatives. This is indeed the case in many environments, notably in political elections. However, one can think of a variety of multiagent settings where the number of alternatives is larger. We would like to achieve a better understanding of the achievable distortion when $n = o(m)$.

References

Blumrosen, L.; Nisan, N.; and Segal, I. 2007. Auctions with severely bounded communication. *Journal of Artificial Intelligence Research* 28:233–266.

Chernoff, H. 1952. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Annals of Mathematical Statistics* 23:493–507.

Conitzer, V., and Sandholm, T. 2005. Communication complexity of common voting rules. In *Proceedings of the 6th ACM Conference on Electronic Commerce (EC)*, 78–87.

Hoeffding, W. 1963. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association* 58(301):13–30.

Jogdeo, K., and Samuels, S. 1968. Monotone convergence of binomial probabilities and a generalization of Ramanujan's equation. *Annals of Mathematical Statistics* 39:1191–1195.

Moon, J. W. 1968. *Topics on Tournaments*. Holt, Reinhart and Winston.

Motwani, R., and Raghavan, P. 1995. *Randomized Algorithms*. Cambridge University Press.

Procaccia, A. D., and Rosenschein, J. S. 2006. The distortion of cardinal preferences in voting. In *Proceedings of the 10th International Workshop on Cooperative Information Agents (CIA)*, volume 4149 of *Lecture Notes in Computer Science (LNCS)*. Springer. 317–331.

Procaccia, A. D.; Rosenschein, J. S.; and Kaminka, G. A. 2007. On the robustness of preference aggregation in noisy environments. In *Proceedings of the 6th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, 416–422.

Appendix: Omitted Proofs

Proof of Theorem 3

Let f be an embedding into Plurality. Consider a utility function $u^* \in \mathcal{U}$ where $u^*(x) = 1/m$ for all $x \in A$. There must exist $x^* \in A$ such that $f(u^*) = x^*$ with probability at most $1/m$.

Let $u_1 \equiv u^*$ be the utility function of agent 1, and let the utility function of agent 2 be defined by $u_2(x^*) = 1$, $u_2(x) = 0$ for all $x \in A \setminus \{x^*\}$. We have that $\text{sw}(x^*, \mathbf{u}) = 1 + 1/m$, $\text{sw}(x, \mathbf{u}) = 1/m$ for any $x \in A \setminus \{x^*\}$.

Now, the probability that $\{x^*\} = \text{win}(f, \mathbf{u})$ is at most $1/m$, since this happens only if both $f(u_1) = x^*$ and $f(u_2) = x^*$, and even the probability of the former event alone is at most $1/m$. We conclude that the distortion of f at $\mathbf{u}$ is at least

$$\text{dist}(f, \mathbf{u}) = \frac{1 + \frac{1}{m}}{\frac{1}{m} \cdot \left(1 + \frac{1}{m}\right) + \frac{m-1}{m} \cdot \frac{1}{m}} = \frac{m+1}{2} = \Omega(m) \ ,$$

hence $\text{dist}(f) = \Omega(m)$. □

Proof of Parts 2 and 3 of Theorem 4

Let $X_1, ..., X_n$ be independent heterogeneous Bernoulli trials. Denote by μ the expectation of the random variable $X = \sum_i X_i$. We employ the following tail inequalities (see, e.g., (Motwani & Raghavan 1995) for the different variations on the Chernoff bounds).

1. (Jogdeo and Samuels (1968)) $\Pr[X < \lfloor \mu \rfloor] < 1/2$.
2. (Lower tail Chernoff bound) For any $\delta \in [0, 1]$,
$$\Pr\left[X \leq (1 - \delta)\mu\right] \leq \exp\left(-\mu\delta^2/2\right) \ . \qquad (2)$$
3. (Upper tail Chernoff bound) For any $\delta \geq 0$,
$$\Pr\left[X \geq (1 + \delta)\mu\right] \leq \left(\frac{e}{1+\delta}\right)^{(1+\delta)\mu} \ . \qquad (3)$$
4. For $\delta \geq 2e - 1$,
$$\Pr\left[X \geq (1 + \delta)\mu\right] \leq 2^{-(1+\delta)\mu} \ . \qquad (4)$$
5. For $\delta < 2e - 1$ we can use the simplified inequality
$$\Pr\left[X \geq (1 + \delta)\mu\right] \leq \exp\left(-\mu\delta^2/5\right) \ . \qquad (5)$$

Proof of Part 2. The proof is similar to Part 1. Given $\mathbf{u} \in \mathcal{U}^n$, we once again denote by x^* the alternative with maximum social welfare, and we let $L \subset A$ be the set of alternatives with social welfare smaller than $\text{sw}(x^*, \mathbf{u})/(3e\sqrt{m})$, that is,

$$L = \left\{x \in A : \ \text{sw}(x, \mathbf{u}) < \frac{\text{sw}(x^*, \mathbf{u})}{3e\sqrt{m}}\right\} \ .$$

If $L = \emptyset$ then the claim follows trivially, hence we can restrict our attention to three cases. In all three cases we demonstrate that with constant probability no alternative in L is among the winners, that is, with probability bounded away from zero an alternative with social welfare at least $\text{sw}(x^*, \mathbf{u})/(3e\sqrt{m})$ is elected, which directly yields the bound on the distortion.

Case 1: $\text{sw}(x^*, \mathbf{u}) < 2$ and $|L| = 1$. Since $n \geq m$ it also holds that $\text{sw}(x^*, \mathbf{u}) \geq 1$ and, hence, by Jogdeo and Samuels the probability that $\text{sc}(x^*, f, \mathbf{u}) = 0$ is at most $1/2$. Let x be the element of L. Since $\text{sw}(x^*, \mathbf{u}) < 2$, it also holds that $\text{sw}(x, \mathbf{u}) < 2/(3e\sqrt{m})$ and, by Markov's inequality, we have that the probability that $\text{sc}(x, f, \mathbf{u}) \geq 1$ is at most $2/(3e\sqrt{m}) < 1/4$.

Case 2: $\text{sw}(x^*, \mathbf{u}) < 2$ and $|L| > 1$. For each $i \in N$, let X_i be a random variable such that $X_i = 1$ if $f(u_i) \notin L$, that is, agent i votes for an alternative not in L. We have that

$$\sum_{i \in N} X_i = \sum_{x \notin L} \text{sc}(x, f, \mathbf{u}) \ .$$

The sum $\sum_{x \notin L} \text{sc}(x, f, \mathbf{u})$ has expectation at least

$$m - \frac{2}{3e\sqrt{m}} \cdot |L| > m - |L| + 1 \ .$$

By Jogdeo and Samuels (using the fact that the X_i are independent) with probability at least $1/2$ it holds that $\sum_{x \notin L} \text{sc}(x, f, \mathbf{u}) \geq m - |L| + 1$. If so then by the Pigeonhole Principle there exists some alternative $x^0 \in A \setminus \{L\}$ which has $\text{sc}(x^0, f, \mathbf{u}) \geq 2$.

Now, consider some alternative $x \in L$. We apply Equation (3) to the random variable $\text{sc}(x, f, \mathbf{u})$ with expectation $\mu = \text{sw}(x, \mathbf{u})$ using $(1 + \delta)\mu = 2$. Since $\mu \leq 2/(3e\sqrt{m})$, it holds that $1 + \delta > 3e\sqrt{m}$. We conclude that

$$\Pr\left[\text{sc}(x, f, \mathbf{u}) \geq 2\right] \leq \left(\frac{1}{3\sqrt{m}}\right)^2 = \frac{1}{9m} \ .$$

By the union bound the probability that some alternative in L is among the winners is at most $(1/9m) \cdot m + 1/2 = 11/18$.

Case 3: $\text{sw}(x^*, \mathbf{u}) \geq 2$. By Jogdeo and Samuels, the probability that $\text{sc}(x^*, f, \mathbf{u}) < \lfloor \text{sw}(x^*, \mathbf{u}) \rfloor$ is at most $1/2$. Next we consider some alternative $x \in L$. We apply (3) to the random variable $\text{sc}(x, f, \mathbf{u})$ with expectation $\mu = \text{sw}(x, \mathbf{u})$ using

$$(1 + \delta)\mu = \lfloor \text{sw}(x^*, \mathbf{u}) \rfloor \ .$$

Since

$$\mu < \frac{\text{sw}(x^*, \mathbf{u})}{3e\sqrt{m}} \leq \frac{\lfloor \text{sw}(x^*, \mathbf{u}) \rfloor}{2e\sqrt{m}} \ ,$$

it holds that $1 + \delta > 2e\sqrt{m}$. We conclude that

$$\Pr\left[\text{sc}(x, f, \mathbf{u}) \geq 2\right] \leq \left(\frac{1}{2\sqrt{m}}\right)^{\lfloor \text{sw}(x*, \mathbf{u}) \rfloor} \leq \left(\frac{1}{2\sqrt{m}}\right)^2 = \frac{1}{4m} \ .$$

Similarly to Case 2, we apply the union bound and conclude that the probability that $\mathrm{sc}(x, f, \mathbf{u}) \geq \mathrm{sc}(x^*, f, \mathbf{u})$ for some $x \in L$ is at most $(1/4m) \cdot m + 1/2 = 3/4$. $\square$

Proof of Part 3. Given $\mathbf{u} \in \mathcal{U}^n$ we consider the alternative $x^* \in A$ with the maximum social welfare. Denote by $L \subset A$ the set of alternatives with social welfare at most $\mathrm{sw}(x^*, \mathbf{u}) - \left(\sqrt{2}+\sqrt{5}\right)\sqrt{\mathrm{sw}(x^*, \mathbf{u}) \ln n}$. We will show that the probability that there exists $x \in L$ such that $\mathrm{sc}(x, f, \mathbf{u}) \geq \mathrm{sc}(x^*, f, \mathbf{u})$ is at most m/n. Specifically, we will establish that the probability that either

$$\mathrm{sc}(x^*, f, \mathbf{u}) \leq \mathrm{sw}(x^*, \mathbf{u}) - \sqrt{2\mathrm{sw}(x^*, \mathbf{u}) \ln n}$$

or

$$\mathrm{sc}(x, f, \mathbf{u}) \geq \mathrm{sw}(x^*, \mathbf{u}) - \sqrt{2\mathrm{sw}(x^*, \mathbf{u}) \ln n}$$

for some $x \in L$ is at most m/n.

We first apply (2) bound to the random variable $\mathrm{sc}(x^*, f, \mathbf{u})$ with $\delta = \sqrt{(2 \ln n)/(\mathrm{sw}(x^*, \mathbf{u}))}$. Since the expectation of $\mathrm{sc}(x^*, f, \mathbf{u})$ is $\mathrm{sw}(x^*, \mathbf{u})$ we have that

$$\Pr\left[\mathrm{sc}(x^*, f, \mathbf{u}) \leq \mathrm{sw}(x^*, \mathbf{u}) - \sqrt{2\mathrm{sw}(x^*, \mathbf{u}) \ln n}\right] \leq \frac{1}{n} \ .$$

Next we consider an alternative $x \in L$ and the random variable $\mathrm{sc}(x, f, \mathbf{u})$. This variable has expectation $\mu < \mathrm{sw}(x^*, \mathbf{u}) - \left(\sqrt{2}+\sqrt{5}\right)\sqrt{\mathrm{sw}(x^*, \mathbf{u}) \ln n}$. We apply the upper tail Chernoff bound with δ such that

$$(1+\delta)\mu = \mathrm{sw}(x^*, \mathbf{u}) - \sqrt{2\mathrm{sw}(x^*, \mathbf{u}) \ln n} \ .$$

Clearly, $\delta\mu > \sqrt{5\mathrm{sw}(x^*, \mathbf{u}) \ln n}$. If $\delta \geq 2e-1$, Equation (4) yields

$$\begin{aligned}
&\Pr\left[\mathrm{sc}(x, f, \mathbf{u}) \geq \mathrm{sw}(x^*, \mathbf{u}) - \sqrt{2\mathrm{sw}(x^*, \mathbf{u}) \ln n}\right] \\
&\leq 2^{-\left(\mathrm{sw}(x^*,\mathbf{u}) - \sqrt{2\mathrm{sw}(x^*,\mathbf{u}) \ln n}\right)} \\
&\leq 2^{-\left(\mathrm{sw}(x^*,\mathbf{u})\left(1-\sqrt{\frac{2 \ln n}{\mathrm{sw}(x^*,\mathbf{u})}}\right)\right)} \\
&\leq 2^{-\left(\frac{n}{m}\left(1-\sqrt{\frac{2m \ln n}{n}}\right)\right)} \leq 2^{-\left(16\left(1-\frac{1}{2\sqrt{2}}\right)\ln n\right)} \leq \frac{1}{n} \ ,
\end{aligned}$$

where the third inequality follows from the fact that $\mathrm{sw}(x^*, \mathbf{u}) \geq n/m$, and the fourth inequality follows from our assumption that $\epsilon(n, m) < 1$.

If $\delta < 2e - 1$, Equation (5) yields

$$\begin{aligned}
&\Pr\left[\mathrm{sc}(x, f, \mathbf{u}) \geq \mathrm{sw}(x^*, \mathbf{u}) - \sqrt{2\mathrm{sw}(x^*, \mathbf{u}) \ln n}\right] \\
&\leq \exp\left(-\frac{(\delta\mu)^2}{5\mu}\right) < \exp\left(-\frac{\mathrm{sw}(x^*, \mathbf{u}) \ln n}{\mu}\right) \leq \frac{1}{n} \ .
\end{aligned}$$

By the union bound, we have that the probability that some of the undesired events happen is at most m/n (there are at most m such events). Hence, with probability at least $1 - m/n$ some alternative $x \in A \setminus L$ is the winner, and the expected score of the worst winner is at least

$$\begin{aligned}
&\left(\mathrm{sw}(x^*, \mathbf{u}) - \left(\sqrt{2}+\sqrt{5}\right)\sqrt{\mathrm{sw}(x^*, \mathbf{u}) \ln n}\right)\left(1 - \frac{m}{n}\right) \\
&= \mathrm{sw}(x^*, \mathbf{u})\left(1 - \left(\sqrt{2}+\sqrt{5}\right)\sqrt{\frac{\ln n}{\mathrm{sw}(x^*, \mathbf{u})}}\right)\left(1 - \frac{m}{n}\right) \\
&\geq \mathrm{sw}(x^*, \mathbf{u})\left(1 - \left(\sqrt{2}+\sqrt{5}\right)\sqrt{\frac{m \ln n}{n}}\right)\left(1 - \sqrt{\frac{m \ln n}{16n}}\right) \\
&\geq \mathrm{sw}(x^*, \mathbf{u})\left(1 - \left(\sqrt{2}+\sqrt{5}+\frac{1}{4}\right)\sqrt{\frac{m \ln n}{n}}\right) \\
&\geq \mathrm{sw}(x^*, \mathbf{u})\left(1 - 4\sqrt{\frac{m \ln n}{n}}\right) \\
&= \mathrm{sw}(x^*, \mathbf{u})\left(1 - \epsilon(n, m)\right) \ .
\end{aligned}$$

The second transition holds since $n \geq 3$ together with $\epsilon(n, m) < 1$ imply that

$$\frac{m}{n} \leq \sqrt{\frac{m}{16n \ln n}} \leq \sqrt{\frac{m \ln n}{16n}}$$

and, furthermore, $\mathrm{sw}(x^*, \mathbf{u}) \geq n/m$. The third transition follows from the inequality $(1-\alpha)(1-\beta) \geq (1-\alpha-\beta)$ for any $\alpha, \beta \in [0, 1]$. $\square$

Proof of Theorem 6

Let t, λ, and k be integers to be defined later. Consider an instance with $N = N' \cup N''$, where $|N'| = t$ and $|N''| = \lambda$ (i.e., $n = t + \lambda$). Furthermore, let $A = A' \cup A''$, where $|A'| = t/2$ and $|A''| = k\lambda/2$ (i.e., $m = t/2 + k\lambda/2$). It also holds that $m = n$, which implies that $m = \lambda(k-1)$. We construct a utility profile $\mathbf{u} \in \mathcal{U}^n$ as follows. Each $x \in A'$ has utility equal to 1 with respect to the utility functions of exactly two agents in N', that is, for all $x \in A'$, $\mathrm{sw}(x, \mathbf{u}) = 2$. Each $x \in A''$ has utility $1/k$ with respect to the utility functions of exactly two of the agents in N'', hence for all $x \in A''$, $\mathrm{sw}(x, \mathbf{u}) = 2/k$.

The probability that an alternative $x \in A''$ satisfies $\mathrm{sc}(x, f, \mathbf{u}) < 2$ is $1 - \frac{1}{k^2}$. Moreover, the probability that $\mathrm{sc}(x, f, \mathbf{u}) < 2$ given that a subset other alternatives in A'' have score less than two is at most $1 - \frac{1}{k^2}$. Therefore (by an implicit application of the chain rule) the probability that no $x \in A''$ has score two is at most

$$\left(1 - \frac{1}{k^2}\right)^{k\lambda/2} \leq \exp\left(-\frac{\lambda}{2k}\right) \ .$$

Selecting $\lambda = 2\lceil k \ln k \rceil$, we have that the probability that no alternative in A'' has score two is at most $1/k$. It follows that the distortion is at least

$$\frac{2}{2 \cdot \frac{1}{k} + \frac{2}{k} \cdot \left(1 - \frac{1}{k}\right)} \geq \frac{k}{2} \ .$$

Clearly $m = \mathcal{O}(k^2 \ln k)$ and, hence, the bound on the distortion follows. $\square$

Proof of Theorem 7

Let $\mathbf{u}$ be a utility profile. Let $x \in \text{win}(f, \mathbf{u})$ be a winning alternative, and let $x^* \in A$ be an alternative which maximizes the social welfare. Alternative x^* has $u_i(x^*) < 1/m$ with respect to $n - \text{sc}(x^*, f, \mathbf{u})$ agents i, and has utility at most one with respect to $\text{sc}(x^*, f, \mathbf{u})$ agents. Hence,

$$\begin{aligned}\text{sw}(x^*, \mathbf{u}) &< \text{sc}(x^*, f, \mathbf{u}) + (n - \text{sc}(x^*, f, \mathbf{u})) \cdot \frac{1}{m} \\ &= \frac{n}{m} + \left(1 - \frac{1}{m}\right) \cdot \text{sc}(x^*, f, \mathbf{u}) \\ &\leq \left(2 - \frac{1}{m}\right) \cdot \text{sc}(x, f, \mathbf{u}) \leq (2m-1) \cdot \text{sw}(x, \mathbf{u}) \ .\end{aligned}$$

The third transition holds since x is a winning alternative (and, hence, $\text{sc}(x^*, f, \mathbf{u}) \leq \text{sc}(x, f, \mathbf{u})$) and also has score at least n/m (since, by the definition of the embedding, at least one alternative is approved by each agent). The last transition follows from the definition of the embedding, which implies that $\text{sc}(x, f, \mathbf{u}) \leq m \cdot \text{sw}(x, \mathbf{u})$. □

Proof of Theorem 8

Let f be a deterministic embedding into Approval. Consider the utility function $u^1 \in \mathcal{U}$ where $u^1(a) = 0$, $u^1(x) = 1/(m-1)$ for all $x \in A \setminus \{a\}$. We can assume that $f(u^1)$ does not approve a and approves at least one $x^* \in A \setminus \{a\}$, otherwise we get infinite distortion by considering the utility profile where $u_i \equiv u^1$ for all $i \in N$. Without loss of generality $f(u^1)$ approves $b \in A \setminus \{a\}$.

Now, let $u^2 \in \mathcal{U}$ be defined by $u^2(a) = 1$, $u^2(x) = 0$ for all $x \in A \setminus \{a\}$ (in particular, $u^2(b) = 0$). We define a utility profile $\mathbf{u} \in \mathcal{U}^n$ by setting $u_i \equiv u^1$ for $\lceil n/2 \rceil$ agents i, and $u_i \equiv u^2$ for $\lfloor n/2 \rfloor$ agents i. By the argument above it holds that $b \in \text{win}(f, \mathbf{u})$, but (using the assumption on the size of n and m) it holds that

$$\text{dist}(f, \mathbf{u}) \geq \frac{\text{sw}(b, \mathbf{u})}{\text{sw}(a, \mathbf{u})} \geq \frac{m-1}{2} \ .$$

We conclude that $\text{dist}(f) \geq (m-1)/2$. □

Proof of Theorem 10

The proof is similar to the proof of Part 3 of Theorem 4. We use the following lemma.

Lemma 12 (Hoeffding (1963))**.** *Let $X_1, ..., X_n$ be independent heterogeneous Bernoulli trials. Denote by μ the expectation of the random variable $X = \sum_i X_i$. Then for any $\lambda > 0$,*

$$\Pr[|X - \mu| \geq \lambda] \leq \exp\left(-\frac{2\lambda^2}{n}\right) \ .$$

Let $\lambda = \frac{1}{2}\sqrt{n \ln n}$. Given $\mathbf{u} \in \mathcal{U}^n$, consider the alternative x^* with maximum social welfare and denote by L the set of alternatives with social welfare less than $\text{sw}(x^*, \mathbf{u}) - 2\lambda(m-1)$. We will show that the probability that either

$$\text{sc}(x^*, f, \mathbf{u}) \leq \mathbb{E}[\text{sc}(x^*, f, \mathbf{u})] - \lambda$$

or

$$\text{sc}(x, f, \mathbf{u}) \geq \mathbb{E}[\text{sc}(x^*, f, \mathbf{u})] - \lambda$$

for some $x \in L$ is at most $m/\sqrt{n}$.

Using the Hoeffding bound for the random variable $\text{sc}(x^*, f, \mathbf{u})$, we have

$$\begin{aligned}&\Pr\left[\text{sc}(x^*, f, \mathbf{u}) \leq \mathbb{E}[\text{sc}(x^*, f, \mathbf{u})] - \lambda\right] \\ &\leq \Pr\left[|\text{sc}(x^*, f, \mathbf{u}) - \mathbb{E}[\text{sc}(x^*, f, \mathbf{u})]\,| \geq \lambda\right] \\ &\leq \exp\left(-\frac{2\lambda^2}{n}\right) = \frac{1}{\sqrt{n}} \ .\end{aligned}$$

Now, Observe that for any $x \in A$,

$$\begin{aligned}\mathbb{E}[\text{sc}(x, f, \mathbf{u})] &= \sum_{i \in N}\left(u_i(x) \cdot 1 + (1 - u_i(x)) \cdot \frac{m-2}{m-1}\right) \\ &= \left(1 - \frac{1}{m-1}\right) n + \frac{\text{sw}(x, \mathbf{u})}{m-1} \ .\end{aligned}$$

Therefore for every alternative $x \in L$ it holds that $\mathbb{E}[\text{sc}(x^*, f, \mathbf{u})] - \lambda \geq \mathbb{E}[\text{sc}(x, f, \mathbf{u})] + \lambda$. Using this observation and the Hoeffding bound for the random variable $\text{sc}(x, f, \mathbf{u})$, we have

$$\begin{aligned}&\Pr\left[\text{sc}(x, f, \mathbf{u}) \geq \mathbb{E}[\text{sc}(x^*, f, \mathbf{u})] - \lambda\right] \\ &\leq \Pr\left[\text{sc}(x, f, \mathbf{u}) \geq \mathbb{E}[\text{sc}(x, f, \mathbf{u}) + \lambda\right] \\ &\leq \Pr\left[|\text{sc}(x, f, \mathbf{u}) - \mathbb{E}[\text{sc}(x, f, \mathbf{u})]| \geq \lambda\right] \\ &\leq \exp\left(-\frac{2\lambda^2}{n}\right) = \frac{1}{\sqrt{n}}.\end{aligned}$$

By the union bound the probability that some of the undesirable events happen is at most $m/\sqrt{n}$. Hence, with probability at least $1 - m/\sqrt{n}$ there is no $x \in L$ among the winners. We conclude that the expected social welfare of the worst winner is at least

$$\begin{aligned}&(\text{sw}(x^*, \mathbf{u}) - 2\lambda(m-1))\left(1 - \frac{m}{\sqrt{n}}\right) \\ &= \text{sw}(x^*, \mathbf{u})\left(1 - \frac{(m-1)\sqrt{n \ln n}}{\text{sw}(x^*, \mathbf{u})}\right)\left(1 - \frac{m}{\sqrt{n}}\right) \\ &\geq \text{sw}(x^*, \mathbf{u})\left(1 - m(m-1)\sqrt{\frac{\ln n}{n}}\right)\left(1 - \frac{m}{\sqrt{n}}\right) \\ &\geq \text{sw}(x^*, \mathbf{u})\left(1 - m(m-1)\sqrt{\frac{\ln n}{n}}\right)^2 \\ &\geq \text{sw}(x^*, \mathbf{u})\left(1 - 2m(m-1)\sqrt{\frac{\ln n}{n}}\right) \\ &= \text{sw}(x^*, \mathbf{u})\left(1 - \epsilon(n, m)\right).\end{aligned}$$

The first transition follows by substituting λ, the second transition holds since $\text{sw}(x^*, \mathbf{u}) \geq n/m$, the third transition easily follows by the condition on $\epsilon(n, m)$ using $n \geq 3$, and the fourth transition follows from $(1-\alpha)^2 \geq 1 - 2\alpha$. This concludes the theorem's proof. □

Proof of Theorem 11

Let f be a randomized embedding into Veto. Let $N = N' \cup N''$, where $|N'| = n - \lambda$ and $|N''| = \lambda$, with λ to be defined later. We define a utility profile $\mathbf{u} \in \mathcal{U}^n$ as follows. For all $i \in N'$ and $x \in A$, $u_i(x) = 1/m$, that is, all the agents in N' have utility $1/m$ for each alternative. Let $x^* \in A$ be the alternative that has the highest probability of being vetoed under f given this utility profile, i.e., for all $i \in N'$,

$$\Pr[f(u_i) = x^*] \geq \Pr[f(u_i) = x] \ . \tag{6}$$

Furthermore, for all $i \in N''$ we have $u_i(x^*) = 1$, $u_i(x) = 0$ for all $x \in A \setminus \{x^*\}$. Note that $\text{sw}(x^*, \mathbf{u}) = \lambda + (n-\lambda)/m$, whereas $\text{sw}(x, \mathbf{u}) = (n-\lambda)/m$ for all $x \in A \setminus \{x^*\}$.

It follows from Equation (6) that the the probability that x^* is among the $\lambda + 1$ alternatives that are vetoed least by the agents in N' is at most $(\lambda+1)/m$. Therefore with probability at least $1 - (\lambda+1)/m$ there are $\lambda + 1$ alternatives that are vetoed at most as many times as x^* by N', and at least one of them is not vetoed by the agents in N'' as there are only λ such agents. We conclude that the distortion of f is at least

$$\begin{aligned} \text{dist}(f, \mathbf{u}) &\geq \frac{\lambda + \frac{n-\lambda}{m}}{\frac{\lambda+1}{m} \cdot \left(\lambda + \frac{n-\lambda}{m}\right) + \left(1 - \frac{\lambda+1}{m}\right) \cdot \frac{n-\lambda}{m}} \\ &= \frac{\lambda + \frac{n-\lambda}{m}}{\frac{\lambda+1}{m} \cdot \lambda + \frac{n-\lambda}{m}} \ . \end{aligned}$$

Taking $\lambda = \Theta(\sqrt{n})$, we get that $\text{dist}(f) \geq \Omega(m/\sqrt{n})$. $\square$