Rules for Choosing Societal Tradeoffs

Vincent Conitzer, Rupert Freeman, Markus Brill, and Yuqian Li

Abstract

We study the societal tradeoffs problem, where a set of voters each submit their ideal tradeoff value between each pair of activities (e.g., "using a gallon of gasoline is as bad as creating 2 bags of landfill trash"), and these are then aggregated into the societal tradeoff vector using a rule. We introduce the family of distance-based rules and show that these can be justified as maximum likelihood estimators of the truth. Within this family, we single out the logarithmic distance-based rule as especially appealing based on a social-choice-theoretic axiomatization. We give an efficient algorithm for executing this rule as well as an approximate hill climbing algorithm, and evaluate these experimentally.

Introduction

There are many actions that we take in life that are generally agreed to have some negative effects on society. For example, consider actions with environmental downsides, such as using gasoline, creating landfill trash, and clearing forest, to name a few. Which of these is worse? To answer this, clearly one would first need to know *how much* gasoline is used, etc. This then suggests the following type of question: how many bags of trash are as bad as using one gallon of gasoline? Knowing the answer to this question could be useful to policy makers as well as to socially minded individuals or companies who are looking to reduce their environmental footprint in the most efficient way. However, since the environmental effects of these actions are different, it seems unlikely that an objective answer to this question exists. Rather, we as a society need to collectively decide what these tradeoffs should be, based on our own subjective opinions.

This suggests a social-choice-theoretic approach, where agents submit their preferences or opinions about what these tradeoff values should be as a *vote*. This social choice problem was suggested by Conitzer, Brill, and Freeman (2015) in an AAMAS 2015 Blue Sky paper. It has close conceptual ties to *judgment aggregation* (List and Pettit, 2002; Endriss, 2015), where the assessments of multiple judges are aggregated into a logically consistent social judgment. One difference is that here the assessments are quantitative rather than logical in nature. Specifically, we assume that each voter expresses for each pair of activities her ideal tradeoff value between those two. For example, a voter may feel that a gallon of gasoline corresponds to two bags of trash.

From a social-choice-theoretic viewpoint, when aggregating numbers, one submitted per voter, choosing the *median* is particularly compelling. When preferences are single-peaked, this results in choosing the Condorcet winner, and the corresponding voting rule is group-strategyproof. However, Conitzer, Brill, and Freeman (2015) pointed out that simply taking the median for each pair of activities can result in the aggregate tradeoffs being *inconsistent*, in the sense that the chosen tradeoff between a and c is not equal to the product of the tradeoff between a and b and the tradeoff between b and c. See the example in Figure 1, where a voter's tradeoffs are represented by a graph with its edges labeled with tradeoff values (e.g., voter 1 believes a gallon of gasoline is as bad as 2 bags of trash). This paradox is reminiscent of judgment aggregation paradoxes where taking majority on all individual issues results in a logically inconsistent aggregate judgment (Kornhauser and Sager, 1993).

So what are we to do? We insist that the aggregate tradeoffs be consistent; if not, then it is not clear how to use them to guide decisions involving three or more activities. That

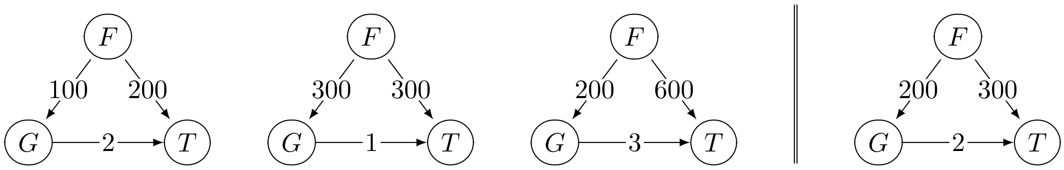

Figure 1: Example from Conitzer, Brill, and Freeman (2015) that illustrates that taking the median on each edge can lead to inconsistent outcomes even when each individual voter is consistent. The activities are *clearing forest* (F), *using gasoline* (G), and *creating trash* (T). Each of the left three graphs illustrates the consistent preferences of a single voter, with the label on each edge indicating the voters ideal tradeoff between the corresponding two activities. The rightmost graph, which results from taking the median on each edge, is inconsistent because $300 \neq 2 \cdot 200$.

means we must judiciously deviate from the median in some cases, but presumably we want to deviate as little as possible. The topic of this paper is how to make this precise.

We introduce a class of rules for this context that we call *distance-based rules*. We prove that these rules choose the median when there are only two activities and can be interpreted as maximum likelihood estimators of the "truth." We also axiomatize this class of rules. We then focus our attention on a particularly natural rule within this class, namely the *logarithmic* distance-based rule, and show it satisfies further nice properties, which allow us to also axiomatize it specifically. We give a linear programming formulation for computing its outcomes, as well as a simple hill-climbing algorithm that can get stuck in local optima but is surprisingly effective in experiments. Generally, our positive results hold even when agents submit inconsistent votes and our negative results hold even when they submit consistent votes.

Preliminaries

Let A be a finite set of activities and $N = \{1, \ldots, n\}$ a finite set of voters. Let E be a set of ordered pairs such that for every pair $a, b \in A$, either $(a, b) \in E$ or $(b, a) \in E$, but not both. For $i \in N$, let t_i^{ab} denote voter i's preferred tradeoff value between activities a and b, and let t^{ab} denote a (potential) aggregate (societal) tradeoff value between a and b. Let $t_i = (t_i^{ab})_{(a,b)\in E}$ denote the vector of all i's preferred tradeoff values (i's *vote*) and $t = (t^{ab})_{(a,b)\in E}$ a (potential) aggregate tradeoff vector. A *profile* P is a collection of n votes (one for each voter).

We assume that all tradeoff values are positive real numbers. A tradeoff vector t is *consistent* if for all $(a, b), (b, c), (a, c) \in E$, $t^{ab}t^{bc} = t^{ac}$, and for all $(a, b), (b, c), (c, a) \in E$, $t^{ab}t^{bc} = 1/t^{ca}$. A *tradeoff rule* f is a function that maps each profile P to a non-empty set of consistent tradeoff vectors $f(P)$. Note that we do not assume votes to be consistent.

An example rule

In this section, we introduce an example tradeoff rule. It is arguably the simplest way to obtain a variant of the Kemeny rule (Kemeny, 1959) for this domain. However, as we will show, it has some very undesirable properties. This will help to motivate the rule that we introduce later in the paper, which avoids these undesirable properties.

Definition 1 (Linear Distance Based Rule (DBR$^{\text{linear}}$)**.** *The linear distance between two tradeoff vectors t_1 and t_2 is $d^{linear}(t_1, t_2) = \sum_{(a,b)\in E} |t_1^{ab} - t_2^{ab}|$. The score of a (potential)*

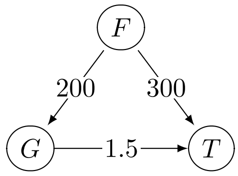

Figure 2: Aggregate tradeoff vector from applying $\mathrm{DBR}^{\mathrm{linear}}$ to the example from Figure 1. There is a disagreement of 100 with each of voters 1 and 2 on the forest-gasoline edge; a disagreement of 100 with voter 1 and 300 with voter 3 on the forest-trash edge; and a disagreement of 1/2 with voters 1 and 2, and 3/2 with voter 3 on the gasoline-trash edge. The total disagreement is thus 602.5, which is minimal.

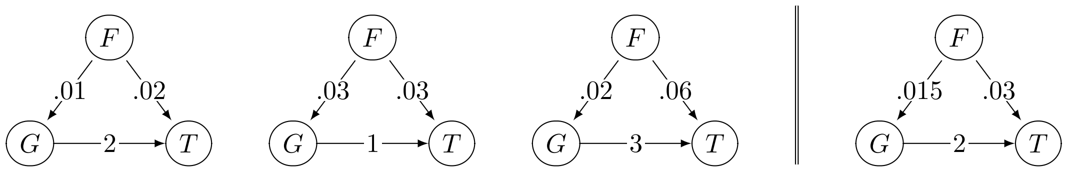

Figure 3: Example illustrating that a change of units can change the outcome on an unrelated edge under $\mathrm{DBR}^{\mathrm{linear}}$. The leftmost three graphs are the votes and the rightmost one is the outcome produced by $\mathrm{DBR}^{\mathrm{linear}}$.

aggregate tradeoff vector t relative to votes $(t_i)_{i\in N}$ is $\sum_{i\in N} d^{linear}(t, t_i)$. The tradeoff rule DBR^{linear} chooses the tradeoff vector(s) with minimum score.

For example, if we apply this rule to the profile from Figure 1, we obtain the aggregate tradeoff vector in Figure 2. Intuitively, the rule chooses to agree with the median on the edges with larger values, because it is more costly to disagree there; instead, it disagrees with the median on the bottom edge. One nice property of $\mathrm{DBR}^{\mathrm{linear}}$ is that when there are only two activities, it necessarily chooses the median. (We will prove a more general result as Proposition 2.)

Unfortunately, the rule has some undesirable properties. Suppose we change the units on the forest clearing activity by a factor of 10,000 (say, we were using m^2 before and are now using cm^2). Then naturally, the voters' ideal tradeoffs on these edges should change accordingly. Unfortunately, as illustrated in Figure 3, this changes the outcome of $\mathrm{DBR}^{\mathrm{linear}}$, even on the unrelated edge from gasoline to trash! Intuitively, the reason is that which edges are important has changed due to the change in units, so now it chooses to agree with the median on the bottom edge. A similar problem occurs if instead of changing units, we change the *direction* of some of the edges. For instance, if we reverse the edges incident to the "forest" node in the example in Figure 1 a vote for an ideal tradeoff of (say) 200 on such an edge would become 1/200 on the reversed edge. Hence, again the bottom edge would end up with the largest numbers, and $\mathrm{DBR}^{\mathrm{linear}}$ will again choose to agree with the median there. These shortcomings of $\mathrm{DBR}^{\mathrm{linear}}$ can be formalized as follows.

Definition 2 (ICU)**.** *A tradeoff rule f satisfies* independence of choice of units (ICU) *if the following holds. Consider an arbitrary profile $(t_i)_{i\in N}$ and let a be an arbitrary activity and k a constant. Let μ be a function modifying tradeoff vectors as follows. For every edge (a, b), $\mu(t)^{ab} = k \cdot t^{ab}$; for every edge (b, a), $\mu(t)^{ba} = k^{-1} \cdot t^{ba}$; and for every edge (b, c) with $a \notin \{b, c\}$, $\mu(t)^{bc} = t^{bc}$. Then $\mu(f(t_1, \dots, t_n)) = f(\mu(t_1), \dots, \mu(t_n))$.*[1]

[1] Note that technically, f is set-valued, so μ is applied to a set of tradeoff vectors, in the natural way.

Definition 3 (IED). *A tradeoff rule f satisfies* independence of edge directions (IED) *if the following holds. Consider an arbitrary profile $(t_i)_{i \in N}$ and let (a, b) be an arbitrary edge. Let μ be a function transforming tradeoff vectors to the modified graph where the edge (a, b) is replaced by (b, a), in the natural way—that is, $\mu(t)^{ba} = 1/t^{ab}$ and $\mu(t)^{cd} = t^{cd}$ for all other (unmodified) edges. Then $\mu(f(t_1, \ldots, t_n)) = f(\mu(t_1), \ldots, \mu(t_n))$.*

Proposition 1. *DBR^{linear} violates both ICU and IED.*

A more general class of rules

We now introduce a broader class of tradeoff rules.

Definition 4. *A* distance-based rule (DBR) *is defined by a function $g : \mathbb{R} \to \mathbb{R}$. The g-distance between two tradeoff vectors t_1 and t_2 is $d^g(t_1, t_2) = \sum_{(a,b)} |g(t_1^{ab}) - g(t_2^{ab})|$. The score of a (potential) aggregate tradeoff vector t relative to votes $t_1, \ldots, t_n$ is $\sum_i d^g(t, t_i)$. DBR^g chooses the tradeoff vector(s) with minimum score.*

We now show that these rules always select the median. (For simplicity, we will only consider the case where the number of voters is odd, but the result extends naturally to even numbers. For our axiomatic results involving the median later, we only need profiles with odd numbers.)

Proposition 2. *For any strictly monotone function g, when there are only two activities, DBR^g chooses the median uniquely.*

Proof. Let $A = \{a, b\}$ and consider the edge (a, b). Consider some potential aggregate tradeoff value t^{ab} that is (without loss of generality) strictly less than the median $t^{ab}_{\text{med}} = \text{med}(t_1^{ab}, \ldots, t_n^{ab})$. For every voter i with $t_i^{ab} \geq t^{ab}_{\text{med}}$, we have $|g(t^{ab}_{\text{med}}) - g(t_i^{ab})| = |g(t^{ab}) - g(t_i^{ab})| - |g(t^{ab}_{\text{med}}) - g(t^{ab})|$. For every voter i with $t_i^{ab} < t^{ab}_{\text{med}}$, we have $|g(t^{ab}_{\text{med}}) - g(t_i^{ab})| \leq |g(t^{ab}) - g(t_i^{ab})| + |g(t^{ab}_{\text{med}}) - g(t^{ab})|$. Because there is at least one more voter in the former category than the latter, it follows that the total score for t^{ab}_{med} is at most the total score for t^{ab}, minus $|g(t^{ab}_{\text{med}}) - g(t^{ab})|$. Because g is strictly monotone, t^{ab}_{med} obtains a strictly lower score than t^{ab}. $\square$

MLE interpretation of distance-based rules

In this section, we show that every distance-based rule can be interpreted as a maximum likelihood estimator of the "correct" tradeoff vector. The interpretation of voting rules as maximum likelihood estimators of the "truth" can be said to date back to Condorcet (1785); Young (1988, 1995) later made this more precise. The assumption is that there is an unobserved correct ranking of the alternatives, and every voter's vote (also a ranking) is a noisy observation of this correct ranking. Then, we can set ourselves the goal of choosing as the aggregate ranking a statistical estimate of the truth, given the votes. It is natural to choose the maximum likelihood estimate, and Young showed that for a particular noise model the Kemeny ranking (Kemeny, 1959) coincides with the maximum likelihood estimate. Other noise models result in MLEs that coincide with other voting rules (Drissi-Bakhkhat and Truchon, 2004; Conitzer and Sandholm, 2005; Truchon, 2008; Conitzer, Rognlie, and Xia, 2009).

Analogously, in our setting, we assume that there exists an unobserved "correct" tradeoff vector, and the votes are noisy observations of this correct vector. We consider the following specific family of noise models.

Definition 5. *Let t_{true} denote the correct tradeoff vector. For a function $g : \mathbb{R} \to \mathbb{R}$, let $P_g^{t_{true}}$ denote the following distribution over votes. Each agent's vote is drawn i.i.d. Moreover, each agent i draws its ideal tradeoffs t_i^{ab} independently across edges. Finally, the probability of a specific value t_i^{ab} is proportional to $e^{-|g(t_{true}^{ab})-g(t_i^{ab})|}$.*

Because tradeoffs are drawn i.i.d. across edges in this model, it will generally not produce consistent votes. This will not matter for our purposes; we can either consider this a feature and treat it as a remarkable accident when voters are in fact consistent, or we can remove the probability on inconsistent votes and renormalize on the consistent votes.[2] We next show that this family produces the distance-based rules as MLEs.

Proposition 3. *DBR^g is the MLE for $P_g^{t_{true}}$.*

Proof. The MLE for the distribution $P_g^{t_{\text{true}}}$ selects $\arg\max \prod_i \prod_{ab} e^{-|g(t_{\text{true}}^{ab})-g(t_i^{ab})|}$. Taking the logarithm results in $\arg\max \sum_i \sum_{ab} -|g(t_{\text{true}}^{ab}) - g(t_i^{ab})| = \arg\min \sum_i \sum_{ab} |g(t_{\text{true}}^{ab}) - g(t_i^{ab})|$, which is also chosen by DBR^g. □

Characterization of distance-based rules

In this section, we give an axiomatic justification for the class of distance-based rules. We first show that monotonicity of g is necessary for selecting the median in two-alternative cases.

Proposition 4. *If DBR^g always uniquely selects the median in profiles with two activities, then g is strictly monotone.*

Proof. Suppose that g is not strictly monotone. There are several cases; all are similar and we present only one here. Suppose there exist $x < y < z$ with $g(x) \geq g(z) \geq g(y)$. Consider a profile with three voters, $t_1^{ab} = x, t_2^{ab} = y$, and $t_3^{ab} = z$. Setting $t^{ab} = z$ gives total score $|g(z) - g(x)| + |g(z) - g(y)| = g(x) - g(y)$. Setting $t^{ab} = y$ gives total score $|g(y) - g(x)| + |g(y) - g(z)| = g(x) + g(z) - 2g(y) \geq g(x) - g(y)$, so $t^{ab} = z$ achieves at least as low a score as $t^{ab} = y = t_{\text{med}}^{ab}$. Thus t_{med}^{ab} is not uniquely chosen. □

By a similar argument, if DBR^g always selects the median (but sometimes not uniquely), then g is weakly monotone.

For the rest of this section, we will take a slightly different view of tradeoff rules, to facilitate the introduction of certain axioms. Let h be a function that takes as input a profile of votes and a tradeoff vector and outputs a nonnegative real number. Further, suppose that h takes value 0 whenever the tradeoff vector exactly matches every vote. We say that h *represents* tradeoff rule f if, for every profile P, $f(P)$ consists exactly of the aggregate tradeoff vectors that minimize the function $h(P, \cdot)$. We note that every unanimous[3] tradeoff rule is represented by at least one such h: simply define $h(P, t) = 0$ whenever $t \in f(P)$ and $h(P, t) = 1$ otherwise. MLE interpretations of rules such as the one given earlier also naturally provide such a score function: see the proof of Proposition 3. Next we show that, subject to two natural conditions, strictly monotone distance-based rules are the only tradeoff rules that choose the median when there are only two activities.

Definition 6 (Agent Separability)**.** *A function h satisfies* agent separability *if $h(P, t) = \sum_{i \in N} h(t_i, t)$ for all profiles P and tradeoff vectors t.*

[2]This is entirely similar to the fact that the simplest way to specify a noise model that produces the Kemeny rule as the MLE is to allow cyclical preferences.

[3]A unanimous rule is one that selects tradeoff vector t (possibly among others) when all votes agree exactly, i.e. $t_1 = \ldots = t_n = t$.

Agent separability implies anonymity, i.e., all voters are treated equally. Under an MLE interpretation, this axiom would correspond to the assumption that votes are drawn independently (conditional on the truth).

Definition 7 (Edge Separability)**.** *For a profile P and a tradeoff vector t, let $P^{ab} = \{t_1^{ab}, \ldots, t_n^{ab}\}$. A function h satisfies* edge separability *if $h(P,t) = \sum_{(a,b)\in E} h(P^{ab}, t^{ab})$ for all profiles P and tradeoff vectors t.*

Edge separability implies a kind of neutrality, i.e., all edges are treated equally. Under an MLE interpretation, this axiom would correspond to the assumption that the tradeoffs on each edge are drawn independently (conditional on the truth).

Theorem 1. *Let f be a tradeoff rule that is represented by function h satisfying agent separability and edge separability, and suppose f uniquely selects the median when there are only two activities. Then $f = DBR^g$ for some strictly monotone function g.*

Proof. To determine f, we need to specify some function h that represents f. By agent separability, $h(P,t) = \sum_{i\in N} h(t_i, t)$ for all P and t, so it is sufficient to specify $h(t_i, t)$ for every possible vote t_i and tradeoff vector t. By edge separability, $h(t_i, t) = \sum_{(a,b)\in E} h(t_i^{ab}, t^{ab})$, so we need only specify the value of h when passed a single vote on a single edge (i.e., a voter's ideal tradeoff for that edge) and a single candidate tradeoff for that edge. For ease of notation, we will write $h(x,y)$ where x is a vote and y a candidate tradeoff value (note that the value of $h(t_i^{ab}, t^{ab})$ does not depend on the voter i or on the edge (a,b)), and $x, y \in \mathbb{R}_{\geq 0}$. Note that $h(x,x) = 0$ by our assumption on h.

Suppose that f uniquely selects the median when there are only two alternatives. We first show that $h(x,y) = h(y,x)$ for all x, y. For contradiction, suppose not. Then without loss of generality there exist x, y such that $h(x,y) < h(y,x)$. Therefore there exists some n such that $(n+1)h(x,y) < nh(y,x)$. Consider a two-alternative profile P on $A = \{a,b\}$, where $n+1$ voters have $t_i^{ab} = x$ and n voters have $t_i^{ab} = y$. We have $h(P,x) = nh(y,x) > (n+1)h(x,y) = h(P,y)$. Thus $x = t_{\text{med}}^{ab} \notin f(P)$, a contradiction.

Next we show that for all $x \leq y \leq z$, $h(x,y) + h(y,z) = h(x,z)$. Suppose that $h(x,y) + h(y,z) > h(x,z)$ for some $x \leq y \leq z$. Then there exists n such that $nh(x,y) + nh(y,z) > nh(x,z) + h(y,z)$. Consider a profile P with n voters with $t_i^{ab} = x$, n voters with $t_i^{ab} = z$, and one voter with $t_i^{ab} = y$. Then $t_{\text{med}}^{ab} = y$ but $h(P,z) = nh(x,z) + h(y,z) < nh(x,y) + nh(y,z) = h(P,y)$, so $y \notin f(P)$. Suppose next that $h(x,y) + h(y,z) < h(x,z)$. Then there exists n such that $(n+1)h(x,y) + nh(y,z) < nh(x,z)$. Consider profile P with $n+1$ voters with $t_i^{ab} = x$ and n voters with $t_i^{ab} = z$. We have $h(P,y) = (n+1)h(x,y) + nh(y,z) < nh(x,z) = h(P,x)$. Thus $x = t_{\text{med}}^{ab} \notin f(P)$, a contradiction.

We can now express f as a distance based rule. Define

$$g(x) = \begin{cases} h(1,x) & \text{if } x \geq 1 \\ -h(1,x) & \text{if } x < 1 \end{cases}$$

Using the "triangle equality" derived above, we now show that $h(x,y) = |g(x) - g(y)|$ for all x, y. There are several cases.
Case 1: $1 < x < y$. Then $|g(x) - g(y)| = |g(y) - g(x)| = |h(1,y) - h(1,x)| = h(x,y) = h(y,x)$.
Case 2: $x < 1 < y$. Then $|g(x) - g(y)| = |g(y) - g(x)| = |h(1,y) + h(1,x)| = h(1,y) + h(x,1) = h(x,y) = h(y,x)$.
Case 3: $x < y < 1$. Then $|g(x) - g(y)| = |g(y) - g(x)| = |-h(1,y) + h(1,x)| = h(x,1) - h(y,1) = h(x,y) = h(y,x)$.

By the definition of h, tradeoff rule f minimizes $h(P,t) = \sum_{i\in N} h(t_i,t) = \sum_{i\in N}\sum_{(a,b)\in E} h(t_i^{ab}, t^{ab}) = \sum_{i\in N}\sum_{(a,b)\in E} |g(t_i^{ab}) - g(t^{ab})|$, therefore $f = \text{DBR}^g$. By

Proposition 4, g must be strictly monotone as f uniquely chooses the median on two activities. □

The logarithmic distance based rule

We will be particularly interested in the logarithmic distance based rule, where $g = \log$.

Proposition 5. *$DBR^{\log}$ is the same regardless of the base of the logarithm.*

Proof. Consider two different bases α and β; we have that $\log_\alpha(x) = \log_\beta(x) \cdot \log_\alpha(\beta)$. Because $\log_\alpha(\beta)$ is a constant, the score of any aggregate tradeoff will only be changed by a constant when we change the base of the logarithm. □

Proposition 6. *$DBR^{\log}$ satisfies ICU and IED.*

Proof. We have $|\log(k \cdot t^{ab}) - \log(k \cdot t_i^{ab})| = |\log k + \log(t^{ab}) - \log k - \log(t_i^{ab})| = |\log(t^{ab}) - \log(t_i^{ab})|$. Therefore, if we perform a change of units (both on the votes and the aggregate tradeoff vectors), no scores change, and hence outcomes remain the same. Similarly, we have $|\log(1/t^{ab}) - \log(1/t_i^{ab})| = |-\log(t^{ab}) + \log(t_i^{ab})| = |\log(t^{ab}) - \log(t_i^{ab})|$. Therefore, if we change the direction of an edge (both in the votes and the aggregate tradeoff vectors), no scores change, and hence outcomes remain the same. □

We now consider a slightly stronger version of ICU that makes sense for the class of distance-based rules. It states that the score on any single edge should be independent of the units chosen for that edge.

Definition 8 (Strong ICU)**.** *A distance-based rule DBR^g satisfies* strong ICU *if, for all $k, x, y \in \mathbb{R}^+$, $|g(kx) - g(ky)| = |g(x) - g(y)|$.*

It is clear that strong ICU implies ICU, since under strong ICU any change of units can not change the score on even a single edge. It remains an open problem whether the converse holds, in general. However, we show that under the condition that the derivative g' is bounded below and above on any closed interval, strong ICU is equivalent to ICU.

Lemma 1. *Let $g : \mathbb{R}^+ \to \mathbb{R}$ be a strictly monotone, differentiable function. Suppose that for any closed interval $[p, q] \subseteq \mathbb{R}^+$, there exist c, C with $0 < c < C$ such that $c < g'(x) < C$ for all $x \in [p, q]$. If DBR^g satisfies ICU then DBR^g satisfies strong ICU.*

Proof. Let g satisfy the conditions of the lemma statement. We will suppose without loss of generality that g is (strictly) increasing. Suppose that DBR^g fails Strong ICU; that is $|g(kx) - g(ky)| \neq |g(x) - g(y)|$ for some k, x, y, and (without loss of generality) that $x > y$, and therefore $g(x) > g(y)$ and $g(kx) > g(ky)$. Let c, C be the lower and upper bounds on g' for the interval $[1, kx]$. Let n be sufficiently large such that $2nc - (C\frac{x}{y} + Cx) > 0$.

We exhibit an instance of the societal tradeoff problem on which DBR^g fails ICU. Consider three activities a, b, c and $2n + 1$ voters who cast the following votes:

$$\begin{array}{ll} n\ \times & t_i^{ab} = 1,\ t_i^{bc} = x,\ t_i^{ac} = x \\ n\ \times & t_i^{ab} = 1,\ t_i^{bc} = y,\ t_i^{ac} = y \\ 1\ \times & t_i^{ab} = \dfrac{x}{y},\ t_i^{bc} = y,\ t_i^{ac} = x \end{array}$$

We first determine the aggregate tradeoff(s) output by DBR^g on this instance. Observe that such a tradeoff t satisfies $t^{ab} \in [1, \frac{x}{y}]$, for the following reason. If $t^{ab} > \frac{x}{y}$ then either $t^{bc} < y$ or $t^{ac} > x$. We can decrease the score of t by adjusting t^{ab} towards $t^{ab}_{\text{med}} = 1$

and simultaneously increasing t^{bc} towards $t^{bc}_{\text{med}} = y$ (in the former case), or decreasing t^{ac} towards $t^{ac}_{\text{med}} = x$ (in the latter case). If $t^{ac} < 1$ then $t^{bc} > t^{ac}$ and we can decrease the score by increasing t^{ab} towards $t^{ab}_{\text{med}} = 1$ while simultaneously either decreasing t^{bc} (if $t^{bc} > y$) or increasing t^{ac} (if $t^{ac} < x$). It can be verified that at least one of these conditions is guaranteed to be true by the relations $x > y$ and $t^{bc} > t^{ac}$. Given that $t^{ab} \in [1, \frac{x}{y}]$, it is also easy to check that $t^{bc}, t^{ac} \in [y, x]$.

We now show that in fact $t^{ab} = 1$. For contradiction, suppose instead that $t^{ab} > 1$ and consider scaling it by some factor ϵ with $\frac{1}{t^{ab}} < \epsilon < 1$ (that is, we shift t^{ab} towards 1 by some absolute amount that is between ϵ and $\epsilon\frac{x}{y}$, depending on the value of t^{ab}). For the $2n$ voters with $t^{ab}_i = 1$, the distance $d^g(t^{ab}_i, t^{ab})$ decreases by at least $c\epsilon$, by the lower bound on g'. For the single voter with $t^{ab}_i = \frac{x}{y}$, the distance $d^g(t^{ab}_i, t^{ab})$ increases by at most $C\epsilon\frac{x}{y}$, by the upper bound on g'. Thus the change in the score on edge (a, b) is at most $-2nc\epsilon + C\epsilon\frac{x}{y}$.

By the consistency constraint, scaling t^{ab} by ϵ requires scaling t^{bc} and/or t^{ac} so that $t^{ab}t^{bc} = t^{ac}$. We will scale t^{ac} by ϵ and obtain an upper bound on the change in score as a result (there are other possiblities here, but we only need to exhibit a single tradeoff vector with lower score than t, so we are free to consider only one case). The scaling results in an absolute change in t^{ac} of at most ϵx. Note that the change in score on this edge for the first n voters is exactly canceled by the change in score for the second set of n voters. So we need only consider the last voter, for whom the distance $d^g(t^{ac}_i, t^{ac})$ increases by at most $C\epsilon x$. Therefore, the total change is at most $-2nc\epsilon + C\epsilon\frac{x}{y} + C\epsilon x$, which is less than zero by the choice of n. Thus no value $t^{ab} > 1$ is optimal.

By consistency, $t^{ab} = 1$ implies that $t^{bc} = t^{ac}$. We note that as long as $t^{bc} \in [x, y]$, the resulting tradeoff vector is optimal. The sum of scores on the two edges is exactly equal to $(2n + 1)|g(x) - g(y)|$. Therefore DBR^g outputs a tie between the tradeoff vectors $(t^{ab} = 1, t^{bc} = x, t^{ac} = x)$ and $(t^{ab} = 1, t^{bc} = y, t^{ac} = y)$ (among others).

We can now prove that DBR^g fails ICU by showing that the output changes when we consider a change of units applied to activity a with constant k. In particular, one of the two tradeoff vectors specified in the previous paragraph (adjusted for change of units) will no longer be chosen. The score of tradeoff $(t^{ab} = k, t^{bc} = x, t^{ac} = kx)$ is now

$$\left| g\left(\frac{kx}{y}\right) - g(k) \right| + (n+1)\,|g(x) - g(y)| + n\,|g(kx) - g(ky)|,$$

and the score of tradeoff $(t^{ab} = k, t^{bc} = y, t^{ac} = ky)$ is

$$\left| g\left(\frac{kx}{y}\right) - g(k) \right| + n\,|g(y) - g(x)| + (n+1)\,|g(ky) - g(kx)|.$$

The difference in the two scores is $|g(ky) - g(kx)| - |g(x) - g(y)| \neq 0$, so there is no longer a tie between the two outcomes, and at least one of them is no longer chosen by DBR^g. □

We are now able to uniquely characterize DBR$^{\log}$.

Theorem 2. *DBR$^{\log}$ is the only distance-based rule that satisfies strong ICU and uniquely selects the median when there are only two activities.*

Proof. By Proposition 4, it is sufficient to show that the logarithm is the only strictly monotone function satisfying $|g(kx) - g(ky)| = |g(x) - g(y)|$ for all $k, x, y \in \mathbb{R}^+$. Let g be a function with this property and assume (without loss of generality) that $x > y$. Rearranging, $g(kx) - g(x) = g(ky) - g(y) = c_k$ for all $x, y \in \mathbb{R}^+$ and some constant c_k that depends on k. So

$$g(kx) = g(x) + c_k \Rightarrow g(k) = g(1) + c_k \Rightarrow c_k = g(k) - g(1).$$

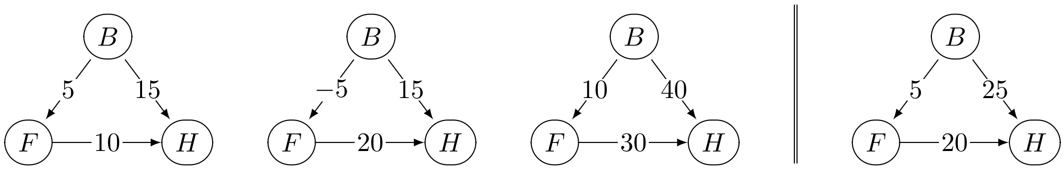

Figure 4: Example illustrating the additive variant with activities *watching basketball* (B), *watching football* (F), and *watching hockey* (H). The leftmost three graphs are the votes and the rightmost one is an outcome produced by the linear distance-based rule with an objective value of 70.

We may assume that $g(1) = 0$, since any distance-based rule is unchanged by the addition of a constant to g. So the condition reduces to

$$g(kx) = g(x) + g(k)$$

for all $x, k \in \mathbb{R}^+$. The only strictly monotone functions satisfying this condition have the form $g(x) = c \log(x)$ for $c \in \mathbb{R}$ (see, e.g., Smítal, 1988, for a proof of this well known fact). The result follows by observing that $\mathrm{DBR}^g = \mathrm{DBR}^{cg}$, which leaves us with only the rule $\mathrm{DBR}^{\log}$. □

The following corollary follows directly from Lemma 1 and Theorem 2.

Corollary 1. *Let g be a strictly monotone, differentiable function. Suppose that for any closed interval $[p, q] \subseteq (0, \infty)$, there exist c, C with $0 < c < C$ such that $c < g'(x) < C$ for all $x \in [p, q]$. Suppose moreover that DBR^g satisfies ICU. Then $DBR^g = DBR^{\log}$.*

Can we efficiently compute outcomes under $\mathrm{DBR}^{\log}$? It turns out that we can. In fact, it turns out that the logarithmic transformation is actually helpful. Intuitively, the reason is that once we apply logarithms to all tradeoff values, the consistency constraint becomes additive. That is, $t^{ab} \cdot t^{bc} = t^{ac}$ is equivalent to $\log(t^{ab}) + \log(t^{bc}) = \log(t^{ac})$. To see more precisely how this is helpful, we first discuss an additive variant of our problem, which may be of independent interest but whose primary purpose is to help us efficiently compute outcomes under $\mathrm{DBR}^{\log}$.

An additive variant

Consider an additive variant of our problem, where we compare activities by saying that a is x units "better" than b. Then, the consistency constraint becomes $t^{ac} = t^{ab} + t^{bc}$. Consider the example in Figure 4, in which, for instance, agent 1 feels that watching basketball is 5 units more enjoyable than watching football.

In this case, we can again define the linear distance based rule, based on the distance

$$d^{\text{linear}}(t_1, t_2) = \sum_{(a,b)} |t_1^{ab} - t_2^{ab}|.$$

Unlike in the original (multiplicative) context, in this additive context using the linear distance seems to make sense – changing units does not seem relevant, and changing the direction of an edge only changes the sign of values on it, rather than their magnitude, so the outcome remains unaffected. The rightmost graph in Figure 4 gives an outcome produced by this rule.

As it turns out, in this variant we can solve for optimal solutions (i.e., the outcomes produced by the rule) in polynomial time, using a linear program. This linear program contains a variable q_a for each activity, representing the aggregate quality of that activity. We will only be interested in *differences* in qualities—e.g., $q_a - q_b = t^{ab}$—so we can normalize an arbitrary one of the activities to have quality 0. The linear program also contains variables d_i^{ab}, denoting the distance $|t_i^{ab} - q_a + q_b|$. The linear program is then as follows:

$$\begin{aligned}
\text{minimize} \quad & \sum_{i \in N} \sum_{(a,b) \in E} d_i^{ab} \\
\text{subject to} \quad & d_i^{ab} \geq q_a - q_b - t_i^{ab} \quad (\forall i, a, b) \\
& d_i^{ab} \geq t_i^{ab} - q_a + q_b \quad (\forall i, a, b)
\end{aligned}$$

Instead of solving the LP directly, there is also a natural *hill-climbing* approach. This involves initializing the q_a variables arbitrarily and then checking them individually to see whether it can be changed to a value that increases the objective. This check has a nice social-choice-theoretic interpretation, as follows. Consider some activity a. Then, for any $b \neq a$ and voter i, define the *implied vote by* (i, b) on a to be $t_i^{ab} + q_b$ (or $q_b - t_i^{ba}$), where q_b is the current setting for b. The reason is that if q_a is set to this value, then there will be no disagreement with t_i^{ab}; more generally, the disagreement with t_i^{ab} resulting from setting q_a to a value will be the distance of that value to the implied vote. Hence, the overall objective value will be maximally improved by setting q_a to the median of these implied votes. (Note that the number of implied votes may be even even if the number of agents is odd, in which case any value between the left and right medians will be optimal.)

Executing DBR$^{\log}$ via the additive model

As it turns out, an algorithm for the additive model (with linear distance) will allow us to directly solve the orginal (multiplicative) model (with logarithmic distance) using a simple transformation. We simply take the logarithm of each t_i^{ab} to obtain $\hat{t}_i^{ab} = \ln(t_i^{ab})$, run an algorithm for the additive model to obtain optimal values $\hat{t}^{ab}$, and exponentiate back to obtain $t^{ab} = e^{\hat{t}^{ab}}$.

Proposition 7. *When using an exact solver for the additive model, the procedure described above results in an optimal solution for DBR*$^{\log}$.

Proof. We first observe that when $t^{ab} = e^{\hat{t}^{ab}}$, t is consistent in the multiplicative model if and only if $\hat{t}$ is consistent in the additive model. This follows simply from the fact that

$$t^{ac} = t^{ab} t^{bc} \Leftrightarrow e^{\hat{t}^{ac}} = e^{\hat{t}^{ab}} e^{\hat{t}^{bc}} = e^{\hat{t}^{ab} + \hat{t}^{bc}} \Leftrightarrow \hat{t}^{ac} = \hat{t}^{ab} + \hat{t}^{bc}.$$

Furthermore, the objective value of t in the multiplicative model is the same as that of $\hat{t}$ in the additive model. This is because $|\ln(t^{ab}) - \ln(t_i^{ab})| = |\ln(e^{\hat{t}^{ab}}) - \ln(e^{\hat{t}_i^{ab}})| = |\hat{t}^{ab} - \hat{t}_i^{ab}|$, so each term in the summation of the objective value is the same. $\square$

Because linear programs can be solved in polynomial time (Khachiyan, 1979), we immediately obtain:

Corollary 2. *We can solve for an outcome under DBR*$^{\log}$ *in polynomial time.*

Of course, we can also use the hill-climbing algorithm described in the previous section to the transformed instance and then transform it back to the multiplicative model to obtain a (possibly suboptimal) solution.

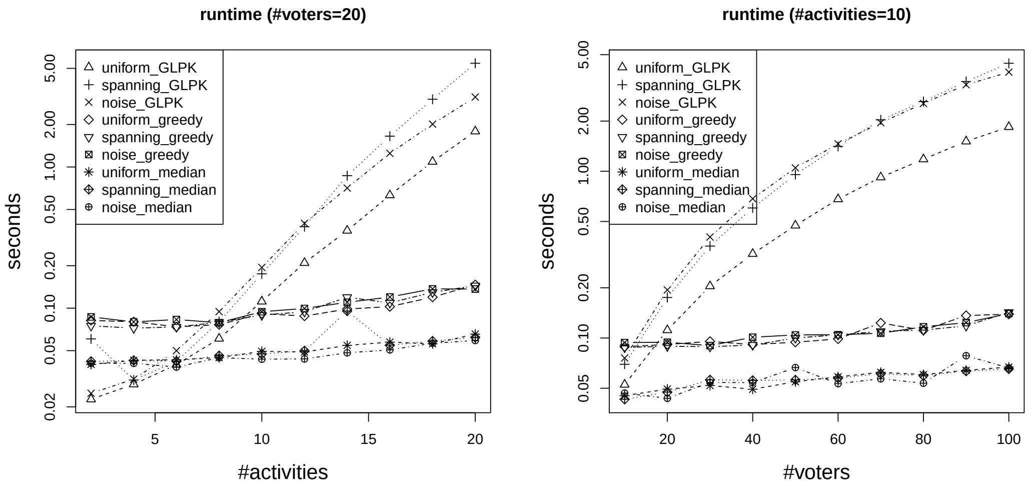

Figure 5: Runtime comparison for different algorithms and distributions over vote profiles. The prefixes "uniform," "spanning," and "noise" specify how votes are generated. The suffixes "GLPK," "greedy," and "median" specify the algorithms: GLPK is the optimal LP solver (using the GNU linear programming kit), greedy is the hill-climbing algorithm, and median simply picks a random spanning tree of activities and uses the median rule for each spanning tree edge.

Experiments

We generated three classes of voting profiles and compare the different algorithms' performances in terms of running time, penalty (LP's objective), and the distance between the aggregated result and the ground truth (if there is one).

For the first class of voting profile (*uniform*), each vote is generated as follows. For every pair of activities, we draw a number $x \in [-1, 1]$ uniformly at random and let the voter's tradeoff between two activities be e^x. Note that this generally generates inconsistent votes.

For the second class (*spanning*), which generates consistent votes, each vote is generated by first generating a random spanning tree among activities. Then, for each pair of activities that forms a spanning tree edge, we draw a number $x \in [-1, 1]$ uniformly at random and let the voter's tradeoff between those two activities be e^x. Finally, we use those spanning tree edges and the consistency constraint to infer the relationships between pairs of activities that do not form a spanning tree edge.

For the third class (*noise*), we first sample a ground truth quality q_a, uniformly at random between -10 and 10 for each activity a. Then for each voter i, we draw noise δ_i^a from a normal distribution with mean 0 and standard deviation 1 for each activity. We then let the tradeoff between two activities a and b be $e^{q_a+\delta_i^a}/e^{q_b+\delta_i^b}$ for that voter.

Results are shown in Figures 5 and 6. Particularly notable is the performance of the hill-climbing algorithm, whose solution quality in the experiments is indistinguishable from that of the LP, while being significantly faster. This is in spite of it being naïvely initialized to 0 and not using random restarts. We have manually constructed an example where hill climbing gets stuck at a local optimum (and verified this with our code), but it appears such instances do not get generated in the experiments.

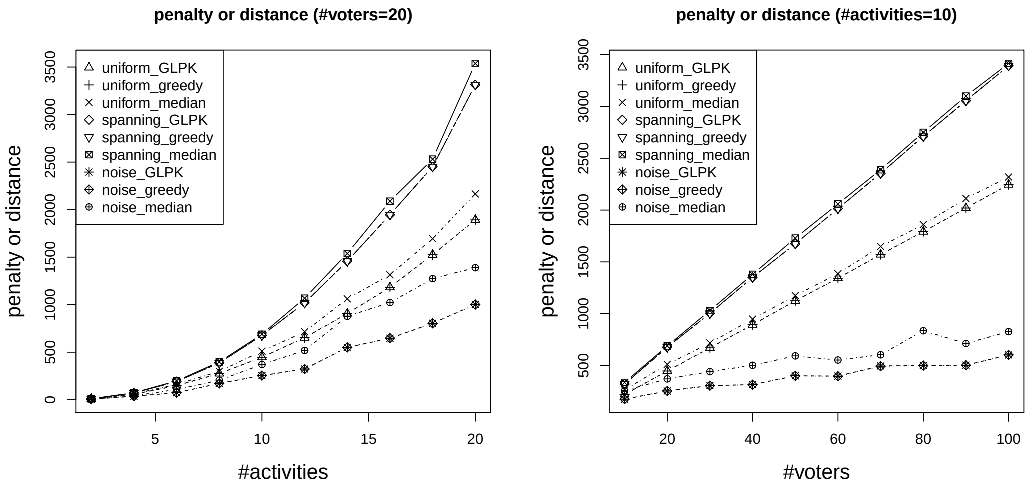

Figure 6: Performance comparison for different algorithms and distributions over vote profiles. "Penalty" is the sum of disagreement between the aggregated tradeoff and all votes (our LP objective): $\sum_{i \in N} \sum_{(a,b) \in E} |\log(t^{ab}) - \log(t_i^{ab})|$. For the noise case, we use distance to the ground truth instead of penalty. "Distance" is $n \cdot |(q_a - q_b) - \log(t_i^{ab})|$ where q is the ground truth (after the log adjustment). We multiply the difference by the number of voters n so that it has the same scale as penalty.

Conclusion

We believe we have made a very strong case for the logarithmic distance-based rule. We have shown that it uniquely satisfies some very desirable properties and can be executed efficiently. Some practical issues would likely need to be addressed before real deployment. For example, one concern may be that agents would have a hard time providing exact ideal tradeoff values; they may, for example, be more comfortable reporting an interval for each edge. Farfel and Conitzer (2011) propose aggregating intervals by taking the median of the lower bounds and the median of the upper bounds; similarly, we could aggregate lower bounds and upper bounds separately. Various other practical issues are discussed by Conitzer, Brill, and Freeman (2015). Still, we believe that the identification of this rule and algorithms for computing it represent a major step forward in this agenda.

We believe this work also generates appealing theoretical questions. Can we say something about the structure of the solutions generated? (We have an example where the optimal solution does not coincide with the median on any edge.) Can we explain the remarkable performance of the hill-climbing algorithm? What about incentives for voters to strategically misrepresent their ideal tradeoffs? Finally, is the societal tradeoffs problem really just one of a larger class of social-choice-theoretic problems? The additive variant suggests so, and one can imagine other variants. For example, the voters may report what they perceive to be the *distances* (i.e., dissimilarities) between the nodes, in which case the consistency constraint may be a triangle inequality on the distance.

Acknowledgements

We are thankful for support from NSF under awards IIS-1527434, IIS-0953756, CCF-1101659, and CCF-1337215, ARO under grants W911NF-12-1-0550 and W911NF-11-1-0332, ERC under StG 639945 (ACCORD), a Guggenheim Fellowship, and a Feodor Lynen research fellowship of the Alexander von Humboldt Foundation. This work was done in part while Conitzer was visiting the Simons Institute for the Theory of Computing.

References

Conitzer, V., and Sandholm, T. 2005. Common voting rules as maximum likelihood estimators. In *Proceedings of the 21st Annual Conference on Uncertainty in Artificial Intelligence (UAI)*, 145–152.

Conitzer, V.; Brill, M.; and Freeman, R. 2015. Crowdsourcing societal tradeoffs. In *Proceedings of the 14th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, 1213–1217.

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.

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.

Drissi-Bakhkhat, M., and Truchon, M. 2004. Maximum likelihood approach to vote aggregation with variable probabilities. *Social Choice and Welfare* 23:161–185.

Endriss, U. 2015. Judgment aggregation. In Brandt, F.; Conitzer, V.; Endriss, U.; Lang, J.; and Procaccia, A. D., eds., *Handbook of Computational Social Choice*. Cambridge University Press.

Farfel, J., and Conitzer, V. 2011. Aggregating value ranges: preference elicitation and truthfulness. *Autonomous Agents and Multi-Agent Systems* 22(1):127–150.

Kemeny, J. 1959. Mathematics without numbers. *Daedalus* 88:575–591.

Khachiyan, L. 1979. A polynomial algorithm in linear programming. *Soviet Math. Doklady* 20:191–194.

Kornhauser, L. A., and Sager, L. G. 1993. The one and the many: Adjudication in collegial courts. *California Law Review* 81(1):1–59.

List, C., and Pettit, P. 2002. Aggregating sets of judgments: An impossibility result. *Economics and Philosophy* 18(1):89–110.

Smítal, J. 1988. *On functions and functional equations.* Adam Hilger.

Truchon, M. 2008. Borda and the maximum likelihood approach to vote aggregation. *Mathematical Social Sciences* 55(1):96–102.

Young, H. P. 1988. Condorcet's theory of voting. *American Political Science Review* 82:1231–1244.

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

Vincent Conitzer
Department of Computer Science
Duke University
Durham, NC, USA
Email: conitzer@cs.duke.edu

Rupert Freeman
Department of Computer Science
Duke University
Durham, NC, USA
Email: rupert@cs.duke.edu

Markus Brill
Department of Computer Science
University of Oxford
Oxford, UK
Email: mbrill@cs.ox.ac.uk

Yuqian Li
Department of Computer Science
Duke University
Durham, NC, USA
Email: yuqian@cs.duke.edu