A Maximum Likelihood Approach towards Aggregating Partial Orders

Lirong Xia
Department of Computer Science
Duke University
Durham, NC 27708, USA
lxia@cs.duke.edu

Vincent Conitzer
Department of Computer Science
Duke University
Durham, NC 27708, USA
conitzer@cs.duke.edu

Abstract

In many of the possible applications as well as the theoretical models of computational social choice, the agents' preferences are represented as partial orders. In this paper, we extend the maximum likelihood approach for defining "optimal" voting rules to this setting. We consider distributions in which the pairwise comparisons/incomparabilities between alternatives are drawn i.i.d. We call such models *pairwise-independent models* and show that they correspond to a class of voting rules that we call *pairwise scoring rules*. This generalizes rules such as Kemeny and Borda. Moreover, we show that Borda is the only pairwise scoring rule that satisfies *neutrality*, when the outcome space is the set of all alternatives. We then study which voting rules defined for linear orders can be extended to partial orders via our MLE model. We show that any weakly neutral *outcome scoring rule* (including any ranking/candidate scoring rule) based on the weighted majority graph can be represented as the MLE of a weakly neutral pairwise-independent model. Therefore, all such rules admit natural extensions to profiles of partial orders. Finally, we propose a specific MLE model π_k for generating a set of k winning alternatives, and study the computational complexity of winner determination for the MLE of π_k.

1 Introduction

In traditional voting theory, it is usually assumed that each agent reports her preferences as a linear order. However, in computational social choice and the associated multiagent applications, it is often desirable to allow for partial orders. There are at least two important reasons for this. First, sometimes an agent is simply unable to decide which of two alternatives should be ranked higher (that is, the two alternatives are *incomparable* [14]). Second, sometimes the number of alternatives is extremely large, so that it is infeasible from a computation/communication standpoint for the agents to fully report their preferences over the whole set of alternatives (that is, the agents' preferences are *incomplete* [14]).

For example, in *multi-issue domains*, the set of alternatives is exponentially large. In this context, researchers have investigated voting methods that are based on the agents representing their preferences by CP-nets [4; 17; 14; 13; 23]. A CP-net corresponds to a specific type of partial order over the alternatives. As another example, perhaps only parts of the agents' preferences have been revealed so far, perhaps through an incremental elicitation process, and in this case the currently available information can be represented by a partial order for each agent. In such a situation, it is useful to know whether a given alternative can still / must certainly win the election at this point (the *possible / necessary winner* problem) [12; 20; 3; 2; 21].

Perhaps because these issues have not been central in the traditional social choice literature, there has not been much research on designing good voting rules that take partial orders as input. It is not difficult to make up such a rule in an adhoc fashion. For example, Rademaker and De Baets [16] proposed a computationally tractable method to aggregate partial orders. Their method in some sense extends a common voting rule called *ranked pairs* [18]. However, we feel that a more principled approach is desirable. For example, we may follow the *axiomatic* approach. This approach reigns supreme in social choice theory, with for example the celebrated example of Arrow's impossibility theorem [1]. Some desirable properties (axioms) of a voting rule are specified, and then the corresponding class of voting rules is characterized. Pini et al. [15] applied the axiomatic approach to voting with partial orders and obtained some significant impossibility theorems.

Much older, and arguably more constructive, is the *maximum likelihood approach*, which was first introduced more than two hundred years ago by Condorcet [7], and has recently been adopted in economics [8; 19], as well as in the artificial intelligence and computational social choice literatures [6; 5; 23]. The idea is to imagine that there is a "correct" outcome, and each voter's preferences constitute a noisy perception of this absolute truth. In Condorcet's model, given the "correct" ranking V, for each pair of alternatives c and c' where $c \succ_V c'$, with probability $p > 1/2$ the voter also prefers c to c' (and with probability $1-p$, she prefers c' to c). This "pairwise independence" assumption might result in intransitive preferences, e.g. $a \succ b$, $b \succ c$, and $c \succ a$. However, this does not cause a problem. We can simply conceptually allow for the possibility that a voter has a vote that violates transitivity.

This gives us a voting rule that is defined on a larger domain that allows for intransitivity, which can then still be applied to the subdomain where transitivity holds. Given a profile, a winning full ranking is defined to be one that maximizes the likelihood of the agents' votes. Condorcet's original MLE model results in Kemeny's rule [26]. It was later shown that some, but not all, other commonly used voting rules are MLEs for some noise models [6]. MLE rules have been characterized as a class of voting rules called *ranking scoring rules* [5]. The MLE approach has also been pursued for voting in multi-issue domains [23].

In this paper, we pursue the MLE approach in the context of voting with partial orders. An illustrative example is as follows. We have a set of candidates for $k \geq 1$ positions (for example, Ph.D. applicants), and a committee that evaluates these candidates. A committee member will generally not give a linear order of all candidates, being both time-constrained and not always comfortable comparing a given pair of candidates (perhaps due to area mismatch). In this context, it is not unreasonable to suppose that some candidates are, in some sense, truly better than others, but of course there is noise in the process. Our goal may be to get an aggregate ranking of all candidates, or to designate the top k alternatives.

2 Preliminaries

Let $\mathcal{X} = \{c_1, \ldots, c_m\}$ be a finite set of *alternatives* (or *candidates*) and O be a finite set of *outcomes*. Despite being finite, the size of $\mathcal{X}$ or O is not constrained. In traditional voting systems, a *vote* V is a linear order over $\mathcal{X}$, i.e., a transitive, antisymmetric, and total relation over $\mathcal{X}$. The set of all linear orders on $\mathcal{X}$ is denoted by $L(\mathcal{X})$. In this paper, we allow a vote to be a *partial order* over $\mathcal{X}$, i.e., a transitive and antisymmetric relation over $\mathcal{X}$. The set of all partial orders on $\mathcal{X}$ is denoted by $\mathrm{PO}(\mathcal{X})$. For any pair of alternatives c, c' and any partial order W, we write $c \sim_W c'$ if c and c' are not comparable in the partial order W. We note that $L(\mathcal{X}) \subseteq \mathrm{PO}(\mathcal{X})$. An n-voter profile P over $L(\mathcal{X})$ (respectively, $\mathrm{PO}(\mathcal{X})$) can be written as $P = (V_1, \ldots, V_n)$, where $V_j \in L(\mathcal{X})$ (respectively, $V_j \in \mathrm{PO}(\mathcal{X})$) for every $j \leq n$. The set of all profiles over $L(\mathcal{X})$ (respectively, $\mathrm{PO}(\mathcal{X})$) is denoted by $\mathrm{F}_L(\mathcal{X})$ (respectively, $\mathrm{F}_{\mathrm{PO}}(\mathcal{X})$). In this paper, a *(voting) rule* r maps any profile to a subset of outcomes. Two example outcome spaces are $\mathcal{X}$, the set of all alternatives, and $L(\mathcal{X})$, the set of all linear orders.[1] A voting rule r_{PO} over profiles of partial orders is an *extension* of a voting rule r_L over profiles of linear orders, if for any $P \in \mathrm{F}_L(\mathcal{X})$, we have $r_{\mathrm{PO}}(P) = r_L(P)$. In this case r_L is called the *restriction* of r_{PO} on $\mathrm{F}_L(\mathcal{X})$. The following common voting rules over $\mathrm{F}_L(\mathcal{X})$ play a role in our paper.

• *Positional scoring rules.* Here we let $O = \mathcal{X}$. Given a *scoring vector* $\vec{v} = (v_1, \ldots, v_m)$ of m integers, for any vote $V \in L(\mathcal{X})$ and any $c \in \mathcal{X}$, let $\mathrm{PS}_{\vec{v}}(V, c) = v_i$, where i is the rank of c in V. For any profile $P = (V_1, \ldots, V_n)$, let $\mathrm{PS}_{\vec{v}}(P, c) = \sum_{j=1}^{n} \mathrm{PS}_{\vec{v}}(V_j, c)$. The rule will select all alternatives $c^* \in \mathcal{X}$ so that $\mathrm{PS}_{\vec{v}}(P, c^*)$ is maximized. One important positional scoring rule is *Borda*, for which the scoring vector is $(m-1, m-2, \ldots, 0)$.

• *Kemeny.* Here we let $O = L(\mathcal{X})$. For any pair of linear orders V and V' and any pair of different alternatives c, c', let $\delta_{c,c'}(V, V')$ equal 1 if V and V' agree on which of c and c' is ranked higher. The Kemeny rule will select all linear orders V^* so that $\sum_{V \in P} \sum_{c \neq c'} \delta_{c,c'}(V, V^*)$ is maximized.

For any profile P composed of linear orders, we let $\mathrm{WMG}(P)$ denote the *weighted majority graph* of P, defined as follows. $\mathrm{WMG}(P)$ is a directed graph whose vertices are the alternatives. For $i \neq j$, let $D_P(c_i, c_j)$ denote the total number of votes in P where $c_i \succ c_j$ minus the total number of votes in P where $c_j \succ c_i$. If $D_P(c_i, c_j) > 0$, then there is an edge (c_i, c_j) with weight $w_{ij} = D_P(c_i, c_j)$. Also, for $i < j$, if $D_P(c_i, c_j) = 0$, then there is an edge (c_i, c_j) with weight $w_{ij} = 0$.

We say a voting rule r (defined over $\mathrm{F}_L(\mathcal{X})$) is *based on the WMG* if for any pair of profiles P_1, P_2 such that $\mathrm{WMG}(P_1) = \mathrm{WMG}(P_2)$, we have $r(P_1) = r(P_2)$. We say a voting r satisfies *non-imposition* if for any alternative c and any natural number n, there exists an n-profile P such that $r(P) = c$. A voting rule r satisfies *neutrality* if r treats the alternatives symmetrically. More precisely, for any permutation M over $\mathcal{X}$ and any profile P, we have $r(M(P)) = M(r(P))$.

Maximum Likelihood Approach to Voting

In the maximum likelihood approach to voting, it is assumed that there is an unobserved correct outcome $o^+ \in O$—for example, a correct ranking $V^+ \in L(\mathcal{X})$ or a correct winner $c^+ \in \mathcal{X}$—and each vote V is drawn conditionally independently given o^+, according to a conditional probability distribution $\pi(V|o^+)$. The independence structure of the noise model is illustrated by the Bayesian network in Figure 1.[2]

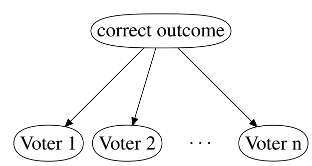

Figure 1: The noise model.

Under this independence assumption, the probability of a profile $P = (V_1, \ldots, V_n)$ given the correct outcome o^+ is $\pi(P|o^+) = \prod_{j=1}^{n} \pi(V_j|o^+)$. Then, the maximum likelihood estimate of the correct winner is $MLE_\pi(P) = \arg\max_{o \in O} \pi(P|o)$. MLE_π is set-valued, as there may be several outcomes o^* that maximize $\pi(P|o^*)$. In this paper, we require all the conditional probabilities to be strictly positive, for technical reasons.

Under the model where outcomes are rankings, it has been shown that a neutral voting rule is an MLE for some noise model if and only if it is a neutral *ranking scoring rule* [5]. More generally, we define an *outcome scoring rule* r_{s^*} to be a voting rule defined by a scoring function $s^* : L(\mathcal{X}) \times O \to \mathbb{R}$, as follows: for any profile P,

[1]The former type of mapping is also known as a *voting correspondence*, and the latter as a *preference function*. In this paper, we simply call them voting rules for convenience.

[2]The use of this independence structure is standard. Moreover, if conditional independence among votes is not required, then any voting rule can be represented by an MLE for some noise model [6], which trivializes the question.

$r_{s^*}(P) = \arg\max_{o\in O} \sum_{V\in P} s^*(V, o)$. If $O = L(\mathcal{X})$, we obtain ranking scoring rules, and if $O = \mathcal{X}$, we obtain candidate scoring rules. An outcome scoring rule is *weakly neutral* if for any pair of outcomes o and o', there exists a permutation M over $\mathcal{X}$ such that for any linear order $V \in L(\mathcal{X})$, we have $s^*(V, o) = s^*(M(V), o')$.

3 MLEs and Pairwise Scoring Rules

We now move on to the contributions of this paper. We recall that the objective of this paper is to study noise models and the associated MLE voting rules for voting with partial orders. Ideally, in the special case where the profiles consist of linear orders, the MLE of such a noise model coincides with a commonly used voting rule. We will mainly study three types of outcome space: (1) an outcome is an alternative ($O = \mathcal{X}$); (2) an outcome is a linear order ($O = L(\mathcal{X})$); and (3) an outcome is a subset of k alternatives ($O = \{S \subseteq \mathcal{X} : |S| = k\}$).

The model that we are about to introduce is similar to Condorcet's noise model, but it is much more general, in terms of both the input profiles and the outcome space. For any outcome $o \in O$ and any pair of alternatives c_i, c_j, let $\pi(c_i \succ c_j|o)$ denote the probability that $c_i \succ c_j$ in a vote given correct outcome o, and let $\pi(c_i \sim c_j|o)$ denote the probability that c_i and c_j are incomparable given correct outcome o. We have $\pi(c_i \succ c_j|o) + \pi(c_j \succ c_i|o) + \pi(c_i \sim c_j|o) = 1$ because these events are disjoint and exhaustive.

We assume that conditional on the correct outcome $o \in O$, for any voter, all the pairwise comparisons are drawn *independently* according to the conditional probability distribution $\pi(\cdot|o)$. That is, for any partial order W over $\mathcal{X}$ and any pair of alternatives c_i, c_j ($i \neq j$), we define:

$$\pi(W_{c_i,c_j}|o) = \begin{cases} \pi(c_i \succ c_j|o) & \text{if } c_i \succ_W c_j \\ \pi(c_j \succ c_i|o) & \text{if } c_j \succ_W c_i \\ \pi(c_i \sim c_j|o) & \text{if } c_i \sim_W c_j \end{cases}$$

Then, for any partial order W over $\mathcal{X}$ and any outcome $o \in O$, we let $\pi(W|o) = \prod_{1\le i<j\le m} \pi(W_{c_i,c_j}|o)$ be the probability of W given correct outcome o. For any profile $P = (W_1, \ldots, W_n)$ of partial orders, we obtain that the probability of this profile given correct outcome o is $\pi(P|o) = \prod_{j=1}^{n} \pi(W_j|o)$ (using the standard assumption of independence across voters). We call such a noise model π a *pairwise-independent model*. For any pairwise-independent model π and any profile P of partial orders, we let $\text{MLE}_\pi(P) = \arg\max_{o\in O} \pi(P|o)$. MLE_π is a voting rule defined over profiles of partial orders.

We say that a pairwise-independent model is *weakly neutral* if for any pair of outcomes $o, o' \in O$, there exists a permutation M over $\mathcal{X}$ such that for any $c_i, c_j \in \mathcal{X}$, $\pi(c_i \succ c_j|o) = \pi(M(c_i) \succ M(c_j)|o')$.

Next, we define pairwise scoring rules, which are voting rules defined by *pairwise scoring functions*. We introduce pairwise scoring rules mainly because they are defined in a similar way as ranking/candidate scoring rules [5], and the relationship between pairwise scoring rules and MLEs of pairwise independence models is similar to the relationship between ranking/candidate scoring rules and MLEs in the linear-orders setting.

Definition 1 *A* pairwise scoring function *is a function $s : \mathcal{X} \times \mathcal{X} \times O \to \mathbb{R}$, where for any $c \in \mathcal{X}$ and any $o \in O$, $s(c, c, o) = 0$. We overload s as follows. For any partial order $W \in \text{PO}(\mathcal{X})$ and any outcome $o \in O$, we let $s(W, o) = \sum_{(c_i,c_j)\in W} s(c_i, c_j, o)$. For any profile $P = (W_1, \ldots, W_n)$ of partial orders over $\mathcal{X}$ and any outcome $o \in O$, we let $s(P, o) = \sum_{j=1}^{n} s(W_j, o)$. The* pairwise scoring rule *r_s over profiles of partial orders is defined by $r_s(P) = \arg\max_{o\in O} s(P, o)$.*

For notational convenience, we write $s(c \succ c', o)$ instead of $s(c, c', o)$.

Definition 2 *A pairwise scoring function s is* weakly neutral *if for any pair of outcomes $o, o' \in O$, there exists a permutation M over $\mathcal{X}$ such that for any $c_i, c_j \in \mathcal{X}$, $s(c_i \succ c_j, o) = s(M(c_i) \succ M(c_j), o')$.*

Theorem 1 *A voting rule is a pairwise scoring rule with a weakly neutral pairwise scoring function if and only if it is the MLE of a weakly neutral pairwise-independent model.*

Proof: For any weakly neutral pairwise-independent model π, we define a pairwise scoring function s as follows. For any $o \in O$ and any pair of different alternatives (c_i, c_j), we let $s(c_i \succ c_j, o) = \log \pi(c_i \succ c_j|o) - \log \pi(c_i \sim c_j|o)$. It follows that for any partial order W, $s(W, o) = \sum_{(c_i,c_j)\in W} \log \pi(c_i \succ c_j|o) - \sum_{(c_i,c_j)\in W} \log \pi(c_i \sim c_j|o) = \sum_{i<j} \log \pi(W_{c_i,c_j}|o) - \sum_{i<j} \log \pi(c_i \sim c_j|o) = \log \pi(W|o) - \sum_{i<j} \log \pi(c_i \sim c_j|o)$. Because π is weakly neutral, for any pair of outcomes $o, o' \in O$, $\sum_{i<j} \log \pi(c_i \sim c_j|o) = \sum_{i<j} \log \pi(c_i \sim c_j|o')$. Hence, $\arg\max_{o\in O} \pi(P|o) = \arg\max_{o\in O} \sum_j \log \pi(W_j|o) = \arg\max_{o\in O} \sum_j s(W_j, o) = \arg\max_{o\in O} s(P, o)$. Therefore, r_s is equivalent to MLE_π. It is easy to verify that s is weakly neutral (using the weak neutrality of π).

Conversely, let r_s be a pairwise scoring rule, where s is weakly neutral. Let $O = \{o_1, \ldots, o_t\}$. For any $2 \le h \le t$, let M_h denote the permutation over $\mathcal{X}$ such that for any pair of different alternatives c_i, c_j, $s(c_i \succ c_j, o_1) = s(M_h(c_i) \succ M_h(c_j), o_h)$. For any alternative c, let $M_1(c) = c$. For any $i < j \le m$, we let $b_{i,j}$ be a constant such that $2^{s(c_i\succ c_j,o)+b_{i,j}} + 2^{s(c_j\succ c_i,o)+b_{i,j}} + 2^{b_{i,j}} = 1$. The existence of such a $b_{i,j}$ is guaranteed by the intermediate value theorem (when $b_{i,j} = \infty$, the left-hand side is $\infty > 1$, and when $b_{i,j} = -\infty$, the left-hand side is $0 < 1$). Now, we define a pairwise-independent model π as follows: for any $h \le t$ and any $i < j$, $\pi(M_h(c_i) \succ M_h(c_j)|o_h) = 2^{s(M_h(c_i)\succ M_h(c_j),o_h)+b_{i,j}}$, $\pi(M_h(c_j) \succ M_h(c_i)|o_h) = 2^{s(M_h(c_j)\succ M_h(c_i),o_h)+b_{i,j}}$, and $\pi(M_h(c_i) \sim M_h(c_j)|o_h) = 2^{b_{i,j}}$. Because s is weakly neutral, it follows that π is weakly neutral. We note that for any $o_h \in O$ and any partial order W, $\log \pi(W|o_h) = \sum_{i<j} \log \pi(W_{c_i,c_j}|o) = s(W, o) + \sum_{i<j} b_{i,j}$, implying that $\text{MLE}_\pi = r_s$. □

We now give some examples of pairwise scoring rules with weakly neutral scoring functions. The first family of pairwise scoring rules consists of extensions of Borda, and the second family consists of extensions of Kemeny.

Proposition 1 (Borda extensions) *Let $O = \mathcal{X}$ and fix some a and b, with $a > b$ and $a \ge 0 \ge b$. For any $c \in O$ and* *any pair of alternatives c_i, c_j ($i \neq j$), we define the weakly neutral pairwise scoring function as follows.*

$$s(c_i \succ c_j, c) = \begin{cases} a & \text{if } c = c_i \\ b & \text{if } c = c_j \\ 0 & \text{otherwise} \end{cases}$$

For any profile P of linear orders, $r_s(P) = \text{Borda}(P)$.

Proposition 2 (Kemeny extensions) *Let $O = L(\mathcal{X})$ and fix some a and b, with $a > b$ and $a \geq 0 \geq b$. For any $l \in O$ and any pair of alternatives c_i, c_j ($i \neq j$), we define the weakly neutral pairwise scoring function as follows.*

$$s(c_i \succ c_j, l) = \begin{cases} a & \text{if } c_i \succ_l c_j \\ b & \text{if } c_j \succ_l c_i \\ 0 & \text{otherwise} \end{cases}$$

For any profile P of linear orders, $r_s(P) = \text{Kemeny}(P)$.

While the various Borda extensions are indeed different from each other—for example, a candidate that is compared to others very often benefits if $|a| > |b|$—the Kemeny extensions in fact all correspond to the same rule. This is because the sum of the number of a's and the number of b's that $l \in L(\mathcal{X})$ receives does not depend on l: it is equal to the total number of pairwise comparisons made by voters.

4 Properties of Pairwise Scoring Rules

In this section, we study some basic properties of pairwise scoring rules. A voting rule is *anonymous* if it treats the voters symmetrically. A voting rule r is *consistent* if for any pair of profiles P_1, P_2 of partial orders, if $r(P_1) \cap r(P_2) \neq \emptyset$, then $r(P_1 \cup P_2) = r(P_1) \cap r(P_2)$. For any profile P of partial orders, we define the *generalized weighted majority graph (GWMG)*, denoted by $\text{GWMG}(P)$, as follows. $\text{GWMG}(P)$ is a directed graph whose vertices are the alternatives, and for each ordered pair of different alternatives c_i, c_j, there is an edge $c_i \rightarrow c_j$ whose weight $w_{i,j}$ is the proportion of agents voting $c_i \succ c_j$ in P. That is, $w_{i,j} = |\{W \in P : c_i \succ_W c_j\}|/|P|$. We say a voting rule r defined over profiles of partial orders is *based on the GWMG* if for any pair of profiles P_1, P_2 of partial orders such that $|P_1| = |P_2|$ and $\text{GWMG}(P_1) = \text{GWMG}(P_2)$, we have $r(P_1) = r(P_2)$. GWMG-based voting rules over profiles of linear orders are defined similarly by requiring that both P_1 and P_2 are profiles of linear orders in the above definition. We note that any WMG-based voting rule (defined over profiles over linear orders) is GWMG-based. The following is straightforward to verify:

Proposition 3 *Any pairwise scoring rule r_s satisfies anonymity and consistency, and is based on GWMG.*

The next proposition characterizes all pairwise scoring rules that, when we consider their restriction on profiles of linear orders, are based on the WMG.

Proposition 4 *For any pairwise scoring rule r_s, its restriction on profiles of linear orders is based on the WMG if and only if for any $o_1, o_2 \in O$, $\sum_{i \neq j} s(c_i \succ c_j, o_1) = \sum_{i \neq j} s(c_i \succ c_j, o_2)$.*

When the set of outcomes is the set of all alternatives, the next theorem characterizes all pairwise scoring rules whose restriction on profiles of linear orders is neutral.

Theorem 2 *Let $O = \mathcal{X}$. For any pairwise scoring rule r_s, its restriction r'_s on profiles of linear orders satisfies neutrality if and only if r'_s is Borda.*

Proof: By Proposition 3, r_s satisfies anonymity and consistency. Therefore, r'_s also satisfies anonymity and consistency. By Young's axiomatic characterization of positional scoring rules [25], a voting rule with $O = \mathcal{X}$ satisfies continuity (which is satisfied by all pairwise scoring rules), anonymity, neutrality, and consistency if and only if it is a positional scoring rule. Therefore, r'_s is a positional scoring rule. Let $(v_1, \ldots, v_m)$ denote the scoring vector. It suffices to prove that for any $2 \leq i' \leq m-1$, $v_1 - v_2 = v_{i'} - v_{i'+1}$ (it should be noted that shifting the Borda score vector by a constant does not change the rule).

For the sake of contradiction, suppose that $2 \leq i' \leq m$ is such that $v_1 - v_2 \neq v_{i'} - v_{i'+1}$. We next derive a contradiction for the case where $v_1 - v_2 > v_{i'} - v_{i'+1}$; the case where $v_1 - v_2 < v_{i'} - v_{i'+1}$ can be handled similarly. Let M_{m-2} denote the cyclic permutation among $\{c_3, \ldots, c_m\}$. That is, $M_{m-2} : c_3 \rightarrow c_4 \rightarrow \cdots \rightarrow c_m \rightarrow c_3$. For any permutation M and any natural number k, we let M^k denote the permutation such that $M^k(x) = M(M^{k-1}(x))$ (where $M^1 = M$). Let M_2 denote the cyclic permutation among $\{c_1, c_2\}$. Let $V = [c_1 \succ c_2 \succ \cdots \succ c_m]$. We define a profile P_1 as follows. $P_1 = (V, M_2(V), M_{m-2}(V), M_{m-2}(M_2(V)), \ldots, M_{m-2}^{m-3}(V), M_{m-2}^{m-3}(M_2(V)))$. It follows that for any $c \in \mathcal{X} \setminus \{c_1, c_2\}$, $\text{PS}(P_1, c_1) = \text{PS}(P_1, c_2) > \text{PS}(P_1, c)$, which means that $r'_s(P_1) = \{c_1, c_2\}$ and $s(P_1, c_1) = s(P_1, c_2) > s(P_1, c)$.

Let V' be an arbitrary linear order where c_1 is ranked in the i'th position and c_2 is ranked in the $(i'+1)$th position. Let $P_2 = (V, M_2(V), V', M_2(V'))$. We have $\text{PS}(P_2, c_1) = \text{PS}(P_2, c_2)$. Hence, there exists a natural number k such that $r'_s(kP_1 \cup P_2) = \{c_1, c_2\}$, where kP_1 represents a profile that is composed of k copies of the votes in P_1. Let $P'_2 = (V, V, M_2(V'), M_2(V'))$. Because $v_1 - v_2 > v_{i'} - v_{i'+1}$, $\text{PS}(kP_1 \cup P'_2, c_1) > \text{PS}(kP_1 \cup P_2, c_1) = \text{PS}(kP_1 \cup P_2, c_2) > \text{PS}(kP_1 \cup P'_2, c_2)$. Moreover, for any alternative $c \in \mathcal{X} \setminus \{c_1, c_2\}$, we have $\text{PS}(kP_1 \cup P'_2, c_1) > \text{PS}(kP_1 \cup P_2, c_1) > \text{PS}(kP_1 \cup P_2, c) = \text{PS}(kP_1 \cup P'_2, c)$. Hence, $r'_s(kP_1 \cup P'_2) = \{c_1\}$. However, because $\text{GWMG}(kP_1 \cup P'_2) = \text{GWMG}(kP_1 \cup P_2)$ and r_s is based on GWMG (Proposition 3), $r'_s(kP_1 \cup P'_2) = r'_s(kP_1 \cup P_2)$, which is a contradiction. $\square$

We recall that Borda can indeed be represented as a pairwise scoring rule with a weakly neutral pairwise scoring function (Proposition 1), which means that it can be naturally extended to profiles of partial orders by our MLE approach. On the other hand, Theorem 2 suggests that, when $O = \mathcal{X}$, no other neutral voting rules can be extended to profiles of partial orders by our MLE approach. (In Section 5, we will see an example where $O \neq \mathcal{X}$.) We note that there is other work on desirable and axiomatizing properties of Borda [10; 24].

We now investigate the relationship between pairwise scoring rules and outcome scoring rules (incl. ranking/candidate scoring rules).

Proposition 5 *For any pairwise scoring rule r_s, there exists a GWMG-based outcome scoring rule r_{s^*} such that for any* *profile P of linear orders, $r_s(P) = r_{s^*}(P)$. Moreover, if s is weakly neutral, then s^* is also weakly neutral.*

The next theorem states that any WMG-based outcome scoring rule r_{s^*} can be extended to a pairwise scoring rule r_s. This is more difficult to prove.

Theorem 3 *For any WMG-based outcome scoring rule r_{s^*} that satisfies non-imposition, there exists a pairwise scoring rule r_s that is an extension of r_{s^*}. Moreover, if s^* is weakly neutral, then s is also weakly neutral.*

Proof: Let $O = \{o_1, \ldots, o_t\}$. For any linear order V, we let $f(V) = (0, s^*(V, o_2) - s^*(V, o_1), \ldots, s^*(V, o_t) - s^*(V, o_1))$. For any profile P of linear orders, we let $f(P) = \sum_{V \in P} f(V)$. We first show the following claim whose proof is omitted.

Claim 1 *For any profile P of linear orders, if the weight of every edge in WMG(P) is zero, then $f(P) = \vec{0}$.*

Next, for any pair of alternatives c_i, c_j ($i \neq j$), we define the following two votes.

$$V_{i,j} = c_i \succ c_j \succ \mathcal{X} \setminus \{c_i, c_j\}$$
$$V'_{i,j} = \text{rev}(\mathcal{X} \setminus \{c_i, c_j\}) \succ c_i \succ c_j$$

Here the alternatives in $\mathcal{X} \setminus \{c_i, c_j\}$ are ranked in an arbitrary way, and in rev($\mathcal{X} \setminus \{c_i, c_j\}$), the alternatives in $\mathcal{X} \setminus \{c_i, c_j\}$ are ranked in the reversed order. Let $\mathcal{M}_{-\{i,j\}}$ denote the set of all permutations over $\mathcal{X} \setminus \{c_i, c_j\}$. We note that $|\mathcal{M}_{-\{i,j\}}| = (m-2)!$. Let $P_{i,j} = \{M(V_{i,j}), M(V'_{i,j}) : M \in \mathcal{M}_{-\{i,j\}}\}$.

Next, we prove that for any linear order V, $f(V) = \sum_{c_i \succ_V c_j} f(P_{i,j})/(2(m-2)!)$. Let $P_1 = \bigcup_{c_i \succ_V c_j} P_{i,j}$. Let $P_2 = \text{rev}(P_1)$. We note that the WMG of $2(m-2)!V$ is the same as the WMG of P_1, and the weight of each edge in the WMG of $P_1 \cup P_2$ is 0. Using Claim 1, we obtain $f(P_1 \cup P_2) = f(P_1) + f(P_2) = \vec{0}$ and $f(\{2(m-2)!V\} \cup P_2) = 2(m-2)!f(V) + f(P_2) = \vec{0}$. Therefore, $f(P_1) = 2(m-2)!f(V)$.

For any $h \leq t$, we let $s(c_i \succ c_j, o_h) = s^*(P_{i,j}, o_h)$. It follows that $f(V) = \sum_{c_i \succ_V c_j} f(P_{i,j})/(2(m-2)!) = \sum_{c_i \succ_V c_j}[(s(c_i \succ c_j, o_1), s(c_i \succ c_j, o_2), \ldots, s(c_i \succ c_j, o_t)) - s(c_i \succ c_j, o_1) \cdot \vec{1}]/(2(m-2)!) = [(s(V, o_1), s(V, o_2), \ldots, s(V, o_t)) - s(V, o_1) \cdot \vec{1}]/(2(m-2)!)$. Therefore, for any $h \leq t$ and any linear order V, $s^*(V, o_h) - s^*(V, o_1) = (s(V, o_h) - s(V, o_1))/(2(m-2)!)$. Hence, for any profile P of linear orders and any $h_1, h_2 \leq t$, $s^*(P, o_{h_1}) \geq s^*(P, o_{h_2})$ if and only if $s(P, o_{h_1}) \geq s(P, o_{h_2})$. It follows that r_s is an extension of r_{s^*}.

Suppose s^* is weakly neutral. Let $M_1, \ldots, M_t$ denote the permutations for the pairs of outcomes $(o_1, o_1), (o_1, o_2), \ldots, (o_1, o_t)$, respectively. For any $i \neq j$ and any $h \leq t$, $M_h(P_{i,j}) = \{M_h(M(V_{i,j})), M_h(M(V'_{i,j})) : M \in \mathcal{M}_{-\{i,j\}}\} = \{M'(V_{M_h(i), M_h(j)}), M'(V'_{M_h(i), M_h(j)}) : M' \in \mathcal{M}_{-\{M_h(i), M_h(j)\}}\} = P_{M_h(i), M_h(j)}$. Therefore, $s(M_h(c_i) \succ M_h(c_j), o_h) = 2(m-2)!s^*(P_{M_h(i), M_h(j)}, o_h) = 2(m-2)!s^*(M_h(P_{i,j}), o_h) = 2(m-2)!s^*(P_{i,j}, o_1) = s(P_{i,j}, o_1)$, which means that s is also weakly neutral. □

Putting Proposition 5 and Theorem 3 together, we obtain an almost complete characterization of the voting rules that can be extended to pairwise scoring rules. The necessary condition (Proposition 5) states that it must be an outcome scoring rule r that is based on the GWMG, and the sufficient condition (Theorem 3) states that if r is based on the WMG (and to satisfy non-imposition), then r can be extended to a pairwise scoring rule. How to close the gap between the GWMG condition and the WGM condition is left as an open question.

5 Selecting a Set of k Winning Alternatives

In this section, we propose a pairwise-independent probability model π_k for the case where the outcome space is composed of all sets of k alternatives (for a given natural number k). For instance, perhaps our goal is to select k candidates to admit/hire based on partial orders given by a committee. We now present the MLE model π_k formally. For any $k \leq m$, let $O_k = \{S \subseteq \mathcal{X} : |S| = k\}$ denote the outcome space. The main idea behind this model is that, given one of the winners and one of the losers in the correct outcome, a voter is more likely to rank the former above the latter than vice versa.

Definition 3 *Let $0 < \bar{p}, p_1, p_2 < 1$ and $(1-\bar{p})/2 < p < 1-\bar{p}$. For any $o \in O_k$, the conditional probabilities in π_k are defined as follows.*

(a) For any pair of alternatives c_i, c_j such that $c_i, c_j \in o$, $\pi_k(c_i \sim c_j | o) = p_1$ and $\pi_k(c_i \succ c_j | o) = \pi_k(c_j \succ c_i | o) = (1-p_1)/2$.

(b) For any pair of alternatives c_i, c_j such that $c_i, c_j \in \mathcal{X} \setminus o$, $\pi_k(c_i \sim c_j | o) = p_2$ and $\pi_k(c_i \succ c_j | o) = \pi_k(c_j \succ c_i | o) = (1-p_2)/2$.

(c) For any $c_i \in o$ and $c_j \in \mathcal{X} \setminus o$, $\pi_k(c_i \sim c_j | o) = \bar{p}$, $\pi_k(c_i \succ c_j | o) = p$, and $\pi_k(c_j \succ c_i | o) = 1 - p - \bar{p}$.

By Theorem 1, MLE$_{\pi_k}$ is a pairwise scoring rule r_s whose scoring function is defined as follows. The base of all logarithms is 2.

$$s(c_i \succ c_j, o) = \begin{cases} \log(p/\bar{p}) & \text{if } c_i \in o \text{ and } c_j \notin o \\ \log((1-p-\bar{p})/\bar{p}) & \text{if } c_j \in o \text{ and } c_i \notin o \\ \log((1-p_1)/p_1) & \text{if } c_i, c_j \in o \\ \log((1-p_2)/p_2) & \text{if } c_i, c_j \in \mathcal{X} \setminus o \end{cases}$$

We now define (the decision variant of) the problem of maximizing the likelihood under π_k. In the LIKELIHOOD-EVALUATION problem (L-EVALUATION for short), we are given a natural number k, a positive number t, and a profile P consisting of partial orders; we are asked whether there exists $o \in O_k$ such that $\pi_k(P|o) \geq t$.

Theorem 4 *The* L-EVALUATION *problem is NP-complete.*

Proof sketch: Given an outcome o and a profile P, it is straightforward to check that it takes polynomial time to compute $\pi_k(P|o)$. Therefore, L-EVALUATION is in NP. We prove NP-hardness by a reduction from the decision variant of MAX-CUT, which is known to be NP-complete [11]. We are given an undirected graph $G = (\mathcal{V}, E)$ and a natural number t'. We are asked whether there exists a partition of $\mathcal{V}$ into two sets, $\mathcal{V}_1$ and $\mathcal{V}_2$, such that the number of edges across $\mathcal{V}_1$ and $\mathcal{V}_2$ is at least t'.

We construct an instance of L-EVALUATION as follows. Let $\mathcal{V}'$ be a set of dummy alternatives with $|\mathcal{V}'| = |\mathcal{V}|$. Let $\mathcal{X} = \mathcal{V} \cup \mathcal{V}'$ and $k = |\mathcal{V}|$. None of the dummy alternatives will ever be compared by a voter to another alternative; the point of these dummy alternatives is merely to make sure that any partition $S_1 \cup S_2 = \mathcal{V}$ corresponds to a partition $S_1' \cup S_2' = \mathcal{V} \cup \mathcal{V}'$ such that $|S_1'| = |S_2'| = |\mathcal{V}| = k$. In π_k, let $p_1 = p_2 = 1/3$, $\bar{p} = 1/4$, and $p = 1/2$.

For each undirected edge $\{c_i, c_j\} \in E$, we define two votes: $W_{i,j}$, which contains only one pairwise comparison, $c_i \succ c_j$; and $W_{j,i}$, which contains only one pairwise comparison, $c_j \succ c_i$. Let P be the profile that consists of all these votes $W_{i,j}, W_{j,i}$ for all $\{c_i, c_j\} \in E$. Let $C = (\log 1/3)|P|(k(k-1) + (m-k)(m-k-1))/2 + (\log 1/4)|P|k(m-k)$, that is, $C = \sum_{c_i \neq c_j} \log \pi(c_i \sim c_j | o)$ for any $o \in O_k$; let $t = 2^{t'+C}$.

For any $o \in O_k$, let $\text{CUT}(o)$ denote the number of edges in E across o and $\mathcal{V} \setminus o$. We can show that the likelihood of o is $2^{\text{CUT}(o)+C}$, which is at least $t = 2^{t'+C}$ if and only if $\text{CUT}(o) \geq t'$—that is, if and only if $(o \cap \mathcal{V}, \mathcal{V} \setminus o)$ is a cut of $\mathcal{V}$ whose size is at least t'. □

We can also show that choosing an outcome to maximize the log-likelihood is APX-hard and MAXSNP-hard, by a PTAS-reduction and an L-reduction, both from the MAX-CUT problem. The reductions are similar to the reduction in the proof of Theorem 4. It is also possible to give a simple integer program with binary variables for computing outcomes of MLE_{π_k}. We omit all this due to the space constraint.

6 Future Work

One direction for future research is to use these techniques in applications. The example from the introduction of committee members giving partial orders of Ph.D. applicants was not fictitious: with the help of an integer program solver, we have already used the extension of the Kemeny rule to partial orders (Proposition 2) to help us rank Ph.D. applicants to our department. There are also open technical questions. For example, what is the exact complexity of computing an arbitrary winning outcome under MLE_{π_k}? (We conjecture it is Θ_2^P-complete.) Can we close the gap between the WMG and GWMG conditions described at the end of Section 4? Do the voting rules proposed in this paper satisfy other properties for voting with partial orders [15]? Can we extend other classes of voting rules to partial orders, for example generalized scoring rules [22] or distance rationalizable models [9]?

Acknowledgements

Lirong Xia is supported by a James B. Duke Fellowship. We acknowledge NSF CAREER 0953756 and IIS-0812113, and an Alfred P. Sloan fellowship for support. We thank all IJCAI-11 reviewers for their helpful comments and suggestions.

References

[1] K. J. Arrow. A difficulty in the concept of social welfare. *Journal of Political Economy*, 58(4):328–346, 1950.

[2] Nadja Betzler and Britta Dorn. Towards a dichotomy for the possible winner problem in elections based on scoring rules. *J. Comput. Syst. Sci.*, 76(8):812–836, 2010.

[3] Nadja Betzler, Susanne Hemmann, and Rolf Niedermeier. A multivariate complexity analysis of determining possible winners given incomplete votes. In *IJCAI*, pages 53–58, 2009.

[4] Craig Boutilier, Ronen Brafman, Carmel Domshlak, Holger Hoos, and David Poole. CP-nets: A tool for representing and reasoning with conditional ceteris paribus statements. *JAIR*, 21:135–191, 2004.

[5] Vincent Conitzer, Matthew Rognlie, and Lirong Xia. Preference functions that score rankings and maximum likelihood estimation. In *IJCAI*, pages 109–115, 2009.

[6] Vincent Conitzer and Tuomas Sandholm. Common voting rules as maximum likelihood estimators. In *UAI*, pages 145–152, 2005.

[7] Marie Jean Antoine Nicolas de Caritat (Marquis de Condorcet). Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix. 1785. Paris: L'Imprimerie Royale.

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

[9] Edith Elkind, Piotr Faliszewski, and Arkadii Slinko. Distance Rationalization of Voting Rules. In *COMSOC*, 2010.

[10] Peter C. Fishburn and William V. Gehrlein. Borda's rule, positional voting, and Condorcet's simple majority principle. *Public Choice*, 28(1):79–88, 1976.

[11] Michael Garey and David Johnson. *Computers and Intractability*. W. H. Freeman and Company, 1979.

[12] Kathrin Konczak and Jérôme Lang. Voting procedures with incomplete preferences. In *Multidisciplinary Workshop on Advances in Preference Handling*, 2005.

[13] Jérôme Lang and Lirong Xia. Sequential composition of voting rules in multi-issue domains. *Mathematical Social Sciences*, 57(3):304–324, 2009.

[14] Maria Silvia Pini, Francesca Rossi, Kristen Brent Venable, and Toby Walsh. Incompleteness and incomparability in preference aggregation. In *IJCAI*, 2007.

[15] Maria Silvia Pini, Francesca Rossi, Kristen Brent Venable, and Toby Walsh. Aggregating partially ordered preferences. *Journal of Logic and Computation*, 19:475–502, 2009.

[16] Michaël Rademaker and Bernard De Baets. A threshold for majority in the context of aggregating partial order relations. In *IEEE FUZZ*, pages 1–4, 2010.

[17] Francesca Rossi, Kristen Brent Venable, and Toby Walsh. mCP nets: Representing and reasoning with preferences of multiple agents. In *AAAI*, pages 729–734, 2004.

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

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

[20] Toby Walsh. Uncertainty in preference elicitation and aggregation. In *AAAI*, pages 3–8, 2007.

[21] Lirong Xia and Vincent Conitzer. Determining possible and necessary winners under common voting rules given partial orders. In *AAAI*, pages 196–201, 2008.

[22] Lirong Xia and Vincent Conitzer. Generalized scoring rules and the frequency of coalitional manipulability. In *EC*, pages 109–118, 2008.

[23] Lirong Xia, Vincent Conitzer, and Jérôme Lang. Aggregating preferences in multi-issue domains by using maximum likelihood estimators. In *AAMAS*, pages 399–406, 2010.

[24] H. Peyton Young. An axiomatization of Borda's rule. *Journal of Economic Theory*, 9(1):43–52, 1974.

[25] H. Peyton Young. Social choice scoring functions. *SIAM Journal on Applied Mathematics*, 28(4):824–838, 1975.

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