Generalized Scoring Rules: A Framework That Reconciles Borda and Condorcet

LIRONG XIA
Harvard University

Generalized scoring rules [Xia and Conitzer 08] are a relatively new class of social choice mechanisms. In this paper, we survey developments in generalized scoring rules, showing that they provide a fruitful framework to obtain general results, and also reconcile the Borda approach and Condorcet approach via a new social choice axiom. We comment on some high-level ideas behind GSRs and their connection to Machine Learning, and point out some ongoing work and future directions.

1. INTRODUCTION

Social choice theory focuses on developing principles and methods for representation and aggregation of individual ordinal preferences. Perhaps the most well-known application of social choice theory is political elections. Over centuries, many social choice mechanisms have been proposed and analyzed in the context of elections, where each agent (voter) uses a linear order over the alternatives (candidates) to represent her preferences (her *vote*). For historical reasons, we will use *voting rules* to denote social choice mechanisms, though we need to keep in mind that the application is not limited to political elections.[1] Most existing voting rules fall into one of the following two categories.[2]

Positional scoring rules: Each alternative gets some points from each agent according to its position in the agent's vote. The alternative with the highest total points wins. For example, *Borda* is a positional scoring rule where for each vote, the alternative ranked at the ith position gets $m - i$ points, where m is the number of alternatives.

Condorcet consistent rules: Whenever there exists a *Condorcet winner*, it must be the unique winner of the election. A Condorcet winner is an alternative that beats every other alternative in head-to-head comparisons. For example,

---

[1]More recently, social choice theory has been adopted in many modern computational systems, including but not limited to recommender systems [Ghosh et al. 1999], meta-search engines [Dwork et al. 2001], belief merging [Everaere et al. 2007], crowdsourcing [Mao et al. 2013].
[2]Some popular voting rules do not fall into the two categories, for example the *Single Transferable Vote (STV)*.

Maximin (a.k.a. *Simpson-Kramer*) is a Condorcet consistent rule, which selects the alternative that has the highest worst-case head-to-head wins.

One key question in social choice theory is: *Which voting rule is the best?* This is not an easy question, and there has been a long debate over even the meaning of optimality between the advocates of the above two categories, with no clear victory claimed by either side. This can date back to the battle in the 18th century between Jean-Charles de Borda, the inventor of the Borda rule, and Marquis de Condorcet, the inventor of Condorcet consistency.

The classical way to evaluate voting rules in social choice theory is to study their satisfiability of *axiomatic properties* (*axioms* in short), which are desired properties measuring various aspects of voting rules. Unfortunately, no voting rule can satisfy the combination of even a few natural axioms, due to the celebrated Arrow's impossibility theorem [Arrow 1963].[3] Specifically, no positional scoring rule is Condorcet consistent [Fishburn 1974]. So at least the Borda advocates and Condorcet advocates can proudly announce "We are different from the opponent".

1.1 Our Approach

Instead of continuing the Borda vs. Condorcet debate and contrasting existing voting rules, we instead seek for a unified approach by asking the following question:

Do most existing voting rules share some common properties?

Notice that this is in fact a "reverse engineering" question. Knowing these common characteristics helps us understand desired properties of voting rules, so that in the future if we want to design a new voting rule, we can focus on these natural properties. More precisely, we ask the following question:

Is there a framework that reconciles the two categories of voting rules?

A straightforward (and uninformative) answer is affirmative, for example "the class of all voting rules". However, a good framework should be general, covering most existing voting rules, but more importantly, needs to have a good mathematical structure that distinguishes it from an arbitrary voting rule. This means that a good framework should not be too general.

In the rest of the paper, we will introduce the class of generalized scoring rules, and show evidences suggesting that it is a good framework for this purpose.

2. GENERALIZED SCORING RULES

We start with an example of rethinking Borda to illustrate the idea behind the definition. Let $\mathcal{A} = \{a_1, \ldots, a_m\}$ denote a set of m alternatives, and let $\mathcal{L}(\mathcal{A})$ denote the set of all linear orders over $\mathcal{A}$. Let $P = (V_1, \ldots, V_n)$ denote a *preference profile*, where each $V_j \in \mathcal{L}(\mathcal{A})$ represents the vote of agent j. A voting rule r is a mapping that chooses a single winner for any preference profile.[4]

---

[3]See [Nurmi 1987] for definitions of some natural axioms and a thorough comparison of voting rules in terms of satisfiability of these axioms.

[4]The definition of GSRs can be much more general, but for better presentation we will focus on the classical election setting in this paper.
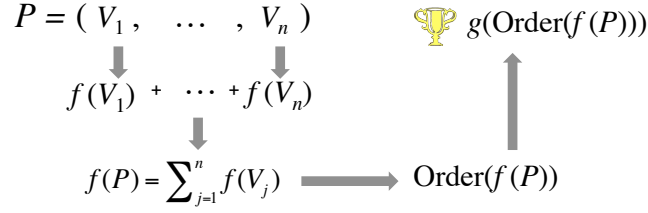
$$P = (\ V_1,\quad \dots\quad ,\ V_n\ ) \qquad\qquad \text{🏆}\ \ g(\text{Order}(f(P)))$$

$$f(V_1)\ +\ \cdots\ +f(V_n)$$

$$f(P) = \sum_{j=1}^{n} f(V_j) \longrightarrow \text{Order}(f(P))$$

Fig. 1. Illustration of generalized scoring rules.

Example 1 (Rethinking Borda) *For each vote $V \in \mathcal{L}(\mathcal{A})$, we map it to an m-dimensional vector $f(V) \in \mathbb{R}^m$, where the ith component is the number of points a_i obtains in V. Given a preference profile P, we let $f(P) = \sum_{j=1}^{n} f(V_j)$. It is not hard to see that $f(P)$ represents the total points obtained by the alternatives in P. Therefore, the winner is a_i where the i is the largest component of $f(P)$.*

In the above example, we clearly see the following pattern in Borda:

(1) Each vote is mapped to a vector via a function f.
(2) The vectors are summed up to produce a total vector.
(3) The winner is determined by the *order* over the components in the total vector via a function g.

This leads to the definition of generalized scoring rules [Xia and Conitzer 2008], illustrated in Figure 1.[5] Slightly more formally, fix the number of alternatives m, we have a number K that represents the dimensionality of the vectors votes are mapped to by f. Then, a *generalized scoring rule (GSR)*, denoted by $GS(f, g)$, is defined by a pair of function f and g. For any input preference profile P, we perform exactly the above three steps.

(1) Each vote V in P is mapped to a vector $f(V) \in \mathbb{R}^K$.
(2) Let $f(P) = \sum_{V \in P} f(V)$.
(3) The winner is $g(\text{Order}(f(P)))$, where $\text{Order}(f(P))$ is the order over the components in $f(P)$.

Remark 2.1 A GSR is defined for a fixed number of alternatives and a variable number of voters.

Remark 2.2 By saying that a voting rule r is a GSR, we mean that there exist f and g such that $r = GS(f, g)$. It is possible that different pairs (f_1, g_1) and (f_2, g_2) correspond to the same voting rule.

We have seen in Example 1 that Borda is a GSR, where $K = m$, f is the function described in Example 1, and g simply selects the alternative whose corresponding component is the largest.[6] The next example shows that Maximin, which is Condorcet consistent, is also a GSR.

---

[5]We use the equivalent definition in [Xia 2012].
[6]Suppose that ties are broken w.r.t. a fixed linear order over $\mathcal{A}$.

Example 2 *To show that Maximin is a GSR, we let*

—$K_M = m(m-1)$; *the components are indexed by pairs (i,j) such that $i,j \leq m$, $i \neq j$.*

—$(f_M(V))_{(i,j)} = \begin{cases} 1 & \text{if } a_i \succ_V a_j \\ 0 & \text{otherwise} \end{cases}$

—$g_M(\cdot)$ *simulates Maximin based on the information contained in* $\text{Order}(f_M(P))$.

3. GENERALITY OF GSRS

GSRs are quite general: It was shown by construction that many commonly studied voting rules using fixed-order tie-breaking are GSRs [Xia and Conitzer 2008], including all positional scoring rules, Maximin, Copeland, ranked pairs, Bucklin, and multi-round voting rules including STV, plurality with runoff, Nanson's rule, and Baldwin's rule. Notice that many of these rules are Condorcet consistent.

GSRs are not too general: GSRs are equivalent to the class of voting rules that satisfy the following two axioms [Xia and Conitzer 2009].

—*Anonymity*: r satisfies anonymity if the winner is insensitive to the names of the agents.

—*Finite local consistency (FLC)*: r satisfies FLC if the set of all preference profiles over $\mathcal{A}$ can be partitioned into T parts $\{S_1, \ldots, S_T\}$, such that for any pair of preference profiles (P_1, P_2) that belong to the same partition and $r(P_1) = r(P_2)$, we have $r(P_1) = r(P_1 \cup P_2)$.

Remark 3.1 FLC implies *homogeneity*, which says that for any preference profile P and any number $k \in \mathbb{N}$, $r(P) = r(nP)$. Therefore, any voting rule that does not satisfy homogeneity is not a GSR. Among commonly studied voting rules, only Dodgson's rule does not satisfy homogeneity, which means that it is not a GSR.[7]

Remark 3.2 Any voting rule that does not satisfy anonymity is not a GSR, including Borda equipped with a non-anonymous tie-breaking mechanism, for example breaking ties using the first voter's vote.

Remark 3.3 FLC is an extension of the *consistency* axiom in social choice theory, which is FLC with $T = 1$. Consistency was only previously known to be satisfied by positional scoring rules.[8]

4. WHY ARE GSRS INTERESTING?

Useful in studying the frequency of manipulability. Suppose there are n' manipulators and their favorite alternative a. Let n non-manipulators votes be generated i.i.d. according to some probability distribution. We are interested in the *frequency of manipulability*, which is the probability that the n' manipulators can make a win by voting collaboratively.

---

[7]However, Dodgson's rule is arguably not a good voting rule since it fails to satisfy many desired axioms, and has a high computational complexity [Brandt 2009].
[8]C.f. Young's insightful axiomatic characterization of positional scoring rules [Young 1975].

For a large class of GSRs, we proved a dichotomy theorem on the frequency of manipulability [Xia and Conitzer 2008]. The theorem states that if the number of manipulators is $o(\sqrt{n})$, then the frequency of manipulability goes to 0 as n goes to infinity; if the number of manipulator is $\omega(n)$, then the frequency of manipulability goes to 1 as n goes to infinity.

The theorem was extended (with slight tweaks) to all GSRs [Mossel et al. 2012], and was also extended to other types of strategic behavior [Xia 2013]. These type of research was generally viewed as negative, because they reconfirm the high-level message that computational complexity is not a strong barrier against manipulation [Faliszewski and Procaccia 2010; Mossel and Rácz 2012]. On the positive side, they suggest that there exists efficient methods for post-election audits by computing the *margin of victory* [Xia 2012].

Reconcile Borda and Condorcet via FLC. Since no Condorcet consistent voting rule satisfies consistency (plus a few other natural axioms), it would be great if a Condorcet consistent voting rule can satisfy a weaker version of consistency. The FLC axiom, which is satisfied by all GSRs, plays such a role, and thus provides a new angle of evaluating Condorcet consistent rules.

At first glance, FLC looks quite abstract, but in fact it has a natural interpretation: each partition S_t can be seen as an abstract "characteristic" of preference profiles. Then, FLC comes down to saying that the voting rule is consistent for preference profiles sharing the same characteristic.

Take Kemeny's rule as an example. It does not satisfy consistency. However, if we define a partition where for every linear order $l \in \mathcal{L}(\mathcal{A})$, S_l is composed of all preference profiles that are closest to l in Kendall tau distance, then Kemeny is consistent within each S_l, since if $P_1, P_2 \in S_l$, then l is the linear order that is closest to $P_1 \cup P_2$ in Kendall tau distance.

Have nice structures and are related to Machine Learning. Mathematically, GSRs are equivalent to *hyperplane rules*, which view all preference profiles in a geometric space and use multiple linear hyperplanes to separate regions for winner determination [Mossel et al. 2012; Xia and Conitzer 2009].

At a high level, GSRs have two interesting connections to Machine Learning. Here a voting rule can be seen as a multi-class classifier, where $\mathcal{A}$ is the set of classes [Procaccia et al. 2009]. A separating hyperplane can be seen as a linear binary classifier.

First, a GSR can be seen as the result of decision making (choosing the winner) based on the position of the input preference profile w.r.t. all hyperplanes. In other words, a GSR classifies a preference profile based on the outputs of all linear binary classifiers. This has been explored in Machine Learning as an effective way to build multi-class classifiers by binary classifiers [Tax and Duin 2002]. Therefore, when designing the g function of a GSR, we may use ideas from the literature on multi-class classifiers.

Another connection is to treat $\mathcal{O}_K$ as the set of features, and f works as the feature abstraction function (though K is not necessarily small). The collective choice is made in an additive manner where "feature values" of the input votes are summed up across the agents. Therefore, when designing the f function of a GSR, we may use techniques developed for feature selection.

How to explore these high-level connections for application is an interesting direction for future research. See [Xia 2013] for some preliminary ideas.

5.   CONCLUSION AND FUTURE DIRECTION

In this paper we surveyed some developments in generalized scoring rules, a relatively new class of voting rules for studying social choice. Given the generality and structure of GSRs, there are many directions for future research. In future/ongoing work, we see at least the following directions.

—Develop more general techniques and results for GSRs, for example post-election audits and compilation complexity [Chevaleyre et al. 2009].
—Explore deeper and more practical relationships between GSRs and Machine Learning.
—Study relationship between GSRs and other classes of voting rules, for example distance-based rules [Meskanen and Hannu 2008; Elkind et al. 2011].

6.   ACKNOWLEDGMENTS

REFERENCES

ARROW, K. 1963. *Social choice and individual values*, 2nd ed. New Haven: Cowles Foundation. 1st edition 1951.

BRANDT, F. 2009. Some remarks on Dodgson's voting rule. *Mathematical Logic Quarterly 55*, 460–463.

CHEVALEYRE, Y., LANG, J., MAUDET, N., AND RAVILLY-ABADIE, G. 2009. Compiling the votes of a subelectorate. In *Proceedings of the Twenty-First International Joint Conference on Artificial Intelligence (IJCAI)*. Pasadena, CA, USA, 97–102.

DWORK, C., KUMAR, R., NAOR, M., AND SIVAKUMAR, D. 2001. Rank aggregation methods for the web. In *Proceedings of the 10th World Wide Web Conference*. 613–622.

ELKIND, E., FALISZEWSKI, P., AND SLINKO, A. 2011. Rationalizations of condorcet-consistent rules via distances of hamming type. *Social Choice and Welfare. To appear*.

EPHRATI, E. AND ROSENSCHEIN, J. S. 1991. The Clarke tax as a consensus mechanism among automated agents. In *Proceedings of the National Conference on Artificial Intelligence (AAAI)*. Anaheim, CA, USA, 173–178.

EVERAERE, P., KONIECZNY, S., AND MARQUIS, P. 2007. The strategy-proofness landscape of merging. *Journal of Artificial Intelligence Research 28*, 49–105.

FALISZEWSKI, P. AND PROCACCIA, A. D. 2010. AI's war on manipulation: Are we winning? *AI Magazine 31,* 4, 53–64.

FISHBURN, P. C. 1974. Paradoxes of voting. *The American Political Science Review 68,* 2, 537–546.

GHOSH, S., MUNDHE, M., HERNANDEZ, K., AND SEN, S. 1999. Voting for movies: the anatomy of a recommender system. In *Proceedings of the third annual conference on Autonomous Agents*. 434–435.

MAO, A., PROCACCIA, A. D., AND CHEN, Y. 2013. Better human computation through principled voting. In *Proceedings of the National Conference on Artificial Intelligence (AAAI)*. Bellevue, WA, USA.

MESKANEN, T. AND HANNU, N. 2008. *Power, freedom, and voting*. Springer-Verlag Berlin Heidelberg, Chapter Closeness counts in social choice.

MOSSEL, E., PROCACCIA, A. D., AND RACZ, M. Z. 2012. A smooth transition from powerlessness to absolute power. http://www.cs.cmu.edu/~arielpro/papers/phase.pdf.

MOSSEL, E. AND RÁCZ, M. Z. 2012. Election Manipulation: The Average Case. *ACM SIGecom Exchanges 11,* 2, 22–24.

NURMI, H. 1987. *Comparing voting systems.* Springer.

PROCACCIA, A. D., ZOHAR, A., PELEG, Y., AND ROSENSCHEIN, J. S. 2009. The learnability of voting rules. *Artificial Intelligence 173*, 1133—1149.

TAX, D. M. AND DUIN, R. P. 2002. Using two-class classifiers for multiclass classification. In *Proceedings of the 16th International Conference on Pattern Recognition.* 124–127.

XIA, L. 2012. Computing the margin of victory for various voting rules. In *Proceedings of the ACM Conference on Electronic Commerce (EC).* Valencia, Spain, 982–999.

XIA, L. 2013. How many vote operations are needed to manipulate a voting system. *Arxiv.*

XIA, L. AND CONITZER, V. 2008. Generalized scoring rules and the frequency of coalitional manipulability. In *Proceedings of the ACM Conference on Electronic Commerce (EC).* Chicago, IL, USA, 109–118.

XIA, L. AND CONITZER, V. 2009. Finite local consistency characterizes generalized scoring rules. In *Proceedings of the Twenty-First International Joint Conference on Artificial Intelligence (IJCAI).* Pasadena, CA, USA, 336–341.

YOUNG, H. P. 1975. Social choice scoring functions. *SIAM Journal on Applied Mathematics 28,* 4, 824–838.