Beyond Electing and Ranking: Collective Dominating Chains, Dominating Subsets and Dichotomies

Jérôme Lang
Université Paris-Dauphine
PSL Research University
CNRS, LAMSADE
75016 Paris, France
lang@lamsade.dauphine.fr

Jérôme Monnot
Université Paris-Dauphine
PSL Research University
CNRS, LAMSADE
75016 Paris, France
monnot@lamsade.dauphine.fr

Arkadii Slinko
The University of Auckland
Private Bag 92019, Auckland
New Zealand
a.slinko@auckland.ac.nz

William S. Zwicker
Department of Mathematics
Union College, Schenectady
NY 12308, USA
zwickerw@union.edu

ABSTRACT

Classical voting rules output a winning alternative (or a nonempty set of tied alternatives). Social welfare functions output a ranking over alternatives. There are many practical situations where we have to output a different structure than a winner or a ranking: for instance, a ranked or non-ranked set of k winning alternatives, or an ordered partition of alternatives. We define three classes of such aggregation functions, whose output can have any structure we want; we focus on aggregation functions that output dominating chains, dominating subsets, and dichotomies. We address the computation of our rules, and start studying their normative properties by focusing on a generalisation of Condorcet-consistency.

1. INTRODUCTION

Most of the work in preference aggregation focuses on social choice rules, also called voting rules, that output a single winner or a set of tied co-winners, and on social welfare functions, that output a ranking of all alternatives. However, in many contexts the desired output has a different structure. We give below a few typical examples:

1. deciding on a ranked list of k candidates for an upcoming election based on party lists (the voters can be, for instance, the caucus members of that party in the legislature);
2. finding a ranked shortlist of k candidates to be invited for a job interview; voters are the members of a recruiting committee;
3. electing a committee of exactly k persons (e.g., a city council);
4. finding an optimal way of partitioning students between two or more groups with homogeneous level of ability in each group.

Appears in: *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2017), S. Das, E. Durfee, K. Larson, M. Winikoff (eds.), May 8–12, 2017, São Paulo, Brazil.*
Copyright © 2017, International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.

In Examples 1 and 2, the output should be a ranking of a subset of alternatives of fixed size, with all remaining alternatives implicitly ranked below them. In Example 3, the output should be a subset of alternatives of fixed size, again with all remaining alternatives implicitly ranked below them. In Example 4, the output should be an ordered partition in a given number of classes, with some flexibility on the cardinality of classes.

The amount of work on single-winner voting rules and social welfare functions overwhelmingly exceed the amount of work on aggregation procedures which output a subset, a partial list or an ordered partition. Myerson [19] pioneered the study of aggregation rules free of the structures social choice typically places on inputs and outputs, with the goal of characterizing axiomatically scoring rules in this widest setting (see the Conclusion for a comparison with our work). We will briefly mention a few other developments. Firstly, there is a number of recent important works on *multi-winner voting rules*, also called *committee selection rules*, where the goal is to output a subset of candidates, i.e., a committee of a predetermined size. Especially, Barberà and Coelho [2] considered methods of selecting a 'non-controversial' list of k names and introduced the notion of stability for the rule which means electing a weak Condorcet set, if there is one; Ratliff [23] extended some Condorcet-consistent rules like Kemeny and Dodgson to multiwinner setting; Elkind *et al.* [9] introduced committee scoring rules and analysed a number of common multiwinner voting rules from axiomatic perspective (see also many bibliographical pointers there, which can guide the reader through the earlier the literature). Zwicker [25] and Duddy et al. [8] study the rules which aggregate preference profiles into dichotomies (we will come back to their work further in the text).

Procaccia et al. [22] consider a setting where votes are seen as noisy signals about the actual ranking of alternatives; the problems considered are to output a set of k alternatives most likely containing the true alternative, or the set of best k alternatives (ranked or not), and they define, through likelihood maximization, a generic class of rules that can be instantiated on each of these three types of output.

Judgment aggregation rules provide a general and abstract setting that can be instantiated on many different collective decision making contexts, including classical voting, but not only. See for instance [20] for examples of rules that can be instantiated for aggregating preferences, committee selection, and also equivalence relations or opinions about budget allocation. (Again, see the Conclusion.) A generic framework for aggregating graphs is studied in [11].

We proceed in a systematic way — as does judgment aggregation — and define, first, a general setting in which the input consists of a classical profile V (a collection of ranking over the set of alternatives). A generalised preference aggregation rule is defined independently from $\mathcal{C}$ and can then be instantiated on every conceivable class $\mathcal{C}$.

We remain, however, at a slightly less general level than judgment aggregation, since we require that the input is a set of linear orders and that the output class $\mathcal{C}$ is a class of relations. For instance, we do not cover the aggregation of equivalence relations. Also, although our methodology can be applied to a large variety of families of rules, here we only consider rules that are C1 or C2 according to Fishburn's classification [13], that is, that are based on the (unweighted or weighted) comparison matrix, and thus takes place in the wider field of the computational study of tournament solutions [4] and of weighted tournament solutions [12].

The outline of our paper is as follows. We introduce a general framework for aggregating profiles into various kinds of relations (dominating k-chains, dominating k-subsets, dichotomies). We discuss two versions of each class, called 'plain' and 'extended', depending on whether or not we care about the relations between the alternatives that are inside the relevant groups (or, for k-chains, between the alternatives dominated by the chain). Then we define three *general*, or *graph* rules that map any profile into an element of the chosen class of graphs: one based on the unweighted majority graph, namely the *minimum Hamming distance* rule, together with an agreement-maximizing variant; and two based on the weighted majority graph, namely the *median* and *generalised ranked pairs* rules. Each such rule, plus each class of outputs, defines a rule whose outputs fits the desired structure (*i.e.*, a k-chain rule, a k-subset rule or a dichotomy rule). We investigate the computation of our rules. Lastly, we introduce a generalised notion of Condorcet-consistency and position our rules with respect to it.

Some of our proofs are omitted due to page limits.[1]

2. PRELIMINARIES AND NOTATION

Let $X = \{x_1, \ldots, x_m\}$ be a set of $m \geq 3$ alternatives. A binary relation E on X is any set of ordered pairs $E \subseteq X \times X$. It can be represented as a directed graph $G = (X, E)$, where an arrow is drawn from x to y iff $(x, y) \in E$; we use the alternative notation $\to_G$ for E, and thus denote $(x, y) \in E$ by $x \to_G y$. We will use the terms 'binary relation' and 'graph' interchangeably and we will often omit the adjective 'directed'. Let also $\mathcal{L}(X)$ be the set of all (strict) linear orders over X. By $Sub_k(X)$ we denote the set of all subsets of X of cardinality k. For any $k \leq m$ let $\mathcal{L}_k(X)$ be the set of all linear orders over k alternatives from X.

A *profile* $V = \langle \succ_1, \ldots, \succ_n \rangle$ is a collection of linear orders on X. *For simplicity, we assume throughout the paper that n is odd.* The set of all n-voter profiles over X is $\mathcal{L}(X)^n$. Let $\mathcal{T}(X)$ be the set of all tournaments over X; a tournament over X is a directed graph (digraph) obtained by assigning a direction for each edge in an undirected complete graph over X. When $(x, y) \in T$ we also write $x \succ_T y$. A *weighted tournament* over X is a function $W\colon X \times X \to \mathbb{Z}$ satisfying (1) $W(x, y) = -W(y, x)$ for all $x, y \in X$, and (2) all $W(x, y)$ have the same parity. Any profile V induces a weighted tournament W_V as follows:

$$W_V(x, y) = \#\{i \in [n] \mid x \succ_i y\} - \#\{i \in [n] \mid y \succ_i x\},$$

i.e., $W_V(x, y)$ is the difference between the number of voters who prefer x to y and the number of voters who prefer y to x. We say that a profile V *generates* weighted tournament W_V. We note that all integers $W_V(x, y)$ have the same parity. The classical (nonweighted) *majority relation* $\succ_V$ on $\mathcal{C}$ is defined by $x \succ_V y \iff W_V(x, y) > 0$. The corresponding graph will be denoted $T(V)$. Since n is odd, $T(V)$ is a tournament.

A vertex $c \in X$ in a weighted (or unweighted) tournament W is a *Condorcet winner* if $W(c, x) > 0$ for all $x \in X$. A vertex $c \in X$ is a *Copeland winner* if its Copeland score $\mathrm{Cop}_V(c) = |\{x \in X, W(c, x) > 0\}|$ is maximal.

We use the notation $x \to y_1 \ldots y_k$ for a graph containing an edge from x to each of $y_1, \ldots, y_k$. For instance, the set of edges $\{(x_1, x_2), (x_2, x_3), (x_3, x_1), (x_1, x_4), (x_2, x_4), (x_4, x_3)\}$ is represented by $[x_1 \to x_2 x_4, x_2 \to x_3 x_4, x_3 \to x_1, x_4 \to x_3]$. We use arrows between subgraphs to denote that there is an arrow from each element on the left to each element on the right. For instance, $[\{x_1 \to x_2, x_2 \to x_3, x_3 \to x_1\} \to \{x_4 \to x_5, x_5 \to x_6, x_6 \to x_4\}]$ denotes the graph containing an upper cycle and a lower cycle, with every element of the upper cycle dominating every element of the lower cycle. Similarly, if $S, S' \subset X$, $S \to S'$ denotes the set of edges (x, y) for all $x \in S$, $y \in S'$. Sometimes we omit brackets.

The *star graph* centered on x, denoted by $star(x)$, is the directed graph whose set of edges is $\{(x, y) \mid y \in X \setminus \{x\}\}$.

3. CONSTRAINTS

A *constraint* is a class of binary relations over X. We define below several particular constraints.

Let $G = (X, E)$ be a directed, irreflexive and asymmetric graph. We say that G is a *plain dominating k-chain* if there exists $A \in Sub_k(X)$ such that

(a) the restriction of $\to_G$ to A is in $\mathcal{L}_k(X)$,

(b) $x \to_G y$ for all $x \in A$ and $y \in X \setminus A$, and

(c) for any two distinct alternatives x, y in $X \setminus A$, neither $x \to_G y$ nor $y \to_G x$.

We denote a plain dominating k-chain this way, here for $m = 5$ and $k = 3$: $x_1 \to x_2 \to x_3 \to x_4 x_5$.

We say that G is an *extended dominating k-chain* if there exists $A \in Sub_k(X)$ such that

(a) the restriction of $\to_G$ to A is in $\mathcal{L}_k(X)$, and

(b) $x \to_G y$ for all $x \in A$ and $y \in X \setminus A$.

(c) G is a tournament.

We denote an extended dominating k-chain this way, here for $m = 5$ and $k = 2$:

$$\{x_1 \to x_2\} \to \{x_3 \to x_4, x_4 \to x_5, x_5 \to x_3\}.$$

Thus the difference between a plain and an extended dominating k-chain lies in the edges below the dominating chain:

[1] They can be found in the long version `http://www.lamsade.dauphine.fr/~lang/papers/JL-JM-AS-WZ-long.pdf`.

in a plain chain, alternatives that are dominated by the chain should be incomparable, while in an extended chain there should be an edge between each pair of such alternatives in one direction or the other.

We say that G is a *plain dominating k-subset* if there is a partition $\{X_1, X_2\}$ of X such that

(a) $|X_1| = k$,

(b) $x \to_G y$ for all $x \in X_1$, $y \in X_2$,

(c) for $i \in \{1, 2\}$ and for any two distinct alternatives x, y in X_i, neither $x \to_G y$ nor $y \to_G x$ hold.

Here is a plain dominating 2-subset: $x_1x_2 \to x_3x_4x_5$.

We say that G is an *extended dominating k-subset*[2] if there is a partition $\{X_1, X_2\}$ of X such that

(a) $|X_1| = k$,

(b) for all $x \in X_1, y \in X_2$ we have $x \to_G y$,

(c) G is a tournament.

Here is an extended dominating 3-subset:
$\{x_1 \to x_2, x_2 \to x_3, x_3 \to x_1\} \to \{x_4 \to x_5, x_5 \to x_6, x_6 \to x_4\}$.

Figure 1 depicts a plain dominating 2-chain, an extended dominating 2-chain, a plain dominating 2-subset and an extended dominating 2-subset.

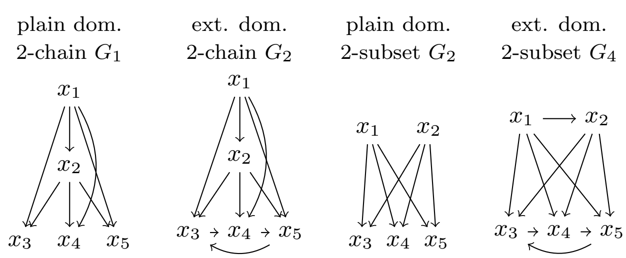

Figure 1: Dominating k-chains and k-subsets

We now define a *plain dichotomy* as a plain dominating k-subset for some $k \in \{1, \ldots, m-1\}$.[3]

Importantly, when k is fixed, for each of the five classes of graphs defined so far (plain/extended dominating k-chain, plain/extended dominating k-subset, plain dichotomy), the subset A or the partition $\{X_1, X_2\}$ of the definition can be uniquely defined from G. Formally, given a plain or extended k-chain G, let $rtop_k(G)$ be the k-chain consisting of the top k elements of G, ranked as in G; given a plain or extended dominating k-subset G, let $top_k(G)$ be the k-subset consisting of the top k elements of G; and given a *plain* dichotomy G, let $top(G)$ be its (unambiguously defined) upper part.

We have not defined yet the notion of an extended dichotomy. First observe that defining an extended dichotomy simply as an extended dominating k-subset for some k leads to the following problem: while for a plain dichotomy it is straightforward to reconstruct the top and the bottom part, for the notion of extended dichotomy the top and the bottom component may not always be uniquely defined. For instance, let $X = \{x_1, x_2, x_3\}$ and G the linear order $x_1 \succ x_2 \succ x_3$. It is an extended dominating k-subset for $k = 1$ and $k = 2$, henceforth would be an extended dichotomy as well, but it would be ambiguous whether the upper part of the dichotomy is $\{x_1\}$ or $\{x_1, x_2\}$. One solution is to define an extended dichotomy as an *augmented graph*, that is, a pair $\tilde{G} = (G, X_1)$ where G is an extended dominating k-subset for some k, and $X_1 \subset X$ denotes the upper part of the dichotomy. Given an extended dichotomy $\tilde{G} = (G, X_1)$, let $top(G) = X_1$. Figure 2 depicts two extended dichotomies corresponding to the same graph G but two different augmented graphs. X_1 is the set of alternatives above the bar.

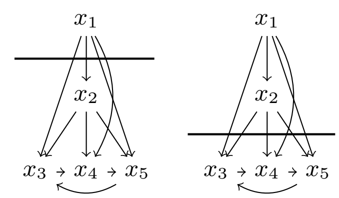

Figure 2: Two dichotomies

We denote by $ExtCh_k$ the set of all extended k-chains; $PlCh_k$ the set of all plain k-chains; $ExtDom_k$ the set of all extended dominating k-subsets; $PlDom_k$ the set of all plain dominating k-subsets; $ExtDich$ the set of all extended dichotomies; and $PlDich$ the set of all plain dichotomies.

Constraints are used to force the output of the rule to be in a particular class of graphs: for instance, we may require that the output is always a dominating k-chain for a fixed k. In this paper, we focus on a few important classes of constraints, namely k-chains, k-subsets, and dichotomies, but some other classes of constraints may be considered. For instance, if a committee has three PhD grants to distribute to applicant students, it makes sense to output a (non-ranked) set of three students, followed by a few ranked students. Yet another example of a constraint would be the set of (plain or extended) *multichotomies*, that is the set of ordered partitions of X into m nonempty classes.

A constraint $\mathcal{C}$ is said to be *cardinality-homogeneous* if any two elements of $\mathcal{C}$ have the same number of edges. While $ExtCh_k$, $PlCh_k$, $ExtDom_k$ and $PlDom_k$ are cardinality-homogeneous, $ExtDich$ and $PlDich$ are not.

An *irresolute $\mathcal{C}$-rule*, for n voters, is a function $F_\mathcal{C}$ mapping a profile (or, simply, a tournament or a weighted tournament) to a nonempty set of relations of class $\mathcal{C}$. When $\mathcal{C}$ is omitted, *i.e.*, when we note F instead of $F_\mathcal{C}$, we assume that F is defined for any possible class $\mathcal{C}$ (F then defines a family of rules – one for each $\mathcal{C}$). A *resolute* $\mathcal{C}$-rule is defined similarly, with the output being a single element of $\mathcal{C}$.

A profile V is *$\mathcal{C}$-majority-consistent* iff $T(V) \in \mathcal{C}$. A $\mathcal{C}$-rule $F_\mathcal{C}$ is *$\mathcal{C}$-consistent* iff, for every $\mathcal{C}$-majority-consistent profile V, we have $F_\mathcal{C}(V) = \{T(V)\}$.

A directed graph G over X is *$\mathcal{C}$-compatible* if there is $G' \supseteq G$ such that $G' \in \mathcal{C}$.

We may need to have aggregated preferences in some specific form. For example, we may need to determine a set of winners of fixed size k or a ranked set of k winners. Our outputs are chosen so as to make this choice simple. For ex-

[2]Also called Condorcet set of size k by Gehrlein [15].

[3]Observe that for dominating k-chains or k-subsets, the value of k is exogenously determined (so we may think of it as a parameter and as part of a pair (P, k) that serves as the input to the rule); while for dichotomies, k is endogenously determined by P alone.

ample, if the class $\mathcal{C}$ is a plain or extended k-chain, then we can select k ranked winners; if our class $\mathcal{C}$ is a dominating k-set, then we can select a winning committee of size k; if our class $\mathcal{C}$ is a plain or extended dichotomy, we can select a winning committee without restriction on the size of the committee. More precisely, an irresolute *k-chain rule* is a function mapping a profile to a nonempty set of k-chains in $\mathcal{L}_k(X)$; an irresolute *k-subset rule* is a function mapping a profile to a nonempty set of k-subsets in $Sub_k(X)$; and an irresolute *dichotomy rule* is a function mapping a profile to a nonempty set of nontrivial partitions of X. Resolute versions are defined similarly.

A $PlCh_k$-rule or an $ExtCh_k$-rule F induces a k-chain rule F^* defined by

$$F^*(V) = rtop_k(F(V)).$$

Similarly, a $PlDom_k$-rule or an $ExtDom_k$-rule F induces a k-dominating subset rule F^* defined by

$$F^*(V) = top_k(F(V)).$$

Finally, a $PlDich$-rule or an $ExtDich$-rule F induces the dichotomy rule F^* defined by

$$F^*(V) = (top(F(V)), X \setminus top(F(V))).$$

We will now define three families of rules, where a family of rules is of the form $\{F_\mathcal{C}\}$ for $\mathcal{C}$ varying.

4. MINIMUM HAMMING DISTANCE AND MAXIMAL AGREEMENT

The symmetric distance between two graphs G and G' [3] is defined by $\Delta(G, G') = |G \setminus G'| + |G' \setminus G|$ (remember that we treat graphs as sets of edges). Given an input tournament $T \in \mathcal{T}(X)$, we define $MH_\mathcal{C}(T)$ to be the set of graphs $G \in \mathcal{C}$ such that $\Delta(G, T)$ is minimal. If V is a profile, then $MH_\mathcal{C}(V) = MH_\mathcal{C}(T(V))$.

Instead of minimising the Hamming distance to T, we could choose instead to maximize the number of agreements with $T(V)$, leading to the *maximal agreement* rule: we define $MA_\mathcal{C}(T)$ to be the set of graphs $G \in \mathcal{C}$ such that $|G \cap T|$ is maximal. A simple but important fact is that *if $\mathcal{C}$ is cardinality-homogeneous then $MA_\mathcal{C} = MH_\mathcal{C}$*. This applies to plain and extended k-chains and k-subsets, but not to dichotomies.

4.1 Dominating k-chains

Informally, $MH_{ExtCh_k}(V)$ (resp. $MH_{PlCh_k}(V)$) is the set of all tournaments obtained by changing a minimum number of edges in $T(V)$ such that the resulting graph is an extended (resp. plain) dominating k-chain.

Observation 1. *(i) $MH_{PlCh_m}(V)$ (and, equivalently, $MH_{ExtCh_m}(V)$) is the set of all Slater rankings for V.*[4]

(ii) $MH_{PlCh_1}(V)$ is the set of all star graphs whose centre is a Copeland winner, i.e.*, the set of graphs $star(c)$ for all Copeland winners c of V.*

(iii) $MH_{ExtCh_1}(V)$ is the set of all graphs G_c containing $\{(c, y) \mid y \in X \setminus \{c\}\}$, where c is a Copeland winner of V, and which, for each $u \neq c, v \neq c$, contains (u, v) if $u \to_T v$ and (v, u) if $v \to_T u$.

[4]A Slater ranking for a tournament T over X is a ranking of X minimising the number of edges that are directed in the opposite way as in T.

Proof. (i) This is the definition of a Slater ranking. (ii) Let G be a plain 1-chain with dominating element x; we have $\Delta(G, T(V)) = 2[(m-1) - \text{Cop}_V(x)] + \frac{1}{2}(m-1)(m-2)$, which implies the result. (iii) Since for an extended 1-chain we have total freedom about the edges between vertices below the dominating element x, minimum distance is obtained by taking the same edges as in $T(V)$ below x; for such an extended 1-chain, we have $\Delta(G, T(V)) = 2[(m-1) - \text{Cop}_V(x)]$, which implies the result. □

Observation 2. *For each $G \in MH_{ExtCh_k}(V)$ there is a $G' \in MH_{PlCh_k}(V)$ such that $rtop_k(G) = rtop_k(G')$, and for each $G' \in MH_{PlCh_k}(V)$ there is a $G \in MH_{ExtCh_k}(V)$ such that $rtop_k(G) = rtop_k(G')$.*

Proof. Recall from above that for extended k-chains, the minimum Hamming distance is obtained for extended chains that are exactly as in the original tournament below the top k candidates. Now, let $G \in MH_{ExtCh_k}(V)$ and let G' be the corresponding plain k-chain (with edges (y, z) suppressed for all y, z below the top k alternatives). Then $\Delta(G', T(V)) = \Delta(G, T(V))$, from which the result follows. □

Therefore, as far as k-chains are concerned, whether we take the 'plain' or the 'extended' notion does not make a difference: the resulting chains or subsets are the same (recall the definition of F^* at the end of Section 3):

Corollary 1. $MH^*_{PlCh_k} = MH^*_{ExtCh_k}$.

The following straightforward observation is a direct consequence of the NP-hardness of winner determination for the Slater rule [1, 7, 5]:[5]

Proposition 2. *Computing MH_{ExtCh_k} and MH_{PlCh_k} are NP-hard, even if $k = m$.*

Let $Cop^*_2(T(V))$ be the set of all ordered pairs (x, y) such that (a) for all $z \neq x, y$, $\min(\text{Cop}_V(x), \text{Cop}_V(y)) \geq \text{Cop}_V(z)$ and (b) $(x, y) \in T(V)$. In other words, $Cop^*_2(T(V))$ is the set of all pairs with the highest two Copeland scores, these two candidates being ordered according to majority.

Proposition 3. *$MH_{PlCh_2}(V)$ consists of the plain 2-chains $x \to y \to X \setminus \{x, y\}$ for $(x, y) \in Cop^*_2(T(V))$.*

Proof. Let $T = T(V)$. Let $G_{x,y}$ be the plain 2-chain $x \to y \to X \setminus \{x, y\}$. If $x \succ_T y$, then $\Delta(G_{x,y}, T) = 2[(m-1) - \text{Cop}_V(x)] + 2[(m-2) - \text{Cop}_V(y)] + \frac{(m-2)(m-3)}{2} = K - (\text{Cop}_V(x) + \text{Cop}_V(y))$ (K being a constant). If $y \succ_T x$ then $\Delta(G_{x,y}, T) = \Delta(G_{y,x}, T) + 2$. Therefore, the minimal value of $\Delta(G_{x,y}, T)$ is obtained when $\text{Cop}_V(x) + \text{Cop}_V(y)$ is maximal and $(x, y) \in T$. □

4.2 Dominating k-subsets

For dominating k-subsets, the following counterparts of Observation 2 and Corollary 1 hold.

Observation 3. *For each $G \in MH_{ExtDom_k}(V)$ there is a $G' \in MH_{PlDom_k}(V)$ such that $top_k(G) = top_k(G')$, and for each $G' \in MH_{PlDom_k}(V)$ there is a $G \in MH_{ExtDom_k}(V)$ such that $top_k(G) = top_k(G')$.*

[5]On the other hand, for a fixed k, $MH_{ExtCh_k}(V)$ and $MH_{PlCh_k}(V)$ can be computed in $O(m^k)$, since there are only $m(m-1)\ldots(m-k+1)$ k-chains to be considered.

COROLLARY 4. $MH^*_{PlDom_k} = MH^*_{ExtDom_k}$.

Let $Cop_k(T(V))$ be the set of all subsets Z of $S_k(X)$ such that for all $x \in Z$ and $t \in X \setminus Z$, $Cop_V(z) \geq Cop_V(t)$. In other words, $Cop_k(T(V))$ is the set of all subsets of k alternatives highest Copeland scores.

THEOREM 1. *$MH_{PlDom_k}(V)$ consists of all plain subsets whose dominating k-set is in $Cop_k(T(V))$.*

PROOF. Let $G_Z = Z \to X \setminus Z$ be a dominating k-subset with $|Z| = k$. Since $\sum_{z \in Z} \text{Cop}_V(z) = \#\{(z,t) \in Z \times (X \setminus Z) | z \to_G t\} - \#\{(z,z') \in Z^2 | z \to_G s'\} = |G_Z \cap T(V)| - \frac{k(k-1)}{2}$ and $\Delta(G_Z, T(V)) = \frac{m(m-1)}{2} - |G_Z \cap T(V)|$, $\Delta(G, T(V))$ is minimal when $\sum_{z \in Z} \text{Cop}_V(z)$ is maximal. □

Thus this rule is the NED rule of Coelho [6] or else k-Copeland rule [9]. As a corollary, MH_{PlDom_k} (and similarly MH_{ExtDom_k}) can be computed in polynomial time.

4.3 Dichotomies

Unlike dominating k-chains and k-subsets, dichotomy constraints are not cardinality-homogeneous; therefore, $MA_{\mathcal{C}}$ does not in general coincide with MH_C. Moreover, it makes a difference whether plain or extended dichotomies are considered. This leads to *four* different rules. It is unclear whether all of them are interesting, or can they be easily characterised. However, using a result from [25], we have the following characterisation of plain dichotomies for the maximum agreement rule.

Given a profile V over X, the average Copeland score w.r.t. V of all alternatives is $m(m-1)/2$. Define $High(V)$ (respectively, $Low(V)$ and $Average(V)$) the set of alternatives whose Copeland score is larger than (respectively, smaller than, equal to) $m(m-1)/2$. Reformulating from [25], a *maximally separating dichotomy* for V is a partition of X in two subsets (X_T, X_B) such that $High(V) \subseteq X_T$ and $Low(V) \subseteq X_B$ (elements of $Average(V)$ can be put either in X_T or in X_B).

THEOREM 2. *[25] $MA^*_{PlDich}(V)$ is the set of all maximally separating dichotomies for V.*

As a consequence, MA_{PlDich} is computable in polynomial time. This is also the case for MH_{PlDich}, which comes from the fact that finding the optimal dichotomy corresponds to finding a minimal cut in a directed graph, which is polynomial [21].

4.4 Example

Consider these two tournaments $W = T(V)$, $W' = T(V')$:

W	a	b	c	d	e
a	$\star$	1	1	−1	−1
b	−1	$\star$	1	1	−1
c	−1	−1	$\star$	1	−1
d	1	−1	−1	$\star$	1
e	1	1	1	−1	$\star$

W'	a	b	c	d	e
a	$\star$	1	1	1	1
b	−1	$\star$	1	−1	1
c	−1	−1	$\star$	1	1
d	−1	1	−1	$\star$	−1
e	−1	−1	−1	1	$\star$

We have

- $MH_{PlCh_1}(V) = \{[e \to abcd]\}$
- $MH_{ExtCh_1}(V) = \{[e \to abcd, a \to bc, b \to cd, c \to d, d \to ae]\}$
- $MH_{PlCh_2}(V) = \{[e \to a \to bcd], [d \to e \to abc]\}$
- $MH_{PlCh_3}(V) = \{[e \to a \to b \to cd], [d \to e \to a \to b, c]\}$
- $MH_{PlDom_2}(V) = \{[a, e \to b, c, d], [de \to abc]\}$
- $MH_{ExtDom_2}(V) = \{[(e \to a) \to (b \to c \to d)], [(d \to e) \to [a \to b \to c]\}$

and

- $MA_{PlDich}(V') = \{[abc \to de], [ab \to cde], [ac \to bde], [a \to bcde]\}$
- $MA_{ExtDich}(V') = MH_{ExtDich}(V') = \{[a \to bcde]\}$

5. THE MEDIAN RULE

Let $\mathcal{C}$ be a class of graphs and $G \in \mathcal{C}$. We define

- $W(G, V) = \sum_{(x,y) \in G} W_V(x, y)$, and
- $MED_{\mathcal{C}}(V) = \text{argmax}_{G \in \mathcal{C}} W(G, V)$.

Thus, the score of a graph $G \in \mathcal{C}$ is the sum of all the weights associated with the edges this graph contains.

5.1 Dominating k-chains

OBSERVATION 4.

(i) $MED_{PlCh_1}(V) = \{star(c) \mid c \in Borda(V)\}$, where $Borda(V)$ denotes the set of winners under Borda rule.[6]

(ii) $MED_{PlCh_m}(V)$ (and equivalently $MED_{ExtCh_m}(V)$) is the set of all Kemeny rankings for V.[7]

PROOF. Every plain 1-chain is a star graph $star(x)$, and $W(star(x), V) = \sum_{y \neq x} W(x, y) = B_V(x)$, where $B_V(x)$ is the Borda score of x for profile V. This proves (i). As to (ii), it follows from the definition of a Kemeny ranking. □

On the other hand, $MED_{ExtCh_1}(V)$ does not always select the Borda winners, as we can see on this example:

	a	b	c	d	e
a	0	+5	−3	+5	−3
b	−5	0	+5	+1	+1
c	+3	−5	0	+1	+1
d	−5	−1	−1	0	+3
e	+3	−1	−1	−3	0

The Borda winner is a. However, $MED_{ExtCh_1}(V)$ contain the following two graphs: $G_c = [c \to abde, b \to de, d \to e, e \to a]$, and $G_e = [e \to abcd, a \to bd, b \to cd, c \to ad]$, with $W(G_c, V) = W(G_2, V) = 18$, while, if $G_a = [a \to bcde, b \to cde, c \to de, d \to e]$, then $W(G_a, V) = 16$.

Since winner determination for the Kemeny rule is NP-hard, winner determination for MED_{ExtCh_k} and MED_{PlCh_1} is NP-hard too, even if $k = m$.

5.2 k-dominating subsets

Unlike plain and extended chains, plain dominating subsets under MED are very easy to characterize.

OBSERVATION 5. *$MED^*_{PlDom_k}(V)$ is the set of the alternatives with the highest k Borda scores (or the sets of sets of alternatives with k highest Borda scores, in case of a tie).*

PROOF. Let $S = \{x_1, \ldots, x_k\}$, and G_S be the associated plain k-dominating set. We have $W(G_S, V) = \sum_{(x,y) \in G_S} W_V(x, y) = \sum_{x \in S, y \in X \setminus S} W_V(x, y) = \sum_{x \in S} (B_V(x) - \sum_{z \in S} W_V(x, z)) = \sum_{x \in S} B_V(x) - \sum_{x \in S} \sum_{z \in S} W_V(x, z) = \sum_{x \in S} B_V(x)$, since $W_V(x, z) + W_V(z, x) = 0$ for all $x, z \in S$. □

THEOREM 3. *Computing $MED_{ExtDom_k}(V)$ is NP-hard, even for $k = \frac{m}{2}$.*

The proof (by reduction from ONEWAY SECTION) is in the long version of the paper.

[6]There are two equivalent ways of defining the Borda rule: either as a positional scoring rule, or as a C2 rule: the Borda score of an alternative x is $\sum_{y \neq x} W_V(x, y)$, and the Borda winners are the alternatives that maximise this score.

[7]A Kemeny ranking for a weighted tournament W is a ranking $\succ$ of X maximising $\sum_{x \succ y} W_V(x, y)$.

5.3 Dichotomies

For plain dichotomies we have a result similar to Theorem 2, with Copeland scored being replaced by Borda scores. Details can be found in the long version.

6. GENERALISED RANKED PAIRS

We now define a family of resolute rules that follows the same construction as the *ranked pairs* rules [24]. This rule works by constructing a ranking of alternatives by considering pairs of alternatives successively, in the decreasing order of their majority margin, and add them to the ranking if they do not introduce a cycle. Here we consider, and we generalise, the *resolute* version of the rule, where ties in the order of pairs are broken as soon as they appear via a priority rule over pairs of alternatives.

Recall that G is $\mathcal{C}$*-compatible* if there is $G' \supseteq G$ such that $G' \in \mathcal{C}$. The GRP rule proceeds as follows.

order all the pairs of alternatives (x, y) in decreasing order of $W_V(x, y)$ (breaking ties, if any, using a predefined priority relation over pairs of alternatives)
initialize G with an empty set of edges
repeat
  take the next pair (x, y) according to the order induced by W_V
  if adding (x, y) to G results in a $\mathcal{C}$-compatible graph **then**
    add (x, y) to G
  end if
until all pairs have been considered
return G

6.1 Dominating k-chains

If $\mathcal{C} = PlCh_k$, then GRP_{PlCh_k} is defined by arranging in a sequence all pairs of alternatives (x, y) in decreasing order of $W_V(x, y)$ and adding the corresponding edge if the resulting graph can still be extended to a graph containing a k-chain. In particular, for the two extreme values of k we recover two well-known rules:

PROPOSITION 5. *Let V be a profile. Then*

(i) $GRP_{PlCh_1} = \{star(x) \mid x \in Maximin(V)\}$*;*

(ii) GRP_{PlCh_m} *is the set of rankings obtained by the* ranked pairs *rules.*

PROOF. (ii) is obvious, let us prove (i). We recap that x is a maximin winner for V if it maximises $\min_{y \neq x} W_V(x, y)$. Let x be a maximin winner, and $\alpha = \min_{y \neq x} W_V(x, y)$. All edges incoming to x have weight at most α. So as long as we consider pairs (u, v) such that $W(u, v) > \alpha$, they can be added to G without violating $PlCh_1$-compatibility. Therefore, after considering all those pairs, all alternatives but the maximin winners have an incoming edge, which implies that the output will be a star centred at a maximin winner. (See Proposition 4 in [18] for a similar result.) □

For example, consider the following weighted tournament:

W	x_1	x_2	x_3	x_4	x_5
x_1	0	+39	+31	−35	+29
x_2	−39	0	+27	+23	−33
x_3	−31	−27	0	+37	−25
x_4	+35	−23	−37	0	+21
x_5	−29	+33	+25	−21	0

The maximin winner is x_5 and the ranked pairs winner (independently of the tie-breaking) is x_3. Let us compute GRP_{Chain_2}. We first add $x_1 \to x_2$, then $x_3 \to x_4$, then $x_4 \to x_1$, then $x_5 \to x_2$; we cannot add $x_1 \to x_3$ and skip it; we add $x_1 \to x_5$; at this point the winning 2-chain is known: $GRP_{Chain_2}(V) = x_3 \to x_4 \to \{x_1, x_2, x_5\}$.

Because the GRP rules repeatedly need $\mathcal{C}$-compatibility checks, we need to know how to check when a directed, asymmetric graph G over X can be completed into a graph satisfying the different types of constraints.

LEMMA 1. *Checking whether G is $ExtCh_k$-compatible can be done in polynomial time.*

PROOF. For $S \subseteq X$ let $undom(S, G)$ be the set of all alternatives $s \in S$ such that there is no $s' \in S$ that dominates s in G. Consider the following algorithm:

begin
$i \leftarrow 1$;
$C \leftarrow \emptyset$;
while $|C| < k$ and $undom(X \setminus C) \neq \emptyset$ **do**
  $C_i := undom(X \setminus C)$;
  $C := C \cup C_i$
end while
if $|C| \geq k$ **then**
  return a chain starting by all elements in C_1 in an arbitrary order, then all elements in C_2, etc. until we have a k-chain
else
  return no
end if

Because at each step C_i is undominated in $X \setminus (C_1 \cup \ldots \cup C_{i-1})$, $C_1 \cup \ldots \cup C_i$ can be extended to a chain in G. Therefore, if we end up with k elements in this union, then G is $ExtCh_k$ consistent. Conversely, if we have to stop before we get a k-chain, then it means that we have reached a step where $undom(X \setminus C) \neq \emptyset$, and then, it is not possible to obtain a k-chain without removing edges: therefore, G is not $ExtCh_k$-consistent. □

Note that G is $ExtCh_1$-compatible if and only if it has an undominated element, and $ExtCh_m$-compatible if and only if it is acyclic.

For plain k-chains, the condition is much simpler:

LEMMA 2. *G is $PlCh_k$-compatible if and only if the following two conditions hold: (1) G is acyclic; and (2) G has at least $m - k$ alternatives with no outgoing edges.*

PROOF. If G is not acyclic, or if it contains less than $m-k$ alternatives without outgoing edge, then it cannot be extended into a plain k-chain. Conversely, if Conditions 1 and 2 are satisfied, then let B be a set of $m-k$ alternatives without outgoing edge, and $S = X \setminus B$. Complete the restriction of G to S into a ranking of S (which is possible since G is acyclic) and add edges from all elements of S to all elements of B; this results in a plain k-chain. □

COROLLARY 6. *For all k, GRP_{ExtCh_k} and GRP_{PlCh_k} are polynomial-time computable.*

6.2 Dominating k-subsets

Checking Dom_k-compatibility is far less easy. We show below that the problem is NP-complete, which in turn will enable us to show that computing GRP_{ExtDom_k} is NP-hard.

LEMMA 3. *G is $ExtDom_k$-compatible if and only if its transitive closure G^* is $ExtDom_k$-compatible.*

PROOF. Note first that G is $ExtDom_k$-compatible if and only if there is some $A \in Sub_k(X)$ such that there is no edge in G from $X \setminus A$ to A. This, in turn, happens if and only if there is no edge from $X \setminus A$ to A in G^*, therefore, the $ExtDom_k$-compatibility of G is equivalent to the $ExtDom_k$-compatibility of G^*. □

Since computing transitive closure is polynomial, Lemma 3 allows us to assume without loss of generality that *G is a partially ordered set*. Now, let $C_1, \ldots, C_q$ be the maximally connected components of G, and let $C_i \succ_G C_j$ if for all $x \in C_i$ and $y \in C_j$ we have $x \rightarrow_G y$. Then

LEMMA 4. *G is $ExtDom_k$-compatible iff there is $J \subset \{1, \ldots, q\}$ such that (1) $j \in J$ and $C_i \succ_G C_j$ implies $i \in J$; and (2) $\sum_{i \in J} |C_i| = k$.*

For instance, assume that G^* contains 5 maximally connected components $C_1, \ldots, C_5$ with $|C_1| = 4$, $|C_2| = 3$, $|C_3| = 4$, $|C_4| = 1$, $|C_5| = 2$, and $C_1 \succ C_2 \succ C_4$, $C_1 \succ C_3 \succ C_4$, $C_2 \succ C_5$. Then G is Dom_k-compatible for $k \in \{4, 7, 8, 9, 11, 12, 13\}$.

The problem is a generalisation of SUBSET SUM with the integers encoded in unary, which we define formally:

CONSTRAINED SUBSET SUM

Input A directed acyclic graph $(\mathcal{V}, E)$ where each vertex $v \in \mathcal{V}$ is associated with a positive integer weight $w(v)$, encoded in unary; and an integer K.

Question Is there $S \subseteq \mathcal{V}$ such that (a) for each $(v, v') \in E$, if $v \in S$ then $v' \in S$, and (ii) $\sum_{v \in S} w(v) = K$?

CONSTRAINED SUBSET SUM is a generalisation of SUBSET SUM, since the latter is obtained when $E = \emptyset$. While SUBSET SUM is weakly NP–complete (Problem [SP13], page 223 in [14]), and thus polynomial if weighted are encoded in unary, this is no longer the case for its constrained version:

THEOREM 4. CONSTRAINED SUBSET SUM *is strongly* NP-*complete, even if the number of distinct weights $w(v)$ is two.*

The proof (consisting of a reduction from CLIQUE) is in the long version of the paper.

PROPOSITION 7. *Computing RS_{ExtDom_k} is* NP-*hard.*

The proof (consisting of a reduction from CONSTRAINED SUBSET SUM) is in the long version of the paper.

For $PlDom_k$, compatibility is much more drastic:

LEMMA 5. *G is $PlDom_k$-compatible if and only if (1) for all $x \in X$, x has only outgoing edges or only ingoing edges in G, and (2) the number of alternatives with at least one outgoing (resp. ingoing) edge is at most k (resp. at most $m - k$).*

PROOF. If 1 or 2 fails to hold then G cannot be completed into a plain dominating k-subset. If 1 and 2 hold then initialize S with the set of p alternatives with an outgoing edge (we know that $p \leq k$), and add to S $k - p$ more alternatives with no ingoing edge; we add edges from all alternatives of S to all alternatives of $X \setminus S$; and we then obtain a plain dominating k-set. □

COROLLARY 8. *For all k, GRP_{ExtCh_k} and GRP_{PlCh_k} are polynomial-time computable.*

6.3 Dichotomies

LEMMA 6. *G is $ExtDich$-compatible if and only if it does not contain a cycle with all the elements of X.*

PROOF. If G contains a cycle with all elements of X then any extension of G does not have a proper dominating set, therefore G is not $ExtDich$-compatible. Conversely, if G does not contain a cycle with all elements of X then there are two vertices x, y such that no path exists in G from x to y. Let S be the set of all vertices which can be reached from x (including x itself). Note that $y \notin S$, therefore $S \neq X$. For any two vertices $u \in S$ and $v \in X \setminus S$, there is either no arrow between them or an arrow from v to u. Adding all missing arrows from $X \setminus S$ to S we obtain a dichotomy. Therefore, G is $ExtDich$-compatible. □

COROLLARY 9. *Checking whether G is $ExtDich$-compatible can be done in polynomial time.*

LEMMA 7. *G is $PlDich$-compatible if and only if every $x \in X$ has only outgoing edges or only ingoing edges in G.*

PROOF. In a plain dichotomy $S \succ X \setminus S$, any x in the upper (resp. lower) part S (resp. $X \setminus S$) has only outgoing (resp. ingoing) edges. Therefore, if for some x there is at least one outgoing edge $x \rightarrow y$ and an ingoing edge $z \rightarrow x$ in G, then G cannot be completed into a plain dichotomy. Now, if for all $x \in X$, x has only outgoing edges or only ingoing edges in G, let U be the set of all $x \in X$ for which there are only outgoing edges: then G can be completed into the plain dichotomy $U \succ X \setminus U$. □

COROLLARY 10. *Checking whether G is $PlDich$-compatible can be done in polynomial time.*

COROLLARY 11. *$RS_{ExtDich}$ and RS_{PlDich} are polynomial-time computable*

7. GENERALISED CONDORCET CONSISTENCY

We first define a uniform definition of Condorcet-consistency, applicable to any constraint $\mathcal{C}$; then we will focus on specific classes of constraints. Recall that we have an odd number of voters, and thus $T(V)$ is a tournament.

A $\mathcal{C}$-rule $F_\mathcal{C}$ is *generalised Condorcet-consistent* (GCC) if for every profile V, whenever $T(V)$ is in $\mathcal{C}$, $F_\mathcal{C}(V) = T(V)$. A family of $\mathcal{C}$-rules F is GCC if it is GCC for all $\mathcal{C}$

We extend GCC-consistency to chain and subset rules as follows: we say that a k-chain rule F^* is GCC if for every profile V, whenever $T(V)$ is an extended k-chain, of the form $x_1 \rightarrow \ldots \rightarrow x_k \rightarrow X \setminus \{x_1, \ldots x_k\}$, then $F^*(V) = x_1 \succ \ldots \succ x_k$. A k-subset rule F^* is said to be GCC if whenever $T(V)$ is an extended k-subset of the form $\{x_1, \ldots, x_k\} \rightarrow X \setminus \{x_1, \ldots x_k\}$, then $F^*(V) = \{x_1, \ldots, x_k\}$. The latter is similar to Condorcet-consistency as defined in [6, 2, 16].

PROPOSITION 12. *MH and GRP are GCC.*

PROOF. For MH, it is a corollary of the fact that the median rule in judgment aggregation is Condorcet-consistent [20]. For GRP, if $T(V)$ is in $\mathcal{C}$ then all edges (u, v) with $W(u, v) > 0$ are added to G without violating $\mathcal{C}$-compatibility; since n is odd, there is no pair (u, v) with $W(u, v) = 0$, therefore, after all edges (u, v) with $W(u, v) > 0$, we have a complete graph. □

MED is not GCC, because, for instance, MED_{PlCh_1} (and $MED_{PlSubset_1}$) corresponds to the Borda rule, which is not Condorcet-consistent. However, it is easy to check that MED_{ExtCh_k} and $MED_{ExtSubset_k}$ are GCC.

Moreover, for dominating chains and subsets, MH satisfies a stronger property, which can be seen as a generalisation of Smith-consistency (a rule is Smith-consistent if it selects from the top cycle). Let T be a tournament and T^* its transitive closure. Let $C_1, \ldots, C_r$ be the maximal indifference classes of T^*, that is, C_i is a maximal subset of alternatives such that T^* contains (x, y) for each $x, y \in C_i$, and the indices being chosen such that for each $x \in C_i$ and $y \in C_j$, $j > i$, T^* contains (x, y). We call $(C_1, \ldots, C_r)$ the ordered decomposition of T. Note that C_1 is the top cycle of T.

Let $(C_1, \ldots, C_r)$ be the ordered decomposition of a tournament T. For $k \leq m$, let s be the smallest integer such that $|C_1|+\ldots+|C_s| \geq k$. A k-subset rule F^* is said to satisfy *generalised Smith-consistency* if $F^*(T)$ contains $C_1 \cup \ldots \cup C_{s-1}$ and is contained in $C_1 \cup \ldots \cup C_s$. A k-chain rule is said to satisfy *generalised Smith-consistency* if (a) the k elements $x_1, \ldots, x_k$ of the chain contain $C_1 \cup \ldots \cup C_{s-1}$, are contained in $C_1 \cup \ldots \cup C_s$, and are such that if $x_i \in C_{h(i)}$, $x_i \in C_{h(j)}$, and $h(i) < h(j)$ then x_i appears above x_j in the k-chain.

Proposition 13. *$MH^*_{\mathcal{C}}$ satisfies generalised Smith-consistency for $\mathcal{C} \in \{ExtCh_k, PlCh_k, ExtDom_k, PlDom_k\}$,*

Proof sketch We give the proof only for $MH^*_{PlDom_k}$ (for $MH^*_{PlCh_k}$ it is similar). Assume $S \in MH^*_{Dom_k}(T)$, $x \in C_i$, $y \in C_j$, $j > i$, $x \notin S$ and $y \in S$. Let $S' = S \cup \{x\} \setminus \{y\}$. The pairs that are in G but not in G' are (y, z) for all $z \notin S$, (t, x) for all $x \in S$, and (y, x). The pairs that are in G' but not in G are (x, z) for all $z \notin S$, (t, y) for all $x \in S$, and (x, y). Also, (a) $(x, y) \in T$ and $(y, x) \notin T$; (b) for all z, $(y, z) \in T$ implies $(x, z) \in T$ and $(z, x) \in T$ implies $(z, y) \in T$. Let G and G' the $PlDom_k$ graphs associated with S and S', respectively. Now,

$$\begin{array}{rl} & |G' \cap T| - |G \cap T| \\ = & |\{(x,z), z \notin S, (x,z) \in T\}| + |\{(t,y), z \in S, (t,y) \in T\}| + 1 \\ & -|\{(y,z), z \notin S, (y,z) \in T\}| - |\{(t,x), z \notin S, (t,x) \in T\}| \\ = & (|\{(x,z), z \notin S, (x,z) \in T\}| - |\{(y,z), z \notin S, (y,z) \in T\}|) \\ & +(|\{(t,y), z \in S, (t,y) \in T\}| - |\{(t,x), z \notin S, (t,x) \in T\}|) + 1 \\ > & 0 \end{array}$$

which contradicts $S \in MH^*_{Dom_k}(T)$.

Since the maximin rule is not Smith-consistent, this result does not carry on to GRP.

8. DISCUSSION

Our contribution is mainly methodological: we give a unified framework that can be instantiated on every meaningful class of graphs. We focused on dominating k-chains, dominating k-subsets and dichotomies because they are particularly interesting, but our notions and some of our results extend beyond. Also, clearly, other rules need to be studied.

We have considered two series of classes, 'plain' and 'extended'. The question of which one makes more sense is of particular interest. It seems that there is no firm answer to this question, which boils down to deciding whether we have to care about what happens inside the clusters of alternatives (such as the set of alternatives below a dominating chain) or not. The fact that well-known rules were sometimes obtained for the 'plain' notion (Borda), some for the 'extended' notion (ranked pairs, maximin) and some for both (Copeland, Slater, Kemeny) can be seen as a sign that so far there is no clear way of choosing between both. This choice may depend both on the application we have in mind, and of the rule (for instance, for GRP the extended versions seem to make more sense than the plain versions).

We have considered only one property, a generalisation of Condorcet-consistency. Other classical properties of voting rules and social welfare functions need to be generalised as well. For some of them, such as monotonicity, this is rather simple, at least for dominating k-chains and k-subsets. For some other, such as strategyproofness, this is less easy.

In the Introduction we mentioned Myerson's work [19], which pioneered the study of aggregation rules free of the structures social choice typically places on inputs and outputs – an important similarity with our work. His (partially met) goal, however, was axiomatic characterization of scoring rules in this widest setting, extending the success of Young and others in the narrower setting. Our ends are structural rather than axiomatic — we aim to show that two highly general aggregation procedures yield a variety of interesting and diverse rules when specialized to particular input/output classes. One procedure – Minimum Hamming Distance – is quite different from scoring rules. The other – the Median Rule – coincides with an important subclass of "generalized" scoring rules in Myerson's sense, but Myerson's focus was not on the median rule, nor on several of the output classes of interest to us.

Also, our work is highly related to judgment aggregation. Judgment aggregation also defines general rules that can be applied to various output structures. The three rules we focus on correspond to rules considered in [20, 17, 11]. Our work is however more specific than corresponding works in judgment aggregation; because we focus on ordered structures (k-chains, k-subsets, dichotomies), we are able to derive specific results that do not necessarily correspond to something in the general case. For instance, while Proposition 12 can be derived as a corollary of a result in [20], Proposition 13 cannot, because Smith-Consistency is specific to ordered structures as considered here. A minor difference, also, is that in judgment aggregation the constraints on the input and the output coincide, which is not the case in our work. In any case, the relation between our methodology and judgment aggregation needs to be explored further.

Our work also takes inspiration from the rationalization of voting rules by a distance and a consensus class. The two frameworks have some similarity but also significant differences. The first major difference is that in [10] the approach was based on consensus classes with a single unambiguous winner (e.g., a profile with a Condorcet winner). From [10] it was unclear how to adopt the approach to multiwinner rules. Our approach allows us to rationalise multiwinner rules and more. Not only can we rationalise rules that output a committee of a predetermined size but also rules that output a committee with structure (*e.g.*, an ordered one). Another major difference is that [10] worked with rules whose input is a profile (since votewise distances can be defined only in this case) while in our framework we mostly work with C1 and C2 rules in Fishburn's classification, which can be either profile-based or tournament-based.

Acknowledgements: Jérôme Lang and Jérôme Monnot were supported by the project ANR-14-CE24-0007-01 CoCoRICo-CoDec. Arkadii Slinko was supported by the Royal Society of NZ Marsden Fund UOA-254.

REFERENCES

[1] N. Alon. Ranking tournaments. *SIAM J. Discrete Math.*, 20(1):137–142, 2006.

[2] S. Barberà and D. Coelho. How to choose a non-controversial list with k names. *Social Choice and Welfare*, 31(1):79–96, 2008.

[3] J.-P. Barthélemy and B. Monjardet. The median procedure in cluster analysis and social choice theory. *Mathematical Social Sciences*, 1(3):235–267, 1981.

[4] F. Brandt, M. Brill, and P. Harrenstein. Tournament solutions. In F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. D. Procaccia, editors, *Handbook of Computational Social Choice*, chapter 3. Cambridge University Press, 2016.

[5] P. Charbit, S. Thomassé, and A. Yeo. The minimum feedback arc set problem is NP-hard for tournaments. *Combinatorics, Probability & Computing*, 16(1):1–4, 2007.

[6] D. Coelho. *Understanding, Evaluating and Selecting Voting Rules Through Games and Axioms*. PhD thesis, Universitat Autònoma de Barcelona, 2004.

[7] V. Conitzer. Computing Slater rankings using similarities among candidates. In *Proceedings of AAAI-06*, pages 613–619, 2006.

[8] C. Duddy, A. Piggins, and W. S. Zwicker. Aggregation of binary evaluations: a Borda-like approach. *Social Choice and Welfare*, 46(2):301–333, 2016.

[9] E. Elkind, P. Faliszewski, P. Skowron, and A. Slinko. Properties of multiwinner voting rules. In *Proceedings of AAMAS '14*, pages 53–60, 2014.

[10] E. Elkind, P. Faliszewski, and A. Slinko. Distance rationalization of voting rules. *Social Choice and Welfare*, 45(2):345–377, 2015.

[11] U. Endriss and U. Grandi. Collective rationality in graph aggregation. In *Proceedings of ECAI 2014*, pages 291–296, 2014.

[12] F. Fischer, O. Hudry, and R. Niedermeier. Weighted tournament solutions. In F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. D. Procaccia, editors, *Handbook of Computational Social Choice*, chapter 4. Cambridge University Press, 2016.

[13] P. C. Fishburn. Condorcet social choice functions. *SIAM Journal on applied Mathematics*, 33(3):469–489, 1977.

[14] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman & Co., New York, NY, USA, 1979.

[15] W. Gehrlein. The Condorcet criterion and committee selection. *Mathematical Social Sciences*, 10(3):199–209, 1985.

[16] E. Kamwa. Stable rules for electing committees and divergence on outcomes. *Group Decision and Negotiation*, pages 1–18, 2016.

[17] J. Lang, G. Pigozzi, M. Slavkovik, L. van der Torre, and S. Vesic. A partial taxonomy of judgment aggregation rules and their properties. *Social Choice and Welfare*, 48(2):327–356, 2017.

[18] J. Lang and M. Slavkovik. Judgment aggregation rules and voting rules. In *Algorithmic Decision Theory - Third International Conference, ADT 2013, Bruxelles, Belgium, November 12-14, 2013, Proceedings*, pages 230–243, 2013.

[19] R. Myerson. Axiomatic derivation of scoring rules without the ordering assumption. *Social Choice and Welfare*, 12(21):59–74, 1995.

[20] K. Nehring, M. Pivato, and C. Puppe. The condorcet set: Majority voting over interconnected propositions. *Journal of Economic Theory*, 151:268–303, 2014.

[21] C. H. Papadimitriou and K. Steiglitz. *Combinatorial Optimization: Algorithms and Complexity*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1982.

[22] A. D. Procaccia, S. J. Reddi, and N. Shah. A maximum likelihood approach for selecting sets of alternatives. In *UAI*, pages 695–704, 2012.

[23] T. C. Ratliff. Selecting committees. *Public Choice*, 126(3-4):343–355, 2006.

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

[25] W. Zwicker. Universal and symmetric scoring rules for binary relations. In *Proceedings of COMSOC*, 2014.