10.1 Introduction

In a typical social choice scenario, agents rank the available alternatives and have to collectively decide on the best alternative, or a ranking of the alternatives. If there are just two alternatives, the decision can be made by a majority vote. However, for three or more alternatives the agents may face a difficult choice. For instance, there can be a cycle in the majority preferences: it may happen that a majority of voters prefer $a$ to $b$, a majority of voters prefer $b$ to $c$, yet a majority of voters prefer $c$ to $a$. Indeed, Arrow (1950) has shown that when there are more than two alternatives, the only voting rule that satisfies a small set of natural axioms is a dictatorship. Moreover, essentially any reasonable voting rule is susceptible to strategic behavior: Gibbard (1973) and Satterthwaite (1975) observed that under any 'fair' voting rule there exists a scenario where some voter benefits from misrepresenting her preferences.

These classic results provide ample evidence that preference aggregation is hard from a conceptual standpoint. On the other hand, preference aggregation is also hard in a very different sense: it can be shown that for many important voting rules computing the winner(s) is NP-hard. In particular, this is the case for the Kemeny rule, which is arguably the most natural method for aggregating a set of preference rankings into a single ranking, as well as for many popular multiwinner rules, such as Proportional Approval Voting, the Chamberlin–Courant rule and the Monroe rule (see Chapter 2 for definitions).

Now, social choice theorists have observed that the first source of hardness can be circumvented by focusing on scenarios where voters’ preferences share some common structure. The most famous result of this type dates back to the important early works of Black (1948) and Arrow (1951). They proved that if voters’ preferences are essentially single-dimensional, then there are no cycles in the majority preferences, and there is a voting rule that is strategyproof. The specific domain of preferences considered by Black and Arrow is that of single-peaked preferences; similar results have been subsequently obtained for other restricted preference domains, such as those of preferences that are single-crossing or single-peaked on a tree (to be formally defined later in this chapter).

It is then natural to ask whether the same approach can be used to circumvent computational complexity issues as well. The first foray in this direction
was made by Walsh (2007), and since 2007 hardness and easiness results for various restricted preference domains have been obtained for a number of problems including winner determination under a variety of voting rules, preference elicitation, as well as several forms of strategic behavior in voting.

Interestingly, while purely social choice-theoretic issues (such as existence of majority cycles) vanish as soon as we assume that voters’ preferences belong to a suitable restricted domain, many of the algorithms for voting-related problems rely on the knowledge of the respective structural relationship among voters/alternatives (such as the order of alternatives witnessing that the profile is single-peaked). Thus, to make use of these algorithms, one also needs an efficient procedure to determine whether a given preference profile has the required structural property and to find a respective witness. Consequently, the problem of designing such procedures has received a considerable amount of attention, too, resulting in polynomial-time algorithms for recognizing preferences that belong to several prominent restricted domains.

In this chapter, we will survey work on four specific topics that concern algorithmic properties of restricted preference domains. After defining the relevant concepts in Section 10.2, in Section 10.3 we discuss two extensions of the single-peaked domain, namely the domains of preferences that are single-peaked on trees and on circles, and show that several positive results for single-peaked preferences extend to these larger domains. In Section 10.4, we look at how the definitions of single-peaked and single-crossing preferences can be adapted to approval voting scenarios, and analyze the resulting preference domains from an algorithmic point of view. In Section 10.5, we review work on the complexity of strategic behavior in settings that are nearly structured. Finally, in Section 10.6, we demonstrate how assuming that voters’ preferences belong to a restricted domain can make preference elicitation more efficient.

10.2 Background

Suppose that citizens of a country X are about to vote on the (flat) tax rate. The set of alternatives is \( A = \{0\%, 1\%, \ldots, 100\%\} \), and it admits a natural ordering \( 0\% < 1\% < \cdots < 100\% \). Consider a voter \( i \) whose most-preferred alternative (the peak) is \( 35\% \). Then it is plausible that \( i \)'s preferences decrease as we move away from this peak: for example, we would expect \( 40\% >_i 50\% \), and \( 30\% >_i 20\% \). Such preferences are called single-peaked with respect to the given ordering \( < \).

Formally, let \( P = (>_1, \ldots,>_n) \) be a preference profile consisting of linear orders over an alternative set \( A \), and let \( N = \{1, \ldots, n\} \). Let \( \text{top}(i) \) denote the most-preferred alternative of voter \( i \). Given a linear order \( < \) of \( A \), we say that \(>_i \) is single-peaked with respect to \( < \) if for all \( a, b \in A \) such that \( \text{top}(i) < a < b \) or \( b < a < \text{top}(i) \) we have \( a >_i b \); we refer to \( < \) as an axis for \( A \). In other words, \(>_i \) is decreasing as we move in either direction from \( i \)'s peak. We say that the profile \( P \) is single-peaked if there exists
some axis $\prec$ such that for each voter $i \in N$ it holds that $\succ_i$ is single-peaked with respect to $\prec$.

The concept of single-peaked preferences was first proposed by Black (1948) and Arrow (1951), who noticed that for every single-peaked profile with an odd number of voters the majority relation is transitive and hence there exists a Condorcet winner. Further, there is a voting rule defined on single-peaked profiles that is strategyproof (see Moulin (1991, p. 263) for details). This result is known as the median voter theorem because of the form of this non-manipulable voting rule: it orders the voters in order of their peaks (according to $\prec$) and then selects the median voter’s peak, which is also the Condorcet winner.

Another notion of structure in voters’ preferences is based on ordering the voters rather than the candidates. A profile $P = (\succ_1, \ldots, \succ_n)$ of linear orders over $A$ is called single-crossing if voters can be ordered so that for all $a, b \in A$, the set of voters who prefer $a$ to $b$ forms an interval of this ordering. Thus, if the very first voter prefers $a$ to $b$, then there is some value $i$, $1 \leq i \leq n$, such that the first $i$ voters prefer $a$ to $b$ and the remaining $n - i$ voters prefer $b$ to $a$, i.e., the voters ‘switch’ from $a \succ b$ to $b \succ a$ at most once. Just like a single-peaked profile, a single-crossing profile is single-dimensional; in this case it is the voters who are ordered on an ‘ideological’ spectrum. Single-crossing profiles with an odd number of voters also enjoy a transitive majority relation. In fact, they have a so-called representative voter property: the majority relation is identical to the preference relation of the median voter with respect to the single-crossing order (Rothstein, 1991).

The class of one-dimensional Euclidean preferences (Coombs, 1950) is defined based on geometric considerations; all preference profiles in this class are both single-peaked and single-crossing. Formally, a profile $P = (\succ_1, \ldots, \succ_n)$ of linear orders over $A$ is called 1-Euclidean if there is a mapping $x : N \cup A \to \mathbb{R}$ which assigns every voter $i \in N$ a position $x(i)$ on the real line, and assigns every alternative $a \in A$ a position $x(a)$ on the real line, so that for all $i \in N$ and all pairs $a, b \in A$ we have $a \succ_i b$ if and only if $|x(i) - x(a)| < |x(i) - x(b)|$. Thus, in a 1-Euclidean profile, voters prefer closer alternatives to those that are further away. It is easy to see that every 1-Euclidean profile is single-peaked and single-crossing; the respective orderings of candidates and voters are given by an embedding $x$ witnessing that $P$ is 1-Euclidean. Yet, there are profiles that are both single-peaked and single-crossing, but fail to be 1-Euclidean (Elkind et al., 2014). The geometric approach extends to higher dimensions: a profile is $d$-Euclidean if there exists an embedding $x : N \cup A \to \mathbb{R}^d$ such that voters’ preferences are consistent with Euclidean distances to alternatives under this embedding.

### 10.2.1 Algorithmic Results

There are polynomial-time algorithms for recognizing single-peaked (Bartholdi III and Trick, 1986; Doignon and Falmagne, 1994; Escoffier et al., 2008), single-crossing (Doignon and Falmagne, 1994; Elkind et al., 2012; Bredereck et al., 2013) and 1-Euclidean (Doignon and Falmagne, 1994; Knoblauch, 2010; Elkind and Faliszewski, 2014) profiles; in contrast, Peters (2017) has shown that rec-
Recognizing $d$-Euclidean profiles is computationally hard. For single-peaked and single-crossing profiles, the recognition problem can be reduced to the consecutive 1s problem, which asks whether the columns of a 0–1 matrix can be permuted so that in each row all 1s appear consecutively; this problem is polynomial-time solvable (Booth and Lueker, 1976). (Section 10.4 provides an example of such a reduction for dichotomous preferences.) Both of these domains also admit direct polynomial-time recognition algorithms; for single-peaked preferences such an algorithm runs in time $O(mn)$, which is linear in the input size.

There are many examples of NP-hard social choice problems that become easy for single-peaked and single-crossing preferences. For instance, with an odd number of voters, both of these preference restrictions guarantee that the majority relation is transitive and, in particular, there exists a Condorcet winner. This implies that for profiles that satisfy these constraints the Kemeny rank aggregation rule can be evaluated in polynomial time (since the transitive majority relation gives an optimal ranking), and winners according to the Dodgson rule and the Young rule can be found efficiently (since the Condorcet winner is the unique winner for both rules). These results can be extended to profiles with an even number of voters (Brandt et al., 2015).

Similar results hold for several NP-hard multiwinner voting rules (see Chapter 2). For example, Betzler et al. (2013) showed that, given a single-peaked profile with $n$ voters and $m$ alternatives, we can find a winning committee according to the Chamberlin–Courant rule (Chamberlin and Courant, 1983) in time $O(m^2n)$ by a dynamic programming algorithm; this result can be extended to single-crossing preferences (Skowron et al., 2015) and to a few other multiwinner rules (Elkind and Ismaili, 2015; Peters, 2016). In essence, the algorithm of Betzler et al. (2013) proceeds along the axis $\prec$ from left to right, deciding whether to add candidates to the committee being constructed; note that this means that the algorithm needs to know such an axis, i.e., it relies on the existence of efficient recognition algorithms discussed earlier in this section. Some of the computational problems associated with various forms of strategic behavior (such as manipulation, control and bribery) also become polynomial-time solvable when voters’ preferences can be assumed to be single-peaked or single-crossing; we survey such results in more detail in Section 10.5.

10.3 Single-Peaked Preferences: Beyond the Line

The positive results for winner determination problems over restricted domains discussed above have a potential drawback: in practice, very few profiles are single-peaked. For example, under the impartial culture model, it is exponentially unlikely that a profile is single-peaked (Lackner and Lackner, 2017), and no real-world profile in PrefLib (see Chapter 15) is single-peaked. One can try to address this issue by extending the existing algorithms to profiles that are “nearly” single-peaked or single-crossing, for an appropriate distance measure; we will survey a sample of such results in Section 10.5. In the rest of this section, we will pursue a different agenda: instead of considering preferences that are single-peaked with respect to an axis, i.e., a path, we consider preferences that are...
single-peaked on more general graphs. We focus on two classes of graphs that
admit positive algorithmic and social choice-theoretic results, namely, trees and
cycles. This approach allows us to capture a broader class of preference profiles
and can be seen as a step towards mapping out the precise boundaries between
tractable and intractable instances of winner determination problems for several
important voting rules.

Preferences Single-Peaked on a Tree

Demange (1982) introduced the notion of preferences that are single-peaked on
a tree. Fix a set of alternatives $A$ and consider a tree $T = (A, E)$. A preference
order $\succ_i$ over $A$ is single-peaked on $T$ if $a \succ_i b$ whenever $a$ lies on the (unique) path
between $\text{top}(i)$ and $b$. Thus, a voter’s preferences decrease as we move away from
her peak along any path in $T$. A profile $P = (\succ_1, \ldots, \succ_n)$ over a set of alternatives
$A$ is said to be single-peaked on a tree if there is some tree $T = (A, E)$ such that
for each voter $i \in N$ the preference order $\succ_i$ is single-peaked on $T$. Note that this
definition is equivalent to the one in Section 10.2 when $T$ is a path.

To make sense of this definition, it is useful to consider the case where $T$ is a
star. Specifically, suppose that $T$ is a star with center $c$, so $E = \{\{c, a\} : a \in
A \setminus \{c\}\}$. Which preference orders are single-peaked on $T$? Consider a voter $i$ with
$\text{top}(i) = c$. No matter how she ranks the candidates in $A \setminus \{c\}$, her preferences
are necessarily single-peaked on $T$. On the other hand, if voter $i$’s peak is a leaf
vertex $a \neq c$, then $c$ lies on the path from $a$ to any other vertex $b$, and so we
must have $c \succ_i b$ for every $b \in A \setminus \{a, c\}$, i.e., $c$ must be $i$’s second-most-preferred
alternative; the remaining alternatives may appear in $\succ_i$ in an arbitrary order.
Thus, a preference order is single-peaked on $T$ if and only if $c$ occurs in first or
second position in that order.

This analysis shows that moving from paths to arbitrary trees gives us many
more profiles: there are only $\Theta(2^m - 1)$ orders that are single-peaked on a given
path, but there are $\Theta((m - 1)!)$ orders that are single-peaked on a given star.
However, this expansion comes at a cost: Demange (1982) shows that profiles
single-peaked on a tree are not guaranteed to have a transitive majority relation.
On the positive side, such profiles still admit a Condorcet winner, and a strategy-
proof voting rule. Moreover, Trick (1989) shows that it is possible to recognize
whether a given profile is single-peaked on a tree and to find a suitable tree in
$O(m^2 n)$ time. A natural next question, then, is whether hard winner determina-
tion problems become easier for profiles that are single-peaked on trees.

For the Dodgson rule and the Young rule, the answer is clearly positive as
long as the number of voters is odd: we can simply output the Condorcet winner.
On the other hand, our characterization of profiles single-peaked on a star shows
that finding a consensus ranking according to the Kemeny rule remains hard.
Indeed, we can transform an arbitrary profile into one that is single-peaked on
a star by adding a dummy candidate and placing it in the first position in every
vote; the consensus ranking for the original profile can be easily extracted from
the one for the new ‘structured’ profile.

For multiwinner rules, the results are somewhat disappointing as well. In par-
ticular, for a profile single-peaked on a tree, while one can efficiently compute a
winning committee under the egalitarian version of the Chamberlin–Courant rule, for the more common utilitarian version, the winner determination problem remains NP-hard (Yu et al., 2013). Interestingly, however, this hardness result does not apply to profiles that are single-peaked on a star. This is because including the center of the star in the committee ensures that each voter is quite well represented; filling the rest of the committee then boils down to choosing candidates that appear most often in the top position. This argument can be generalized to show that the problem of winner determination under the Chamberlin–Courant rule for preferences that are single-peaked on trees is fixed-parameter tractable with respect to the number of non-leaf vertices of the tree (Peters and Elkind, 2016). In a similar vein, Yu et al. (2013) show that dynamic programming can be used for trees that are ‘path-like’, in the sense of having a few leaves; they place the Chamberlin–Courant winner determination problem into the class XP with respect to the number of leaves. Both the algorithm for trees with few leaves and the algorithm for trees with few internal nodes rely on knowing a suitable tree; Yu et al. (2013) and Peters and Elkind (2016) show that it is indeed possible to efficiently decide whether a given profile is single-peaked on some such tree.

Preferences Single-Peaked on a Circle

Peters and Lackner (2017) initiate the algorithmic study of preferences that are single-peaked on a circle. A preference profile is said to be single-peaked on a circle if the alternatives can be arranged on a circle in such a way that for each voter we can cut this circle so that her preferences are single-peaked on the resulting path.

An intriguing property of this class of profiles is that it is closed under preference reversal: if an order is single-peaked on a circle, then so is the reverse of this order. In particular, a profile that combines orders that are single-peaked with respect to some axis and ones that are single-caved with respect to the same axis is single-peaked on a circle. Thus, in a political context, this model allows for voters with a preferred point along the ideological left-to-right spectrum as well as for ‘extremists’ who dislike centrist alternatives. It can also capture other application scenarios, including some that are more explicitly cyclic, such as scheduling international meetings across time zones or placing a facility (e.g., an airport) somewhere on the boundary of a city.

Profiles that are single-peaked on a circle do not inherit nice axiomatic properties of profiles that are single-peaked on a path; indeed, the (in)famous Condorcet cycle (i.e., the three-voter profile over \{x, y, z\} given by \(x \succ_1 y \succ_1 z, y \succ_2 z \succ_2 x,\) and \(z \succ_3 x \succ_3 y\)) is single-peaked on a circle. This means, in particular, that profiles single-peaked on a circle do not necessarily have a Condorcet winner. In fact, every majority relation can be realized by a profile that is single-peaked on a circle, as we can implement the construction in the proof of McGarvey’s theorem using a profile in this domain (Peters and Lackner, 2017). This implies that the Kemeny rule remains hard to evaluate on such profiles. Furthermore, the Gibbard–Satterthwaite Theorem can be proven using only profiles single-peaked on a circle (Kim and Roush, 1980), which means that there is no analogue of the median voter procedure for circles.
However, from an algorithmic perspective, this domain restriction turns out to be quite useful. For example, given a profile single-peaked on a circle, a greedy algorithm can efficiently compute winners according to the Young rule (Peters and Lackner, 2017). Also, for such profiles we can efficiently compute a winning committee under the Chamberlin–Courant rule and its variants, by reducing this problem to solving integer linear programs with totally unimodular constraint matrices (Peters, 2016; Peters and Lackner, 2017).

Preferences Single-Peaked on Arbitrary Graphs?

In principle, for any graph $G = (A, E)$, one can formally define what it means for a preference order over $A$ to be single-peaked on $G$: one can require that for each $b \in B$ the upper-contour set $\{a \in A : a \succ b\}$ is connected in $G$. Note that under this definition, every profile is single-peaked on the complete graph. However, as we try to move beyond trees and circles, we cannot expect many positive results: any class of graphs that contains circles would inherit the negative social choice-theoretic results for circles, and any class of graphs that contains trees would inherit the computational hardness results for trees. Moreover, the associated recognition problem may be difficult as well: e.g., a result of Gottlob and Greco (2013) implies that it is NP-hard to decide whether a profile is single-peaked on a graph of treewidth at most 3.

There are other possibilities for definitions of single-peaked preferences on arbitrary graphs, such as ones based on shortest paths (Nehring and Puppe, 2007). It would be interesting to compare them in terms of algorithmic usefulness.

10.4 Structure in Dichotomous Preferences

Approval Voting is one of social choice theorists’ favorite voting rules (Brams and Fishburn, 2007; Laslier and Sanver, 2010). It asks voters to report dichotomous preferences, i.e., to split the alternatives into approved and disapproved choices—a dichotomy. It then selects the alternative(s) with the maximum number of approvals. This voting rule has many desirable axiomatic properties, but to a large extent its attraction stems from its input format: it is easy for voters to make up their mind about which preferences to report, it is easy to elicit such preferences, and it is easy to reason mathematically about them. However, some attractive voting rules for dichotomous preferences are still hard to evaluate, particularly in the multiwinner setting. It is therefore natural to ask what it means for dichotomous preferences to be essentially one-dimensional, and whether the respective preference restrictions are algorithmically useful.

Building on earlier work of List (2003), Dietrich and List (2010) and Faliszewski et al. (2011), a recent paper by Elkind and Lackner (2015) considers several ways of extending the definitions of single-peaked, single-crossing and 1-Euclidean preferences to the dichotomous setting. The paper studies algorithmic properties of the resulting domains, focusing on the complexity of recognizing profiles that belong to these domains and computing the outputs of well-known approval-based multiwinner rules.
VEI: Voter Extremal Interval
CEI: Candidate Extremal Interval
DUE: Dichotomous Uniform Euclidean
VI: Voter Interval
CI: Candidate Interval
DI: Dichotomous Euclidean
PSP: Possibly Single-Peaked
PE: Possibly Euclidean
PSC: Possibly Single-Crossing

Figure 10.1: Relations between notions of structure for dichotomous preferences as established by Elkind and Lackner (2015). Arrows indicate containment; more restrictive domains are at the top.

**Defining Domain Restrictions for Dichotomous Preferences**

One approach that allows us to adapt any preference restriction defined for linear orders to the realm of dichotomous orders is to view dichotomies as weak orders: we can ask if it is possible to refine each weak order in a given dichotomous profile so as to obtain a profile of linear orders with a given structural property (Lackner, 2014; Elkind et al., 2015). Formally, we say that a linear order \( \succ \) extends an approval ballot \( B \) if for every pair of candidates \((a, b)\) such that \( a \) is approved in \( B \) and \( b \) is not approved in \( B \) we have \( a \succ b \). We then say that a dichotomous profile belongs to the domain of possibly single-peaked (PSP) profiles if it can be extended to a profile of linear orders that is single-peaked. Possibly single-crossing (PSC) and possibly 1-Euclidean (PE) dichotomous profiles can be defined in a similar manner.

For linear orders, it is known that the single-peaked and the single-crossing domain overlap, but neither is contained in the other, and that the 1-Euclidean domain is strictly contained in their intersection (see, e.g., Elkind et al., 2014). Interestingly, the relationship among their approval-based cousins is different: Elkind and Lackner (2015) show that PSP coincides with PE, whereas PSC is a strict subdomain of PSP/PE (see Figure 10.1).

A more direct approach is based on the idea of contiguity: we could say that a dichotomous profile is single-peaked if there exists an ordering of candidates such that each voter’s approval set forms an interval of this ordering. This definition is used by Faliszewski et al. (2011); we will say that such profiles belong to the candidate interval (CI) domain. Similarly, a profile belongs to the voter interval (VI) domain if the voters can be ordered so that for every candidate \( c \), the set of voters approving \( c \) forms an interval of that ordering. Stronger variants of both properties require every interval to contain the leftmost or the rightmost element of the candidate/voter ordering; this yields the candidate/voter extremal interval (CEI/VEI) domains.

For single-peaked preferences, the two approaches result in the same class of profiles: Elkind and Lackner (2015) show that the CI domain coincides with the PSP domain (and hence with the PE domain). In contrast, for single-crossing
preferences, this is not the case: VI is strictly contained in PSC. Further, for the interval-based approach, the relationship between CI and VI (direct analogues of the single-peaked and the single-crossing domains) is similar to that for linear orders: CI and VI do not contain each other and have a non-empty intersection, which strictly contains a dichotomous analogue of 1-Euclidean preferences (to be defined in the next paragraph).

The interval-based approach can also be applied to the 1-Euclidean domain. We say that a profile belongs to the dichotomous Euclidean (DE) domain if voters and candidates can be positioned on the real line so that for every voter \( i \) there exists a radius \( r_i \) such that all candidates within a distance \( r_i \) of \( i \) are approved by \( i \). We can also require the radius \( r \) to be the same for all voters; the resulting domain is called the dichotomous uniform Euclidean (DUE) domain. Remarkably, DE turns out to coincide with CI: the order of candidates in a DE embedding witnesses that the profile belongs to CI, and for the converse direction we can place the candidates on the real line in a way that respects the CI ordering, and then pick a suitable position for each voter. In contrast, the DUE domain is much smaller; in particular, every DUE profile belongs to the VI domain (and similarly to the case of linear orders, there are profiles that are CI and VI, but not DUE).

So far in this section, we focused on one-dimensional preference domains. However, the approaches based on Euclidean distances can be easily generalized to higher dimensions. Let us say that a profile belongs to the \( d \)-DE domain for \( d \in \mathbb{N} \) if voters and candidates can be placed in \( \mathbb{R}^d \) so that for every voter \( i \), there exists a radius \( r_i \) such that \( i \) approves exactly the candidates in the \( r_i \)-ball around \( i \); the \( d \)-DUE domain is defined similarly, with the additional restriction that \( r_i = 1 \) for each voter \( i \).

**Recognition Algorithms**

Elkind and Lackner (2015) show that almost all one-dimensional restricted domains defined earlier in this section can be recognized in polynomial time; the only exception is PSC, for which the complexity is open. All the polynomial-time algorithms except for the one for DUE are based on reductions to the consecutive 1s problem, defined in Section 10.2.

To illustrate the proof technique, we will now show how to reduce the problem of deciding if a given dichotomous profile belongs to the CI domain to an instance of the consecutive 1s problem; our reduction is illustrated in Figure 10.2. Given
a dichotomous profile, we construct a binary matrix that contains a row for each voter and a column for each candidate; the entry associated with voter \( v \) and candidate \( c \) is set to 1 if \( v \) approves \( c \) and to 0 otherwise. By construction, a permutation of the columns that results in 1s appearing consecutively in each row corresponds to a permutation of candidates witnessing that the input profile belongs to the CI domain. Similar reductions work for VI, CEI, and VEI. For DUE, there is a reduction to recognizing bipartite permutation graphs.

Peters (2017) shows that detecting whether a given profile belongs to \( d \)-DE or to \( d \)-DUE is NP-hard for \( d \geq 2 \); more precisely, he shows that these problems are \( \exists \text{R} \)-complete. In this respect, dichotomous orders behave like linear orders.

### Algorithms for Approval-Based Multiwinner Rules

Let us now turn to applications of the preference restrictions considered in this section. We consider two multiwinner voting rules that are defined for dichotomous preferences (see Chapter 2 for a more general discussion of multiwinner rules), namely Maximin Approval Voting (MAV) and Proportional Approval Voting (PAV). For both of these rules computing a winning committee is NP-hard (LeGrand et al., 2007; Skowron et al., 2016; Aziz et al., 2015). Hence, it is natural to ask whether focusing on restricted domains, such as CI, VI, CEI, VEI, etc. allows for faster algorithms.

We will first define the MAV rule. Let \( A_i \) denote the set of candidates approved by voter \( i \). Given a target committee size \( k \), MAV returns a set of candidates \( W \), \(|W| = k\), that minimizes \( \max_{i \in N} |W \setminus A_i| + |A_i \setminus W| \), i.e., the maximum Hamming distance between a voter’s preferences and the committee (both viewed as 0/1 strings). Liu and Guo (2016) prove that a winning committee under MAV can be computed in polynomial time for preference profiles that belong to CI or VI. This is achieved by dynamic programming algorithms that exploit the structure of the respective preferences. As a consequence, winner determination is also easy for DUE, VEI, and CEI preferences (cf. Figure 10.1).

PAV is a less egalitarian, but more proportional rule than MAV. It returns a set of candidates \( W \), \(|W| = k\), that maximizes \( \sum_{i \in N} h(|A_i \cap W|) \), where \( h(1) = 1 \), \( h(2) = 1 + \frac{1}{2} \), \( h(3) = 1 + \frac{1}{2} + \frac{1}{3} \), etc. Elkind and Lackner (2015) showed that for preference profiles that belong to CEI or VEI, a winning committee under PAV can be computed in polynomial time via dynamic programming. Recently, Peters (2016) extended this result to the CI domain, using a very different approach: he shows that this problem reduces to solving an integer linear program with a totally unimodular constraint matrix. Whether a polynomial-time algorithm is also possible for the VI domain is an open problem.

### 10.5 Nearly Structured Preferences

While definitions in Sections 10.2 and 10.3 are mathematically appealing, we cannot expect real-world preference data to satisfy them. Indeed, for all domains we consider, the presence of a single voter with an unorthodox opinion, or a few minor errors made during the preference elicitation process, may result in
a profile that does not belong to the target domain. At the same time, we do not encounter arbitrary combinations of preference orders in real-life preference data, and in many cases we expect the voters’ preferences to be essentially one-dimensional. One way to formalize this intuition is to define what it means for a profile to be nearly single-peaked or nearly single-crossing, and then verify whether these definitions are satisfied by the available preference data. A related question is whether tractability results for structured domains can be extended to nearly structured domains.

**Defining Nearly Structured Preferences**

This research agenda was put forward by Faliszewski et al. (2014), who focused on single-peaked preferences. They proposed several measures of distance to the single-peaked domain, including, in particular, the number of voters that have to be deleted from a given profile so as to make it single-peaked. Alternatively, one can ask how many candidates need to be removed to make a profile single-peaked; this measure was suggested by Escoffier et al. (2008). Another approach shares motivation with the definitions of the Dodgson rule and the Kemeny rule: we ask how many swaps of adjacent candidates are needed to arrive to a single-peaked profile (Erdélyi et al., 2017). An egalitarian variant of this measure, which asks what is the smallest number of swaps such that a given profile can be made single-peaked by performing at most candidate swaps per vote, was proposed by Faliszewski et al. (2014). We can also try to partition voters (Escoffier et al., 2008) or candidates (Erdélyi et al., 2017) into a small number of sets, so that each component forms a single-peaked profile. Yet another closeness measure is based on the idea of **decloning**. Recall that a set of candidates forms a **clone set** if each voter ranks these candidates consecutively in her vote. To make a given profile single-peaked, we can ‘collapse’ one or more clone sets by replacing each such set with a single candidate; the ‘cost’ of this operation can be measured as the overall reduction in the number of candidates (Elkind et al., 2012) or the size of the largest clone set that we collapsed (Cornaz et al., 2012). Of course, each of these approaches can also be used to measure how close a given profile is to being single-crossing, 1-Euclidean, or single-peaked on a tree.

The suitability of each of these closeness measures depends on the kind of errors we expect: for instance, the swap-based approach implicitly assumes that the preferences are fundamentally single-peaked, but small errors have been made during the elicitation process, whereas the decloning-based approach is based on the intuition that the set of available options is one-dimensional, yet some of the options are represented by several virtually indistinguishable alternatives. When several types of errors can be present, it may be useful to combine several closeness measures, e.g., to allow, say, a few candidate deletions and a small number of swaps.

**Recognition of Nearly Structured Preferences**

It is natural to ask if we can efficiently determine whether a given profile is nearly structured. Technically, we are interested in computing the number of modifi-
cations of a given type that are necessary to make a given profile single-peaked or single-crossing. The complexity of this task has been considered by several authors. Erdélyi et al. (2017) focus on computing the distance to the single-peaked domain, for many of the distance measures listed above. Bredereck et al. (2016) consider both the single-peaked domain and the single-crossing domain, but limit themselves to two types of modifications, namely, voter deletion and candidate deletion. The complexity of optimally decloning a given profile so as to make it single-peaked or single-crossing was investigated by Cornaz et al. (2012) and Elkind et al. (2012). Most of the results in these papers are negative: checking if a given profile is close to being single-peaked or single-crossing is typically NP-hard. However, there are several notable exceptions: it can be efficiently decided how many candidates have to be deleted to make an election single-peaked (Erdélyi et al., 2017) or how many voters need to be deleted to make an election single-crossing (Bredereck et al., 2016); also, there are several positive results for optimal decloning (Cornaz et al., 2012; Elkind et al., 2012). Moreover, Elkind and Lackner (2014) provide efficient constant-factor approximation algorithms for all computational problems considered by Bredereck et al. (2016).

**Manipulation and Control with Nearly Structured Preferences**

We have already seen a few examples of hard computational social choice problems that become polynomial-time solvable for single-peaked or single-crossing preferences. One may then wonder if such results extend to preferences that are nearly single-peaked or nearly single-crossing. Faliszewski et al. (2014) were the first to ask this question for coalitional manipulation and control.

An instance of the constructive coalitional manipulation problem is given by an election, a distinguished candidate $p$ and a positive integer $k$; we ask if we can add $k$ new voters (manipulators) to the election to make $p$ an election winner. In the weighted variant of this problem (CCWM), each of the (old and new) voters is associated with an integer voting weight (encoded in binary). For $k > 1$, finding a successful manipulation is typically NP-hard (see, e.g., Conitzer and Walsh, 2016). However, if the existing voters’ preferences are known to be single-peaked with respect to a given axis, and the manipulators’ votes are required to be single-peaked with respect to the same axis, CCWM becomes polynomial-time solvable for several voting rules (Faliszewski et al., 2011). This is viewed as a negative result, since NP-hardness results for manipulation are often interpreted as ‘barriers’ to strategic behavior; thus, for single-peaked preferences these barriers may disappear. However, as argued above, while real-life preferences may be close to single-peaked, they are unlikely to be single-peaked; does this mean that we can expect manipulation to be NP-hard in practical scenarios?

Faliszewski et al. (2014) show that easiness results for CCWM with single-peaked preferences can be fragile. For instance, they identify a class of 3-candidate scoring rules for which CCWM has been shown to be NP-hard for general profiles and polynomial-time solvable for single-peaked profiles (Faliszewski et al., 2011), and show that it remains NP-hard for profiles that can be made single-peaked by deleting a single voter. They obtain a similar result for profiles that can be made single-peaked by swapping at most one pair of alternatives in
each vote; the hardness proof is based on the observation that any preference order over three candidates can be made single-peaked with respect to a given axis by a single swap. Further results of this type have been obtained by Erdélyi et al. (2015) for other measures of closeness to the single-peaked domain. However, Erdélyi et al. (2015) also present an easiness result: for \( k \)-approval with \( m \) candidates, CCWM is polynomial-time solvable for profiles that can be made single-peaked by deleting \( \ell \) voters as long as \( \ell < \frac{2k-m}{m-k} \). CCWM for nearly single-peaked profiles is also considered by Menon and Larson (2016), who are interested in the complexity of this problem when voters are allowed to submit partial orders of a certain form, namely top-truncated ballots.

Another type of strategic behavior considered by Faliszewski et al. (2014) is (constructive) control: here, the goal is to make a certain candidate an election winner by adding or deleting a given number of voters or candidates; the difference between control by adding voters and coalitional manipulation is that in the former problem the voters to be added have to be selected from a given pool of voters. Faliszewski et al. (2014) investigate the complexity of these forms of control for several voting rules, including Plurality and \( t \)-approval; they identify several scenarios where a control problem is hard for general preferences, but can be solved in polynomial time for preferences that are single-peaked or can be made single-peaked by removing a constant number of voters or performing at most a constant number of candidate swaps in each vote. Yang and Guo (2014a,b, 2015) continue this line of inquiry for some other measures of closeness to the single-peaked domain, and obtain several fixed parameter tractability results with respect to the number of modifications needed to make a profile single-peaked. See also the survey by Hemaspaandra et al. (2016).

### 10.6 Elicitation of Structured Preferences

While much of the work in (computational) social choice deals with aggregating the collective preferences into a joint decision, sometimes the goal is simply to elicit the voters’ preferences over the alternatives. It is typically assumed that we know the number of voters \( n \) and the set of alternatives \( A, |A| = m \), and have access to an oracle that, given a triple \((i, a, b) \in \mathbb{N} \times A \times A\), outputs 1 if the \( i \)th voter prefers \( a \) to \( b \) and 0 otherwise. The goal is then either to fully determine the preference order of each voter or to obtain enough information to determine the winner(s) under a given voting rule. For unrestricted preferences, the complexity of the former task can be easily seen to be \( \Theta(nm \log m) \): effectively, we have to ‘sort’ the \( m \) alternatives in the correct order for each of the \( n \) voters.

Perhaps unsurprisingly, this problem, too, becomes easier when voters’ preferences belong to a restricted domain. In this section, we provide a brief summary of three papers on this topic: an early paper by Conitzer (2009), who considers single-peaked and 1-Euclidean preferences, and two very recent papers by Dey and Misra (2016a,b), which deal with, respectively, single-crossing preferences and preferences that are single-peaked on a tree.

We start by considering the single-peaked domain. Suppose first that the axis is known: assume without loss of generality that it is given by \( a_1 \triangleleft a_2 \triangleleft \cdots \triangleleft a_m \).
Consider a voter $i$. We know that her least preferred alternative is either $a_1$ or $a_m$. Thus, by asking the oracle whether $i$ prefers $a_1$ to $a_m$, we can determine the alternative ranked last in $i$’s preference order. We can then continue recursively, building up $i$’s preference order from the bottom to the top. Clearly, $m - 1$ queries suffice to elicit the full preference order, so for $n$ voters the number of queries reduces from $O(nm \log m)$ to $O(nm)$. Conitzer (2009) describes an alternative $O(m)$ elicitation algorithm. His algorithm uses binary search to identify the voter’s top alternative $a^*$. If $a^* = a_t$ for some $t \in [m]$, we know that the voter orders the alternatives in $\{a_1, \ldots, a_{t-1}\}$ and $\{a_{t+1}, \ldots, a_m\}$ as $a_{t-1} \succ \ldots \succ a_1$ and $a_{t+1} \succ \ldots \succ a_m$, respectively, so it remains to merge these two orders; this can be accomplished in linear time.

To see that the number of queries for a single voter cannot be reduced to $o(m)$, suppose that $m = 2t - 1$ and consider the weak order $a_t \succ \{a_{t-1}, a_{t+1}\} \succ \ldots \succ \{a_1, a_m\}$. Every linear order that refines this weak order is single-peaked with respect to $a_1 \cup \cdots \cup a_m$. Thus, to identify a specific linear order from this set, we would have to query the oracle about each of the $\frac{m-1}{2} = \Omega(m)$ pairs $(a_{t-1}, a_{t+1}), \ldots, (a_1, a_m)$.

Now, suppose that the axis is not known. Then there is not much we can do for a single voter: saying that her preferences are single-peaked with respect to some axis provides no information whatsoever. However, Conitzer (2009) shows that if the number of voters is large, the number of queries can be essentially as low as in the case where the axis is known. His algorithm elicits the ranking of a single voter (using the trivial $O(m \log m)$ algorithm) and then uses it as a guiding order to elicit the preferences of the remaining voters; each additional ranking can be elicited in $O(m)$ queries given the first ranking. The overall number of queries is then $O(m \log m + nm)$.

For the 1-Euclidean domain, knowing the positions of the alternatives on the axis provides an impressive reduction in the number of queries: Conitzer (2009) demonstrates that eliciting a single voter’s preferences only requires $2\lceil \log m \rceil$ queries. To see why this is the case, suppose that alternative $a_j$ appears in position $x_j$ on the axis, with $x_1 < \cdots < x_m$. Then voter $i$ prefers $a_j$ to $a_k$, $j < k$, if and only if she is positioned to the left of $\frac{x_j + x_k}{2}$. Thus, to determine a voter’s ranking, it suffices to determine her position with respect to each of the $\binom{m}{2}$ points of the form $\frac{x_j + x_k}{2}$, $j, k \in [m], j < k$. These points divide the axis into $\binom{m}{2} + 1$ intervals, with voters in each interval having the same preference order. The appropriate interval for each voter can be identified by asking $\lceil \log \binom{m}{2} + 1 \rceil \leq 2\lceil \log m \rceil$ queries using binary search. However, Conitzer (2009) shows that if the embedding of the alternatives into the line is not known, then it is not possible to do better than in the single-peaked case.

For single-crossing preferences, the relevant additional information is the single-crossing order of the voters. Dey and Misra (2016a) observe that when this order is known and we can query the voters in any order, all we need to do is to elicit the preferences of the first voter and then find a ‘crossing point’ for each pair of alternatives (i.e., if the first voter ranks $a$ above $b$, we need to find the first voter in the single-crossing order who ranks $b$ above $a$). Indeed, before the crossing point, all voters agree on that pair of alternatives with the first voter, and from that point on they disagree with her on that pair. The first voter’s ranking
can be elicited using $O(m \log m)$ queries, and the crossing point for each of the $\binom{m}{2}$ pairs can be found using binary search over $N$. Altogether, we need $O(m^2 \log n)$ queries; while this bound is incomparable with the $O(nm \log m)$ bound for general preferences, it provides a significant improvement for the setting where the number of voters is much larger than the number of alternatives.

The analysis above assumes that one has full control over the order of queries. However, it may be the case that the voters arrive one by one (in the single-crossing order or in an arbitrary order) and one has to elicit a voter’s preferences when she arrives (i.e., by the time we start querying voter $i$, we must have elicited the full rankings of all voters who arrived before $i$). For the sequential model, Dey and Misra (2016a) propose an algorithm that ‘expects’ the preference order being elicited to be similar to the nearest ranking among the ones elicited so far. If this is indeed the case, the current order can be elicited quickly, as the number of disagreements with the neighboring order will be small. Each disagreement contributes to the elicitation cost, but the total number of disagreements can be bounded by above for any single-crossing profile. The resulting algorithm asks $O(nm + m^2)$ queries if the voters arrive in the single-crossing order and $O(nm + m^2 \log n)$ queries if the arrival order can be arbitrary.

If the single-crossing order is not known, one can use the fact that the number of distinct preference orders in a given single-crossing profile is bounded by $\binom{m}{2} + 1$ (Bredereck et al., 2013). Thus, throughout the elicitation process, there may be at most $\binom{m}{2} + 1$ voters whose preference orders are different from all rankings elicited so far (and are therefore costly to elicit), and in all other cases we can quickly find a ‘match’ among the already elicited rankings. This approach leads to an algorithm that makes $O(nm + m^3 \log m)$ queries. Dey and Misra (2016a) also provide lower bounds for each of their models; for most (though not all) models these bounds are tight when $n = \Omega(m^3 \log m)$.

We now consider the case when voters’ preferences are single-peaked on a tree. If there are no additional constraints on the structure of the tree, we cannot expect to have an asymptotic improvement over the general case, even if the tree and the assignment of the alternatives to the vertices of that tree are known. Indeed, observe that all $(m - 1)!$ rankings that place some alternative $a \in A$ first are single-peaked on a star with $a$ in the center, so it may take $\log((m - 1)!)$ = $O(m \log m)$ queries to identify a specific ranking in this set. However, Dey and Misra (2016b) show that one can obtain improved bounds when the tree is, in some sense, close to a path. Specifically, for trees that can be covered with $k$ paths, they bound the number of queries by $O(nm \log k)$; in particular, this implies an upper bound of $O(nm \log \ell)$ for trees with $\ell$ leaves. The algorithm proceeds by eliciting each voter’s preferences along each path in the cover, and then merging the results using the standard $k$-way merging algorithm. A similar argument shows that if a tree can be turned into a path by removing $d$ vertices, the query complexity can be bounded by $O(nm + nd \log d)$. However, Dey and Misra (2016b) show that we still need $\Omega(nm \log m)$ queries if the tree has bounded degree (in fact, the lower bound holds even if the degree of each vertex is at most 3). Moreover, our analysis for the star shows that the same lower bound applies if the tree in question has bounded pathwidth or bounded diameter.
10.7 Further Directions and Trends

At the end of this chapter we would like to highlight a few research directions that we consider promising. First, most of this chapter has focused on domain restrictions that are in some sense one-dimensional: single-peaked, single-crossing and 1-Euclidean preferences are all defined by a linear order or an embedding into the real line. Multidimensional analogues of these notions have received much less attention in the computational social choice literature. In particular, little is known about computational benefits of such higher-dimensional restrictions. For example, it is not known whether the Kemeny rule is computable in polynomial time on two-dimensional single-peaked profiles (for definitions, see Sui et al., 2013). Other natural higher-dimensional restricted domains arise from 1-Euclidean preferences—their definition can easily be extended to more dimensions. More dimensions also make the choice of metric interesting: apart from the Euclidean $\ell_2$-metric, the $\ell_1$- or $\ell_\infty$-metrics are sensible choices as well (Peters, 2017). Even if NP-hard voting problems remain hard for these domains, it might be that better approximation algorithms can be found than for general preferences. Multidimensional domain restrictions offer many challenging research questions, but faster algorithms for these classes are very desirable: these algorithms would be applicable to a much larger class of preferences than algorithms for one-dimensional restrictions.

We have presented a number of results for one-dimensional dichotomous preference domains. More broadly, one can consider trichotomous or even $k$-chotomous preferences (see, e.g., Ju, 2005; Zwicker, 2016). An example for trichotomous preferences would be the distinction between satisfying, acceptable, and unsatisfying candidates, thus allowing for the indication of compromise outcomes. Notions of structure specifically for $k$-chotomous preferences have not yet been studied, but some of the concepts discussed in this chapter can easily be adapted to this setting.

Another direction is to consider completely new domain restrictions. Domains suggested in the social choice literature usually guarantee the existence of a Condorcet winner, but this is not a necessarily relevant property for algorithmic purposes. Inspiration could be found by adapting structural concepts from graph theory, such as restrictions resembling treewidth. For a systematic study of domain restrictions, the framework of forbidden subprofiles (Ballester and Haeringer, 2011; Bredereck et al., 2013) could prove to be valuable. Preference profiles (sets of linear orders) are mathematically rich structures and there is hope for a similarly diverse and powerful classification of structure as exists for graph classes—along with algorithmic applications of these structural restrictions.

Finally, the work on structured preferences has mostly focused on voting-related topics: winner determination, manipulation, control, etc. Given the advances that have been made in these fields, it could prove to be worthwhile to investigate the impact of structured preferences in other fields of social choice; fair division and judgment aggregation are natural candidates.
**Acknowledgments**

This work was supported by ERC Starting Grant ACCORD (GA 639945). A shorter version of this survey, with a slightly different focus, is Elkind et al. (2016).

**Bibliography**


