Social Choice Theory and Multicriteria decision aiding

D. Bouyssou [1] CNRS

Th. Marchant [2] Universiteit Gent

P. Perny [3] LIP6

17th June 2008

[1] LAMSADE, Université Paris Dauphine, Place du Maréchal de Lattre de Tassigny, F-75775 Paris Cedex 16, France, tel : +33 1 44 05 48 98, fax : +33 1 44 05 40 91, e-mail : bouyssou@lamsade.dauphine.fr.

[2] Universiteit Gent, H. Dunantlaan 1, B-9000 Gent, Belgique, tel : +32 9-264.63.73, fax : +32 9-264.64.87, e-mail : Thierry.Marchant@UGent.be.

[3] LIP6 – Université Pierre et Marie Curie, 8 rue du Capitaine Scott 75015 Paris, FRANCE, tel : + 33 1 44 27 70 04, fax : + 33 1 44 27 70 00 e-mail : Patrice.Perny@lip6.fr.

Abstract

The aim of this chapter is to present some important results in Social Choice Theory in a simple way and to discuss their relevance for multicriteria decision aiding. Using some classical examples of voting problems (Section 2), we will show some fundamental difficulties arising when aggregating preferences. We will then present some theoretical results that can help us better understand the nature of these difficulties (Section 3). We will then try to analyze the consequences of these results for multicriteria decision aiding (Section 4). A long references list will help the interested reader to deepen his understanding of these questions.

Keywords : aggregation, multiciteria decision aiding, Social Choice Theory, elections

1 Introduction

Many organisations face such complex and important management problems that they sometimes want their decisions to be based on a 'scientific preparation', sometimes called a decision analysis. The expert in charge of this preparation faces many diverse tasks: stakeholders identification, problem statement, elaboration of a list of possible actions, definition of one or several criteria for evaluating these actions, information gathering, sensitivity analysis, elaboration of a recommendation (for instance a ranking of the actions or a subset of 'good' actions), etc. The desire or necessity to take multiple conflictual viewpoints into account for evaluating the actions often makes his task even more difficult. In that case, we speak of multicriteria decision aiding (see Pomerol and Barba-Romero, 1993 ; Vincke, 1989 ; Roy, 1985). The expert must then try to synthetize the partial preferences (modelled by each criterion) into a global preference on which a recommendation can be based. This is called preference aggregation.

A very similar aggregation problem has been studied for long in the framework of voting theory. It consists in searching a 'reasonable' mechanism (we call it voting system or aggregation method in the sequel) aggregating the opinions expressed by several voters about the candidates in an election, in order to determine a winner or to rank all candidates in order of preference. This problem is of course very old but its modern analysis dates back to the end of the eighteenth century with works by Borda (1781) and Condorcet (1785).

The diversity of voting systems actually used in the world shows that this problem is still important. In the 1950s, the works of Arrow (1963), Black (1958) and May (1952) have initiated a huge literature (see Kelly, 1991) forming what is today called social choice theory. It analyzes the links that (should) exist between the individual preferences of the members of a society and the decisions made by this group when these decisions are supposed to reflect the collective preference of the group.

The many results obtained in social choice theory are valuable for multicriteria decision aiding. There are indeed links between these two domains: it is easy to go from one to the other by replacing the words 'action', 'criterion', 'partial preference' and 'overall preference' by 'candidate', 'voter', 'individual preference' and 'collective preference' (see Arrow and Raynaud, 1986).

The aim of this chapter is to present some important results in Social Choice Theory in a simple way and to discuss their relevance for multicriteria decision aiding. Using some classical examples of voting problems (Section 2), we will show some fundamental difficulties arising when aggregating preferences. We will then present some theoretical results that can help us better understand the nature of these difficulties (Section 3). We will then try to analyze the consequences of these results for multicriteria decision aiding (Section 4). A long references list will help the interested reader to deepen his understanding of these questions.

2 Introductory examples

Choices made by a society often impact the individuals making up this society. It therefore seems reasonable to ground these choices on the preferences of the individuals. The choice of a candidate (law, project, social state, etc.) then depends on the outcome of an election in which the individuals (voters) express their preferences. A voting system (or aggregation method) uses the information provided by the voters in order to determine the elected candidate or, more generally, the decision made by the group.

In such conditions, how should we conceive a 'good' voting system? 'Common sense' tells us that such a system must be democratic, i.e., it must yield collective preferences reflecting as much as possible the individual preferences. In many countries (groups, companies, committees), this is operationalized by the majority rule (or some variant of it): candidate a wins against b if the majority of the voters prefer a to b. This simple rule is very intuitive. As we will later see, when there are only two candidates, this rule raises almost no problem (see May, 1952).

This rule can be adpated in many ways to face situations with more than two candidates. These adaptations can lead to surprising outcomes. This will be illustrated by a few examples in this section. We will begin with uninominal voting systems, where each voter chooses one candidate (section 2.1), before moving to other systems where the voters can express their preferences in more complex ways (section 2.2).

In all examples, we will assume that each voter is able to rank (possibly with ties) all candidates in order of preference, i.e., he can express his preferences by means of a weak order. If a voter prefers a to b and b to c (thereby prefering a to c), we write '$a \succ b \succ c$'. Except otherwise stated, we will suppose that the voters are sincere, i.e., they express their 'true' preferences. Finally, notice that most examples presented here are classic. Many more examples and the analysis of many voting systems can be found in Moulin (1980, 1988), Dummet (1984), Fishburn (1977) and Nurmi (1987).

2.1 Uninominal systems

Example 1 (Dictatorship of majority)
Let $\{a, b, c, \ldots, z\}$ be the set of 26 candidats for an election with 100 voters whose preferences are :

51	voters have preferences	$a \succ b \succ c \succ \ldots \succ y \succ z$,
49	voters have preferences	$z \succ b \succ c \succ \ldots \succ y \succ a$.

It is clear that 51 voters will vote for a while 49 vote for z. Thus a has an absolute majority and, in all uninominal systems we are aware of, a wins. But is a really a good candidate ? Almost half of the voters perceive a as the worst one. And candidate b seems to be a good candidate for everyone. Candidate b could be a good compromise. As shown by this example, a uninominal election combined with the majority rule allows a 'dictatorship of majority' and doesn't favour a compromise. A possible way to avoid this problem might be to ask the voters to provide their whole ranking instead of their preferred candidate. We will see some examples in Section 2.2. $\diamond$

The possibility of a dictatorship of the majority was already acknowledged by classic greek philosophers. The following examples show that many other strange phenomena can occur with uninominal voting systems.

Example 2 (Respect of majority in the British system)
The voting system in the United Kingdom is plurality voting, i.e. the election is uninominal and the aggregation method is simple majority. Let $\{a, b, c\}$ be the set of candidates for a 21 voters election (or 21×10^6 voters if one wishes a more realistic example). Suppose that

10	voters have preferences	$a \succ b \succ c$,
6	voters have preferences	$b \succ c \succ a$,
5	voters have preferences	$c \succ b \succ a$.

Then a (resp. b and c) obtains 10 votes (resp. 6 and 5) so that a is chosen. Nevertheless, this might be different from what a majority of voters wanted. Indeed, an absolute majority of voters prefers any other candidate to a (11 out of 21 voters prefer b and c to a). $\diamond$

Let us see, using the same example, if such a problem would be avoided by the two-stage French system. After the first stage, as no candidate has an absolute majority, a second stage is run between candidates a and b. We suppose that the voters keep the same preferences on $\{a, b, c\}$. So,

10	voters have preferences	$a \succ b$,
11	voters have preferences	$b \succ a$.

Thus a obtains 10 votes and b, 11 votes so that candidate b is elected. This time, none of the beaten candidates (a and c) are preferred to b by a majority of voters. Nonetheless we cannot conclude that the two-stage French system is superior to the British system from this point of view, as shown by the following example.

Example 3 (Respect of majority in the two-stage French system)
Let $\{a, b, c, d\}$ be the set of candidates for a 21 voters election. Suppose that

10	voters have preferences	$b \succ a \succ c \succ d$,
6	voters have preferences	$c \succ a \succ d \succ b$,
5	voters have preferences	$a \succ d \succ b \succ c$.

After the first stage, as no candidate has absolute majority, a second stage is run between candidates b and c. Candidate b easily wins with 15 out of 21 votes though an absolute majority (11/21) of voters prefer a and d to b. $\diamond$

Because it is not necessary to be a mathematician to figure out such problems, some voters might be tempted not to sincerely report their preferences as shown in the next example.

Example 4 (Manipulation in the two-stage French system)
Let us continue with the example used above. Suppose that the six voters having preferences $c \succ a \succ d \succ b$ decide not to be sincere and vote for a instead of c. Then candidate a wins after the first stage because there is an absolute majority for him (11/21). If they had been sincere (as in the previous example), b would have been elected. Thus, casting a non sincere vote is useful for those 6 voters as they prefer a to b. Such a system, that may encourage voters to falsely report their preferences, is called manipulable. $\diamond$

This is not the only weakness of the French system as attested by the three following examples.

Example 5 (Monotonicity in the two-stage French system)
Let $\{a, b, c\}$ be the set of candidates for a 17 voters election. A few days before the election, the results of a survey are as follows:

6	voters have preferences	$a \succ b \succ c$,
5	voters have preferences	$c \succ a \succ b$,
4	voters have preferences	$b \succ c \succ a$,
2	voters have preferences	$b \succ a \succ c$.

With the French system, a second stage would be run, between a and b and a would be chosen obtaining 11 out of 17 votes. Suppose that candidate a, in order to increase his lead over b and to lessen the likelihood of a defeat, decides to strengthen his electoral campaign against b. Suppose that the survey did exactly reveal the preferences of the voters and that the campaign has the right effect on the last two voters. Hence we observe the following preferences.

$$\begin{array}{lll} 8 & \text{voters have preferences} & a \succ b \succ c, \\ 5 & \text{voters have preferences} & c \succ a \succ b, \\ 4 & \text{voters have preferences} & b \succ c \succ a. \end{array}$$

After the first stage, b is eliminated, due to the campaign of a. The second stage opposes a to c and c wins, obtaining 9 votes. Candidate a thought that his campaign would be beneficial. He was wrong. Such a method is called non monotonic because an improvement of a candidate's position in some of the voter's preferences can lead to a deterioration of his position after the aggregation. It is clear with such a system that it is not always interesting or efficient to sincerely report one's preferences. You will note in the next example that some manipulations can be very simple. $\diamond$

Example 6 (Participation in the two-stage French system)
Let $\{a, b, c\}$ be the set of candidates for a 11 voters election. Suppose that

$$\begin{array}{lll} 4 & \text{voters have preferences} & a \succ b \succ c, \\ 4 & \text{voters have preferences} & c \succ b \succ a, \\ 3 & \text{voters have preferences} & b \succ c \succ a. \end{array}$$

Using the French system, a second stage should oppose a to c and c should win the election obtaining 7 out of 11 votes. Suppose that 2 of the first 4 voters (with preferences $a \succ b \succ c$) decide not to vote because c, the worst candidate according to them, is going to win anyway. What will happen? There will be only 9 voters.

$$\begin{array}{lll} 2 & \text{voters have preferences} & a \succ b \succ c, \\ 4 & \text{voters have preferences} & c \succ b \succ a, \\ 3 & \text{voters have preferences} & b \succ c \succ a. \end{array}$$

Contrary to all expectations, candidate c will loose while b will win, obtaining 5 out of 9 votes. Our two lazy voters can be proud of their abstention since they prefer b to c. Clearly such a method does not encourage participation. $\diamond$

Example 7 (Separability in the two-stage French system)
Let $\{a, b, c\}$ be the set of candidates for a 26 voters election. The voters are located in two different areas: countryside and town. Suppose that the 13 voters located in the town have the following preferences.

$$\begin{array}{lll} 4 & \text{voters have preferences} & a \succ b \succ c, \\ 3 & \text{voters have preferences} & b \succ a \succ c, \\ 3 & \text{voters have preferences} & c \succ a \succ b, \\ 3 & \text{voters have preferences} & c \succ b \succ a. \end{array}$$

Suppose that the 13 voters located in the countryside have the following preferences.

$$\begin{array}{lll} 4 & \text{voters have preferences} & a \succ b \succ c, \\ 3 & \text{voters have preferences} & c \succ a \succ b, \\ 3 & \text{voters have preferences} & b \succ c \succ a, \\ 3 & \text{voters have preferences} & b \succ a \succ c. \end{array}$$

Suppose now that an election is organised in the town, with 13 voters. Candidates a and c will go to the second stage and a will be chosen, obtaining 7 votes. If an election is organised in the countryside, a will defeat b in the second stage, obtaining 7 votes. Thus a is the winner in both areas. Naturally we expect a to be the winner in a global election. But it is easy to observe that in the global election (26 voters) a is defeated during the first stage. Such a method is called non separable. $\diamond$

The previous examples showed that, when there are more than 2 candidates, it is not an easy task to imagine a system that would behave as expected. Note that, in the presence of 2 candidates, the British system (uninominal and one-stage) is equivalent to all other systems and it suffers none of the above mentioned problems (May, 1952). Thus we might be tempted by a generalisation of the British system (restricted to 2 candidates). If there are two candidates, we use the British system; if there are more than two candidates, we arbitrarily choose two of them and we use the British system to select one. The winner is opposed (using the British system) to a new arbitrarily chosen candidate. And so on until no more candidates remain. This would require $n - 1$ votes between 2 candidates. Unfortunately, this method suffers severe drawbacks.

Example 8 (Influence of the agenda in sequential voting)
Let $\{a, b, c\}$ be the set of candidates for a 3 voters election. Suppose that

1	voter has preferences	$a \succ b \succ c$,
1	voter has preferences	$b \succ c \succ a$,
1	voter has preferences	$c \succ a \succ b$.

The 3 candidates will be considered two by two in the following order or agenda: a and b first, then c. During the first vote, a is opposed to b and a wins with absolute majority (2 votes against 1). Then a is opposed to c and c defeats a with absolute majority. Thus c is elected.

If the agenda is a and c first, it is easy to see that c defeats a and is then opposed to b. Hence, b wins against c and is elected.

If the agenda is b and c first, it is easy to see that, finally, a is elected. Consequently, in this example, any candidate can be elected and the outcome depends completely on the agenda, i.e. on an arbitrary decision. Let us note that sequential voting is very common in different parliaments. The different amendments to a bill are considered one by one in a predefined sequence. The first one is opposed to the status quo, using the British system; the second one is opposed to the winner, and so on. Clearly, such a method lacks neutrality. It doesn't treat all candidates in a symmetric way. Candidates (or amendments) appearing at the end of the agenda are more likely to be elected than those at the beginning. We say that such a method is not neutral. Notice that the British and French systems are neutral because they do not favour any candidate. $\diamond$

Example 9 (Violation of unanimity in sequential voting)
Let $\{a, b, c, d\}$ be the set of candidates for a 3 voters election. Suppose that

1	voter has preferences	$b \succ a \succ d \succ c$,
1	voter has preferences	$c \succ b \succ a \succ d$,
1	voter has preferences	$a \succ d \succ c \succ b$.

Consider the following agenda: a and b first, then c and finally d. Candidate a is defeated by b during the first vote. Candidate c wins the second vote and d is finally elected though all voters unanimously prefer a to d. Let us remark that this cannot happen with the French and British systems. $\diamond$

Example 10 (Tie-breaking president)
Suppose we use the two-stage French system and, at the second stage, the two candidates have the same number of votes. This is very unlikely in a national election but can often occur in small scale elections (board of trustees, court jury, Ph.D. jury, ...). It is then usual to use the president's vote to break the tie. In this case, the opinions of all voters are not treated in the same way.

We then say that the voting system is not anonymous, unlike all systems we have seen so far. Note that using the president's vote is not the only possibility: we could break the tie by choosing, for instance, the oldest of the two candidates (this would not respect neutrality). ◇

Up to now, we have assumed that the voters are able to rank all candidates from best to worse without ties but the only information that we collected was the best candidate. We could try to palliate the many encountered problems by asking voters to explicitly rank the candidates in order of preference (some systems, like approval voting, use another kind of information; see Brams and Fishburn, 1982). This idea, though interesting, will lead us to many other pitfalls that we discuss just below.

2.2 Systems based on rankings

In this kind of election, each voter provides a ranking without ties of the candidates. Hence the task of the aggregation method is to extract from all these rankings the best candidate or a ranking of the candidates reflecting the preferences of the voters as much as possible.

Condorcet (1785) suggests to compare all candidates pairwise in the following way.

Condorcet method (or majority method) Candidate a is preferred to b if and only if the number of voters ranking a before b is larger than the number of voters ranking b before a. In case of tie, candidates a and b are indifferent.

He then states the following principle:

Condorcet principle If a candidate is preferred to each other candidate, using the majority rule, then he should be chosen. The candidate, the Condorcet winner, is necessarily unique.

Note that none of the British or French system respect this principle. Indeed, in example 2, the British system leads to the election of a while b is the Condorcet winner and, in example 3, the French system elects b while a is the Condorcet winner.

The Condorcet principle seems very sensible and close to the intuitive notion of democracy (yet it can be criticized, as suggested in example 1 where candidate a is a Condorcet winner). It is not always operational: in some situations, there is no Condorcet winner; this is the so-called Condorcet paradox. Indeed, in example 8, a is preferred to b, b is preferred to c and

c is preferred to a. No candidate is preferred to all others. In such a case, the Condorcet method fails to elect a candidate. One might think that example 8 is very bizarre and unlikely to happen. Unfortunately it isn't. If you consider an election with 25 voters and 11 candidates, the probability of such a paradox is significantly high as it is approximately 1/2 (Gehrlein, 1983) and the more candidates or voters, the higher the probability of such a paradox. Note that, in order to obtain this result, all rankings are supposed to have the same probability. Such an hypothesis is clearly questionable (Gehrlein, 1983).

We must find how to proceed when there is no Condorcet winner. We may, for example, choose a candidate such that no other candidate defeats him according to the majority rule (weak Condorcet principle) but such a candidate does also not always exist (as in example 8). Many methods have been proposed for exploiting the relation constructed using the majority method. A lot of them can be found in Fishburn (1977), Nurmi (1987) and Laslier (1997).

An alternative approach has been proposed by Borda (1781). He suggests to associate a global score to each candidate. This score is the sum of his ranks in the rankings of the voters.

Borda method Candidate a is preferred to b if the sum of the ranks of a in the rankings of the voters is strictly smaller than the corresponding sum for b (we now assume that the rankings are without tie and we assign rank 1 to the best candidate in the ranking, rank 2 to the second best candidate, and so on; as we will see, the method can be easily generalized for handling ties).

Example 11 (Borda and Condorcet methods)
Let $\{a, b, c, d\}$ be the set of candidates for a 3 voters election. Suppose that

2	voters have preferences	$b \succ a \succ c \succ d$,
1	voters have preferences	$a \succ c \succ d \succ b$.

The Borda score of a is $5 = 2{\times}2{+}1{\times}1$. For b, it is $6 = 2{\times}1{+}1{\times}4$. Candidates c and d receive 8 and 11. Thus a is the winner and the collective ranking is $a \succ b \succ c \succ d$. Using the Condorcet method, the conclusion is different: b is the Condorcet winner. Furthermore, the collective preference obtained by the Condorcet method is transitive and yields the ranking $b \succ a \succ c \succ d$. The two methods diverge; the Borda method does not verify the Condorcet principle. Nevertheless, it can be shown that the Borda method never chooses a Condorcet looser, i.e. a candidate that is beaten by all other candidates by an absolute majority (contrary to the British system, see Example 2). $\diamond$

The Borda method has an important advantage with respect to the Condorcet method. In any situation, it selects one or several winners (those with the lowest sum of ranks). Furthermore, it always yields a ranking of the candidates, from best to worse. The Condorcet method, on the contrary, sometimes yields non-transitive preferences and it is then impossible to rank the candidates or even to choose a subset of 'good' candidates (see example 8). It is easy to verify that the Borda method is neutral, anonymous, separable, monotonic and encourages participation.

The Borda method nevertheless sometimes behaves in a strange way. Indeed, consider example 11 and suppose that candidates c and d decide on the eve of the election not to compete because they are almost sure to lose. With the Borda method, the new winner is b. Thus b now defeats a just because c and d dropped out. Thus the fact that a defeats or is defeated by b depends not only one the relative positions of a and b in the rankings of the voters but is also contingent upon the presence of other candidates and on their position with respect to all other candidates. This can be a problem as the set of the candidates is not always fixed. It is even more a problem in decision aiding because the set of actions is seldom given and is to a large extent the outcome of a modelling process.

After all these examples, we would like to propose a democratic method with the advantages of the Borda method (transitivity of the collective preferences) and those of the Condorcet method (Condorcet principle and absence of contingency problems). We will see in Section 3 that it is hopeless.

Let us mention that we limited this discussion to voting systems aimed at choosing a candidate and not a subset of candidates. The reader might then be tempted to conclude that those systems are inferior to systems aimed at choosing a representative body with some 'proportional' method. But this is too simple, for at least two reasons. First, the definition of what constitues a fair or democratic proportional representation is complex and most proportional systems lead to paradoxical situations (see Balinski and Young, 1982). Second, representative bodies must make decisions and, to this end, they need voting systems aimed at choosing a single action.

3 Some theoretical results

Based on the preceding examples, we now have the intuition that conceiving 'good' preference aggregation methods raises serious problems. This is confirmed by some celebrated results in Social Choice Theory.

3.1 Arrow's theorem

Arrow's theorem is central in Social Choice Theory. It is about voting systems aimed at aggregating n ($n \geq 3$) weak orders (rankings possibly with ties) in a collective weak order. Just as in Section 2.2, each voter ranks all the candidates, possibly with ties.

Formalization
A binary relation R on a set A is a subset of $A \times A$. We often write $a\ R\ b$ instead of $(a, b) \in R$. A weak order on A is a complete (for all $a, b \in A$ we have $a\ R\ b$ and/or $b\ R\ a$) and transitive (for all $a, b, c \in A$, $a\ R\ b$ and $b\ R\ c$ imply $a\ R\ c$) binary relation on A. Let $\mathcal{WO}(A)$ denote the set of all weak orders on the set A. The asymmetric part of R is the binary relation P defined by $a\ P\ b \Leftrightarrow [a\ R\ b$ and $Not[b\ R\ a]]$. The symmetric part of R is the binary relation I defined by $a\ I\ b \Leftrightarrow [a\ R\ b$ and $b\ R\ a]$.

Let $N = \{1, 2, \ldots n\}$ represent the set of voters and A the set of candidates. We assume that voter $i \in N$ expresses his preferences by means of a weak order $R_i \in \mathcal{WO}(A)$ on the set A. We write P_i (resp. I_i) for the asymmetric (resp. symmtric) part of R_i. •

Arrow was interested in the aggregation methods satisfying the following conditions:

Universality every configuration of rankings is admissible.

Formalization
We want to find an aggregation function F yielding a result (a collective weak order) for every element $(R_1, R_2, \ldots, R_n)$ of $\mathcal{WO}(A)^n$. •

This condition excludes any constraint on the set of admissible rankings. The examples of previous section have shown that some problems are caused by some specific rankings or configurations of rankings. A possible way out would then consist in proposing a method that works only with 'simple' configurations. Imposing restrictions on the admissible configurations is sometimes reasonable. For instance, one may sometimes assume that all voters and candidates are located on a right-left axis and that each voter ranks the candidates in order of increasing distance between himself and the candidates. The preferences of the voters are then single-peaked and (Black, 1958) showed that a Condorcet winner then necessarily exists. But, such restrictions imply, for instance, the absence of atypical voters. Yet, this cannot be excluded a priori. With a non-universal aggregation method, some ballots would be impossible to analyze.

Transitivity The outcome of the aggregation method must always be a complete ranking, possibly with ties.

Formalization
The aggregation function takes its values in $\mathcal{WO}(A)$.
When there is no ambiguity, we write $R = F(R_1, R_2, \ldots, R_n)$ and P (resp. I) the asymmetric part (resp. symmetric) of R. •

This condition imposes that the outcome be transitive irrespective of the preference of the voters. So, whenever the society prefers a to b and b to c, it must prefer a to c. We have seen that the Condorcet method does not satisfy this condition. It is sufficient (but not necessary) to ensure that the method will, in all cases, designate one or several best candidates (those with the best positions in the ranking). We will later see that weakening this condition does not much improve the situation formalized by Arrow's Theorem.

Unanimity The outcome of the aggregation method may not contradict the voters when they vote unanimously.

Formalization
The aggregation function F must be such that, for all $a, b \in A$, if $a\ P_i\ b$ for all $i \in N$, then $a\ P\ b$. •

If a is ranked before b in each ranking, then it must be before b in the collective ranking. This condition is very sensible; Example 9 nevertheless shows that some methods violate it.

Independence The relative position of two candidates in the collective ranking only depends on their relative position in the individual rankings.

Formalization
For all $(R_1, R_2, \ldots, R_n), (R'_1, R'_2, \ldots, R'_n) \in \mathcal{WO}(A)^n$ and all $a, b \in A$, if $a\ R_i\ b \Leftrightarrow a\ R'_i\ b$ and $b\ R_i\ a \Leftrightarrow b\ R'_i\ a$, then $a\ R\ b \Leftrightarrow a\ R'\ b$. •

This condition is more complex than the previous ones. When comparing a and b, it forbids

- taking preference intensities into account. The only thing that matters is that a is ranked by the voters before or after b.
- taking other candidates into account.

Let us illustrate this condition with an example.

Example 12 (The Borda method and Independence)
Let $\{a, b, c, d\}$ be the set of candidates. Suppose there are three voters, with the following preferences :

$$\begin{array}{lll} 2 & \text{voters have preferences} & c \succ a \succ b \succ d, \\ 1 & \text{voters has preferences} & a \succ b \succ d \succ c. \end{array}$$

The Borda method yields the ranking : a, c, b, d with the respective scores 5, 6, 8 and 11.

Suppose now that :

$$\begin{array}{lll} 2 & \text{voters have preferences} & c \succ a \succ b \succ d, \\ 1 & \text{voters has preferences} & a \succ c \succ b \succ d. \end{array}$$

The Borda method yields the ranking : c, a, b, d with the respective scores 4, 5, 9 and 12.

Note that, in each individual ranking, the relative position of a and c did not vary across ballots : one voter prefers a to c while two voters prefers c to a. Independence then imposes that the position of a and c in the collective ranking be identical. This is not the case with the Borda method. Indeed, this method uses the fact that the 'distance' between a and c seems larger in the ranking $a \succ b \succ d \succ c$ than in the ranking $a \succ c \succ b \succ d$, because b and d lie between a and c in the first case.

The dependence of the relative position of a and c with respect to b and d is ruled out by the Independence condition. It also excludes any method using, in addition to the rankings, some information regarding preference intensities. ◇

The last condition used by Arrow states that no voter can impose, in all circumstances, his preferences to the society. This condition is extremely sensible for anyone willing to use a 'democratic' method.

Non-dictatorship there is no dictator.

Formalization
For all $i \in N$ and all $a, b \in A$, there is a profile $(R_1, R_2, \ldots, R_n) \in \mathcal{WO}(A)^n$ such that $a \; P_i \; b$ and $b \; R \; a$. •

We are now ready to state the celebrated :

Theorem 1 (Arrow, 1963)
If the number of voters is finite and at least equal to three, no aggregation method can simultaneously satisfy universality, transitivity, unanimity, independence and non-dictatorship.

Proof

The proof of Arrow's Theorem uses the following definitions. A subset $I \subseteq N$ of voters is almost decisive for the pair of candidates $(a, b) \in A^2$ if, for all $(R_1, R_2, \ldots, R_n) \in \mathcal{WO}(A)^n$, $[a\ P_i\ b, \forall i \in I$ and $b\ P_j\ a, \forall j \notin I] \Rightarrow a\ P\ b]$. Similarly, The subset $I \subseteq N$ of voters is decisive for the pair of candidates $(a, b) \in A^2$ if, for all $(R_1, R_2, \ldots, R_n) \in \mathcal{WO}(A)^n$, $[a\ P_i\ b, \forall i \in I] \Rightarrow a\ P\ b$.

We first show that, if I is almost decisive for the pair (a, b), then I is decisive for all pairs of candidates.

Let c be a candidate distinct from a and b (such a candidate always exists because we assumed $n \geq 3$). Let $(R_1, R_2, \ldots, R_n) \in \mathcal{WO}(A)^n$ be a profile such that $a\ P_i\ c$, $\forall i \in I$. Let $(R'_1, R'_2, \ldots, R'_n) \in \mathcal{WO}(A)^n$ be a profile such that

- $a\ P'_i\ b\ P'_i\ c$, $\forall i \in I$,
- $b\ P'_j\ a$ and $b\ P'_j\ c$, $\forall j \notin I$.

Since I is almost decisive for the pair (a, b), we have $a\ P'\ b$. Unanimity imposes $b\ P'\ c$. Transitivity then implies $a\ P'\ c$. Since the relation between a and c for the voters outside I in the profile $(R'_1, R'_2, \ldots, R'_n)$ has not been specified, Independence implies $a\ P\ c$. So, we have proved that whenever I is almost decisive for the pair (a, b), then I is decisive for any pair of candidates (a, c) such that $c \neq a, b$. This reasoning is easily generalized to the case where c is not distinct from a or b.

We now show that there is always a voter $i \in N$ almost decisive for some pair of candidates. As shown above, this voter will be decisive for all pairs of candidates and will therefore be a dictator.

By unanimity, N is almost decisive for all pairs of candidates. Since N is finite, there is at least one subset $J \subseteq N$ almost decisive for the pair (a, b) with a minimal cardinality. Suppose $|J| > 1$ and consider a profile $(R_1, R_2, \ldots, R_n) \in \mathcal{WO}(A)^n$ such that:

- $a\ P_i\ b\ P_i\ c$, for $i \in J$,
- $c\ P_j\ a\ P_j\ b\ \forall j \in J \setminus \{i\}$,
- $b\ P_k\ c\ P_k\ a\ \forall k \notin J$.

Since J is almost decisive for the pair (a, b), we have: $a\ P\ b$. It is impossible that $c\ P\ b$. Indeed, by independence, this would imply $J \setminus \{i\}$ is almost decisive for the pair (c, b) and, hence, decisive for all pairs, contrary to our hypothesis. So, we have $b\ R\ c$ and transitivity implies $a\ P\ c$. This implies that $\{i\}$ is almost decisive for the pair (a, c). $\square$

This negative result applies only when there are more than two candidates. It is easy to verify that the majority method satisfies the five conditions of Arrow's Theorem with two candidates. Arrow's Theorem explains to a large extent the problems we met in Section 2 when we were trying to find a 'satisfying' aggregation procedure. Observe, for instance, that the Borda method verifies universality, transitivity, unanimity and non-dictatorship. Hence, it cannot verify independence, as shown in example 12. The Condorcet method respects universality, unanimity, independence and non-dictatorship. It can therefore not be transitive, as shown in example 8.

Notice that Arrow's Theorem uses only five conditions. In addition to these, we might wish to impose also neutrality, anonymity, monotonicity, non-manipulability, separability or Condorcet's principle. What makes Arroww's Theorem so strong is precisely that it uses only five conditions, all seemingly reasonable, and this is enough to prove an impossibility.

Arrow's Theorem initiated a huge literature, a good overview of which can be found in Campbell and Kelly (2002), Kelly (1978), Fishburn (1987) and Sen (1986). Let us mention that weakening transitivity does not solve the problem revealed by Arrow's Theorem. For instance, if we impose quasi-transitivity (i.e., transitivity of the asymmetric part) instead of transitivity, then we can always determine one or several winners; but it is possible to prove that replacing transitivity by quasi-transitivty in Arrow's Theorem leads to an oligarchy instead of a dictatorship. An oligarchy is a subset of voters that can impose their preferences when they are unanimous and such that each of them can vetoe any strict preference, that is, if a member of the oligarchy strictly prefers a to b, then b cannot be strictly better than a in the collective preference (see Gibbard, 1969 ; Mas-Colell and Sonnenschein, 1972).

Example 13
Let us consider six voters numbered from $i = 1$ to 6 and an aggregation method yielding the relation $R = F(R_1, R_2, \ldots, R_6)$ by means of:

$$\begin{array}{l} x \; P \; y \quad \Leftrightarrow \sum_{\{i: x P_i y\}} w_i > \lambda, \\ x \; I \; y \quad \text{otherwise}, \end{array}$$

with $w_1 = w_2 = 0,4$, $w_3 = w_3 = w_5 = w_6 = 0,05$ and $\lambda = 0,7$. This method is oligarchic. Indeed, consider the set O containing voters 1 and 2. It is easy to verify that, for any profile of preferences,

$$\begin{array}{c} [x \; P_1 \; y \text{ and } x \; P_2 \; y] \Rightarrow x \; P \; y, \\ [x \; P_1 \; y \text{ or } x \; P_2 \; y] \Rightarrow Not[y \; P \; x]. \end{array} \qquad \diamond$$

The existence of an oligarchy is as problematic as the existence of a dictator. Indeed, if the oligarchy contains all voters (this is the only possibility if we want a democratic method), then, because of the veto right of each voter, the collective preference will often contain many indifferences (and/or incomparabilities if the collective preference is not complete). It will not discriminate among candidates. On the contarry, an oligarchy containing only one voter is a dictatorship. Between these two extreme cases, no solution is satisfactory.

We can weaken transitivity even more and impose that there is no circuit in the asymmetric part of the collective preference relation. This condition is necessary and sufficient to guarantee the existence of maximal elements in any finite set of candidates (Sen, 1970). But it is then possible to prove the existence of a voter with an absolute veto (Mas-Colell and Sonnenschein, 1972). So, this does not really help.

Arrow's Theorem and fuzzy preferences. Why is it impossible to aggregate voters' preferences in a satisfactory way (i.e., while respecting Arrow's conditions) ? There are mainly two reasons :

- because the information contained in the weak orders describing the voters' preferences is too poor : it is ordinal. If we use richer structures, we can hope to escape Arrow's Theorem. In particular, if we represent the voters' preferences by means of fuzzy relations, we can not only speak of the preference of a over b but also of the intensity of this preference.
- because the global preference must be a weak order and this is a strong constraint. If we weaken this condition, we may consider aggregation methods yielding relations with more flexibility, like fuzzy relations.

Some authors (for instance, Barrett, Pattanaik et Salles, 1986, 1992 ; Leclerc, 1984 ; Perny, 1992a) have analyzed the consequences of imposing that the outcome of the aggregation be a fuzzy relation, that is a mapping R from A^2 to $[0, 1]$. Their findings are unfortunately largely negative : if we impose that the fuzzy relation has some properties permitting to easily designate a winner or construct a ranking, then we find that the only possible aggregation methods give very different power to the various voters (as in oligarchies or dictatorships). In particular, it is the case if we impose that the collective preference relation verifies min-transitivity, that is for all $a, b, c \in A$:

$$R(a, c) \geq \min(R(a, b), R(b, c)).$$

This condition guarantees that the relation R_λ defined by

$$aR_\lambda b \Leftrightarrow R(a,b) \geq \lambda,$$

is transitive for any value of λ. Hence, starting from a min-transitive relation, it is not difficult to designate a winner or to rank the candidates.

Yet, there are some positive results in the literature, using weaker transitivity conditions (for instance, Ovchinnikov, 1991). It is then tempting to believe that Arrow's Theorem does not hold with fuzzy relations. But these apparently positive results are misleading: the transitivity condition they use is so weak that is not incompatible with Condorcet cycles, as shown in the following example.

Example 14
The transitivity condition used by Ovchinnikov (1991) can be expressed as follows: for all $a, b, c \in A$:

$$R(a,c) \geq R(a,b) + R(b,c) - 1. \tag{1}$$

Suppose we want to aggregate the preferences of n voters. We can define the collective fuzzy preference relation by

$$R(a,b) = \frac{1}{n} \#\{i \in A : a\ R_i\ b\}.$$

It is easy to show that it satisfies Condition (1). Let us now consider $3k$ voters with the following preferences:

$$\begin{array}{lll} k & \text{voters have preferences} & a \succ b \succ c, \\ k & \text{voters have preferences} & b \succ c \succ a, \\ k & \text{voters have preferences} & c \succ a \succ b. \end{array}$$

We obtain : $R(a,b) = 2/3$, $R(b,c) = 2/3$ and $R(c,a) = 2/3$; this is indeed compatible with (1). But notice that this relation is in some sense cyclic and does not permit us to designate a winner or to rank the candidates. So, this does not solve the problem raised by Arrow's Theorem. ◇

In summary, unless we consider a very weak transitivity relation (without any practical interest), aggregation methods yielding fuzzy relations do not escape Arrow's Theorem.

3.2 Some other results

Arrow's Theorem and its many extensions represent only a part of the numerous results in Social Choice Theory. For a comprehensive overview of this field, see Sen (1986) and Campbell and Kelly (2002). In this paper, we will roughly group the results in three categories:

- impossibility results, like Arrow's Theorem, showing that some conditions are incompatible. These results help us better understand why it is difficult to find a 'good' aggregation method;
- characterization results. These present a set of conditions that a given aggregation method and only this one simultaneously respects. Such results help us better understand the essential characteristics of a method. It is then easier to compare it with other methods;
- 'analysis' results. Given a set of desirable conditions, these results compare different methods in order to see which one satisfies most axioms. This can help to find a satisfactory method (within the limits revealed by impossibility results).

This distinction is of course to some extent arbitrary and the three kinds of results are not contradictory. They often use the same conditions.

We will now informally mention some results that we find important or interesting for understanding some phenomena presented in the examples of Section 2.

3.2.1 Impossibility results

Among the impossibility results in Social Choice Theory, two are particularly important:

- Gibbard-Satterthwaite's Theorem (Gibbard, 1973 ; Satterthwaite, 1975). This result shows that there is no aggregation method (for choosing a single candidate) verifying universality, non-dictatorship and non-manipulability when there at least three candidates. The French electoral system is clearly non-dictatorial and satisfies universality. If we neglect the ties than can happen during the second stage, Gibbard-Satterthwaite's Theorem tells us that there is at least one situation where a voter would benefit from voting not sincerely. We have seen such a situation in example 4. Note that this result initiated a huge literature analyzing voting problems in terms of non-cooperative games (see Dummet, 1984 ; Moulin, 1980, 1988 ; Peleg, 1984).

- Sen's 'Theorem of the Paretian liberal' (Sen, 1970). Suppose a society must vote to choose one of several social states. These are defined in such a way that they concern the private sphere of an individual. Clearly, there are conflicts between the majority principle, possibly yielding to a dictatorship of majority (see example 1), and the respect, for this individual, of his private sphere, in which he should decide alone. The Theorem of the Paretian liberal tells us much more than this : it proves that the respect of a private sphere is incompatible with universality and unanimity. This result initiated a large literature, a good overview of which can be found in Sen (1983, 1992).

3.2.2 Characterizations

Among the many characterization results (many such results are presented in (Sen, 1986)), those about the Borda method (see Section 2.2) are particularly interesting. Indeed, this method satisfies most conditions encountered so far and it is very easy to implement.

A characterization of the Borda method In this paragraph, we present a characterization of the Borda method proved by Young (1974). He considers this method as a choice procedure, i.e., a procedure mapping each profile of weak orders on A to a non-empty subset of A. In this context, the Borda method works as follows : for each candidate a, we calculate a score (Borda score) $B(a)$ equal to the sum of the ranks of candidate a in the weak orders of the voters. In case of tie, one uses the mean rank. The choice set then contains the candidate(s) with the largest score(s). Example 11 illustrates how the scores are computed. Notice that, in this example, the Borda method is used to rank and not to choose.

Formalization

A choice procedure is a function $f : \mathcal{WO}(A)^n \rightarrow 2^A \setminus \varnothing$. To each n-uple of weak orders, f associates a non-empty subset of A, interpreted as the set of the best candidates. The Borda method is defined by :

$$f(R_1, R_2, \ldots, R_n) = \{a \in A : B(a) \leq B(b),\ \forall b \in A\},$$

where $B(a)$ is the Borda score of candidate a and is defined by :

$$B(a) = \sum_{i=1}^{n} \left[\#\{b \in A : b\ R_i\ a\} - \#\{b \in A : a\ R_i\ b\}\right]. \qquad (2)$$

This formalization is not exactly the sum of the ranks but the reader will easily check that $B(a)$, defined by (2), is an affine transformation of the sum of the ranks and, hence, using (2) or the sum of the ranks always yields the same result. We will use (2) because it is more convenient than the sum of the ranks. •

In order to characterize the Borda method, Young (1974) uses four conditions.

Neutrality the choice set depends only on the position of the candidates in the preferences of the voters and not, for instance, on the name of the candidates or on their age.

Formalization
Let $\mathcal{P}$ be the set of all permutations on A, π an element of $\mathcal{P}$ and R a binary relation on A. We write $\pi(R)$ for the binary relation such that $\pi(a)\ \pi(R)\ \pi(b) \Leftrightarrow a\ R\ b$. A choice method is neutral if and only if $f(R_1, \ldots, R_n) = \pi(f(\pi(R_1), \ldots, \pi(R_n)))$ for any permutation π in $\mathcal{P}$. •

This condition imposes that all candidates be treated in the same way. It excludes, for instance, methods where, in case of tie, the older candidate wins. Similarly, sequential voting (example 8) is ruled out.

Faithfulness if there is only one voter, then the choice set must contain the best candidates according to this unique voter.

Formalization
$f(R_1) = \{a \in A : a\ R_1\ b,\ \forall b \in A\}$. •

This condition is extremely intuitive. Inedeed, if there is only one voter, why not respect his/her preferences ?

Consistency Suppose, as in example 7, that the voters are divided in two groups. We use the same choice method in both groups. If some candidates belong to both choice sets, then these candidates and only them should belong to the choice set resulting from applying the same choice method to the whole set of voters.

Formalization
$f(R_1, \ldots, R_m) \cap f(R_{m+1}, \ldots, R_n) \neq \varnothing \Rightarrow$
$\quad f(R_1, \ldots, R_n) = f(R_1, \ldots, R_m) \cap f(R_{m+1}, \ldots, R_n)$. •

Consistency is quite sensible. If two groups agree that some candidate, say a, is one of the best, then it is difficult to understand why a would not be a winner when both groups vote together.

Many such conditions, involving two groups of voters, have been used in the literature. They are often called separability. Consistency is one of these conditions.

Cancellation Let us consider two candidates a and b and suppose the number of voters preferring a to b is equal to the number of voters preferring b to a. This is not very particular. Suppose now this is true not only for a and b but for all pairs of candidates, simultaneously. We then face a very particular situation. In such a situation, Cancellation requires that the choice set contains all candidates.

Formalization
$\forall a, b \in A,\ \#\{i \in N : a\ R_i\ b\} = \#\{i \in N : b\ R_i\ a\} \Rightarrow f(R_1, \ldots, R_n) = A.$ •

Among the four conditions used by Young, Cancellation is probably the most questionable one. In some sense, it is reasonable : when, for each pair a, b of candidates, there are as many voters in favour of a as in favour of b, one can indeed prudently consider that no candidate is better than the other ones. But there are other situations where prudence recommends to consider all candidates tied. For instance, when the majority relation is cyclic (see above, Condorcet paradox). Choosing cancellation rather than another condition imposing a complete tie in case of a cyclic majority relation or in another case is quite arbitrary.

The reader will easily verify that the Borda method verifies neutrality, faithfulness, consistency and cancellation. The following theorem, proved by Young, tells us much more.

Theorem 2 (Young, 1974)
One and only one choice method verifies neutrality, faithfulness, consistency and cancellation : the Borda method.

The proof of this theorem being quite long, we do not present it in this chapter. Notice that a similar characterization exists for the borda method used to rank (Nitzan and Rubinstein, 1981). Moreover, different generalizations of this result have been proved for the Borda method used to aggregate many different kinds of binary relations and even fuzzy binary relations (see Debord, 1987 ; Marchant, 1996, 1998, 2000 ; Ould-Ali, 2000).

Generalizations of the Borda method The Borda method is a particular case of a general family of aggregation methods called scoring rules. These rules associate a number (a score) to each position in a binary relation. In order to aggregate n preference relations, one computes, for each candidate, the sum of its scores in the preference relations of the n voters. The winner is the the candidate with the smallest total score. The Borda method is a particular scoring rule where the numbers associated to each rank are equally spaced. The British system is also a scoring rule where the best candidate in a preference relation receives 1 point and all the other ones receive the same score, say 2.

Smith (1973) and Young (1974, 1975) have shown that scoring rules are essentially characterized by neutrality, anonymity and separability (if we then add cancellation, we obtain a characterization of the Borda method). For an overview of many results about scoring rules, see Saari (1994). The French system is not a scoring rule because of the second stage. Yet, it is neutal and anonymous. It is therefore not separable, as shown in example 7. We have noticed in Section 2 that the British system and the Borda method do not satisfy the Condorcet principle (see examples 2 and 10). This is not a surprise: indeed, it is possible to prove that no scoring rule can satisfy the Condorcet principle (see Moulin, 1988).

The French system can be considered as a scoring rule with iteration: at the first stage, it uses the British system for selecting two candidates; the same system is then used at the second stage. Notice that there are many ways to iterate a scoring rule (one could for example use more than two stages). A result by Smith (1973) shows that no iterated scoring rule is monotonic. The violation of monotonicity by the French system (example 5) is just a consequence of this.

A characterization of simple majority In this paragraph, we present the characterization of simple majority of May (1952), for two candidates. In this case, the distinction between choosing and ranking is no longer meaningful but, in order not to use a new formalism, we adopt here the choice formalism. May considers a choice procedure, i.e., a method designating one or several winners, based on th epreferences of the voters. A formal definition of a choice method was presented above, in relation with the Borda method.

A candidate belongs to the choice set with a simple majority if the number of voters supporting him is not smaller than the number of voters supporting his contender.

Formalization

The simple majority choice method is defined by: $a \in f(R_1, \dots, R_n)$ iff

$$\#\{i \in N : a\ R_i\ b\} \geq \#\{i \in N : b\ R_i\ a\}. \qquad \bullet$$

Notice that voters that are indifferent between a and b have no effect on the outcome of the election. Their votes are counted on both sides of the inequality. The outcome would be the same if they would not exist. In order to characterize simple majority, May (1952) used three conditions.

Anonymity The choice set depends only on the preferences of the voters and not, for instance, on their name or age.

Formalization
Let $\mathcal{S}$ be the set of all permutations on $N = \{1, \ldots, n\}$. A choice method is anonymous if and only if $f(R_1, \ldots, R_n) = f(R_{\sigma(1)}, \ldots, R_{\sigma(n)})$ for any permutation σ in $\mathcal{S}$. •

This condition rules out, for example, the methods where some voters weigh more than others and methods where a voter (usually the president of the committee) has the power to decide in case of a tie.

Neutrality see above.

Strict monotonicity given the preferences of the voters, if the candidtes a and b are chosen and if one of the voters changes his preferences in favour of a (the other voters do not change anything), then only a is chosen. If, at the beginning, only a was chosen, then a stays alone in the choice set.

Formalization
Consider two identical weak orders R_i and R'_i except that there is a pair of candidates (a, b) such that:

- $Not[a\ R_i\ b]$ and $a\ R'_i\ b$ or
- $b\ R_i\ a$ and $Not[b\ R'_i\ a]$.

Strict monotonicity then imposes:

$$f(R_1, \ldots, R_i, \ldots, R_n) = \{a\} \Rightarrow f(R_1, \ldots, R'_i, \ldots, R_n) = \{a\},$$

and

$$f(R_1, \ldots, R_i, \ldots, R_n) = \{a, b\} \Rightarrow f(R_1, \ldots, R'_i, \ldots, R_n) = \{a\}. \quad \bullet$$

A consequence of this condition is that, in case of a tie, a single voter changing his mind is enough to break the tie.

Simple majority clearly verifies the three above-mentioned conditions. Moreover, no other method satisfies them all.

Theorem 3 (May, 1952)
When there are exactly two candidates, the only choice method satisfying neutrality, anonymity and strict monotonicity is simple majority.

To understand why this theorem only applies to the case of two candidates, notice that many different choice methods coincide when there are only two candidates. In particular, the Borda method and many scoring methods always yield the same result as simple majority with two candidates. You may then question the interest of this characterization. Actually, Arrow's Theorem has shown us that simple majority cannot be extended to more than two candidates (without deeply modifying it). The characterization with two candidates is therefore essential.

3.2.3 Analysis

The few aggregation methods presented so far are just a small sample of all the methods proposed in the literature. In particular, we did not speak of the methods using the majority relation (constructed by the Condorcet method) to arrive at a choice set or a ranking. Similarly, the few properties (like neutrality or monotonicity) presented so far are also a very small subset of all those studied in the literature. For an overview of methods and properties, see De Donder, Le Breton and Truchon (2000), Felsenthal and Moaz (1992), Fishburn (1977), Levin and Nalebuff (1995), Nurmi (1987) and Richelson (1975, 1978a,b, 1981).

4 Multicriteria decision aiding and social choice theory

4.1 Relevance and limits of social choice results for multicriteria decision aiding

We have seen in section 1 that aggregation problems in multicriteria decision aiding and social choice are formally very close to each other. The examples of section 2 and the results of section 3 taught us that conceiving a satisfactory aggregation method is a challenging task. Some authors (see, for instance, Gargaillo, 1982) have then concluded that muticriteria decision aiding is doomed to failure. For a detailed answer to this objection, see Roy and Bouyssou (1993). We nonetheless mention :

- such a conclusion flows from a biased and too radical interpretation of the available results in social choice theory. There are some impossibility results but this does not means that resorting to an aggregation method for trying to find a collective decision is a futile exercise. It is a demanding task requiring to make compromises between several exigencies that are in general not compatible. These results, when combined with characterization and analysis results, provide a good support to motivate the choice of a method. There is no ideal method but some are perhaps more satisfactory than others. As an example, see Saari (1994) for a convincing plea in favor of the Borda method or Brams and Fishburn (1982) for approval voting;

- the formal proximity between both problems does not imply that both problems are identical. In particular,
  - the goal of a multicriteria decision aiding process is not always to choose one and only one action. There are many other kinds of outcomes, unlike in social choice theory (see Roy, 1985);
  - some conditions look intuitive in social choice theory but are questionable in multicriteria decision aiding, and conversely. Let us mention, for example, that anonymity is not relevant in multicriteria decision aiding as soon as one wishes to take the different importances of the criteria into account. Conversely, the set of potential actions to be evaluated is seldom given in multicriteria decision aiding (contrary to the set of candidates in social choice theory). It can evolve. The conditions telling us how an aggregation method should behave when this set changes (some actions are added or removed) are therefore more important in multicriteria decision aiding than in social choice theory;
  - the preferences to be aggregated in multicriteria decision aiding are the outcome of a long modelling phase along each criterion (see Bouyssou, 1990). This modelling phase can sometimes lead to incomplete preferences, fuzzy preferences, preferences such that indifference is not transitive (see Fodor and Roubens, 1994 ; Perny and Roubens, 1998 ; Perny and Roy, 1992 ; Roubens and Vincke, 1985). In some circumstances, it is possible to finely model preference intensities or even to compare preference differences on different criteria (see Keeney and Raiffa, 1976 ; von Winterfeldt and Edwards, 1986). Let us mention that handling uncertainty, imprecision or indeterminacy is often necessary to arrive at a recommendation in multicriteria decision aiding (Bouyssou, 1989), contrary to social choice theory.
  - in multicriteria decision aiding, contrary to social choice, it is not always necessary to completely construct the global preference. Indeed, it can happen that the decision-maker can express his global preference with respect to some pairs of alternatives. For example, he is able to state that he prefers x to z and y to z but he

hesitates between x and y. If he then uses an aggregation method, it is in order to construct the preference only between x and y and not on the whole set of alternatives. Of course, these preferences that we construct on some pairs of alternatives must be based on the single-criterion preferences of the decision-maker but also on the global preferences that he stated. So, in multicriteria decision aiding, we have a new element at our disposal: the global preferences. These do not exist in social choice theory. They are of course (very) incomplete but they can nevertheless help construct the global preference relation. In practice, these global preferences are often used by analysts in order to set the value of some parameters of the aggregation method they use. For instance, with the methods based on multi-attribute value theory (MAVT), the decision-maker must compare (sometimes fictitious) alternatives in order to determine the shape of the value functions. The existence of these global preferences, totally inexisting in social choice theory, breaks the symmetry between multicriteria decision aiding and social choice theory. Few theoretical results have so far taken the global preferences of the decision-maker into account. More research is needed (see however Marchant, 2003).

Even if both domains are formally close to each other and if some conditions used in social choice theory can also be found in multicriteria decision aiding, we must beware of crude transpositions because of the many specificities of multicriteria aggregation.

Conversely, we must not conclude that both domains are unrelated and that the examples and results of Sections 2 and 3 are of no consequence for multicriteria anlysis. Vansnick (1986a) has clearly shown that it is possible and useful to consider multicriteria aggregation methods in the light of social choice theory. Let us mention that, for example, the difference between the Condorcet and the Borda method can be found in multicriteria anlysis between the ordinal methods (outranking methods, see Roy, 1991 ; Roy and Bouyssou, 1993) and the cardinal ones where the idea of preference difference is central (methods based on multi-attribute value theory, see Keeney and Raiffa, 1976 ; von Winterfeldt and Edwards, 1986). In the light of Arrow's Theorem, it is not surprising that ordinal methods often lead to global preference relations from which a recommendation is not always easy to derive (Vanderpooten, 1990).

Many results of social choice theory still need to be adapted and/or extended to make them relevant to multicriteria analysis. Among the works in this direction, let us mention:

- impossibility results (see Arrow and Raynaud, 1986 ; Bouyssou, 1992a ; Perny, 1992b),
- characterization results (see Bouyssou, 1992b ; Bouyssou and Vansnick, 1986 ; Bouyssou and Perny, 1992 ; Marchant, 1996 ; Pirlot, 1995, 1997) and
- analysis results (see Bouyssou and Vincke, 1997 ; Lansdowne, 1996, 1997 ; Pérez, 1994 ; Pérez and Barba-Romero, 1995 ; Pirlot, 1997 ; Vincke, 1992).

But there is still much to do (see Bouyssou, Perny, Pirlot, Tsoukiàs and Vincke, 1993).

4.2 Some results in close relation with multicriteria analysis

So far, we tried to sketch a global overview of social choice theory and to show the links with multicriteria decision aiding and the limits of this analogy. In this last section, we mention some results of social choice theory that are directly relevant for the analysis of some popular aggregation methods in multicriteria decision aiding.

TACTIC (Vansnick, 1986b) The first relevant result is the characterization of simple majority with two alternatives by May (1952), presented higher. And this aggregation method can be seen as a particular case of TACTIC, with a concordance threshold equal to 1, without weights and without discordance. For the case of two alternatives, a result by Fishburn (1973) characterizes simple majority with weights.

Another article worth mentioning here is Marchant (2003). It presents two characterizations of weighted simple majority with any number of alternatives. It is therefore slightly more general than the results of May and Fishburn. It corresponds to a particular case of TACTIC with a concordance threshold equal to 1 and no discordance.

Multi-attribute value theory (MAVT) (Keeney and Raiffa, 1976 ; von Winterfeldt and Edwards, 1986) The methods of this family are usually analyzed in the framework of measurement theory (Krantz, Luce, Suppes et Tversky, 1971 ; Wakker, 1989). There are nonetheless some relevant results in social choice theory and, in particular, in cardinal social choice theory. In this part of social choice theory, the information to be aggregated is not ordinal (not a binary relation) but cardinal : it consists in utilities, that is, numbers representing preferences. An interesting article in this respect is Roberts (1980).

As far as we know, none of these results have been transposed in multicriteria decision aiding.

Weighted sum The weighted sum is a particular case of MAVT methods. The previous paragraph is therefore relevant for the weighted sum. Yet, let us point out a particular result : Theorem 2 in Roberts (1980) characterizes the weighted sum (see also Blackwell and Girshik, 1954 ; d'Aspremont and Gevers, 1977).

ELECTRE and PROMETHEE (Roy, 1991 ; Roy and Bouyssou, 1993 ; Vincke, 1989) With ELECTRE and PROMETHEE, each alternative is represented by a vector of $\mathbb{R}^n$, $x = (x_1, \ldots, x_n)$ where x_i represents the performance of x on criterion i (we suppose that all criteria are to be maximized). The first step in PROMETHEE consists in choosing, for each criterion, a preference function f_i (Mareschal and Brans, 1988) used to compute, for each pair of alternatives x, y, a number between 0 and 1 representing a preference degree, denoted by $P_i(x, y)$ and defined by $P_i(x, y) = f_i(x_i, y_i)$. So, at the end of the first step, we have a fuzzy preference relation for each criterion, P_i being the fuzzy relation associated to criterion i and $P_i(x, y)$ the value of this relation for the pair x, y. In the next step, these fuzzy relations are aggregated by means of a generalization of the Borda method. This generalization has been characterized by Marchant (1996). Some variants of this characterization are presented in Marchant (1998, 2000) and Ould-Ali (2000).

The ELECTRE methods use a somehow similar construction but with veto effects (see Roy, 1991 ; Roy and Bouyssou, 1993). The preference relation constructed at the end of the aggregation phase uses some functions f_i and g_i with values in $[0; 1]$ in order to define, on the one hand, concordance indices $C_i(x, y) = f_i(x_i, y_i)$ representing to what extent x_i is at least as good as y_i and, on the other hand, discordance indices $D_i(x, y) = g_i(x_i, y_i)$ expressing to what extent the difference $y_i - x_i$ is compatible with a global preference of x over y. When $y_i - x_i$ exceeds a certain threshold (veto threshold), $D_i(x, y)$ equals 1 and the aggregation method then forbids a preference of x over y (for more details on the aggregation of the fuzzy relations C_i and D_i, see Perny and Roy (1992)).

The ELECTRE and PROMETHEE methods thus use aggregation procedures based on the construction and aggregation of fuzzy relations. They therefore do not escape the impossibility results mentioned in section 3.1, about the aggregation of fuzzy relations (for more details, see Perny, 1992b). This is why a last phase (exploitation) is necessary in order to reach a recommendation (see, for instance, Roy and Bouyssou, 1993 ; Vanderpooten, 1990) This last phase is often difficult and the problems it raises can also be analyzed in the light of axiomatic results about ordinal aggregation of preferences. For instance, some non-monotonicity phenomena arising with exploitation procedures based on an iterated choice function(see Fodor, Orlovski, Perny et Roubens, 1998 ; Perny, 1992a) can be explained by Smith's Theorem presented in paragraph 3.2.2 or by more recent axiomatic analyses in the same direction (see Bouyssou, 2004 ; Juret, 2003).

Let us finally mention that Bouyssou (1996) has extended to ELECTRE and PROMETHEE the classic results of McGarvey (1953) regarding simple majority.

References

Arrow, K. J. (1963), Social choice and individual values, 2^{e} édition, Wiley, New York.

Arrow, K. J. and Raynaud, H. (1986), Social choice and multicriterion decision-making, MIT Press, Cambridge.

Balinski, M. L. and Young, H. P. (1982), Fair representation, Yale University Press, New Haven.

Barrett, C. R., Pattanaik, P. K. and Salles, M. (1986), On the structure of fuzzy social welfare functions, Fuzzy Sets and Systems 19, 1–10.

Barrett, C. R., Pattanaik, P. K. and Salles, M. (1992), Rationality and aggregation of preferences in an ordinally fuzzy framework, Fuzzy Sets and Systems 49, 9–13.

Black, D. (1958), The theory of committees and elections, Cambridge University Press.

Blackwell, D. and Girshik, M. A. (1954), Theory of games and statistal decisions, Wiley, New York.

Borda, J.-Ch. de. (1781), Mémoire sur les élections au scrutin, Comptes Rendus de l'Académie des Sciences. Traduction anglaise par Alfred de Grazia "Mathematical derivation of an election system", Isis, Vol. 44, 1953, pp. 42–51.

Bouyssou, D. (1989), Modelling inaccurate determination, uncertainty, imprecision using multiple criteria, in A. G. Lockett and G. Islei (Eds), Improving Decision Making in Organisations, Springer-Verlag, Heidelberg, pp. 78–87.

Bouyssou, D. (1990), Building criteria: A prerequisite for mcda, in C. A. Bana e Costa (Ed.), Readings in multiple criteria decision aid, Springer-Verlag, pp. 58–80.

Bouyssou, D. (1992a), On some properties of outranking relations based on a concordance-discordance principle, in L. Duckstein, A. Goicoechea and S. Zionts (Eds), Multiple criteria decision making, Springer-Verlag, pp. 93–106.

Bouyssou, D. (1992b), Ranking methods based on valued preference relations: A characterization of the net flow method, European Journal of Operational Research 60, 61–67.

Bouyssou, D. (1996), Outranking relations: do they have special properties ?, Journal of Multiple Criteria Decision Analysis 5, 99–111.

Bouyssou, D. (2004), Monotonicity of 'ranking by choosing' procedures: A progress report, Social Choice and Welfare 23(2), 249–273.

Bouyssou, D. and Perny, P. (1992), Ranking methods for valued preference relations: A characterization of a method based on entering and leaving flows, European Journal of Operational Research 61, 186–194.

Bouyssou, D., Perny, P., Pirlot, M., Tsoukiàs, A. and Vincke, Ph. (1993), The manifesto of the new mcda era, Journal of Multi-Criteria Decision Analysis 2, 125–127.

Bouyssou, D. and Vansnick, J.-C. (1986), Noncompensatory and generalized noncompensatory preference structures, Theory and Decision 21, 251–266.

Bouyssou, D. and Vincke, Ph. (1997), Ranking alternatives on the basis of preference relations: a progress report with special emphasis on outranking relations, Journal of Multi-Criteria Decision Analysis 6, 77–85.

Brams, S. J. and Fishburn, P. C. (1982), Approval voting, Birkhäuser, Basel.

Campbell, D. E. and Kelly, J. S. (2002), Impossibility theorems in the Arrovian framework, in K. J. Arrow, A. K. Sen and K. Suzumura (Eds), Handbook of social choice and welfare, Vol. 1, North-Holland, Amsterdam, pp. 35–94.

Condorcet, J. A. N. Caritat, marquis de. (1785), Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix, Imprimerie Royale, Paris.

d'Aspremont, C. and Gevers, L. (1977), Equity and the informational basis of collective choice, Review of Economic Studies 44, 199–209.

De Donder, Ph., Le Breton, M. and Truchon, M. (2000), Choosing from a weighted tournament, Mathematical Social Sciences 40, 85–109.

Debord, B. (1987), Axiomatisation de procédures d'agrégation de préférences, Thèse de doctorat, Université Scientifique, Technique and Médicale de Grenoble.

Dummet, M. (1984), Voting Procedures, Oxford University Press, Oxford.

Felsenthal, D. S. and Moaz, Z. (1992), Normative properties of four single-stage multi-winner electoral procedures, Behavioral Science 37, 109–127.

Fishburn, P. C. (1973), The Theory of Social Choice, Princeton University Press, Princeton, New-Jersey.

Fishburn, P. C. (1977), Condorcet social choice functions, SIAM Journal on Applied Mathematics 33, 469–489.

Fishburn, P. C. (1987), Interprofile Conditions and Impossibility, Harwood Academic Publishers, Chur.

Fodor, J., Orlovski, S., Perny, P. and Roubens, M. (1998), The use of fuzzy preference models in multiple criteria choice, ranking and sorting, in R. Słowiński (Ed.), Fuzzy sets in decision analysis, operations research and statistics, Kluwer, Dordrecht, pp. 69–101.

Fodor, J. and Roubens, M. (1994), Fuzzy preference modelling and multiple criteria decision support, Kluwer, Dordrecht.

Gargaillo, L. (1982), Réponse à l'article "Le plan d'extension du métro en banlieue parisienne, un cas type de l'analyse multicritère", Les Cahiers Scientifiques de la Revue Transports 7, 52–57.

Gehrlein, W. V. (1983), Condorcet's paradox, Theory and Decision 15, 161–197.

Gibbard, A. (1969), Intransitive social indifference and the Arrow dilemma. Mimeographed.

Gibbard, A. (1973), Manipulation of voting schemes: A general result, Econometrica 41, 587–601.

Hudry, O. (2003), Votes et paradoxes : les élections ne sont pas monotones, Mathématiques et Sciences Humaines (163), 9–39.

Juret, X. (2003), Conditions suffisantes de monotonie des procédures de rangement itératives, Mathématiques et Sciences Humaines (161), 59–76.

Keeney, R. L. and Raiffa, H. (1976), Decisions with multiple objectives: Preferences and value tradeoffs, Wiley, New York.

Kelly, J. S. (1978), Arrow impossibility theorems, Academic Press, New York.

Kelly, J. S. (1991), Social choice bibliography, Social Choice and Welfare 8, 97–169. Voir aussi www.maxwell.syr.edu/maxpages/faculty/jskelly/O.htm.

Krantz, D. H., Luce, R. D., Suppes, P. and Tversky, A. (1971), Foundations of measurement, vol. 1: Additive and polynomial representations, Academic Press, New-York.

Lansdowne, Z. L. (1996), Ordinal ranking methods for multicriterion decision making, Naval Research Logistics Quarterly 43, 613–627.

Lansdowne, Z. L. (1997), Outranking methods for multicriterion decision making: Arrow's and Raynaud's conjecture, Social Choice and Welfare 14, 125–128.

Laslier, J.-F. (1997), Tournament solutions and majority voting, Springer-Verlag, Berlin.

Leclerc, B. (1984), Efficient and binary consensus functions on transitively valued relations, Mathematical Social Sciences 8, 45–61.

Levin, J. and Nalebuff, B. (1995), An introduction to vote-counting schemes, Journal of Economic Perspectives 9, 3–26.

Marchant, Th. (1996), Valued relations aggregation with the Borda method, Journal of Multi-Criteria Analysis 5, 127–132.

Marchant, Th. (1998), Cardinality and the Borda score, European Journal of Operational Research 108, 464–472.

Marchant, Th. (2000), Does the Borda rule provide more than a ranking ?, Social Choice and Welfare 17, 381–391.

Marchant, Th. (2003), Towards a theory of mcdm; stepping away from social choice theory, Mathematical Social Sciences 45, 343–363.

Mareschal, B. and Brans, J.-P. (1988), Geometrical representations for MCDA, European Journal of Operational Research 34, 69–77.

Mas-Colell, A. and Sonnenschein, H. F. (1972), General possibility theorems for group decisions, Review of Economic Studies 39, 185–192.

May, K. O. (1952), A set of independent necessary and sufficient conditions for simple majority decisions, Econometrica 20, 680–684.

McGarvey, D. C. (1953), A theorem on the construction of voting paradoxes, Econometrica 21, 608–610.

Moulin, H. (1980), La stratégie du vote, Monographie du séminaire d'économétrie n ° 14, Éditions du CNRS, Paris.

Moulin, H. (1988), Axioms of cooperative decision making, Cambridge University Press, Cambridge.

Nitzan, S. and Rubinstein, A. (1981), A further characterization of Borda ranking method, Public Choice 36(1), 153–158.

Nurmi, H. (1987), Comparing voting systems, D. Reidel, Dordrecht.

Ould-Ali, S. (2000), Variations autour de la méthode de Borda : une approche axiomatisée, Thèse de doctorat, Université Joseph Fourier, Grenoble.

Ovchinnikov, S. (1991), Social choice and Łukasiewicz logic, Fuzzy Sets and Systems 43, 275–289.

Peleg, B. (1984), Game theoretic analysis of voting in committees, Cambridge University Press, Cambridge.

Pérez, J. (1994), Theoretical elements of comparison among ordinal discrete multicriteria methods, Journal of Multi-Criteria Decision Analysis 3, 157–176.

Pérez, J. and Barba-Romero, S. (1995), Three practical criteria of comparison among ordinal preference aggregating rules, European Journal of Operational Research 85, 473–487.

Perny, P. (1992a), Modélisation, agrégation et exploitation de préférences floues dans une problématique de rangement, Thèse de doctorat, Université Paris Dauphine.

Perny, P. (1992b), Sur le non-respect de l'axiome d'indépendance dans les méthodes de type electre, Cahiers du CERO 34, 211–232.

Perny, P. and Roubens, M. (1998), Fuzzy preference modelling, in R. Słowiński (Ed.), Fuzzy sets in decision analysis, operations research and statistics, Kluwer, Dordrecht, pp. 3–30.

Perny, P. and Roy, B. (1992), The use of fuzzy outranking relations in preference modelling, Fuzzy Sets and Systems 49, 33–53.

Pirlot, M. (1995), A characterization of "min" as a procedure for exploiting valued preference relations and related results, Journal of Multi-Criteria Decision Analysis 4, 37–56.

Pirlot, M. (1997), A common framework for describing some outranking procedures, Journal of Multi-Criteria Decision Analysis 6, 86–93.

Pomerol, J.-Ch. and Barba-Romero, S. (1993), Choix multicritère dans l'entreprise, Hermès. Traduction anglaise : Multicriterion Decision in Management, Principles and Practice, Kluwer, Dordrecht, 2000.

Richelson, J. T. (1975), A comparative analysis of social choice functions, Behavioral Science 20, 331–337.

Richelson, J. T. (1978a), A comparative analysis of social choice functions ii, Behavioral Science 23, 38–44.

Richelson, J. T. (1978b), A comparative analysis of social choice functions iii, Behavioral Science 23, 169–176.

Richelson, J. T. (1981), A comparative analysis of social choice functions iv, Behavioral Science 35, 346–353.

Roberts, K. W. S. (1980), Interpersonal comparability and social choice theory, Review of Economic Studies 47, 421–439.

Roubens, M. and Vincke, Ph. (1985), Preference modelling, Springer-Verlag, Berlin.

Roy, B. (1985), Méthodologie multicritère d'aide à la décision, Economica, Paris. Traduction anglaise : Multicriteria methodology for decision aiding, Kluwer, Dordrecht, 1996.

Roy, B. (1991), The outranking approach and the foundations of electre methods, Theory and Decision 31, 49–73.

Roy, B. and Bouyssou, D. (1993), Aide multicritère à la décision : Méthodes et cas, Economica, Paris.

Saari, D. G. (1994), Geometry of voting, Springer-Verlag, Heidelberg.

Satterthwaite, M. A. (1975), Strategyproofness and Arrow's conditions: existence and correspondence theorems for voting procedures and social welfare functions, Journal of Economic Theory 10, 187–217.

Sen, A. K. (1970), The impossibility of a paretian liberal, Journal of Political Economy 72, 152–157.

Sen, A. K. (1983), Liberty and social choice, Journal of Philosophy 80, 5–28.

Sen, A. K. (1986), Social choice theory, in K. J. Arrow and M. D. Intriligator (Eds), Handbook of mathematical economics, Vol. 3, North-Holland, Amsterdam, pp. 1073–1181.

Sen, A. K. (1992), Minimal liberty, Economica 59, 139–159.

Smith, J. H. (1973), Aggregation of preferences with a variable electorate, Econometrica 41, 1027–1041.

Vanderpooten, D. (1990), The construction of prescriptions in outranking methods, in C. A. Bana e Costa (Ed.), Readings in Multiple Criteria Decision Aid, Springer-Verlag, pp. 184–215.

Vansnick, J.-C. (1986a), De Borda et Condorcet à l'agrégation multicritère, Ricerca Operativa (40), 7–44.

Vansnick, J.-C. (1986b), On the problem of weights in multiple criteria decision making: the noncompensatory approach, European Journal of Operational Research 24, 288–294.

Vincke, Ph. (1989), L'aide multicritère à la décision, Éditions de l'Université de Bruxelles - Éditions Ellipses, Bruxelles-Paris. Traduction anglaise : Multi-criteria decision aid, Wiley, New York, 1992.

Vincke, Ph. (1992), Exploitation of a crisp relation in a ranking problem, Theory and Decision 32, 221–240.

von Winterfeldt, D. and Edwards, W. (1986), Decision analysis and behavorial research, Cambridge University Press, Cambridge.

Wakker, P. P. (1989), Additive representations of preferences: A new foundation of decision analysis, Kluwer, Dordrecht.

Young, H. P. (1974), An axiomatization of Borda's rule, Journal of Economic Theory 9, 43–52.

Young, H. P. (1975), Social choice scoring functions, SIAM Journal on Applied Mathematics 28, 824–838.