Mathematics of Social Choice

Mathematics of Social Choice

Voting, Compensation, and Division

Christoph Börgers

Tufts University
Medford, Massachusetts

siam.
Society for Industrial and Applied Mathematics
Philadelphia

10 9 8 7 6 5 4 3 2 1

Cover photo courtesy of the Chicago History Museum.

Library of Congress Cataloging-in-Publication Data

Börgers, Christoph.
Mathematics of social choice : voting, compensation, and division / Christoph Börgers.
p. cm.
Includes index.
ISBN 978-0-898716-95-5
1. Voting–Mathematical models. 2. Elections–Mathematical models. 3. Social choice–Mathematical models. I. Society for Industrial and Applied Mathematics. II. Title.
JF1001.B65 2010
302'.13015195–dc22

2009038044

Contents

Preface

The topics of this book are election systems and fair division of resources. The book is not intended for readers who want a broad and sophisticated introduction to these large fields of mathematical economics, in which I am not by any means an expert. I have in mind a reader without any college-level mathematical background who will enjoy seeing examples of elementary mathematics applied to interesting nonmathematical questions. A few passages in small print are addressed to more advanced readers; see below. With the exception of those paragraphs, none of the material requires knowledge beyond elementary algebra, and most of it does not even require that. I do assume, however, that the reader is willing to make an effort to work through mathematical arguments which, while elementary, are not always easy.

The book grew out of notes for a course for liberal arts majors that I have taught at Tufts University several times. Mathematics has been applied to a vast array of important subjects, so why this particular choice of topics for a course aimed toward this group of students? My answer is that unlike most applications of mathematics, the topics in this book can be studied without extensive background knowledge. Nevertheless the reader will find some ideas and results that are less than fifty years old.

When I teach out of this text, I don't cover all chapters in a semester, but I do cover more than half. As laid out in detail below, many chapters are independent of each other, and Parts I, II, and III are independent of each other altogether. This makes it easy to be flexible, spend as much time as needed on a given chapter, give extra examples where needed, and compensate by omitting other chapters. It is my experience that liberal arts majors, even when they take the course for no reason other than to fulfill a mathematics distribution requirement, are entirely capable of absorbing and even enjoying many of the more theoretical topics and proofs, as long as the instructor goes slowly enough. The material is certainly a challenge for many of them; that is intentional.

In some places in the book, there are passages in small print. These require more knowledge, or more stamina in following mathematical reasoning, than I expect of students taking my course. They can be omitted without any loss of continuity, and they must definitely be omitted when teaching undergraduate liberal arts students with little mathematical background.

A small number of exercises are labeled with asterisks ($*$). These are exercises that I would consider too difficult to assign in my course for liberal arts majors.

I imagine that one could also teach a course for beginning undergraduate mathematics majors out of this text, although I have not yet tried that myself. One could then go faster, cover more chapters, put stronger emphasis on proofs, include some of the material in small print, and not shy away from assigning the exercises labeled with asterisks.

The chapters are grouped as follows.

Chapters 1–8: Elections for the purpose of selecting a single winner from a field of candidates. A fairly wide range of winner selection methods is introduced. A list of intuitively reasonable requirements narrows down the options, until we arrive at a single preferred method (preferred among those discussed in this book, not among all conceivable methods), Schulze's beatpath method. These chapters are best worked through in the order in which they are presented, without omissions.

Chapters 9 and 10: These chapters are among the hardest in the book, and can be omitted without losing continuity. They give examples of criteria that may seem attractive at first sight, but cannot be satisfied by any reasonable winner selection method. Theorem 10.3 is an easy result in this direction, and its statement and proof are easy to follow without reading anything else in Chapters 9 and 10, except for Definition 10.1.

Chapters 11 and 12: Elections for the purpose of ranking a field of candidates. Chapter 11 is a theoretical discussion of the relation between ranking and winner selection. Chapter 12 culminates in a proof of Arrow's famous dictatorship theorem.

Chapters 13–15 form the second part of the book. They are about "fair compensation," i.e., problems of the following kind. Two or more people jointly own some indivisible object, such as a house, a valuable painting, or a rabbit. They would like to transfer ownership to one person, who will have to compensate the others monetarily. One might think of a divorce or an inheritance. Who should receive ownership of the object, and what should the compensation payments be? This is obviously not a mathematical question, but one can formulate general principles that one might want to adhere to, and see, using mathematics, what follows from them. Chapter 13–15 preview, in this simple context, the notions of fairness, envy-freeness, Pareto-optimality, and equitability, which are fundamental to the following chapters on fair division. However, the same notions are introduced again, in the context of cake cutting, in Chapter 16, and therefore the third part of the book (Chapters 16–26) does not depend on the second.

Chapter 16 is an introduction to the "cake cutting" problem, that is, the problem of sharing a divisible resource among people with different tastes, preferences, or needs. The chapter introduces the fundamental notions of fairness, envy-freeness, Pareto-optimality, and equitability for cake divisions.

Chapters 17–20 describe generalizations of the "I cut, you choose" method of cake division. In the "I cut, you choose" method known to children in kindergarten, one child cuts the cake into two halves that she considers of equal value, and the other child chooses her half. The generalizations discussed in Chapters 17–20 have in common with "I cut, you choose" that the procedure starts with one person dividing the cake into pieces that she considers equal in value. A method of this sort for three people was proposed by Hugo Steinhaus in the 1940s (Chapter 17). Steinhaus's idea was extended to the case of arbitrarily many participants by Harold Kuhn in 1967 (Chapter 19). Chapter 18, on Hall's marriage theorem, is needed for Chapter 19. Chapter 20, which is independent of all other chapters in the book except for Chapter 16, describes a method for *envy-free* cake division among three people, invented by John Selfridge and John Conway.

In the rest of the book, Chapters 21–26, we assume that the cake consists of finitely many homogeneous pieces—some cheesecake and some chocolate cake, for instance. The important point about homogeneous pieces is that they are divisible *objectively*: We can disagree over whether chocolate mousse cake is more valuable than cheesecake or vice versa, but we cannot reasonably disagree over how much is one-quarter of eight ounces of chocolate mousse.

In Chapters 21–24, we describe and analyze the “adjusted winner method” of Steven Brams and Alan Taylor, quite arguably *the* right way for two people to divide a cake composed of finitely many homogeneous pieces. This method combines the three desirable properties of envy-freeness, Pareto-optimality, and equitability, and it is the only method with these three properties.

In Chapter 25, we study proportional allocation, an alternative way for two people to share a cake composed of finitely many homogeneous pieces. It is less easily manipulated by a dishonest person than the adjusted winner method, but it is not Pareto-optimal. It is best to read Chapters 21–24 prior to Chapter 25.

In Chapter 26, we study (necessarily imperfect) generalizations of the idea of the adjusted winner method to the case of more than two participants. It turns out that the three desirable properties of the adjusted winner method (envy-freeness, Pareto-optimality, and equitability) cannot all be guaranteed at the same time when there are more than two participants, but any two of the three can be. Chapter 26 answers questions that arise naturally from Chapters 21–23, but it is also among the most difficult chapters in the book, and can certainly be omitted.

There are appendices on set notation, logic, and mathematical induction. It is probably best to consult those when needed while reading the text, instead of studying them in advance. I also provide solutions to many of the homework exercises.

I am grateful to my wonderful colleague Martin Guterman, who created the course on Mathematics of Social Choice at Tufts. This book reflects my own way of teaching the course. While I don’t know whether Marty, who passed away in 2004, would have liked it, I do know that he believed in teaching rigorous mathematics to students who major in fields far from mathematics, and I adopted this approach here. I am sure Marty would have been delighted to know that the course continues to draw large numbers of Tufts students every year.

I also owe gratitude to many students who have commented on my course notes and thereby helped improve them. I would like to mention one student by name, Sally Greenwald, who took the trouble of reading through the entire text one more time after the semester had ended (and she had already graduated) and gave me a long list of helpful comments.

Finally I would like to thank the excellent staff at SIAM’s book program, especially Elizabeth Greenspan (senior acquisitions editor) and Lou Primus (production editor). They made the process of preparing the book for publication smooth and efficient. It was a great pleasure to work with them.

Christoph Börgers

Chapter 1

Winner Selection

Elections can serve various purposes. One may want to select a single winner from a field of candidates, select a committee, or rank the entire field of candidates. We begin by thinking about elections aimed at the selection of a single winner from a field of candidates. There are many different ways of organizing such elections; we refer to them as *winner selection methods*. In this chapter, we describe several winner selection methods and demonstrate, using examples, that they can easily produce different outcomes.

In the literature on this subject, one often sees the expression *single-winner method* in place of *winner selection method*. I prefer the term *winner selection method*, leaving open the possibility of a tie that results in the selection of several winners. (Think, for instance, of a three-way race in which two candidates receive 28 votes, and the third receives 12 votes.) I will use the expression *single-winner method* only when I really mean a method that never produces any ties.

A friend whom I told about this book was puzzled, and asked, "What is wrong with just the normal way of voting?" The trouble, of course, is that there is no "normal" way. Different societies organize elections in different ways. I will begin with an example illustrating that different perfectly reasonable winner selection methods can in fact produce different election outcomes.

Example 1.1. In a small village, a mayor is to be elected. There are 18 voters, and 4 candidates for mayor, whom I will call A, B, C, and D, unimaginatively but simply. The votes are cast and counted. The outcome is as follows:

$$A : 6 \text{ votes}$$
$$B : 5 \text{ votes}$$
$$C : 4 \text{ votes}$$
$$D : 3 \text{ votes}$$

Candidate A has a *plurality* of the vote: No candidate has more votes than A. He is therefore declared the winner. Making the candidate(s) with a plurality of the votes the winner(s) is called the *plurality method*. It is the most common winner selection method in political elections in the United States. (Perhaps that is what my friend meant by "the normal way"?)

Note that A does not have a *majority*: 6 is not larger than 18/2. The distinction between *plurality* and *majority* is important.

B challenges the result. She thinks that most of those who voted for C or D would prefer her over A. In an election featuring only A and B as candidates, she argues, she would have won. The village elders therefore decide to conduct a *runoff* election, with only the two leading candidates, A and B, on the ballot. In the runoff, B beats A, 12 to 6, so she is now declared the winner. This winner selection method, involving a runoff election in which only the two leading candidates appear on the ballot, is called the *runoff method*. Runoff voting is used in the presidential elections of many countries, including France, Russia, and Brazil.

The presence of the losing candidates C and D spoiled the election for B. "Spoiler effects" can occur in real life. In the 2000 presidential election in the United States, Al Gore might well have won Florida, and thereby the presidency, had Ralph Nader not been in the race. In the 1992 presidential election, George H. W. Bush might well have beaten Bill Clinton and won a second term had Ross Perot not been in the race.

In our example, C now challenges the result. He feels that he should have been included in the runoff. Wearily, the village elders grant his wish and conduct yet another election, featuring A, B, and C on the ballots. The outcome is as follows:

$$A : 6 \text{ votes}$$
$$B : 5 \text{ votes}$$
$$C : 7 \text{ votes}$$

Since B is now the weakest of the three candidates, it is decided that in fact B should be dropped from the ballot, and another runoff election featuring only A and C should be conducted. In this runoff election, C beats A, 12 to 6. So C is declared the winner. We will call this winner selection method, in which only the candidate with the lowest number of votes is dropped in each round, the *elimination method*. It is often called the *instant runoff method*. The International Olympic Committee uses this method to select the city in which the games are held. ∎

How is it possible that three seemingly reasonable winner selection methods lead to three different outcomes in Example 1.1? To understand, it helps to know how each voter *ranks* the candidates. We will assume that each voter has a *strict* ranking, without any ties. A ballot asking voters to rank the candidates is called a *preference ballot*.

Preference ballots that forbid ties require extremely knowledgeable voters. For instance, in the 2008 presidential election in the United States, the leading candidates were Barack Obama and John McCain. However, the field included others, for instance—in a few of the states—a man named Thomas Stevens (the candidate of the Objectivist Party), and another named Charles Jay (the candidate of a political party called the Boston Tea Party). Both appeared on the ballot in Colorado and in Florida. It would have been unreasonable to require all voters in Colorado and Florida to make up their minds whether they preferred Stevens over Jay or vice versa. Preference ballots become much more practical when one allows voters to rank a few candidates at the top, and rank all others equally and at the bottom. Nevertheless, we will assume here that ties in individual voters' preferences are not allowed. This avoids technical issues that may not always be particularly difficult to resolve, but are cumbersome.

Let us assume that in Example 1.1, the voters' preferences are as follows:

1	2	3	4	5	6	7	8	9	10
A	D	B	A	A	A	B	A	B	B
B	C	C	B	B	B	C	B	C	C
C	B	A	C	C	C	A	C	A	A
D	A	D	D	D	D	D	D	D	D

11	12	13	14	15	16	17	18
B	C	C	C	D	C	D	A
C	B	B	B	C	B	C	B
A	A	A	A	B	A	B	C
D	D	D	D	A	D	A	D

. (1.1)

For instance, voter 1 ranks the candidates A, B, C, D (A is her favorite candidate, and D is her least favorite candidate), while voter 17 ranks them in the reverse order D, C, B, A. We call the above table the *detailed preference schedule*. We can summarize it more compactly like this:

6	5	4	3
A	B	C	D
B	C	B	C
C	A	A	B
D	D	D	A

. (1.2)

There are 6 voters who rank the candidates A, B, C, D, 5 who rank them B, C, A, D, and so on. We call this the *reduced preference schedule*, or simply the *preference schedule*. Notice that the numbers in the first row do not have the same meaning in (1.2) as in (1.1). In (1.1), they simply label voters, whereas in (1.2), they indicate the total numbers of voters ranking the candidates in certain ways.

In a democracy, the outcome of an election usually depends only on the reduced preference schedule, not on the detailed preference schedule. We give this principle a name.

Definition 1.2 (principle of one person, one vote). *A winner selection method satisfies the principle of one person, one vote if the outcome of the election never depends on anything other than the reduced preference schedule.*

According to this principle, it should not matter *who* ranks the candidates C, B, A, D, for instance; all that should matter is *how many* rank them that way. There are situations in which it might in fact be reasonable to use a method that violates this principle. For instance, suppose that each voter belonged to one of two ethnic groups. The constitution of such a society might then impose the rule that the winner must not be a candidate who is overwhelmingly unpopular among one of the two ethnic groups. With that rule, one might have to know more than the reduced preference schedule to determine the winner.

It is easy to verify that the preference schedule (1.2) leads to the outcomes described earlier. For instance, if each voter is asked to name his or her favorite candidate, then 6 voters name A, 5 name B, 4 name C, and 3 name D. So A wins by the plurality method. On the other hand, if C and D are removed from the ballot, the voters' preferences among the remaining two candidates A and B become

6	5	4	3
A	B	B	B
B	A	A	A

,

or briefly

6	12
A	B
B	A

.

(Here we have assumed that voters' relative rankings of A and B are not affected by the presence of C and D on the ballots.) So B beats A, 12 to 6, in a runoff election featuring the two. You can check easily that C wins by the elimination method. I don't think there is any reasonable winner selection method that would make D the winner in Example 1.1.

Many other winner selection methods have been proposed. A simple method explicitly based on preference schedules is *Borda count*, named after Jean-Charles de Borda, who invented it in 1770. (It was in fact also invented about 500 years earlier by the Mallorcan philosopher Ramon Llull.) We will describe this method now.

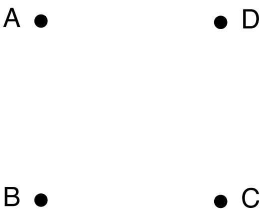
From here on, and throughout Part I of this book, we make the convention that the letter n denotes the number of candidates, and N the number of voters. In Borda count, a candidate is assigned *n Borda points* for each voter who ranks her first, $n-1$ Borda points for each voter who ranks her second, etc. (A candidate gets one Borda point even for a candidate who ranks her last.) The total number of Borda points that a candidate receives will be called the candidate's *Borda score*. The candidate with the greatest Borda score wins. In our example, with the preference schedule (1.2), the Borda scores are as follows.

$$\begin{aligned}
A &: 6\times 4+5\times 2+4\times 2+3\times 1=45\\
B &: 6\times 3+5\times 4+4\times 3+3\times 2=56\\
C &: 6\times 2+5\times 3+4\times 4+3\times 3=52\\
D &: 6\times 1+5\times 1+4\times 1+3\times 4=27
\end{aligned} \tag{1.3}$$

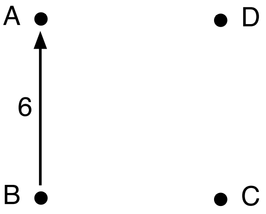
So B wins by Borda count.

Borda count is used in politics in only a very few places, including the election of ethnic minority members of the Slovenian National Assembly. In the United States, a variation of Borda count is used to determine the Most Valuable Player in Major League Baseball.

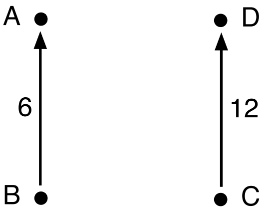
Other winner selection methods are based on the *pairwise comparison graph*. We will use our example to explain the definition of the pairwise comparison graph. Each of the four candidates is represented by a dot:



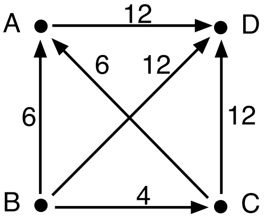
Then we ask, for each *pair* of candidates, who would win if these two were the only candidates on the ballot. In other words, we ask who of the two would win in *head-to-head competition*. For example, if only A and B were on the ballot, B would win by a margin of 6, namely, 12:6. We indicate this by drawing an arrow from B to A, and writing the number 6 (the margin of victory) next to it:



For another example, if only C and D were on the ballot, C would win by a margin of 12, namely, 15:3. We indicate this by drawing an arrow from C to D, and writing the number 12 next to it:



In this way, each of the six possible head-to-head competitions gives rise to an arrow with a number next to it:



There is a potential subtlety that might arise, though it does not in our example: If two candidates were to tie in head-to-head competition, we would not connect the corresponding two dots with any arrow. Ties are possible only if the number N of voters is even. In this case, all margins of victory (the numbers appearing in the pairwise comparison graph) are even, whereas if N is odd, all margins of victory are odd; see Exercise 1.5.

We say that the pairwise comparison graph represents the voters' *collective preferences*, or the *societal preferences*. Note that the societal preferences are not necessarily linear; that is, it is not necessarily clear from the pairwise comparison graph which candidate is liked best by society, which is liked second best, etc. In fact, it is possible for a majority to prefer A to B, a (different) majority to prefer B to C, and a (yet different) majority to prefer C to A. Thus societal preferences can be *circular*. We will return to this important point in Chapter 2.

The simplest winner selection method based on the pairwise comparison graph is called the *method of pairwise comparison*. It is also known as the *Copeland method*, in honor of the mathematician Arthur Herbert Copeland, who proposed and analyzed it in a seminar lecture at the University of Michigan in 1951. In this method, the numbers next to the arrows are ignored. Each candidate gets one *pairwise comparison point* for each other candidate whom he would beat in head-to-head competition. In our example, the points are distributed like this:

$$A:1 \qquad B:3 \qquad C:2 \qquad D:0 \tag{1.4}$$

We call this the outcome of the *pairwise comparison tournament*. If there were a tie between two candidates, each of them would derive half a point from that. In our example, B wins by the method of pairwise comparison.

Notice that B would beat each of the other three candidates in head-to-head competition. Therefore she wins the largest possible number of points in the pairwise comparison tournament, namely, 3. In general, when there are n candidates, the largest possible number of pairwise comparison points is $n-1$; to achieve this number, a candidate must beat every other candidate in head-to-head competition.

Definition 1.3 (Condorcet candidate). *If a candidate X beats every other candidate Y in head-to-head competition (that is, with all but X and Y removed from the ballots), then X is called a Condorcet candidate.*

This notion is named after Marie Jean Antoine Nicolas de Caritat, Marquis de Condorcet, an 18th century philosopher, mathematician, and political scientist.[1] (Just like the Borda count, the idea of a Condorcet candidate was actually proposed in the 13th century by Ramon Llull.)

There can be at most one Condorcet candidate (see Exercise 1.6). If X is a Condorcet candidate, then it seems reasonable to say that X should be the sole winner of the election, since for any other candidate Y, a majority prefer X to Y. Thus in our example, there is a solid reason why B *should* really become the mayor—she is a Condorcet candidate.

We note that a *majority candidate*, that is, a candidate who gets more than half the (first-place) votes, is always a Condorcet candidate (see Exercise 1.7). However, Example 1.1

[1] Condorcet and Thomas Jefferson were close friends during Jefferson's years as the United States ambassador to France. However, in a 1992 article titled "Did Jefferson or Madison understand Condorcet's Theory of Social Choice?" in the journal *Public Choice*, Ian McLean and Arnold Urken wrote, "We have examined all the known primary documents which appear to bear on the question, none of which shows any hint that either Jefferson or Madison understood the social choice components of Condorcet's work."

shows that a Condorcet candidate need not be a majority candidate. The following example illustrates this point more dramatically.

Example 1.4. A family of five decide to purchase a pet. The possibilities that they would like to consider are a dog, a cat, a tortoise, and a goldfish. They rank the options as follows:

Mom	Dad	Peter	Paul	Mary
goldfish	goldfish	dog	dog	tortoise
tortoise	cat	cat	cat	cat
cat	tortoise	tortoise	goldfish	dog
dog	dog	goldfish	tortoise	goldfish

In head-to-head competition, the cat beats each of the other possibilities. Thus the cat is the Condorcet candidate here. On the other hand, the cat does not even have a single first-place vote. ∎

In all of the methods discussed in this chapter, ties are possible. For instance, if the preference schedule in the mayoral election had been

6	6	4	2
A	B	C	D
B	C	B	C
C	A	A	B
D	D	D	A

(1.5)

then A and B would have tied if the plurality method were used. In general, we use the letter W to denote the set of winners. If (1.5) is the preference schedule and the plurality method is used, then

$$W = \{A, B\}.$$

It is less clear how ties should be dealt with in the runoff and elimination methods. What if there are three leading candidates with the same number of first-place votes? Should they all get into the runoff round? We will not spell out detailed (and by necessity arbitrary) rules answering these questions.

Exercises

1.1. In an election involving five voters and three candidates, the preference schedule is

2	2	1
A	B	C
B	C	A
C	A	B

Determine the set W of winners using (a) the plurality method, (b) the runoff method, (c) the elimination method, (d) Borda count, and (e) the method of pairwise comparison. Is there a Condorcet candidate?

1.2. In an election involving seven voters and four candidates, the preference schedule is

1	2	3	1
A	B	D	A
B	C	A	C
C	A	C	B
D	D	B	D

Prove: (a) A wins against every other candidate in head-to-head competition. That is, A is the Condorcet candidate. (b) D loses against every other candidate in head-to-head comparison. We are tempted to call D the *anti-Condorcet candidate*. (c) D is the winner when the plurality method is used.

1.3. The sum of all Borda scores in the mayoral election is $45+56+52+27=180$; see (1.3). Explain why the sum of all Borda scores is 180 in any election involving eighteen voters and four candidates, no matter what the preference schedule is. What would be the sum of all Borda scores in an election involving twenty voters and three candidates?

1.4. The total number of points in the pairwise comparison tournament in the mayoral election is $1+3+2+0=6$; see (1.4). Explain why the total number of pairwise comparison points is six in any election involving four candidates, no matter what the pairwise comparison graph is—even when there are ties in head-to-head competitions. What would be the total number of pairwise comparison points in an election involving five candidates?

1.5. Explain: If the number N of voters is even, all margins of victory (the numbers appearing in the pairwise comparison graph) are even, whereas if N is odd, all margins of victory are odd.

1.6. Explain why there could never be two different Condorcet candidates in a given election.

1.7. Explain why a majority candidate is a Condorcet candidate.

1.8. Consider an election involving three candidates A, B, and C and four voters. (Think of a family of four choosing a restaurant.) Suppose that A is the first choice of three out of the four voters. Prove that A wins if Borda count is used.

1.9. Consider an election involving three candidates A, B, and C and five voters. (Unannounced, as always, Uncle Charles has dropped by for a visit and joins the dinner crowd of Exercise 1.8.) Give an example of a preference schedule in which A is the first choice of three out of the five voters, but A is not the winner when Borda count is used.

Chapter 2

Rule of the Majority

Anybody would agree that in a democracy "the majority should rule," at least when it comes to selecting, say, the head of government. Unfortunately, however, it is not at all clear what exactly that should mean. In this chapter, we will try to make the "rule of the majority" requirement precise. Here is the simplest idea.

Definition 2.1 (majority criterion). *A winner selection method satisfies the majority criterion if it makes any majority candidate—if there is one—the sole winner of the election.*

(Remember that a *majority candidate* is one who is placed first by more than half the voters.) We will call a method *majority-fair* if it satisfies this criterion. It is easy to see that the plurality, runoff, and elimination methods are majority-fair. Pairwise comparison, too, is majority-fair, though it may take a moment's thought to convince yourself: If X is a majority candidate, she wins *every* head-to-head competition, so she gets $n-1$ points in the pairwise comparison tournament. Everybody else loses at least one head-to-head competition—namely, the one against X—and can therefore get at most $n-2$ points. So X is indeed the sole winner of the pairwise comparison tournament.

On the other hand, Borda count violates the majority criterion. This is demonstrated by the preference schedule

3	2
A	B
B	C
C	A

.

Here A is the majority candidate (he has three out of five first-place votes, and $3 > 5/2$), but A gets only 11 Borda points, while B gets 12. (This is the solution to Exercise 1.9.)

In summary, four of the five winner selection methods considered in Chapter 1 are majority-fair. Thus the majority criterion is not very stringent. The following criterion is a far more stringent way of giving precise meaning to the principle of the "rule of the majority."

Definition 2.2 (Condorcet criterion). *A winner selection method satisfies the Condorcet criterion if it makes any Condorcet candidate—if there is one—the sole winner.*

Intuitively, it seems sensible to require the Condorcet criterion: If X is a Condorcet candidate, but a different candidate Y wins the election, then a majority of voters could be pleased more by making X the winner. A winner selection method that satisfies the Condorcet criterion is called *Condorcet-fair*.

Proposition 2.3. *Any Condorcet-fair winner selection method is majority-fair. However, the plurality, runoff, and elimination methods, all of which are majority-fair, are not Condorcet-fair.*

Proof. Suppose that a winner selection method is Condorcet-fair, and that X is a majority candidate. Then X is also a Condorcet candidate (see Exercise 1.7). Since the method is Condorcet-fair, it must then make X the sole winner of the election. Thus the method indeed satisfies the majority criterion.

To see that the plurality, runoff, and elimination methods are not Condorcet-fair, consider the following preference schedule:

5	6	7
A	B	C
B	A	A
C	C	B

. (2.1)

A is the Condorcet candidate here, but the plurality method makes C the winner, and the runoff method makes B the winner. For three candidates, the elimination method is the same as the runoff method. Therefore the plurality, runoff, and elimination methods are Condorcet-unfair. □

Example 2.4. In the presidential elections in the United States in 2000, the official end result in Florida was as follows:

George W. Bush	48.847%
Al Gore	48.838%
Ralph Nader	1.634%
Patrick J. Buchanan	0.293%
Harry Browne	0.275%
other	0.112%

.

It seems fairly likely that Al Gore would have beaten George W. Bush in Florida if the two had been the only candidates on the ballot. (This is based on the assumption that a large majority of Nader voters would have voted for Al Gore if he and George W. Bush had been the only choices. This is of course not certain. Many might, for instance, not have voted at all.) Assuming that this is correct, we have here a real-life example of a Condorcet candidate who lost the election, illustrating that the plurality method is not Condorcet-fair. ■

Example 2.5. Presidential elections were held in France on April 21–22, 2007, with a runoff election on May 5–6, 2007. The following table shows results of the first round, with percentages rounded to the nearest integer.

Nicolas Sarcozy	31%
Ségolène Royal	26%
François Bayrou	19%
Jean-Marie LePen	10%

In the second, runoff round, Sarcozy beat Royal, 53% to 47%, and thereby became president of France. A poll of March 15, 2007, had predicted the result of the (at that time, hypothetical) Sarcozy/Royal runoff with remarkable accuracy: The prediction was 54% to 46%. The same poll predicted that in a Sarcozy/Bayrou runoff, Bayrou would beat Sarcozy, also 54% to 46%. Further it predicted that in a Bayrou/Royal runoff, Bayrou would win 60% to 40%. It appears therefore that Bayrou was probably a Condorcet candidate who was eliminated in the first round. This illustrates that the runoff method is Condorcet-unfair. Pyotr Romanov, a commentator for the Russian news agency Novosti, put it this way:

> The less obvious winner of the first stage is centrist François Bayrou, who gained 18.5%. According to exit polls, Bayrou, had he made it to the second round, would have won the election hands down. He is a compromise politician who suits the majority of the French.... The French race has shown that the election system in the "democratic West" is not as democratic as it may seem at first glance.

Incidentally, Russia also uses the runoff method in presidential elections. ∎

Borda count violates even the majority criterion, so it certainly violates the Condorcet criterion. Pairwise comparison, on the other hand, satisfies the Condorcet criterion. To see this, suppose that X is a Condorcet candidate. Then X wins $n-1$ points in the pairwise comparison tournament, while all other candidates win at most $n-2$. Thus X is the sole winner of the pairwise comparison tournament. In summary, of the five methods introduced in Chapter 1, only pairwise comparison is Condorcet-fair.

It is easy to construct examples in which there is no Condorcet candidate. For instance, think about the preference schedule

2	5	4
A	B	C
B	C	A
C	A	B

(2.2)

Here a majority of voters prefer A to B, a majority prefer B to C, and a majority prefer C to A. The voters' collective preferences are circular, and there is no Condorcet candidate. Notice that none of the individual voters have circular preferences. Society can, in the sense illustrated by this example, have circular preferences even if none of the individuals does. The simplest example in which there is a circle in societal preferences is

1	1	1
A	B	C
B	C	A
C	A	B

(2.3)

Again, A beats B, B beats C, and C beats A in head-to-head competition.

It is not hard to see that there must be a circle in the voters' collective preferences whenever there is no Condorcet candidate, provided that there are no ties in head-to-head competitions (see Exercise 2.4). The Condorcet criterion says nothing about what the outcome of the election should be when there is no Condorcet candidate.

If the will of the *majority* is to be respected, then certainly the *unanimous* will of the voters is to be respected. Any reasonable winner selection method should satisfy the following criterion.

Definition 2.6 (unanimity criterion). *A winner selection method satisfies the unanimity criterion if it guarantees that candidate Y is not among the winners of the election if there is a candidate X who is preferred to Y by every voter.*

Example 2.7. A family of five plan their annual family vacation. They have already narrowed down the options to the following four possibilities: camping in Maine, hiking in the Adirondack Mountains, a week in Yellowstone National Park, or a week in San Francisco. They each rank the options:

Mom	Dad	Peter	Paul	Mary
Maine	S. F.	Yellowstone	Yellowstone	S. F.
Adirondacks	Maine	Maine	Maine	Yellowstone
S. F.	Adirondacks	Adirondacks	Adirondacks	Maine
Yellowstone	Yellowstone	S. F.	S. F.	Adirondacks

.

It is not clear what they should do, based on these preferences. What *is* clear, however, is that they should not go hiking in the Adirondacks, since *each* of them would rather go camping in Maine. A winner selection method that satisfies the unanimity criterion would not include the Adirondacks in the set of winners in this example. ∎

The five methods introduced in Chapter 1 satisfy the unanimity criterion. For example:

Proposition 2.8. *Pairwise comparison satisfies the unanimity criterion.*

Proof. Suppose that every voter ranks X above Y. If Y beats a candidate Z in head-to-head competition, then so does X. If Y ties with Z, then X either beats Z or ties with Z (see Exercise 2.5). So for every point or half-point that Y gets in the pairwise comparison tournament, X gets at least the same. However, X gets at least one extra point, one that Y does not get: X beats Y. So X gets more points than Y in the tournament, and therefore Y cannot be among the winners. □

A simple example of a winner selection method violating the unanimity criterion is the *method of sequential comparison*. Sequential comparison requires that the candidates be put into some specific a priori order. You might order them alphabetically by name, for instance. In a first round, candidates 1 and 2 compete head-to-head. The winner of the first round proceeds to the second round, in which she competes head-to-head against candidate 3. The winner of the second round proceeds to the third round, in which he competes head-to-heat against candidate 4, and so on. In the last round, the winner of the second-to-last round competes against candidate n. The winner of the last round is the overall winner. We will refer to this process as the *sequential comparison tournament*. To define the method unambiguously, we must specify what is to happen when one of the competitions of the tournament results in a tie. We make the (arbitrary) convention that in that case, the candidate who comes earlier in the a priori order wins.

Nobody would seriously propose sequential comparison as a winner selection method. It is evidently unfair: Candidate n has a far greater chance of winning than candidate 1. Sequential comparison violates the following basic principle.

Definition 2.9 (principle of independence of candidate names). *A winner selection method satisfies the principle of independence of candidate names if the outcome of the election does not change when the candidates are renamed or renumbered.*

Proposition 2.10. *Sequential comparison satisfies the Condorcet criterion, but violates the unanimity criterion.*

Proof. Sequential comparison is Condorcet-fair: A Condorcet candidate will win every head-to-head competition, and will therefore prevail in sequential comparison. To see that sequential comparison violates the unanimity criterion, consider the preference schedule

1	1	1
A	C	B
B	A	D
D	B	C
C	D	A

. (2.4)

Let us order the candidates alphabetically. Here A beats B, then loses against C, and D wins against C. So D is the overall winner, even though every voter prefers B to D. ☐

In Chapter 4, we will encounter a method that looks much more reasonable than sequential comparison at first glance, yet violates the unanimity criterion as well.

Exercises

2.1. True or false? (a) Any Condorcet-fair winner selection method is majority-fair. (b) Any majority-fair winner selection method is Condorcet-fair. (c) Any winner selection method that satisfies the unanimity criterion also satisfies the majority criterion. (d) Any winner selection method that satisfies the majority criterion also satisfies the unanimity criterion.

2.2. Explain why the elimination method is majority-fair.

2.3. What is the outcome of the election if (2.2) is the preference schedule (thus there is a circle in societal preferences), and the winner selection method is (a) the plurality method, (b) the runoff method, (c) the elimination method, (d) Borda count, and (e) pairwise comparison?

2.4. (a) Assume that there is no Condorcet candidate, and there are no ties in head-to-head competitions. (One easy way of ruling out ties in head-to-head competitions is to assume that the number of voters is odd.) Prove that there must then be a circle in societal preferences in the following sense. There are candidates $X_1, X_2, \ldots, X_k$ (k could be smaller than n) so that a majority of voters prefer X_1 to X_2, a majority prefer X_2 to X_3, and so on, and finally a majority prefer X_{k-1} to X_k, but also a majority prefer X_k to X_1.

(b) Give an example that demonstrates the following: When there is a circle in societal preferences, in the sense explained in part (a), there may still be a Condorcet candidate.

2.5. This exercise refers to a point in the proof of Proposition 2.1. Give an example in which every voter ranks X above Y, and in head-to-head competition, Y ties with Z and X also ties with Z (but does not beat Z).

2.6. Assume that there are no ties in head-to-head competitions, for instance because the number of voters is odd. Prove that then in the method of sequential comparison, the candidate who comes first in the a priori order wins if and only if she is a Condorcet candidate.

2.7. Consider the following preference schedule:

3	2	2	1	1
A	B	C	C	C
B	A	B	A	D
C	D	D	D	A
D	C	A	B	B

(a) Is there a Condorcet candidate? (b) By the plurality method, who wins? (c) What are the Borda scores? Who wins by Borda count? (d) Who wins by the method of pairwise comparison?

2.8. Explain: When there are only two candidates, all winner selection methods that we have discussed yield the same winner, namely, the candidate with the larger number of first-place votes.

2.9. The *Coombs method* is just like the elimination method, except that in each round we eliminate not the candidate with the fewest first-place votes, but the one with the most last-place votes. Suppose that the preference schedule is

4	3	6	2
B	A	C	C
A	B	A	B
C	C	B	A

(a) Who wins by the elimination method? (b) Who wins by the Coombs method? (c) Is there a majority candidate? If so, who is it? (d) Explain: Parts (b) and (c) demonstrate that the Coombs method violates the majority criterion.

2.10. Explain why Borda count satisfies the unanimity criterion.

Chapter 3

Election Spoilers

In Chapter 1, we saw that candidates C and D *spoiled* the election for B when the plurality method was used. With them in the running, A won, whereas without them in the running, B would have won. Can a winner selection method be designed to guarantee that this sort of *spoiler effect* will never occur? On the web page of an organization called the New America Foundation, I found this sentence:

> It so happens that there's a way to fix the system to do away with the spoiler problem.... It's called Instant Runoff Voting.

Instant runoff voting is what I call the elimination method. In this chapter, you will see a simple argument showing that the elimination method does *not* do away with the spoiler problem entirely, although it unquestionably makes it less likely to occur in practice. In fact, the argument shows that no reasonable winner selection method based on preference ballots can entirely do away with the spoiler problem in a strict sense.

We must first make completely precise what we mean by an *election spoiler*. (We will use the notation of elementary set theory here. If you are not familiar with this notation, please read Appendix A before reading this chapter.) Suppose that a candidate X joins the field of candidates just before the election. Assume that X's entry does not affect the order in which individual voters rank the *other* candidates. (That seems eminently reasonable.) Denote by W' the set of candidates who would have been the winners prior to X's entry, and by W the set of winners after X enters. In an ideal world, one of the following three statements should be true:

(i) $W = \{X\}$—that is, X beats all other candidates.

(ii) $W = W' \cup \{X\}$—that is, X joins the set of winners.

(iii) $W = W'$—that is, X's entry does not affect the outcome of the election.

If anything else happens, we say that X is an *election spoiler*. For instance, suppose that W contains all candidates in W', X, and one extra candidate Y. You might argue then that X has not really *spoiled* the election for anybody—everybody who was a winner prior to X's entry is still a winner. That is true, but the candidates in W' must now share their victory not only with X (nobody could have an issue with that, since the electorate ranks X equal to the candidates in W') but also with Y, who was ranked lower than the candidates in W' prior to X's entry.

Definition 3.1 (no-spoiler criterion). *A winner selection method satisfies the no-spoiler criterion if it guarantees that* (i), (ii), *and* (iii) *are the only possibilities when X joins the field of candidates.*

The no-spoiler criterion refers to the preferences of *society* (as determined by an election based on preference ballots). It is an interesting question of psychology, though not of mathematics, to which extent individuals satisfy the no-spoiler criterion in their decision making. The following story, attributed to Sidney Morgenbesser, a philosopher who taught at Columbia University, illustrates the notion of a spoiler in *individual* (not *social*) choice.

> After finishing dinner, Professor Morgenbesser decides to order dessert. The waitress offers two choices: apple pie or blueberry pie. Morgenbesser orders the apple pie. After a few minutes the waitress comes back to the table and says, "I forgot to tell you, we also have cherry pie tonight." Morgenbesser replies, "Oh, in that case I'll have the blueberry pie."

One can also think about the no-spoiler criterion in the reverse direction: Suppose that candidate X participates in the election, but one day after the election, it is discovered that X is a convicted felon who should never have been on the ballot in the first place. Therefore X is retroactively disqualified. Assume that the presence or absence of X on the ballot does not affect the order in which individual voters rank the other candidates. Denote by W the set of winners with X on the ballot, and by W' the set of winners after X is retroactively disqualified. In an ideal world, there should be only three possibilities:

(i)$'$ $W = \{X\}$—that is, X was the sole winner of the election prior to his retroactive disqualification; in this case, nothing can be said about what W' ought to be.

(ii)$'$ $W' = W - \{X\}$—that is, retroactive disqualification of X simply removes X from the set of winners; W must contain more than one candidate in this case.

(iii)$'$ $W' = W$—that is, retroactive disqualification of X does not affect the election outcome.

Definition 3.2 (retroactive disqualification criterion). *A winner selection method satisfies the retroactive disqualification criterion if it guarantees that* (i)$'$, (ii)$'$, *and* (iii)$'$ *are the only possibilities when X is retroactively disqualified.*

The retroactive disqualification criterion is nothing other than an alternative way of stating the no-spoiler criterion. However, in some situations I find it more natural to think in terms of retroactive disqualification than in terms of a candidate joining the field.

The following example demonstrates that the method of pairwise comparison violates the retroactive disqualification criterion.

Example 3.3. Four partners in a law firm would like to hire a new secretary. They interview three candidates named A, B, and C. After the interviews, they vote on whom to hire. The preference schedule is as follows:

2	1	1
A	C	B
B	A	C
C	B	A

.

A beats B, B beats C, and C and A tie, so A wins the pairwise comparison tournament. But a day after the ballots have been cast, B calls to announce that he will be unavailable for 18 to 24 months because of "a huge misunderstanding" between him and a judge. The partners decide that this disqualifies him from working in a law firm. He is removed from the ballots. This yields

2	1	1
A	C	C
C	A	A

,

or more compactly

2	2
A	C
C	A

.

Now A ties C. This means that retroactive disqualification of B (who was not a winner) changed the set of winners: From $\{A\}$ to $\{A, C\}$. Before B is disqualified, A wins unequivocally. After B is disqualified, A ties with C. Since only one position is available, the partners now flip a coin to determine whether A or C should get the job, and it so happens that C wins. ∎

Example 3.3 also proves that the plurality method violates the retroactive disqualification criterion. Before B's jail sentence becomes known and he gets disqualified, A is the sole winner by the plurality method. After B gets disqualified, A and C tie. It is, in fact, not hard to verify that *all* winner selection methods that we have encountered thus far violate the retroactive disqualification criterion (see Exercise 3.1). We seem to be back to square one: We don't know any method that is truly satisfactory, since they all allow the possibility of spoilers. We will prove now that this is mathematically unavoidable: Spoilers, in the strict sense in which we use the word here, cannot be ruled out by any reasonable winner selection method.

To see why and in which sense this is true, we make the following assumption about our winner selection method: When there are two candidates X and Y and three voters, and when at least two of the three voters prefer X to Y, then X will be the sole winner. This assumption holds, for instance, if the winner selection method satisfies the majority criterion; but it really holds for *any* reasonable winner selection method, including Borda count, which violates the majority criterion.

With this assumption, a single example proves that spoilers are unavoidable. The example is the simplest case in which societal preferences are circular:

1	1	1
A	B	C
B	C	A
C	A	B

.

Because of the symmetry of the situation, the only reasonable thing to do would be to make all three candidates the winners: $W = \{A, B, C\}$. Suppose that that is what our method does. Then retroactive disqualification of A turns the set of winners into $\{B\}$, which is not $\{A, B, C\} - \{A\}$, as it would have to be if the retroactive disqualification criterion were satisfied.

The only potential way out would be to assume that our method does something rather strange, for instance, that it selects two out of the three candidates as the winners. Let us consider, for example, the possibility $W = \{A, B\}$. (Because of symmetry, similar reasoning applies to any choice of two candidates as the winners.) When C—not an element of W—is retroactively disqualified, then by the retroactive disqualification criterion, the set of winners should stay the same—but it does not, since without C, A becomes the sole winner. So even this rather strange possibility does not offer a way out.

Or our method could select only one candidate as the winner—a very strange thing to do, but let us assume $W = \{A\}$ for argument's sake. (Again, because of symmetry, similar reasoning applies to any choice of one candidate as the winner.) If we retroactively disqualify B, then C becomes the winner—in violation of the retroactive disqualification criterion.

There is no way out: The retroactive disqualification criterion (or, equivalently, the no-spoiler criterion) is incompatible with the majority criterion, and even with the weaker assumption that in an election with two candidates, a candidate with a 2:1 majority wins.

There does exist a winner selection method that rules out election spoilers, namely, *dictatorship of the kth voter*: Always declare the top choice of voter k the (sole) winner of the election. This method evidently satisfies the no-spoiler criterion, but nobody other than voter k likes it.

All winner selection methods analyzed in this book are based on preference ballots. There are, however, winner selection methods not based on preference ballots, and among these, there are some that, in a certain sense, rule out election spoilers. The simplest example is *approval voting*. In elections using this method, just as in elections based on the plurality method, the voter sees on the ballot a list of candidates. However, the instruction is not "Check one," but "Check as many as you like." The candidate(s) with the largest overall number of checks win(s). This method of voting was first proposed and analyzed in the 1970s by several people, including Robert J. Weber, to whom the name "approval voting" is attributed.

Let us make the assumption that a voter will not take away checks from other candidates just because a new candidate X joins the field. The voter may, however, put a check next to X as well. With this assumption, of course either (i) $W = \{X\}$ (X gets more checks than any other candidate), or (ii) $W = W' \cup \{X\}$ (X ties for first place in the number of checks she gets), or (iii) $W = W'$ (X gets fewer checks than some of the others).

You can quite validly object to the entire discussion in this chapter that I should have defined *election spoiler* more narrowly. Not all election spoilers in the sense of my definition are really all that objectionable. In Exercise 3.2, you will find a definition of a particularly objectionable kind of spoiler called a *losing spoiler*. Even losing spoilers are not ruled out by any of the methods we have studied so far; see Exercises 3.2 and 3.3. However, we will later encounter an example of a Condorcet-fair method that does rule out losing spoiler (Proposition 4.13).

Exercises

3.1. Using the preference schedule

2	4	3
C	B	A
A	C	B
B	A	C

,

verify that the (a) plurality, (b) runoff, (c) elimination, and (d) Coombs methods (see Exercise 2.9), (e) Borda count, (f) pairwise comparison, and (g) sequential comparison all violate the retroactive disqualification criterion.

3.2. In this chapter, we have used the term *election spoilers* for two different kinds of candidates, whom we will now call *losing spoilers* and *winning spoilers*. A *losing spoiler* is a candidate $X \notin W$ whose retroactive disqualification changes the set of winners. A *winning spoiler* is a candidate X who belongs to W but is not the only candidate in W (that is, not the only winner), and whose retroactive disqualification leads to a new set of winners $W' \neq W - \{X\}$. Using the preference schedule in Exercise 3.1, show that the plurality, runoff, elimination, and Coombs methods (see Exercise 2.9), Borda count, and the method of sequential comparison allow losing spoilers. (The example does not work to prove the same for pairwise comparison, but see Exercise 3.3.)

3.3. Give an example showing that the method of pairwise comparison allows losing spoilers. (See Exercise 3.2 for the definition of *losing spoiler*.)

3.4. Consider the following preference schedule:

4	3	4	2
B	A	C	C
A	B	A	B
C	C	B	A

.

(a) Prove that A is a Condorcet candidate. (b) Prove that A is a losing spoiler when the plurality method is used. (See Exercise 3.2 for the definition of *losing spoiler*.)

3.5. Approval voting is not of the same form as the other winner selection methods considered in this book. We can therefore not ask whether approval voting is majority-fair without defining first what we mean by that. For argument's sake, let us assume that each voter has a unique favorite candidate. (It is in fact an attractive aspect of approval voting that it does *not* require this assumption, and more generally that it does not force voters to order candidates linearly.) Let us further assume that each voter assigns a check to his or her favorite candidate. Each voter may or may not also assign checks to other candidates. Give an example showing that approval voting is not majority-fair in the following sense: If a majority of voters have the same favorite candidate, that candidate may still not win the election.

Chapter 4

The Smith Set

In Chapter 3, we argued that no reasonable winner selection method based on preference ballots can guarantee that there will never be any election spoilers. This is a regrettable result. It is therefore natural to try to refine the notion of an election spoiler in such a way that, with the new notion, it is possible to guarantee that there will never be any spoilers.

Exercise 3.2 presented an attempt at such a refinement. The idea was to impose the additional requirement that a "spoiler" be a "weak" candidate, for when a weak candidate changes the outcome, that is more disturbing than when a strong candidate does. In Exercise 3.2, candidates who lose (according to the given winner selection method) were considered weak, and ones who win were considered strong. This did not turn out to be very fruitful: Most winner selection methods do not just permit spoilers in the sense defined in Chapter 3, but also losing spoilers in the sense defined in Exercise 3.2.

In this chapter, we will modify this idea. We will make the distinction between "weak" and "strong" candidates independently of the winner selection method, based just on the pairwise comparison graph. This will lead to a notion of "weak spoilers." We will see in Chapter 5 that there are many winner selection methods which guarantee that there will never be any weak spoilers.

When there is a Condorcet candidate X, it is reasonable to say that X is a strong candidate, and all others are weak candidates. The trouble is that there is not always a Condorcet candidate. We will therefore define a notion of "generalized Condorcet candidates," a set S consisting of the strongest candidates. When there is a Condorcet candidate X, then $S = \{X\}$. When there is no Condorcet candidate, S is still defined, but contains at least two candidates. The set S is called the *Smith set*, after John H. Smith, a mathematician at Boston College.[2]

To define S, we begin with the auxiliary concept of a *dominating* set of candidates. A nonempty set $\mathcal{D}$ of candidates is called dominating if the candidates inside $\mathcal{D}$ are stronger than those outside $\mathcal{D}$.

Definition 4.1 (dominating set). *A nonempty set $\mathcal{D}$ of candidates is called dominating if every candidate $X \in \mathcal{D}$ beats every candidate $Y \notin \mathcal{D}$ in head-to-head competition.*

[2] John H. Smith, "Aggregation of preferences with variable electorate," *Econometrica* 41, issue 6, pages 1027–1041 (1973).

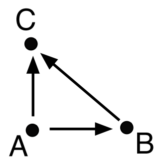
(I use the font $\mathcal{D}$, not D, in this context to avoid confusing the dominating set $\mathcal{D}$ with the candidate D.)

Example 4.2. Consider the preference schedule

4	7	5
A	B	C
B	A	A
C	C	B

.

Here A beats B and C, and B beats C:



The set $\mathcal{D} = \{A,B\}$ is dominating, since every candidate in $\mathcal{D}$ (candidates A and B) beats every candidate outside $\mathcal{D}$ (only C is outside $\mathcal{D}$) in head-to-head competition. Another example of a dominating set is $\mathcal{D} = \{A\}$, since every candidate inside (just A) beats every candidate outside (B and C) in head-to-head competition. There is a silly third example: $\mathcal{D} = \{A,B,C\}$. Every candidate inside $\mathcal{D}$ beats every candidate outside $\mathcal{D}$, which is not at all hard to do, since there are no candidates outside this set $\mathcal{D}$. (If this seems strange to you, please read the last paragraph in Section B.5 of Appendix B.) ∎

In fact, the set of *all* candidates is *always* an example of a dominating set, albeit a silly one. Recall that by definition a dominating set must be nonempty; otherwise the empty set would be another silly example.

Definition 4.3 (Smith set). *The Smith set S is the smallest dominating set.*

We refer to members of the Smith set as *Smith candidates*, and to the others as *non-Smith candidates*. It is easy to see that $S = \{X\}$ if X is a Condorcet candidate (see Exercise 4.4). On the other hand, the following is an example in which there are three Smith candidates.

Example 4.4. If the preference schedule is

4	3	5
A	B	C
B	C	A
D	D	D
C	A	B

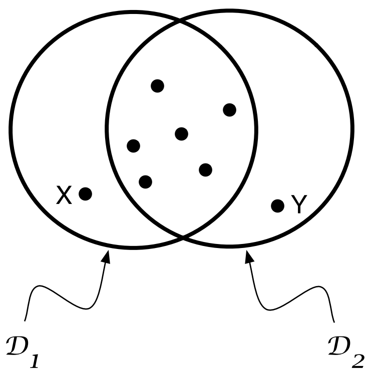
,

then A beats B, B beats C, C beats A, and everybody beats D. Therefore $S = \{A,B,C\}$. You cannot narrow it down any further. ∎

You may (and in fact you should) wonder whether there could be two or more smallest dominating sets. For instance, suppose that there is no Condorcet candidate, so there is no dominating set containing only one candidate, but both $\{A,B\}$ and $\{C,D\}$ are dominating. Then which of these two is the Smith set? Fortunately, this situation is impossible: If $\{A,B\}$ is dominating, then A and B beat C and D in head-to-head competition; but if $\{C,D\}$ is also dominating, then C and D also beat A and B in head-to-head competition. Both cannot be true. A very similar argument shows that dominating sets must be contained within each other.

Lemma 4.5. *If $\mathcal{D}_1$ and $\mathcal{D}_2$ are dominating sets, then either $\mathcal{D}_1 \subseteq \mathcal{D}_2$ or $\mathcal{D}_2 \subseteq \mathcal{D}_1$.*

Proof. Suppose neither were true. Then there would be a candidate $X \in \mathcal{D}_1 - \mathcal{D}_2$, and also a candidate $Y \in \mathcal{D}_2 - \mathcal{D}_1$.



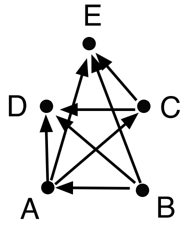
In head-to-head competition, X would have to beat Y (because X is in $\mathcal{D}_1$ and Y is not), but also Y would have to beat X (because Y is in $\mathcal{D}_2$ and X is not). Both cannot be true. The only way out of this contradiction is to assume that either $\mathcal{D}_1$ is wholly contained in $\mathcal{D}_2$, or vice versa. $\square$

Lemma 4.5 implies that there can never be two *different* dominating sets with exactly 7 elements, say. If the total number of candidates is 10, there can be *at most* 10 dominating sets, one with 10 elements, one with 9 elements, etc. There is a unique smallest dominating set; it is the Smith set S.

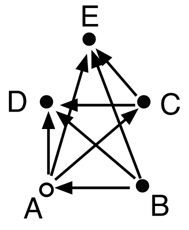
From the definition of S, it would appear that to compute S is no easy task in general. One would seem to have to compute and list all dominating sets first, and that seems hard at first sight. However, I will now show that it is not really so difficult to compute all dominating sets, and that, besides, one need not even compute all of them to compute the Smith set. Let us begin by thinking about how to compute a special kind of dominating sets, which we call *minimal*.

Definition 4.6 (primitive dominating set). *If X is a candidate, we denote by $\mathcal{D}_X$ the smallest dominating set containing X. A dominating set $\mathcal{D}$ is called primitive if $\mathcal{D} = \mathcal{D}_X$ for some candidate X.*

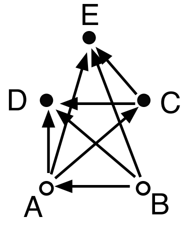
Example 4.7. Let us assume that the pairwise comparison graph looks like this:



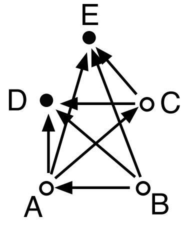
(To compute dominating sets, one needs the pairwise comparison graph, but one does not need the margins of victory.) We want to compute $\mathcal{D}_A$, the smallest dominating set that contains A. We will indicate candidates that belong to it by open circles instead of dots. We start by indicating A itself by an open circle:



$\mathcal{D}_A$ must also contain all candidates whom A does not beat—in the example, that is just B:



$\mathcal{D}_A$ must also contain the candidates whom B does not beat. That is just C:



Now we must add those whom C does not beat—but that yields nothing new, and therefore the process ends here, and we conclude $\mathcal{D}_A = \{A, B, C\}$. We can similarly find $\mathcal{D}_B$, $\mathcal{D}_C$, $\mathcal{D}_D$, and $\mathcal{D}_E$:

$$\begin{aligned}
\mathcal{D}_A &= \{A, B, C\},\\
\mathcal{D}_B &= \{A, B, C\},\\
\mathcal{D}_C &= \{A, B, C\},\\
\mathcal{D}_D &= \{A, B, C, D, E\},\\
\mathcal{D}_E &= \{A, B, C, D, E\}.
\end{aligned}$$

So in this example, there are exactly two primitive dominating sets, namely, $\{A, B, C\}$ and $\{A, B, C, D, E\}$. ∎

How do we find dominating sets that are not primitive? Fortunately, we need not worry about that.

Proposition 4.8. *Every dominating set is primitive.*

Proof. Consider an arbitrary dominating set $\mathcal{D} = \{X_1, X_2, \ldots, X_k\}$. (Here k is the number of candidates in $\mathcal{D}$. Of course, $k \le n$, where n is, as always, the total number of candidates.) Lemma 4.5 implies that the union

$$\tilde{\mathcal{D}} = \bigcup_{j=1}^{k} \mathcal{D}_{X_j} \tag{4.1}$$

equals $\mathcal{D}_{X_i}$ for some i with $1 \le i \le k$. We will prove now that $\mathcal{D} = \tilde{\mathcal{D}}$.

First, it follows immediately from (4.1) that $\mathcal{D} \subseteq \tilde{\mathcal{D}}$, that is, that all X_j (the elements of $\mathcal{D}$) belong to $\tilde{\mathcal{D}}$. Second, we prove that $\tilde{\mathcal{D}} \subseteq \mathcal{D}$. This is true because $\tilde{\mathcal{D}} = \mathcal{D}_{X_i}$ for some i with $1 \le i \le k$, so $\tilde{\mathcal{D}}$ is the smallest dominating set containing X_i, and $\mathcal{D}$ is a dominating set containing X_i.

So $\mathcal{D} = \mathcal{D}_{X_i}$, and thus $\mathcal{D}$ is a primitive dominating set. □

So in Example 4.7, we did not merely find all primitive dominating sets, we found all dominating sets.

Lemma 4.9. *If $X \in S$, then $S = \mathcal{D}_X$.*

Proof. S is the smallest of all dominating sets, and it contains X. $\mathcal{D}_X$ is the smallest dominating set containing X. These two must be equal. □

Lemma 4.9 shows that the Smith set is easy to compute once we know one Smith candidate. In fact it is also easy to find one Smith candidate, as the following proposition shows.

Proposition 4.10. *When the method of sequential comparison is used, the winner is a member of the Smith set.*

Proof. As soon as the first contest between a Smith candidate and a non-Smith candidate occurs in the sequential comparison tournament, the Smith candidate will win. Every one of the subsequent head-to-head competitions will be won by a Smith candidate. Therefore the overall winner is a Smith candidate. □

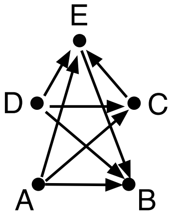
In summary, we have arrived at a method for computing the Smith set S: First, determine the candidate X who wins by the method of sequential comparison; then compute $\mathcal{D}_X$, and that is S.

Example 4.11. Let us consider the preference schedule

2	2	4
A	A	D
B	E	A
D	D	C
C	B	E
E	C	B

.

The pairwise comparison graph is



With alphabetical ordering of the candidates, A wins the sequential comparison tournament, and therefore $A \in S$. If A belongs to a dominating set, then so does D, since A does not beat D. Both A and D beat all other candidates. So $S = \mathcal{D}_A = \{A, D\}$. ■

We return now to the topic of election spoilers. Recall from Chapter 3 that a *spoiler* is a candidate who is not the sole winner of the election, but whose retroactive disqualification would affect other candidates, turning them from winners into losers, or vice versa. In a similar spirit, we now ask whether the retroactive disqualification of a non-Smith candidate might affect the Smith set. The answer is no.

Lemma 4.12. *If a non-Smith candidate is removed from the ballots, and if this does not affect individual voters' rankings of the remaining candidates, the Smith set S does not change.*

Proof. S is still a dominating set. Furthermore, it must still be the smallest dominating set, for if there were a dominating set $\mathcal{D} \subseteq S$, $\mathcal{D} \neq S$, then $\mathcal{D}$ would have been dominating even before the non-Smith candidate was removed, and therefore S would not have been the smallest dominating set even before removal of the non-Smith candidate. □

As explained in Exercise 3.2, one may distinguish two kinds of election spoilers, *losing* and *winning* ones. A *losing spoiler* is a candidate $X \notin W$ whose retroactive disqualification changes the set of winners. A *winning spoiler* is a candidate $X \in W$ who is not the only candidate in W (that is, not the only winner), and whose retroactive disqualification leads to a new set W' of winners that is *not* equal to $W - \{X\}$. We proved in Chapter 3 that any winner selection method that satisfies the majority criterion must admit election spoilers—losing or winning ones. One can in fact verify easily that all methods that we have considered so far, with the exception of dictatorship, even allow *losing* spoilers; see Exercises 3.2 and 3.3. There is, however, a method that satisfies the majority criterion (and in fact even the Condorcet criterion) and rules out losing spoilers—namely, the method that simply sets $W = S$. I will call this the *Smith method*.

Proposition 4.13. *The Smith method is Condorcet-fair and does not allow losing spoilers.*

Proof. To see that the Smith method is Condorcet-fair, suppose that X is a Condorcet candidate. Then $S = \{X\}$, and therefore the Smith method selects X as the sole winner: $W = S = \{X\}$. It does not allow losing spoilers by Lemma 4.12. □

Nevertheless, the Smith method is seriously flawed. (My apologies to John Smith, who certainly never suggested this winner selection method!) The Smith set S is often fairly large, so it is desirable to select a set W of winners smaller than the Smith set. The following proposition demonstrates a closely related problem with the Smith method.

Proposition 4.14. *The Smith method violates the unanimity criterion.*

Proof. The following preference schedule was used in Chapter 2 to demonstrate that the method of sequential comparison violates the unanimity criterion:

1	1	1
A	C	B
B	A	D
D	B	C
C	D	A

Here every voter ranks B over D, yet the method of sequential comparison, with alphabetic a priori ordering of the candidates, makes D the winner. By Proposition 4.10, this implies $D \in S$. Thus the Smith method makes D a winner even though every voter ranks B above D. □

In fact, any example that proves that the method of sequential comparison violates the unanimity criterion is also an example that proves that the Smith method violates the unanimity criterion (see Exercise 4.8).

Exercises

4.1. Suppose the preference schedule is

3	2	2	1	1
A	B	C	A	B
B	A	B	C	C
C	C	A	B	A

.

List all dominating sets. What is the Smith set? By pairwise comparison, who wins the election?

4.2. Suppose the preference schedule is

2	2	4
A	A	D
B	E	A
D	D	C
C	B	E
E	C	B

.

List all dominating sets. What is the Smith set? By pairwise comparison, who wins the election?

4.3. This exercise refers to Example 2.7. (a) Order the options as follows: Maine, Yellowstone, San Francisco, Adirondacks. By the method of sequential comparison, which option wins? (b) What is the Smith set?

4.4. Explain: If X is a Condorcet candidate, then the Smith set is $S = \{X\}$.

4.5. Give an example in which the method of pairwise comparison selects a set of winners W not equal to the Smith set S.

4.6. Explain: A candidate X belongs to the Smith set if and only if she belongs to every dominating set.

4.7. Verify that $S = \{A, B, C, D\}$ in the example used in the proof of Proposition 4.14.

4.8. Explain: Any example that demonstrates that the method of sequential comparison violates the unanimity criterion also demonstrates that the Smith method violates the unanimity criterion.

Chapter 5

Smith-Fairness and the No-Weak-Spoiler Criterion

The notion of the Smith candidates is a generalization of that of a Condorcet candidate. It naturally leads to a strengthening of the notion of Condorcet-fairness, which we will call Smith-fairness. The requirement of Smith-fairness is compatible with the requirement that there be no "weak spoilers," that is, no non-Smith candidates that affect the election outcome.

Definition 5.1 (Smith-fairness). *A winner selection method is said to be Smith-fair, or to satisfy the Smith criterion, if it guarantees that the set W of winners is always a subset of the Smith set S: $W \subseteq S$.*

The simplest Smith-fair method, of course, is the Smith method introduced in Chapter 4, which simply sets $W = S$.

Proposition 5.2. *Every Smith-fair winner selection method is Condorcet-fair, but not every Condorcet-fair method is Smith-fair.*

Proof. Exercise 5.3 asks the reader to explain why Smith-fair methods are Condorcet-fair. It is not hard, on the other hand, to construct examples of methods that are Condorcet-fair, but not Smith-fair. For instance, suppose that our method is to determine whether or not there is a Condorcet candidate first, and if there is one, make her the sole winner of the election; if there is no Condorcet candidate, however, then we use the plurality method. There is no good reason for doing this—it is a silly method. However, it is clearly Condorcet-fair, and it is not Smith-fair (see Exercise 5.4). □

The plurality, runoff, and elimination methods and Borda count are not even Condorcet-fair, and therefore are certainly not Smith-fair. The lack of Smith-fairness of the plurality method is illustrated strikingly by the following example.

Example 5.3. Consider this preference schedule:

1	2	1	1	2	2
D	D	D	A	B	C
A	B	C	B	C	A
B	C	A	C	A	B
C	A	B	D	D	D

In head-to-head competition, A beats B, B beats C, C beats A, and all three of them beat D. So D is the only non-Smith candidate here. Nevertheless, D is the winner according to the plurality method. ∎

Proposition 5.4. *The method of pairwise comparison is Smith-fair.*

Proof. Suppose there are k Smith candidates and $n-k$ non-Smith candidates. A non-Smith candidate cannot beat any Smith candidate in head-to-head competition. Therefore a non-Smith candidate can get at most $n-k-1$ points in the pairwise comparison tournament. On the other hand, a Smith candidate gets at least $n-k$ points for beating every non-Smith candidate. This proves that non-Smith candidates cannot win the pairwise comparison tournament. □

The method of pairwise comparison often selects a proper subset of S as the set of winners; that is, it is not the same as the Smith method. The preference schedule

1	1	2
A	B	C
B	A	A
C	C	B

(5.1)

shows this. Here $S = \{A, B, C\}$, but $W = \{A\}$ according to the method of pairwise comparison. (This is a solution to Exercise 4.5.)

Proposition 5.5. *The method of sequential comparison is Smith-fair.*

Proof. This is the content of Proposition 4.10. □

It is easy to find other examples of Smith-fair methods. In fact, we can obtain a Smith-fair method simply by first computing S, removing the non-Smith candidates from the ballots, then doing whatever we wish to select a set of winners $W \subseteq S$—for instance the plurality, runoff, or elimination methods, the Coombs method, Borda count, even dictatorship. We call any such method *a priori Smith-fair*. So a winner selection method is a priori Smith-fair if its definition begins (or could begin) with the sentence "Remove all non-Smith candidates from the ballots." Or, in slightly different words:

Definition 5.6 (a priori Smith-fairness). *A winner selection method is called a priori Smith-fair if the presence or absence of non-Smith candidates never affects the election outcome.*

Any of the winner selection methods that we have discussed can be turned into a priori Smith-fair methods by starting the process of evaluating the election outcome by removing all non-Smith candidates from the ballots. For instance, the method of removing all non-Smith candidates from the ballots first, then determining the winner(s) based on plurality will be called the *a priori Smith-fair plurality method*. Similarly, when Borda count is used after removal of the non-Smith candidates, we refer to *a priori Smith-fair Borda count*, and so on.

Example 5.7. For the preference schedule

1	1	1	1
A	E	C	C
B	B	B	A
C	C	D	B
D	A	A	D
E	D	E	E

, (5.2)

$S = \{A, B, C\}$ (see Exercise 5.6). After removing the non-Smith candidates from the ballots, the preference schedule becomes

1	1	1	1
A	B	C	C
B	C	B	A
C	A	A	B

.

If we now apply, say, the plurality method, we obtain $W = \{C\}$. So the a priori Smith-fair plurality method makes C the winner. (As it happens, the plain plurality method would also make C the winner.) If instead we apply dictatorship of the second voter (the one who ranked the candidates E, B, C, A, D), we obtain $W = \{B\}$. So a priori Smith-fair dictatorship of the second voter makes B the winner. (Note that B is not the dictator's first choice—just her first choice among the Smith candidates.) ∎

Example 5.8. What happens when we apply the a priori Smith-fair elimination method to the preference schedule in Example 5.3? First, the non-Smith candidate D is removed. The resulting preference schedule is

2	4	3
A	B	C
B	C	A
C	A	B

.

In the elimination method, A is eliminated now because she has the fewest first-place votes. Then B beats C, 6:3, and therefore B is the winner. ∎

There are Smith-fair methods that are not a priori Smith-fair. As an example, consider the following winner selection method, which I will call a posteriori Smith-fair Borda count: Compute the Borda scores of all candidates (leaving non-Smith candidates on the ballots), then define the winner(s) to be the Smith candidate(s) with the highest Borda score—even if there is a non-Smith candidate with a higher Borda score. This method is clearly Smith-fair, since $W \subseteq S$. However, it is not a priori Smith-fair:

Proposition 5.9. *A posteriori Smith-fair Borda count is not a priori Smith-fair.*

Proof. You may think that this is obvious, since the definition of the method does not start with "Remove all non-Smith candidates from the ballots." However, might it be possible that there is a different but equivalent definition of the method that does start with "Remove all non-Smith candidates from the ballots"? The answer is no, that is not possible. To see why not, consider the preference schedule

2	1	1
B	A	A
A	B	C
C	C	B

. (5.3)

Here $S = \{A, B\}$, a posteriori Smith-fair Borda count selects $W = \{A\}$, but if the non-Smith candidate C were removed from the ballots from the start, then $W = \{A, B\}$ (see Exercise 5.7). Thus the presence of the non-Smith candidate C affects the outcome of a posteriori Smith-fair Borda count in this example. ∎

Definition 5.10 (a posteriori Smith-fairness). *A Smith-fair method that is not a priori Smith-fair is called a posteriori Smith-fair.*

The definition of pairwise comparison does not begin with "Remove all non-Smith candidates from the ballots" either. However, it *could* begin that way: If we removed all non-Smith candidates from the ballots, then did pairwise comparison within the Smith set, the outcome would be the same as that of doing pairwise comparison without paying attention to the Smith set.

Proposition 5.11. *The method of pairwise comparison is a priori Smith-fair.*

Proof. See Exercise 5.8. ∎

Smith-fairness is a particularly stringent "rule of the majority" requirement: It is stronger than Condorcet-fairness, which in turn is stronger than majority-fairness. However, a priori Smith-fair methods also satisfy a weak form of the no-spoiler criterion: They do not rule out spoilers (in Chapter 3 we saw that that would be impossible), nor do they necessarily rule out losing spoilers (pairwise comparison allows losing spoilers—see Exercise 3.3). However, they do, in a certain sense, rule out the worst kind of election spoilers, as we will now explain.

Definition 5.12 (weak spoiler). *We call a candidate X a weak spoiler if $X \notin S$, but removal of X from the ballots affects the set W of winners.*

Example 5.13. In an election with four candidates, A, B, C, and D, the preference schedule is

2	2	1	3	2
A	A	B	C	D
B	B	C	A	C
D	C	A	B	A
C	D	D	D	B

.

In head-to-head competition, A beats B, B and C tie, C beats A, and everybody beats D. Therefore the Smith set is $S = \{A, B, C\}$. If the plurality method is used, A is the sole winner, but if the non-Smith candidate D were removed from the ballots, then the preference schedule would become

2	2	1	3	2
A	A	B	C	C
B	B	C	A	A
C	C	A	B	B

,

so C would win. Therefore D is a weak spoiler here. ∎

Definition 5.14 (no-weak-spoiler criterion). *A winner selection method is said to satisfy the no-weak-spoiler criterion if it guarantees that there will never be any weak spoilers.*

Proposition 5.15. *A winner selection method satisfies the no-weak-spoiler criterion if and only if it is a priori Smith-fair.*

Proof. This is true by definition of a priori Smith-fairness (Definition 5.1). □

Thus the ideas of majority rule and absence of election spoilers, which clashed in Chapter 3, come together here: A priori Smith-fairness, a strong "majority rule" criterion, is equivalent to a weak no-spoiler criterion.

Unfortunately, there are methods that are a priori Smith-fair yet still patently unreasonable: Sequential comparison is one example (see Exercise 5.9), and a priori Smith-fair dictatorship is another. Further fairness requirements rule out many of these methods. For instance, the principle of independence of candidate names rules out sequential comparison. The principle of one person, one vote rules out a priori Smith-fair dictatorship. The unanimity criterion rules out the Smith method (see Proposition 4.14). In Chapter 7, we will discuss another requirement that further narrows down the class of acceptable a priori Smith-fair methods.

Exercises

5.1. In the *a posteriori Smith-fair plurality method*, the Smith candidate with the largest number of first-place votes (among all Smith candidates) is declared the winner. The Smith candidates are not removed a priori from the preference schedule. (a) Who wins if the a posteriori Smith-fair plurality method is applied to the preference schedule in Example 5.3? (b) Who wins if the a priori Smith-fair plurality method is applied to the same preference schedule?

5.2. In Example 5.13, is D a weak spoiler if we use (a) the elimination method, (b) Borda count, (c) pairwise comparison?

5.3. Explain why any Smith-fair method is Condorcet-fair.

5.4. Prove that the silly method in the proof of Proposition 5.2 is not Smith-fair.

5.5. In the proof of Proposition 5.4, we stated that every Smith candidate gets at least $n-k$ points in the pairwise comparison tournament. Prove that every Smith candidate in fact gets at least $n-k+1/2$ points.

5.6. Prove that in Example 5.7, $S = \{A, B, C\}$.

5.7. Verify that for the preference schedule (5.3), (a) $S = \{A, B\}$, (b) $W = \{A\}$ if a posteriori Smith-fair Borda count is used, and (c) $W = \{A, B\}$ if the non-Smith candidate C is removed first, and Borda count is then applied only within the Smith set.

5.8. Prove that the method of pairwise comparison is a priori Smith-fair. (Note that this is different from Proposition 5.4, which left open the possibility that the method of pairwise comparison might be merely a posteriori Smith-fair.)

5.9. Prove that the method of sequential comparison is a priori Smith-fair.

5.10. (a) Give an example of a weak spoiler who is not a spoiler in the sense of Chapter 3. (b) Give an example of a spoiler in the sense of Chapter 3 who is not a weak spoiler.

Chapter 6

Schulze's Beatpath Method

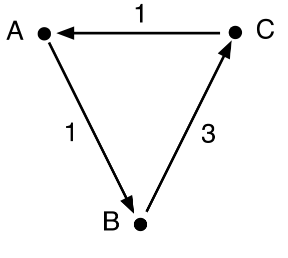
In this chapter, we describe a particularly attractive a priori Smith-fair winner selection method, proposed by Markus Schulze in 1997 and called the *beatpath method*. In this method, one considers not only who beats whom in head-to-head competition, but also the *margins* of victory—that is, the full pairwise comparison graph. As a result, this method makes ties very unlikely when the number of voters is large, much less likely than they are, for instance, with the pairwise comparison method; see Chapter 8 for more on this point.

Example 6.1. To illustrate the idea of the beatpath method, consider this preference schedule:

2	2	1
A	B	C
B	C	A
C	A	B

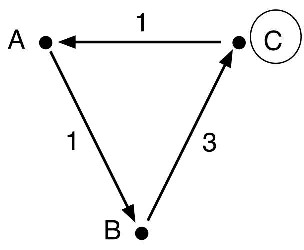
. (6.1)

The pairwise comparison graph looks like this:



The numbers next to the arrows indicate the margins of victory. For instance, in head-to-head competition, B beats C 4:1, thus the margin of victory is 3.

However, in an indirect sense, C also beats B, since C beats A and A, in turn, beats B. We say that $B \to C$ is a *beatpath* from B to C, and $C \to A \to B$ is a *beatpath* from C to B. The *strength* of the beatpath from B to C equals 3, and that of the beatpath from C to B equals 1. In general, the strength of a beatpath is the lowest margin of victory encountered along the path. For instance, there is a beatpath from B to A: $B \to C \to A$. The strength of this beatpath equals 1, since the margins of victory are 3 (the margin by which B beats C) and 1 (the margin by which C beats A)—so the lowest margin of victory encountered along the path is 1. We say that the beatpath $B \to C$ (of strength 3) is *not matched* by the beatpath $C \to A \to B$ (only of strength 1). Therefore we declare C a loser, and record this fact by circling C:



It is important that we do *not* take C out of the graph. So even after we have circled C, thereby indicating that $C \notin W$, there is still a beatpath from B to A, namely, $B \to C \to A$, as before. This beatpath is of strength 1, and matches the beatpath $A \to B$, which is also of strength 1. A does not have an unmatched beatpath against B, and B does not have an unmatched beatpath against A. Therefore we say $W = \{A, B\}$.

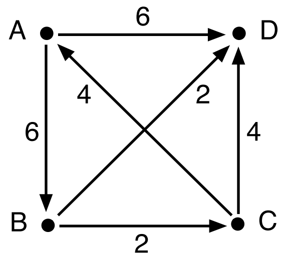
The method of pairwise comparison would have called all three candidates winners. The beatpath method declares C a loser as a result of the high margin with which he loses the head-to-head competition against B. ∎

Example 6.2. Consider now the preference schedule

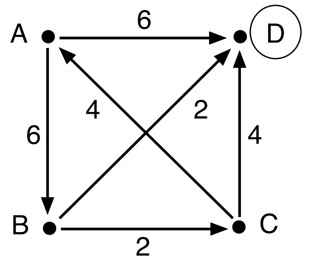
4	3	5
A	B	C
B	C	A
D	D	D
C	A	B

. (6.2)

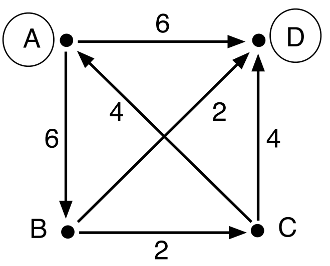
It is not so easy to have an intuition about who should win here. The pairwise comparison graph looks like this:



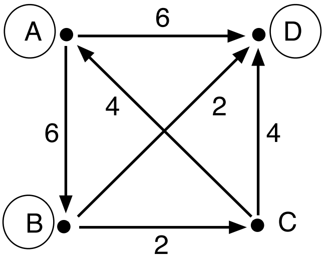
B has a beatpath against D of strength 2. (Even though it is irrelevant here, we note that B has two further, indirect beatpaths against D, both of strength 2 as well: $B \to C \to D$ and $B \to C \to A \to D$.) On the other hand, D does not have any beatpath against B (nor, for that matter, against anybody, since there are no arrows starting at D). Therefore D is not a winner. We indicate this by circling D:



C has a direct beatpath against A, of strength 4. A also has a beatpath against C, namely, $A \to B \to C$, but that path is only of strength 2, and A has no other beatpath against C. Therefore A is not able to match the beatpath that C has against it, and is not a winner:



C has a beatpath of strength 4 against B: $C \to A \to B$. (Remember that A is not removed from the graph; even though A is circled, she can still be part of a beatpath.) B cannot match this beatpath; his only beatpath against C is $B \to C$, and that is only of strength 2. Therefore B is not a winner:



The only candidate left standing is C; so $W = \{C\}$. ∎

You might wonder if this method guarantees that anybody is left standing in the end. Could it happen that for *every* candidate X, there is some candidate who has a beatpath against X which X cannot match? We will prove now that this is impossible, so the beatpath method always singles out at least one winner.

Let us use the notation $X \triangleright Y$ (or $Y \triangleleft X$) to indicate that there is a beatpath from X to Y which Y cannot match. We use the symbol "$\triangleright$" because it looks a little bit like "$>$". Of course, it isn't the same as "$>$". The symbol "$>$" stands between two numbers, whereas the symbol "$\triangleright$" stands between two candidates.

Proposition 6.3. *If $X \triangleright Y$ and $Y \triangleright Z$, then $X \triangleright Z$.*

Proof. We denote beatpaths by capital letters, such as P or Q. If P is a beatpath from X to Y, and Q a beatpath from Y to Z, we denote by PQ the beatpath from X to Z obtained by putting the two individual beatpaths together—first P and then Q.

Suppose that P is an unmatched beatpath from X to Y, and Q an unmatched beatpath from Y to Z. We will prove that PQ is an unmatched beatpath from X to Z, so $X \triangleright Z$.

We denote by p the strength of P, and by q the strength of Q. So p and q are positive integers. Then clearly PQ is a beatpath from X to Z, and its strength is the smaller of the two numbers p and q. We will prove that PQ is unmatched. We will proceed *indirectly* by assuming that the beatpath PQ were matched, then deriving that in that case, either P or Q would have to be matched as well—a contradiction, since P and Q were assumed to be unmatched. We distinguish the two cases $p \leq q$ and $p > q$.

Case 1: $p \leq q$.

In this case, the strength of PQ equals p. We assume that there were a matching beatpath from Z to X, that is, a beatpath R from Z to X of strength $\geq p$. Then QR is a beatpath from Y to X of strength $\geq p$. But this means that the beatpath P from X to Y is matched by QR, which is a contradiction—P was assumed to be an unmatched beatpath from X to Y.

Case 2: $p > q$.

In this case, PQ is a beatpath from X to Z of strength q. We assume again that there were a matching beatpath from Z to X, that is, a beatpath R from Z to X of strength $\geq q$. Then RP is a beatpath from Z to Y of strength $\geq q$. But this means that the beatpath Q from Y to Z is matched by RP, which is a contradiction—Q was assumed to be an unmatched beatpath from Y to Z. ∎

We summarize Proposition 6.3 by saying that unmatched beatpaths order the candidates in a *transitive* way. (*Transitive* means that $X \triangleright Y$ and $Y \triangleright Z$ imply $X \triangleright Z$.) Proposition 6.3 implies quickly that the beatpath method always singles out a nonempty set W of winners.

Proposition 6.4. *There is always at least one candidate against whom there is no unmatched beatpath.*

Proof. Suppose that there were an unmatched beatpath against *any* candidate. Then pick any candidate X_1. By our assumption, there is a candidate X_2 with $X_1 \triangleleft X_2$. Again by our assumption, there is a candidate X_3 with $X_2 \triangleleft X_3$. We can continue in this way, so we get a chain of the following kind:

$$X_1 \triangleleft X_2 \triangleleft X_3 \triangleleft X_4 \triangleleft \cdots .$$

No candidate can appear twice in this chain, for if X did appear twice, then we could conclude (because of Proposition 6.3) that $X \lhd X$, which obviously is false. But then there must be infinitely many candidates—and of course the number of candidates is finite! This contradiction can only be resolved by assuming that there is at least one candidate against whom there is no unmatched beatpath. $\square$

Proposition 6.5. *The beatpath method is a priori Smith-fair.*

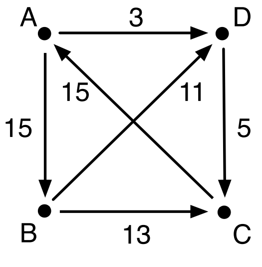
Proof. It is clear that the beatpath method is Smith-fair: Any Smith candidate has an unmatched beatpath against any non-Smith candidate (see Exercise 6.5(a)), so no non-Smith candidate can win. To prove that it is a priori Smith-fair, we must prove that the presence of the non-Smith candidates never affects the set W of winners. This is indeed the case because any beatpath against a Smith candidate can never pass through any non-Smith candidate (see Exercise 6.5(b)). $\square$

Proposition 6.6. *The beatpath method satisfies the unanimity criterion.*

Proof. Suppose that candidate X is preferred to candidate Y unanimously, that is, by a margin of N, where N denotes the number of voters. There is then a beatpath from X to Y of strength N. To prove that Y is not among the winners, we will prove that this beatpath cannot be matched. Suppose it were matched, that is, suppose that there were a beatpath of strength N (the strength of a beatpath could not be greater than N) from Y to X. So there would then be candidates $X_1, X_2, \ldots, X_r$ (with $r \geq 1$) so that every voter would prefer Y to X_1, X_1 to X_2, X_2 to X_3, $\ldots$, X_{r-1} to X_r, and X_r to X. This would mean that every voter would have circular preferences. However, a preference ballot requires voters to express linear, noncircular preferences. This contradiction proves that the beatpath from X to Y of strength N is in fact unmatched. $\square$

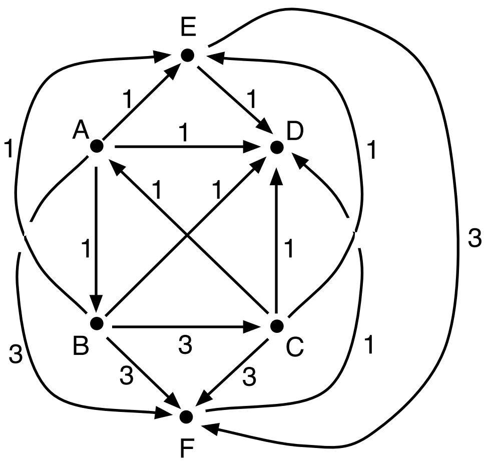
Exercises

6.1. Suppose that the pairwise comparison graph were as follows:



Who would win the election by the beatpath method?

6.2. Suppose that he pairwise comparison graph were as follows:



(Exercise 7.2 gives a preference schedule that gives rise to this pairwise comparison graph, but that is not relevant to this exercise.) Who would win the election by the beatpath method?

6.3. Four partners of a law firm interview three candidates for the job of secretary, then vote to determine who should be offered the job. The outcome of the vote is summarized by the following preference schedule.

2	1	1
A	C	B
B	A	C
C	B	A

Who should be the winner(s) according to the beatpath method?

6.4. The proof of Proposition 6.6 shows that a circle in the pairwise comparison graph in which each link has strength N (that is, the outcome of each of the head-to-head competitions along the circle is unanimous) is impossible. Using similar arguments, one can see that circles in the pairwise comparison graph in which the margins of victory are "too large" (the exact meaning of "too large" depends on how many candidates are involved in the circle, and how many voters there are) are impossible. Here is an example. Suppose that there are 4 voters and 3 candidates, A, B, and C. Prove that it is not possible that 3 voters prefer A to B, 3 voters prefer B to C, and 3 voters prefer C to A. (The argument is similar to that in the proof of Proposition 6.6.)

6.5. Explain the following two points from the proof of Proposition 6.3. (a) Any Smith candidate has an unmatched beatpath against any non-Smith candidate. (b) No beatpath against a Smith candidate can ever pass through any non-Smith candidate.

6.6. Use the preference schedule of Exercise 3.1 to prove that the beatpath method allows losing spoilers.

Chapter 7

Monotonicity

In the preceding chapters, I argued that a good winner selection method should satisfy a sequence of criteria:

a. the principle of one person, one vote;

b. the principle of independence of candidate names;

c. the no-weak-spoiler criterion (which is equivalent to a priori Smith-fairness, and implies Smith-fairness and therefore Condorcet-fairness);

d. the unanimity criterion.

Here is a list of winner selection methods satisfying all of these criteria:

1. the a priori Smith-fair plurality method;
2. the a priori Smith-fair runoff method;
3. the a priori Smith-fair elimination method;
4. a priori Smith-fair Borda count;
5. pairwise comparison;
6. the beatpath method.

In this chapter, we will consider yet another desirable property of winner selection methods: *monotonicity*. It will turn out that methods 5 and 6 on the above list have this property, but 1 through 4 do not. We begin with an example illustrating the notion of monotonicity.

Example 7.1. In a small village, a mayor is to be elected using the runoff method. There are 35 voters and 3 candidates, named A, B, and C. The preference schedule is

14	3	8	10
A	B	B	C
B	A	C	A
C	C	A	B

. (7.1)

The votes are cast, and the ballots locked away for the night, to be counted the following morning. Since you already know the ballots, you know the outcome: A and B will get into the runoff. In the runoff, A will beat B handily (24 to 11), so A is the winner.

This is what *should* happen, if it weren't for two tough-looking gentlemen who are on the payroll of A. These two gentlemen break into the room where the ballots are kept overnight and change three of the ballots in favor of their employer: They change the BAC ballots into ABC ballots, turning the preference schedule into

17	8	10
A	B	C
B	C	A
C	A	B

. (7.2)

The following morning, the ballots are counted. A and C get into the runoff, and in the runoff, C beats A just barely—18 to 17. So C is the winner. (The two tough-looking gentlemen are never seen again.) ∎

The disturbing fact about this example is that the changes introduced by the two tough guys were clearly favorable to A: A was moved up on three of the ballots. Nonetheless, A turned from a winner into a loser as a result of these changes. Because the runoff method allows this to happen, we say that it is *not monotonic*. Here are the formal definitions.

Definition 7.2 (ballot change favorable to X). *If X is moved upwards in a ballot, while the ordering of all other candidates on the altered ballot, as well as the preference rankings on all other ballots, are left unchanged, we call the change favorable to X.*

Example 7.3. A change from the preference ballot

B
D
A
C
E

(7.3)

to

B
D
E
A
C

is favorable to E in the sense of Definition 7.2. On the other hand, a change from (7.3) to

B
D
E
C
A

(7.4)

is not the sort of change that Definition 7.2 refers to, since in going from (7.3) to (7.4), E was not only moved upwards, but the ordering of A and C was reversed as well. ∎

Definition 7.4 (monotonic). *A winner selection method is called monotonic if a ballot change favorable to X can never turn X from a winner into a loser.*

If X can be turned from a winner into a loser by making changes favorable to X in *several* ballots, there must in fact be an example in which the same is accomplished even with a change in just one ballot. For if changes favorable to X in several ballots can turn X from a winner into a loser, those changes could be made one ballot at a time, and at some point, as a result of a change favorable to X made in one single ballot, X would have to turn from a winner into a loser. For a similar reason, if X can be turned from a winner into a loser by moving X up by several places on a single ballot, then in fact there must be an example in which X turns from a winner into a loser by being moved up by just one place on a single ballot, that is, by being swapped with a candidate Y who is immediately above X prior to the change, and immediately below X after the change:

$$\begin{array}{|c|}\hline \cdot \\ \hline \cdot \\ \hline Y \\ \hline X \\ \hline \cdot \\ \hline \cdot \\ \hline \end{array} \quad \rightarrow \quad \begin{array}{|c|}\hline \cdot \\ \hline \cdot \\ \hline X \\ \hline Y \\ \hline \cdot \\ \hline \cdot \\ \hline \end{array} \; . \tag{7.5}$$

We will now investigate which of the six methods listed at the beginning of this chapter are monotonic.

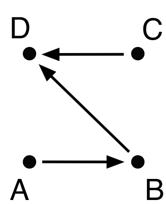
Proposition 7.5. *The a priori Smith-fair plurality method is not monotonic.*

Proof. Consider the preference schedule

3	2	1	1	1
A	D	C	C	B
B	C	D	A	D
C	A	A	B	A
D	B	B	D	C

. (7.6)

The pairwise comparison graph looks like this:



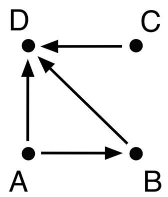
The Smith set is $S = \{A, B, C, D\}$, and the winner by the a priori Smith-fair plurality method (which is, in this example, the same as the plain plurality method) is A.

Suppose now that we replace the ballot $CDAB$ by $CADB$, a change that is clearly favorable to A. This turns (7.6) into

3	2	1	1	1
A	D	C	C	B
B	C	A	A	D
C	A	D	B	A
D	B	B	D	C

. (7.7)

After this change, the pairwise comparison graph looks like this:



Now $S = \{A, B, C\}$. After removal of the non-Smith candidate D, the preference schedule becomes

3	4	1
A	C	B
B	A	A
C	B	C

.

Now $W = \{C\}$. So a change that was favorable to A turned A from a winner into a loser. $\square$

Interestingly (and quite obviously), the plain plurality method *is* monotonic (see Exercise 7.3). Thus by making the plurality method a priori Smith-fair, an important flaw (its lack of Condorcet fairness) is removed, but another flaw (lack of monotonicity) is newly introduced.

Proposition 7.6. *The a priori Smith-fair runoff method is not monotonic.*

Proof. Although Example 7.1 showed that the *plain* runoff method is not monotonic, the same example also shows that the a priori Smith-fair runoff method is not monotonic (see Exercise 7.5). $\square$

Proposition 7.7. *The a priori Smith-fair elimination method is not monotonic.*

Proof. When there are only three candidates, the runoff and elimination methods are the same. Therefore Example 7.1 also shows that the elimination method and the a priori Smith-fair elimination method are not monotonic. $\square$

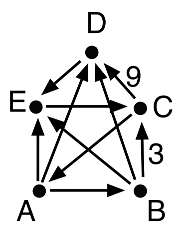
Proposition 7.8. *A priori Smith-fair Borda count is not monotonic.*

Proof. I will give a rather complicated example proving that a priori Smith-fair Borda count is not monotonic. The example involves 15 voters and 5 candidates. The preference schedule is

3	2	1	1	2	1	1	1	1	2
A	B	B	B	D	D	E	B	B	E
C	C	A	C	A	A	C	E	E	C
D	D	C	A	B	E	D	C	C	A
E	E	E	E	E	B	B	A	D	B
B	A	D	D	C	C	A	D	A	D

. (7.8)

The pairwise comparison graph looks like this:



All margins of victory are 1, with the exception of the two margins indicated in the figure: B beats C by a margin of 3, and C beats D by a margin of 9.

You will now ask, “Where in the world did this example come from?” One answer is, “From trial and error.” But I will try to give a better answer. Notice that the pairwise comparison graph consists of two circles: $A \to B \to C \to A$ and $C \to D \to E \to C$. Candidate C glues the two circles together. The stronger candidates are A, B, and C: They each beat D and E in head-to-head competition, with one important exception: C (narrowly) loses against E.

The Smith set is

$$S = \{A, B, C, D, E\}.$$

However, notice that D and E make it into the Smith set only “just barely”: The only reason E makes it in is that she beats C by a margin of 1, the slimmest possible margin. And the only reason why D makes it into the Smith set is that E is in the Smith set, and D beats E. The change we will introduce now is to move C, on a single ballot, from (immediately) below E to (immediately) above E, thereby reversing the arrow between E and C, causing E and, as a crucial side effect, D to drop out of the Smith set.

Notice that C beats D by a comparative landslide, namely, by a margin of 9. This is so by design: Because C is ahead of D in so many ballots, D is a rich source of Borda points for C as long as D remains on the ballots; see Exercise 7.6. With D on the ballots, C wins the Borda count. Without him, she loses. This is the idea, and I will verify now that in fact it works.

As we have noted, for the preference schedule (7.8), the Smith set is $S = \{A, B, C, D, E\}$. Therefore a priori Smith-fair Borda count is the same as plain Borda count in this example. The Borda scores are as follows;

$$A:46, \quad B:47, \quad C:48, \quad D:40, \quad E:44,$$

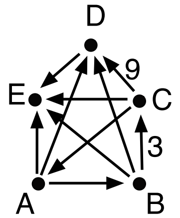
so C wins.

Now suppose that we replace one of the $DABEC$ ballots by $DABCE$, a change that is favorable to C. The new preference schedule is

3	2	1	1	1	1	1	1	1	1	2
A	B	B	B	D	D	D	E	B	B	E
C	C	A	C	A	A	A	C	E	E	C
D	D	C	A	B	B	E	D	C	C	A
E	E	E	E	E	C	B	B	A	D	B
B	A	D	D	C	E	C	A	D	A	D

.

Nothing in the pairwise comparison graph changes as a result, except for the arrow from E to C, which reverses:



As a result, D and E drop out of the Smith set:

$$S = \{A, B, C\}.$$

In a priori Smith-fair Borda count, one must now first remove D and E from the preference schedule:

3	2	1	1	1	1	1	1	1	1	2
A	B	B	B	A	A	A	C	B	B	C
C	C	A	C	B	B	B	B	C	C	A
B	A	C	A	C	C	C	A	A	A	B

,

or more compactly,

3	5	1	3	1	2
A	B	B	A	C	C
C	C	A	B	B	A
B	A	C	C	A	B

.

Now the Borda scores are as follows:

$$A:30, \quad B:31, \quad C:29.$$

Thus C has indeed turned from the (sole) winner into a loser as a result of a change in one ballot that was favorable to C. □

It is straightforward to prove that Borda count itself *is* monotonic (see Exercise 7.7). Thus, as in the case of the plurality method, restricting the competition to the Smith set a priori removes one flaw (lack of Condorcet-fairness), but introduces another (lack of monotonicity).

Proposition 7.9. *The method of pairwise comparison is monotonic.*

Proof. Suppose that $X \in W$, and that a change in the ballots is made that is favorable to X. Arrows in the pairwise comparison graph that do not involve X are left unaffected by such a change. Arrows pointing away from X may get strengthened, i.e., the margin of victory may rise, but they cannot disappear or reverse directions. Arrows pointing towards X may get weakened, i.e., the margin of victory may decrease, or they may disappear or reverse direction. Thus X may gain points in the pairwise comparison tournament as a result of a change, whereas any other candidate cannot gain, but can lose a point. Therefore if $X \in W$ before the change, then $X \in W$ after the change. □

Proposition 7.10. *The beatpath method is monotonic.*

Proof. Suppose that $X \in W$, and that a change in the ballots is made that is favorable to X. Arrows in the pairwise comparison graph that do not involve X are left unaffected by such a change. Arrows pointing away from X may get strengthened, i.e., the margin of victory may rise. Arrows pointing towards X may get weakened, i.e., the margin of victory may decrease, or they may disappear altogether or reverse direction. Thus as a result of the change, any beatpath that X has against any other candidate can only be strengthened (if it is affected at all), and any beatpath that any other candidate has against X can only be weakened or broken altogether. Since $X \in W$ before the change, there is no unmatched beatpath against X before the change. There can then certainly not be any unmatched beatpath against X after the change, so $X \in W$ still. □

In summary, of the six methods listed at the beginning of this chapter, only two are monotonic: pairwise comparison and the beatpath method. In the remainder of this chapter, we prove that the nonmonotonicity of a priori Smith-fair Borda count is only a very minor flaw, namely, that it is visible only in examples in which the Smith set has at least five members.

We begin with a lemma that describes how the Smith set itself may change as a result of a change in a ballot that is favorable to X.

Lemma 7.11. *Suppose that for a given preference schedule, a candidate X belongs to the Smith set S. Suppose that on a single ballot, X's position is swapped with that of some other candidate Y who is, prior to the swap, ranked immediately above X on the ballot; see* (7.5). *Let S' denote the Smith set after the change. Then either*

(i) *$S' = S$, or*

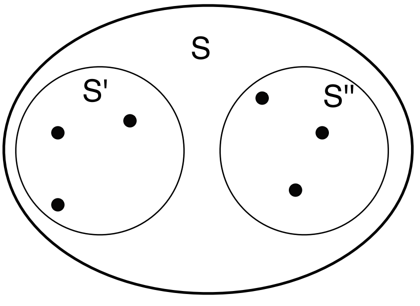
(ii) *$S' \subseteq S$ but $S' \neq S$, $X \in S'$, and $Y \in S - S'$.*

Proof. The only head-to-head competition affected by the change is that between X and Y. If prior to the change there was a tie between X and Y, then after the change, X may beat Y in head-to-head competition by a margin of 2, the smallest possible margin of victory if the number N of voters is even. If prior to the change Y beat X in head-to-head competition, then after the change, there may be a tie (if N is even), or X may beat Y by a margin of 1 (if N is odd).

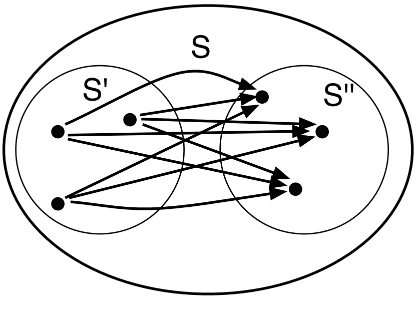
We will prove now that S is a dominating set even after the change. Suppose $A \in S$ and $B \notin S$. To prove that S is dominating after the change, we must prove that after the change, A beats B in head-to-head competition. Note that $B \neq X$, since $B \notin S$ but $X \in S$. If also $A \neq X$, then the head-to-head competition between A and B is not affected by the change; since A beats B before the change, so it does after the change. Suppose now that $A = X$. Since X beats B in head-to-head competition prior to the change, and since the change is favorable to X, X beats B in head-to-head competition after the change.

Thus S is indeed dominating even after the change. Since the Smith set is contained in any dominating set, this implies $S' \subseteq S$.

Now suppose that $S' \neq S$, and write $S'' = S - S'$:



After the change, S' is the Smith set, hence there are arrows in the pairwise comparison graph from every candidate in S' to every candidate in S'':



Prior to the change, not all of these arrows can have been present, otherwise S' would have been dominating even prior to the change, and therefore S could not have been the Smith set. But the only arrow that can have changed is between X and Y. Therefore $X \in S'$ and $Y \in S''$. $\square$

Corollary 7.12. *The Smith method is monotonic.*

Proof. This follows immediately from Lemma 7.11. □

However, the Smith method is not among the six that we focus on in this chapter, since it violates the unanimity criterion.

We will now think about how complicated an example proving Proposition 7.8 has to be. The strategy is to try to find as small an example as possible.

Recall first the comment following Definition 7.4: If a winner selection method is not monotonic, then the lack of monotonicity can always be demonstrated by an example in which X turns from a winner into a loser as a result of exchanging, in a single ballot, X with the candidate Y immediately above X (see (7.5)). We will now think in general about examples in which X turns from a winner into a loser in a priori Smith-fair Borda count as a result of exchanging, on a single ballot, X with a candidate Y who is immediately above X prior to the change, and immediately below X after the change. One such example was given in the proof of Proposition 7.8; there $X = C$ and $Y = E$.

It is clear that the ballot change must alter the Smith set for the example to work: If it did not, then X could only gain Borda points as a result of the change, Y could only lose, and no other candidate's Borda scores be affected, so X could not possibly turn from a winner into a loser.

By Lemma 7.11, some candidates drop out of the Smith set as a result of the change, but no new ones join the Smith set. We will use the same notation as in the proof of Lemma 7.11: S is the Smith set prior to the change, S' is the Smith set after the change, and $S'' = S - S'$ is the set of candidates who drop out of the Smith set as a result of the change. By Lemma 7.11, $X \in S'$ and $Y \in S''$. In the example in the proof of Proposition 7.8, $S' = \{A, B, C\}$ and $S'' = \{D, E\}$.

We will now argue that there is no *very small* example in which swapping X, on a single ballot, with the candidate Y immediately above X turns X from a winner into a loser by a priori Smith-fair Borda count. Specifically, we will prove that in any such example, S'' must contain at least two candidates (D and E in the example in the proof of Proposition 7.8), and S' must contain at least three (A, B, and C in the example).

First, we prove that Y cannot be the only candidate who drops out of the Smith set, i.e., that S'' must contain at least two candidates. To see this, suppose that in fact Y were the only candidate in S''. As a result of Y's departure, the remaining Smith candidates, i.e., the candidates in S', would get fewer points in the a priori Smith-fair Borda count than before. Each candidate would lose one Borda point for each ballot on which he or she ranks above Y; see Exercise 7.6. In other words, the better a candidate does in head-to-head competition against Y, the more Borda points he or she would lose as a result of Y's removal. But after the change, X beats Y just barely, with the smallest possible margin. Therefore the departure of Y from the Smith set would hurt X's Borda score minimally, no more than the Borda scores of any of the other candidates in S'. Since X wins prior to the change, X would still have to win after the change. But our assumption is that the change turns X into a loser. This contradiction proves that S'' must contain at least two candidates.

Second, we will argue that S' must contain at least three candidates. Again, the proof is indirect. Suppose first that S' contained only a single candidate. That candidate would have to be X, by Lemma 7.11. Then X would be the winner even after the change, in contradiction with our assumption that the change turns X into a loser. Suppose next that S' contained exactly two candidates. In head-to-head competition, these two candidates would have to tie, otherwise the winner would be a Condorcet candidate and the sole member of the Smith set after the change. But the two candidates would then also tie for victory after the change, in the a priori Smith-fair Borda count, again in contradiction with our assumption that the change turns X into a loser.

We have now concluded that S' must contain at least three candidates, and S'' must contain at least two. Thus S must contain at least five candidates.

We have thereby proved the following proposition.

Proposition 7.13. *A priori Smith-fair Borda count is monotonic if there are fewer than five Smith candidates.*

In particular, a priori Smith-fair Borda count is monotonic when the total number n of candidates is smaller than 5.

Exercises

7.1. Consider the following preference schedule:

9	6	2	7
A	C	C	B
C	B	A	A
B	A	B	C

(a) Who wins by the elimination method? (b) If the two voters who vote

C
A
B

change their minds and instead vote

A
C
B

a move that clearly seems to be in A's favor, then who wins by the elimination method?

7.2. Consider the preference schedule

3	2	1	1
D	E	F	F
A	B	C	E
B	C	A	C
C	A	B	B
E	F	E	A
F	D	D	D

(a) By a priori Smith-fair Borda count, who wins? (b) By the beatpath method, who wins?

7.3. Prove that the plurality method is monotonic.

7.4. ($*$) Prove that the a priori Smith-fair plurality method is monotonic if there are fewer than four Smith candidates.

7.5. Explain why Example 7.1 proves not only that the *plain* runoff method is not monotonic, but also that the a priori Smith-fair runoff method is not monotonic.

7.6. This exercise clarifies a point that plays a role in the proof of Proposition 7.8. Let Y and Z be two candidates in an election, and let v denote the number of voters who prefer Z to Y. Explain: If Y is removed from the ballots, the Borda score of Z drops by v. (Thus the better Z does against Y in head-to-head competition, the more her Borda score decreases when Y is dropped from the ballots.)

7.7. Prove that Borda count is monotonic.

Chapter 8

Elections with Many or Few Voters

When the number N of voters is very large, one hopes that the probability of a precise tie in the election outcome is very small. This is indeed the case for most of the winner selection methods that we have thought about, for instance, for the simplest among them, the plurality method. However, for some methods, for instance, for the method of pairwise comparison, the likelihood of a tie is substantial even for large N. The following example shows a simple special case.

Example 8.1. We assume here that the number n of candidates is 3, and the number N of voters is odd, so that exact ties in head-to-head competitions are impossible. We call the three candidates A, B, and C. Suppose that the preference schedule is

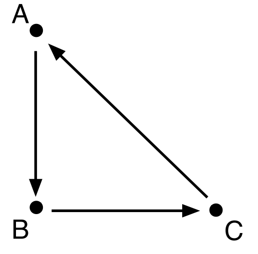
N_1	N_2	N_3	N_4	N_5	N_6
A	A	B	B	C	C
B	C	A	C	A	B
C	B	C	A	B	A

,

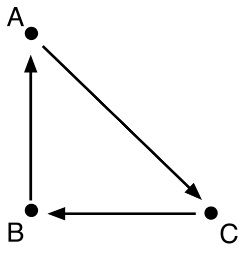
where of course we must assume

$$N_1 + N_2 + N_3 + N_4 + N_5 + N_6 = N. \tag{8.1}$$

For which values of N_1, $N_2, \ldots$ does the method of pairwise comparison result in a tie? There is a tie if and only if there is a circle in the societal preferences, and there are two possible circles here (see Exercise 8.2). We will refer to the first possible circle,



as "circle 1," and to the second,



as "circle 2."

Circle 1 is the pairwise comparison graph if and only if

$$\begin{aligned} N_1+N_2+N_5 &> N_3+N_4+N_6 \quad \text{and} \\ N_1+N_3+N_4 &> N_2+N_5+N_6 \quad \text{and} \\ N_4+N_5+N_6 &> N_1+N_2+N_3. \end{aligned} \tag{8.2}$$

Similarly, circle 2 is the pairwise comparison graph if and only if

$$\begin{aligned} N_1+N_2+N_5 &< N_3+N_4+N_6 \quad \text{and} \\ N_1+N_3+N_4 &< N_2+N_5+N_6 \quad \text{and} \\ N_4+N_5+N_6 &< N_1+N_2+N_3. \end{aligned} \tag{8.3}$$

The fraction of all possible choices of $N_1, N_2, \ldots$ satisfying (8.1) that also satisfy (8.2) or (8.3) is substantial and does not decrease to zero with increasing N. As a result, the likelihood of a tie in the pairwise comparison tournament is substantial no matter how large N is.

On the other hand, the winner in the beatpath method depends not only on the pairwise comparison graph, but also on the margins of victory. The beatpath method cannot lead to a tie unless (1) there is a circle in the pairwise comparison graph, and (2) at least two of the three margins of victory equal each other (see Exercise 8.3). For large N, it is very unlikely that any two of the three margins of victory are equal to each other.

Similarly, ties in a priori Smith-fair Borda count can occur only if (1) there is a circle, and (2) at least two of the three Borda scores are equal to each other. Again, this is very unlikely for large N. ∎

Here is a way in which one might say this precisely. Suppose that the voters' rankings are random choices from the $n!$ possible rankings of the n candidates. Suppose that different voters' rankings are chosen with the same probability distribution μ, independently of each other. The distribution μ certainly need not be uniform, but we assume that each of the $n!$ possible rankings has a positive probability. Fixing n and μ, we denote by P_N the probability of a tie when N is the number of voters. Then

$$\lim_{N\to\infty} P_N = 0$$

for the beatpath method and a priori Smith-fair Borda count, but not for pairwise comparison.

It is also interesting to think about cases when the number N of voters is very small, and check whether the various winner selection methods yield results that seem intuitively reasonable.

Example 8.2. The smallest interesting case is that of $N = 2$, with $n = 3$ candidates, A, B, and C; think of a couple choosing among three restaurants. In this case, we can easily list all possible (detailed) preference schedules. Each voter can put any candidate first (there are three options), then any candidate second (there are two options left after the voter has decided whom to put first), and the remaining candidate will be third in the voter's ranking. So each voter gets to choose from $3 \times 2 = 6$ different ways of ranking the candidates. Since they get to make their choices independently of each other, there are $6 \times 6 = 36$ different preference schedules. This still may seem like a painfully large number. However, notice that we might as well assume that the first voter's ranking is

$$\begin{array}{|c|}\hline A \\ \hline B \\ \hline C \\ \hline \end{array}\,.$$

If it is not, we simply rename the candidates, calling the candidate whom the first voter places first "A", the one whom she places second "B", and the one whom she places last "C". With this convention, there are only six different possible (detailed) preference schedules:

$$(1)\ \begin{array}{|c|c|}\hline A & A \\ \hline B & B \\ \hline C & C \\ \hline \end{array}\,, \qquad (2)\ \begin{array}{|c|c|}\hline A & A \\ \hline B & C \\ \hline C & B \\ \hline \end{array}\,, \qquad (3)\ \begin{array}{|c|c|}\hline A & B \\ \hline B & A \\ \hline C & C \\ \hline \end{array}\,,$$

$$(4)\ \begin{array}{|c|c|}\hline A & B \\ \hline B & C \\ \hline C & A \\ \hline \end{array}\,, \qquad (5)\ \begin{array}{|c|c|}\hline A & C \\ \hline B & A \\ \hline C & B \\ \hline \end{array}\,, \qquad (6)\ \begin{array}{|c|c|}\hline A & C \\ \hline B & B \\ \hline C & A \\ \hline \end{array}\,.$$

We can now determine, for each of our winner selection methods, whom it would make the winner in each of these six cases. The following table lists the answers for three winner selection methods.

	a priori Smith-fair Borda count	beatpath method	pairwise comparison
(1)	A	A	A
(2)	A	A	A
(3)	A, B	A, B	A, B
(4)	B	A, B	B
(5)	A	A, C	A
(6)	A, B, C	A, B, C	A, B, C

The only cases in which the three methods in this table do not all agree on the set of winners are (4) and (5). But it seems eminently reasonable to make B the winner for preference schedule (4), and A the winner for (5). Thus in the smallest possible cases, a priori Smith-fair Borda count and the method of pairwise comparison seem more in line with our intuition than the beatpath method. ∎

Our search for good winner selection methods comes to an end here. Of all the methods we have considered, the beatpath method and a priori Smith-fair Borda count are the only ones that satisfy the following five criteria:

a. the principle of one person, one vote;

b. the principle of independence of candidate names;

c. the no-weak-spoiler criterion (which is equivalent to a priori Smith-fairness, and implies Smith-fairness and therefore Condorcet-fairness);

d. the unanimity criterion;

f. the requirement that precise ties be very unlikely when the number of voters is large.

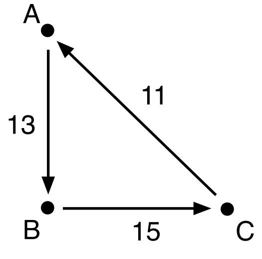
I have deliberately left out

e. monotonicity.

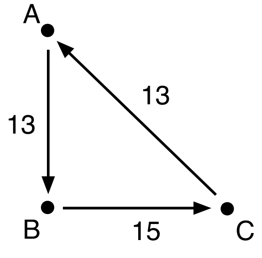
If this requirement is added, then the beatpath method becomes the only one, among the winner selection methods that we have studied, satisfying all criteria on the list. However, it seems very unclear whether the lack of monotonicity is really an important flaw of a priori Smith-fair Borda count. The example that we gave in Chapter 7 to prove the nonmonotonicity of a priori Smith-fair Borda count is quite contrived, and the method *is* in fact monotonic as long as there are fewer than 5 Smith candidates (Proposition 7.13). A priori Smith-fair Borda count has the advantage of being considerably simpler than the beatpath method, and as we have seen, when $N = 2$ and $n = 3$, it sometimes yields more plausible results than the beatpath method.

Exercises

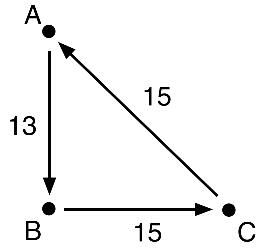
8.1. Consider an election with three candidates, A, B, and C. (a) If the pairwise comparison graph is



who wins by the beatpath method? (b) If the pairwise comparison graph is



who wins by the beatpath method? (c) If the pairwise comparison graph is



who wins by the beatpath method?

8.2. This exercise refers to a point in Example 8.1. Assume that there are three candidates, named A, B, and C, and N voters, where N is an odd number. Explain: The method of pairwise comparison leads to a tie if and only if there is a circle in societal preferences. The only two possible circles are $A \to B \to C \to A$ and $A \to C \to B \to A$. (See the two pairwise comparison graphs drawn in Example 8.1.)

8.3. This exercise refers to another point in Example 8.1. It is helpful to think about Exercise 8.1 before doing this exercise. Assume that there are three candidates, named A, B, and C, and N voters, where N is an odd number. Explain: The beatpath method cannot lead to a tie unless (1) there is a circle in societal preferences, and (2) at least two of the three margins of victory equal each other.

8.4. What does the plurality method do in the six cases of Example 8.2?

8.5. ($*$) In this chapter, we discussed the smallest interesting elections, involving $N = 2$ voters and $n = 3$ candidates. Here we will turn to the case of $N = 3$ voters, still with $n = 3$ candidates. (a) Explain: If $N = 3$ and $n = 3$, and if there is no Condorcet candidate, then up to renaming candidates and renumbering voters, the (detailed) preference schedule cannot be different from

A	B	C
B	C	A
C	A	B

.

(b) From part (a), conclude: When there are three voters and three candidates, all Condorcet-fair winner selection methods obeying the principles of one person, one vote and independence of candidate names yield the same results in all cases.

Chapter 9

Irrelevant Comparisons and the Muller–Satterthwaite Theorem

We started out, in Chapter 1, with a list of winner selection methods. By imposing more and more requirements, we narrowed the list, until only the beatpath method and a priori Smith-fair Borda count were left. There are, of course, further requirements that one might want to impose on winner selection methods. In Chapters 9 and 10, we study two examples of requirements that *no* reasonable winner selection method can satisfy.

In the present chapter, our focus is on what we will call "irrelevant comparisons." The following example illustrates the idea.

Example 9.1. Three people vote to make a choice among three possible restaurants—Afghan, Chinese, or Italian. The first person's favorite kind of food is Chinese. The second person prefers Afghan over Chinese food, and Chinese over Italian food. The third prefers Italian over Chinese food, and Chinese over Afghan food. Is this enough information to determine whether or not the Chinese restaurant wins, assuming that the plurality method is used?

The first person places the Chinese restaurant first, the second places the Afghan restaurant first, and the third places the Italian restaurant first. If we use the plurality method, there is a three-way tie. We would say that all three restaurants "win" (although that doesn't help the three people decide where to eat). So yes, the given information is, in this case, enough to determine that the Chinese restaurant is among the winners. ∎

In general, if all we want to know is whether or not a candidate X wins, then one might think at first that it should be enough to know how each voter ranks X in comparison with each other candidate. After all, X is among the winners if and only if society considers X as good as, or better than, all other candidates. Why should that depend on anything other than how people rank X in comparison with others? Comparisons among other candidates should be irrelevant, shouldn't they? But a moment's thought proves that this reasoning is wrong, as shown in the following example.

Example 9.2. Modify Example 9.1 by assuming that the third person's least favorite food is Chinese, with the preferences of the others unchanged. Now whether or not the Chinese restaurant is among the winners, according to the plurality method, will depend on whether the third person ranks Italian over Afghan food (in which case each restaurant gets one first-place vote), or whether she ranks Afghan over Italian food (in which case the Afghan restaurant is the sole winner, with two first-place votes out of three). ∎

If the winner selection method is dictatorship, of course, all we need to know is whether or not the dictator prefers X to each other candidate. So in this case, comparisons among other candidates are indeed irrelevant.

If one can determine whether or not X wins based on nothing other than knowledge of how each voter compares X with each of the other candidates, then we say that the winner selection method satisfies the independence-of-irrelevant-comparisons (IIC) criterion.

Definition 9.3 (independence of irrelevant comparisons, or IIC). *A winner selection method satisfies the independence-of-irrelevant-comparisons (IIC) criterion*[3] *if, for each candidate X, it is possible to determine whether or not X is among the winners if one merely knows, for each voter and each candidate Y different from X, whether the voter ranks X above or below Y.*

Example 9.4. In an election with 7 voters and 3 candidates, the preference schedule is

3	2	2
A	B	C
C	A	B
B	C	A

. (9.1)

By a priori Smith-fair Borda count, which is the same here as plain Borda count since $S = \{A, B, C\}$, A is the sole winner. But if the preference schedule were

3	2	2
A	B	B
C	A	C
B	C	A

, (9.2)

then B would be a Condorcet candidate and therefore the sole winner by a priori Smith-fair Borda count. Note that every voter ranks A, in comparison with the others, in the same way in (9.1) and (9.2). To determine whether A wins, it is therefore not sufficient to know how every voter ranks A in comparison with each other candidate: A priori Smith-fair Borda count violates the IIC criterion. ∎

Preference schedules (9.1) and (9.2) can be used to demonstrate violations of the IIC criterion for other methods that we haves studied as well; see Exercise 9.1. In fact, the central result of this chapter, Theorem 9.7, tells us that the IIC criterion is incompatible with other desirable properties discussed earlier. To understand the statement of this theorem, one has to understand the notion of Pareto-efficiency, named after the Italian economist Vilfredo Pareto (1848–1923).

Definition 9.5 (Pareto-efficiency). *A winner selection method is Pareto-efficient if it guarantees $W = \{X\}$ if X is ranked first by every voter.*

This is somewhat reminiscent of, but much weaker than, the unanimity criterion stated in Definition 2.6, which I will repeat here for the reader's convenience.

[3] Instead of "independence of irrelevant comparisons," one should perhaps say, more accurately, "independence of comparisons which perhaps at first sight look as though they might be irrelevant."

Definition 2.6 (unanimity criterion). *A winner selection method satisfies the unanimity criterion if it guarantees that candidate Y is not among the winners of the election if there is a candidate X who is preferred to Y by* every *voter.*

Example 9.6. Suppose that the preference schedule is

3	5	1
B	C	B
C	B	A
A	A	C
D	E	D
E	D	E

.

(i) Suppose we know that our winner selection method is Pareto-efficient, and nothing more. We can then conclude nothing at all, since there is no candidate who is placed first on every ballot.

(ii) Suppose we know that our winner selection method satisfies the unanimity criterion. We can then conclude that A, D, and E lose, since they are ranked below B by every voter. ∎

Any winner selection method that satisfies the unanimity criterion is Pareto-efficient (see Exercise 9.2). However, the Smith method and the method of sequential comparison, which violate the unanimity criterion, are still Pareto-efficient (see Exercise 9.3). A method that would not be Pareto-efficient would have to be very strange indeed.

Theorem 9.7 refers to *single-winner methods*, that is, winner selection methods that never lead to any ties in the outcomes. Any winner selection method can be turned into a single-winner method by making the convention that in the event of a tie, the candidate in the set W who comes first in the alphabet will be made the sole winner; for instance, if the winner selection method yields $W = \{B, D, E\}$, then the single-winner method will say $W = \{B\}$. This seems like an excellent convention if your name is Aaronson, but it seems terribly unfair if your name is Zysman. However, if A, B, $C, \ldots$ are not the candidates' real names, but labels assigned to them at random for the purpose of evaluating the election outcome, then resolving ties alphabetically amounts to nothing other than resolving them at random. That, of course, is what is typically done, and should be done, in the event of a tie.

Theorem 9.7 (E. Muller and M. A. Satterthwaite, 1977). *When there are more than two candidates, the only Pareto-efficient, monotonic single-winner method satisfying the IIC criterion is dictatorship of the kth voter for some k between 1 and N (the number of voters).*[4]

Proof. It is easy to verify that dictatorship of the kth voter is a Pareto-efficient single-winner method satisfying the monotonicity and IIC criteria (see Exercise 9.4). Assume, conversely, that we are given a Pareto-efficient, monotonic single-winner method that satisfies the IIC

[4] This is equivalent to "Theorem A" of Philip J. Reny, "Arrow's theorem and the Gibbard–Satterthwaite theorem: A unified approach," *Economics Letters* 70, issue 1, pages 99–105 (2001). The proof given in this chapter is taken from this paper.

criterion. We want to prove that the method is in fact dictatorship of the kth voter for some k.

As a preliminary step, we prove that the method must satisfy the unanimity criterion. Suppose it did not satisfy the unanimity criterion; thus suppose there were a preference schedule in which each voter ranked X above Y, but Y were nevertheless the winner. By the IIC criterion, Y would still be the winner if X were moved to first place on every single ballot, while preserving the rankings of Y relative to any candidate on any of the ballots. But since X would then be in first place on every single ballot, X would have to be the winner by Pareto-efficiency, so Y would *not* be the winner anymore. This contradiction proves our assertion that the method must satisfy the unanimity criterion.

Now pick an arbitrary candidate, let's say A. (The argument can be done with any other candidate in place of A, and this point will be important later.) Let us think about a situation in which we know what must happen: Suppose A is placed first by every voter. By Pareto-efficiency, A is then the winner. We indicate this as follows:

1	...	$k-1$	k	$k+1$	...	N
A	...	A	A	A	...	A
.	...	.	.	.	...	.
.	...	.	.	.	...	.
.	...	.	.	.	...	.
.	...	.	.	.	...	.
.	...	.	.	.	...	.

$$\rightarrow \quad A\,. \tag{9.3}$$

This is a *detailed* preference schedule, not a reduced one. Each column refers to a single voter. The numbers at the top of the columns simply label voters.

Of course, A *should* be the winner in (9.3). He is extremely popular, every voter places him first! The strategy is now to change the preference schedule gradually, making A look worse and worse, yet still ensuring that, by the assumptions of Theorem 9.7, he must remain the winner. In the end, A will be utterly unpopular, yet still the winner, and the only possible explanation will be that the winner selection method is dictatorship.

To make A look less popular, let's bring a specific competitor into play. It does not matter which candidate we choose. Let's call her B. As long as A is in first place on each ballot, A wins, since the method is Pareto-efficient. For instance,

1	...	$k-1$	k	$k+1$	...	N
A	...	A	A	A	...	A
.	...	.	.	.	...	.
.	...	.	.	.	...	.
.	...	.	.	.	...	.
.	...	.	.	.	...	.
B	...	B	B	B	...	B

$$\rightarrow \quad A\,. \tag{9.4}$$

Now let's make B look better by moving her up on the first ballot. A remains the winner as long as B does not reach the top position, by the IIC criterion (and also by Pareto-efficiency). Even when B reaches the second position in the first ballot, A is still the winner:

1	2	...	$k-1$	k	$k+1$	...	N
A	A	...	A	A	A	...	A
B	.	...	.	.	.	...	.
.	.	...	.	.	.	...	.
.	.	...	.	.	.	...	.
.	.	...	.	.	.	...	.
.	B	...	B	B	B	...	B

$\rightarrow \quad A\,.$ (9.5)

The moment, however, when B rises above A on the first ballot, it is unclear what will happen:

1	2	...	$k-1$	k	$k+1$	...	N
B	A	...	A	A	A	...	A
A	.	...	.	.	.	...	.
.	.	...	.	.	.	...	.
.	.	...	.	.	.	...	.
.	.	...	.	.	.	...	.
.	B	...	B	B	B	...	B

$\rightarrow \quad ?\,.$ (9.6)

You may be tempted to say, "A is still in first place on all ballots except the first, and in second place on the first ballot, so A should still be the winner." This will of course be true for most single-winner methods, but not for all. If the single-winner method is dictatorship of the first voter, for instance, then going from (9.5) to (9.6) changes the winner from A to B.

The winner in (9.6) is, in any case, either A or B. To see this, suppose that some third candidate C were the winner in (9.6). Then, by IIC, C would have to be the winner even in (9.5). But in (9.5), we know the winner is A, not C. So we have concluded the following:

1	2	...	$k-1$	k	$k+1$	...	N
B	A	...	A	A	A	...	A
A	.	...	.	.	.	...	.
.	.	...	.	.	.	...	.
.	.	...	.	.	.	...	.
.	.	...	.	.	.	...	.
.	B	...	B	B	B	...	B

$\rightarrow \quad A$ or $B\,.$ (9.7)

Suppose the winner is still A. Then we now move B upward on the second ballot. As long as B does not reach the top in the second ballot, A remains the winner by IIC. When B reaches the top, and A falls to second place on the second ballot, either A remains the winner or B becomes the winner, by the argument we just gave. If A remains the winner, we proceed to the third ballot and let B rise to the top there. We continue in this way until, at some point, the winner changes from A to B when B is moved to the top of one of the ballot. This *must* happen eventually, for if we move B into the top position in *all* ballots, then B is the winner, since the method is Pareto-efficient. Let's say that B becomes the winner the moment she is moved into the top position in the kth ballot. So

1	...	$k-1$	k	$k+1$	...	N
B	...	B	A	A	...	A
A	...	A	B	.	...	.
.	...	.	.	.	...	.
.	...	.	.	.	...	.
.	...	.	.	.	...	.
.	...	.	.	B	...	B

$\rightarrow \quad A\,,$ (9.8)

but

1	...	$k-1$	k	$k+1$	...	N
B	...	B	B	A	...	A
A	...	A	A	.	...	.
.	...	.	.	.	...	.
.	...	.	.	.	...	.
.	...	.	.	.	...	.
.	...	.	.	B	...	B

$\rightarrow \quad B\,.$ (9.9)

In this example, therefore, voter k plays a pivotal role: As long as he places A above B, the overall winner is A, but as soon as he places B above A, the overall winner is B. Not much seems surprising so far. You might think, for instance, that k is probably approximately $N/2$—the winner changes from A to B as soon as B, not A, has the majority.

However, we can make these examples more surprising. To start with, we will prove that in (9.8), we can move A to the bottom of ballots 1 through $k-1$, and A will still be the winner overall:

1	...	$k-1$	k	$k+1$	...	N
B	...	B	A	.	...	.
.	...	.	B	.	...	.
.	...	.	.	.	...	.
.	...	.	.	.	...	.
.	...	.	.	A	...	A
A	...	A	.	B	...	B

$\rightarrow \quad A\,.$ (9.10)

To see this, you have to make two observations. First, the winner in (9.10) is certainly *not* B. For if the winner were B in (9.10), then the winner would have to be B in (9.8), by IIC; but we know that in (9.8), the winner is A. Second, the winner in (9.10) cannot be some third candidate C either. For if

1	...	$k-1$	k	$k+1$	...	N
B	...	B	A	.	...	.
.	...	.	B	.	...	.
.	...	.	.	.	...	.
.	...	.	.	.	...	.
.	...	.	.	A	...	A
A	...	A	.	B	...	B

$\rightarrow \quad C\,,$ (9.11)

then

1	...	$k-1$	k	$k+1$	...	N
B	...	B	B	.	...	.
.	...	.	A	.	...	.
.	...	.	.	.	...	.
.	...	.	.	.	...	.
.	...	.	.	A	...	A
A	...	A	.	B	...	B

$\rightarrow \quad C$ (9.12)

by IIC, but (9.9) implies

1	...	$k-1$	k	$k+1$	...	N
B	...	B	B	.	...	.
.	...	.	A	.	...	.
.	...	.	.	.	...	.
.	...	.	.	.	...	.
.	...	.	.	A	...	A
A	...	A	.	B	...	B

$\rightarrow \quad B$, (9.13)

again by IIC.

We have now proved (9.10), and that is certainly a rather surprising conclusion—A is at or near the bottom on all ballots except the kth, yet A is still the winner! It would become truly astonishing if we could move A all the way to the bottom in ballots $k+1$ through N. However, it is not clear that we can do that:

1	...	$k-1$	k	$k+1$	...	N
B	...	B	A	.	...	.
.	...	.	B	.	...	.
.	...	.	.	.	...	.
.	...	.	.	.	...	.
.	...	.	.	B	...	B
A	...	A	.	A	...	A

$\rightarrow \quad ?$. (9.14)

One thing is clear, though: In (9.14), the winner cannot be anybody other than A or B. For if a third candidate C were the winner, then C would have to be the winner even in (9.10), by IIC; but in (9.10), the winner is A. So we would like to find an argument that rules out that B is the winner in (9.14).

We do this by introducing a third candidate, C. (Here the argument uses the assumption that there are at least three candidates.) We insert C into (9.10) in a way that, by IIC, does not alter the fact that A is the winner:

1	...	$k-1$	k	$k+1$	...	N
C	...	C	A	.	...	.
B	...	B	C	.	...	.
.	...	.	B	.	...	.
.	...	.	.	C	...	C
.	...	.	.	A	...	A
A	...	A	.	B	...	B

$\rightarrow \quad A$. (9.15)

If in this preference schedule we swap the positions of A and B in the $(k+1)$st through Nth ballots, the winner is still A or B, as argued earlier, but now the winner *cannot be* B anymore, by the unanimity criterion, since C is ahead of B on every single ballot. So

1	...	$k-1$	k	$k+1$	...	N
C	...	C	A	.	...	.
B	...	B	C	.	...	.
.	...	.	B	.	...	.
.	...	.	.	C	...	C
.	...	.	.	B	...	B
A	...	A	.	A	...	A

$$\rightarrow \quad A\,. \tag{9.16}$$

So now we have an example of a preference schedule where A is ranked *last* by every voter except voter number k, who ranks A first, and A wins anyway!

Any preference schedule in which voter k ranks A first can be obtained from (9.16) by first reordering (if necessary) the candidates other than A, while keeping A in its position, then moving A upwards in ballots 1 through $k-1$ and $k+1$ through N (if necessary). By IIC and monotonicity, this does not turn A from a winner into a loser. So whenever voter k ranks A first, A is the winner. We say that voter k is the *A-dictator*. An *A-dictator* is a voter who does not necessarily determine who wins the election, *unless* she puts A first—but in that case, A wins the election.

Now recall that A was chosen arbitrarily at the beginning. So we have proved that for *any* candidate X, there is an X-dictator. If X and Y are two different candidates, could the X-dictator and the Y-dictator be two different voters? If so, what would happen if the X-dictator placed X first on her ballot, and the Y-dictator placed Y first on his? Both X and Y would have to be winners, but we are discussing *single-winner* methods here. So in fact the X- and Y-dictators must be the same.

This proves that there is a *single* dictator whose top choice is the winner of the election. □

In the literature on voting, the word "monotonic" is often used to denote winner selection methods that are monotonic in the sense of Definition 7.4 *and* satisfy the IIC criterion. I will use the expression *strongly monotonic* for such methods.

Definition 9.8 (strongly monotonic). *A winner selection method is called strongly monotonic if it is monotonic and satisfies the IIC criterion.*

Proposition 9.9. *A winner selection method is strongly monotonic if and only if it has the following property. If X is among the winners, and if the ballots are changed in such a way that on each ballot, every candidate who ranked below X prior to the changes still ranks below X after the changes, then X is still among the winners after the changes.*

Proof. See Exercise 9.6. □

Using the notion of strong monotonicity, we can then briefly state the Muller–Satterthwaite theorem as follows.

Theorem 9.7′ (Muller–Satterthwaite theorem, second version). *When there are more than two candidates, dictatorship*[5] *is the only Pareto-efficient, strongly monotonic single-winner method.*

Thus strong monotonicity is not a goal that can be achieved by any acceptable single-winner method. It doesn't seem to be so clear that we should be very disappointed (or surprised, for that matter) to learn this. However, the Muller–Satterthwaite theorem almost immediately implies the Gibbard–Satterthwaite theorem, an astonishing statement that will be the subject of Chapter 10.

Exercises

9.1. Use the preference schedules (9.1) and (9.2) in Example 9.4 to prove that the following winner selection methods violate the IIC criterion: (a) pairwise comparison, (b) the beatpath method.

9.2. Explain: A winner selection method that satisfies the unanimity criterion is Pareto-efficient.

9.3. Prove that any Condorcet-fair method is Pareto-efficient. (For example, the Smith method and the method of sequential comparison are Condorcet-fair and therefore Pareto-efficient, even though they violate the unanimity criterion.)

9.4. Explain why dictatorship is Pareto-efficient, satisfies the monotonicity criterion, and satisfies the IIC criterion.

9.5. Give an example demonstrating that the Muller–Satterthwaite theorem does not hold when there are only two candidates.

9.6. Prove Proposition 9.9.

[5] I use the singular *dictatorship* here, even though there are of course N different dictatorial single-winner methods, one for each voter.

Chapter 10

Strategic Voting and the Gibbard–Satterthwaite Theorem

In the United States, supporters of the Green Party would often (not, of course, always) rather see a Democrat win than a Republican. Realizing that the Green Party candidate is unlikely to win, say, the presidency, they sometimes cast their vote for the Democrat, even if that does not reflect their true first choice. For similar reasons, Libertarians sometimes vote Republican. It would be nice to design a winner selection method that would guarantee that such strategizing could never be beneficial, thereby assuring voters that they need not strategize, but can simply vote according to their true judgment. Unfortunately, no reasonable winner selection method can give such a guarantee. This is the content of the Gibbard–Satterthwaite theorem, which we will prove in this chapter.

I will say that a voter casts a *dishonest* vote if he casts a vote not reflecting his true preference. I don't mean to suggest that voters who cast votes not reflecting their true preferences do something that they ought not do. I simply use the word *dishonest* for brevity, without meaning to imply a moral judgment.

Sometimes a voter can change the outcome of an election by voting dishonestly. Suppose that as a result of one voter's dishonest vote, the winner of the election changes from X to Y. If the voter who changed the outcome of the election by casting a dishonest vote actually, according to his true preferences, preferred Y to X, then we say that he has voted *strategically*. Thus, in the terminology of this chapter, *strategic voting* means *successful* dishonest voting—it's called *strategic* only if it actually causes the winner to change, and the change is one that the dishonest voter likes to see.

Definition 10.1 (strategy-proof). *A single-winner method is called strategy-proof if it makes strategic voting impossible in all cases.*

Let us think about the simplest of all winner selection methods, plurality voting, from this point of view. To make it a single-winner method, we make the convention that in the event of a tie, the candidate who comes first in the alphabet wins. Suppose that the preference schedule is as follows:

4	3	2
B	A	C
A	B	A
C	C	B

.

The winner is B. Two voters like C the best, and B the worst. Suppose one of them decided—dishonestly—to rank A first and C second. Then the preference schedule would become

4	3	1	1
B	A	C	A
A	B	A	C
C	C	B	B

.

There is now a tie between A and B in the number of first-place votes, and since ties are resolved alphabetically, A wins—an outcome that the voter who voted dishonestly is pleased to see, since he likes A better than B.

You could now check every single-winner method that we have discussed. None of them is strategy-proof—except dictatorship. (It is very easy to see that dictatorship *is* strategy-proof; see Exercise 10.2.) In fact, as the following theorem shows, no reasonable single-winner method is strategy-proof.

Theorem 10.2 (A. Gibbard, 1973, and M. A. Satterthwaite, 1975). *When there are more than two candidates, the only Pareto-efficient, strategy-proof single-winner method is dictatorship.*[6]

Proof. The proof that dictatorship of the kth voter ($1 \le k \le N$, with $N =$ number of voters) is Pareto-efficient and strategy-proof is Exercise 10.2. Assume now, conversely, that we are given a single-winner method that is Pareto-efficient and strategy-proof. We will prove that the method is strongly monotonic. It will then follow, by the (second version of the) Muller–Satterthwaite theorem (Theorem 9.7′), that the single-winner method is dictatorship.

So suppose we are given a Pareto-efficient, strategy-proof single-winner method. Let X be the election winner. Suppose we now make changes in one of the ballots, say the ith, in such a way that any candidate who was ranked below X on the ith ballot before the changes is still ranked below X after the changes. Will X still be the winner? If so, the method is strongly monotonic (see Proposition 9.9).

We denote the original ballot of the ith voter by B_i, and the changed ballot by B_i'. Suppose that after replacing B_i by B_i', Y is the winner. Our goal is to prove that $X = Y$.

Suppose that the original ballot B_i reflects the ith voter's honest opinion. If she casts the ballot B_i, then X is the winner. But if she dishonestly casts the ballot B_i', then Y is the winner. Since the method is strategy-proof, Y does not rank above X on B_i, and therefore Y does not rank above X on B_i' either.

But now suppose that in fact B_i' represents the ith voter's honest opinion. If she casts the ballot B_i', then Y is the winner. If she (dishonestly, now that we assume that B_i' reflects

[6]As earlier, I use the singular *dictatorship* even though there are N different dictatorial single-winner methods, one for each voter.

her honest opinion) casts the ballot B_i, however, then X is the winner. Since the method is strategy-proof, X cannot rank above Y on B_i'.

We have now concluded that on B_i', X does not rank above Y, nor does Y rank above X. Thus indeed $Y = X$. This completes the proof that our single-winner method is strongly monotonic. Therefore, by the Muller–Satterthwaite theorem, it is dictatorship of the kth voter for some k. □

The assumption of Pareto-efficiency in Theorem 10.2 can be replaced by the assumption that the method *gives every candidate a chance to win*. By this we mean that for every candidate X, there exists a preference schedule for which X becomes the winner. The following is the Gibbard–Satterthwaite theorem in the form in which it is often stated in the literature on voting.

Theorem 10.2′ (Gibbard–Satterthwaite theorem, second version). *When there are more than two candidates, the only single-winner method that gives every candidate a chance to win and is strategy-proof is dictatorship.*

Proof. In the proof of Theorem 10.2, we argued that a strategy-proof single-winner method is strongly monotonic. The assumption of Pareto-efficiency was not needed in this argument. (It played a role only at the end, when we used the Muller–Satterthwaite theorem.) It is very easy to see that any strongly monotonic single-winner method that gives every candidate a chance to win is Pareto-efficient. Therefore the assumptions of Theorem 10.2′ imply those of Theorem 10.2. □

We conclude this chapter with a much weaker statement that is much easier to understand, but which still implies that any strategy-proof single-winner method is flawed.

Theorem 10.3. *When the number n of candidates is greater than two, and the number N of voters equals n, any strategy-proof single-winner method violates the majority criterion.*

Proof. I will give the proof for $n = 4$ candidates. It will be clear that a similar argument applies for any $n \geq 3$ (but not for $n = 2$). Suppose that there are four candidates A, B, C, and D, and four voters. Consider the following detailed preference schedule:

A	B	C	D
B	C	D	A
C	D	A	B
D	A	B	C

. (10.1)

Notice that the second column is obtained from the first by moving the first candidate (A) into last place. The third column is obtained from the second in the same way, and the fourth from the third.

In spite of the perfect symmetry of the preference schedule, one of the four candidates must be the winner, since we are talking about single-winner methods here. Let us suppose that A is the winner. (Analogous arguments work when B, C, or D is the winner; see Exercise 10.4.) Now suppose that the second voter dishonestly moves D into first place on his ballot, like this:

A	D	C	D
B	B	D	A
C	C	A	B
D	A	B	C

.

Since the method is strategy-proof, the winner is still A, since the voter would be pleased with any change (he ranks A last). Next, suppose that the third voter dishonestly moves D to the top of her ballot:

A	D	D	D
B	B	C	A
C	C	A	B
D	A	B	C

.

The winner is now either A or B, since the third voter would otherwise be pleased with the change in outcome. But D now has 3 out of 4 first-place votes; therefore the single-winner method violates the majority criterion. □

Exercises

10.1. Suppose that the preference schedule is

3	2	3
C	B	A
B	A	C
A	C	B

,

and the election is to be evaluated using Borda count, with the convention that ties are resolved alphabetically. Suppose that one of the two voters whose ranking is

B
A
C

dishonestly votes

A
B
C

instead. Show that this will change the election outcome in a way that the dishonest voter likes.

10.2. Explain why dictatorship is strategy-proof.

10.3. Prove by example that the Gibbard–Satterthwaite theorem does not hold for $n = 2$.

10.4. In the proof of Theorem 10.3, we made the assumption that the method makes A the winner when (10.1) is the preference schedule. Repeat the argument assuming that D is made the winner (but not assuming the principle of one person, one vote, nor the principle of independence of candidate names).

Chapter 11

Winner Selection versus Ranking

We have considered *winner selection* methods so far. We will now study *ranking* methods, that is, methods aimed at ranking all candidates in the field. Why would we want to do this? Here are some examples of situations in which it would be good to know how to construct *societal rankings* from *individual rankings*:

(1) Imagine that management tells a team in a company, "You will be able to hire two to five new employees into your group. How many you are allowed to hire is not certain yet, it depends on how much money we will make this quarter, but please submit a list of people whom you would like to hire." In that case, the group might want to submit a ranked list of five candidates, with the idea that the first two, three, or four, or perhaps all five will be hired.

(2) Suppose we want to elect a committee of five people, and have eight candidates. One good way of proceeding would be to rank all eight candidates, then pick the first five as members of the committee.

(3) A real-life variation on the second example: I serve on a university committee that selects the student speaker at commencement. In a first round, approximately 12 students apply. The committee selects approximately 6 of these 12 for a second round. A good way of doing this is to rank all 12 applicants, then admit the top 6 to the second round.

(4) In some sports, for instance, in diving, competitors are ranked by judges. The competitors are the "candidates," and the judges are the "voters."

In this chapter, we will examine the link between ranking and winner selection. First, a subtle issue of terminology. In Chapters 9 and 10, we tacitly assumed that both the number n of candidates and the number N of voters were fixed and given. By contrast, when we talk about a *winner selection method* in this chapter, we mean a method that assigns to *any* (detailed) preference schedule for *any number n of candidates* a set W of winners. (We can still think of the number N of voters as fixed and given; that does not matter.) The plurality method is the simplest example. Another example, albeit a contrived one, would be: "Use pairwise comparison if $n > 3$ and the plurality method if $n \leq 3$." Likewise, when we talk about a *ranking method* in this chapter, we mean a method that assigns to any preference schedule *for any n* a societal ranking.

Given a winner selection method, one can derive a ranking method. One picks the winners among the n candidates, ranks them first, then eliminates them from the ballots, and re-evaluates the ballots, etc. This is called *recursive ranking*. For example, the ranking method derived in this way from the plurality method is called the *recursive plurality method*.

Example 11.1. Suppose that 23 voters want to rank 5 candidates, and that the preference schedule is

5	3	4	2	3	1	5
C	C	B	A	A	D	E
B	A	C	E	E	B	D
A	D	D	C	B	C	A
D	E	A	B	D	E	B
E	B	E	D	C	A	C

. (11.1)

How would the recursive plurality method rank the candidates? To find out, we apply the plurality method, note the winner(s), remove the winner(s) from the preference schedule, then apply the plurality method again, etc.

The plurality method selects C as the winner, so the recursive plurality method ranks C first. After removing C from the ballots, we are left with

5	3	4	2	3	1	5
B	A	B	A	A	D	E
A	D	D	E	E	B	D
D	E	A	B	B	E	A
E	B	E	D	D	A	B

,

or, more compactly,

5	3	4	5	1	5
B	A	B	A	D	E
A	D	D	E	B	D
D	E	A	B	E	A
E	B	E	D	A	B

.

Now the winner by the plurality method is B, so the recursive plurality method ranks B second. Removing B from the ballots, we obtain

5	3	4	5	1	5
A	A	D	A	D	E
D	D	A	E	E	D
E	E	E	D	A	A

.

Again, we can write this a bit more compactly:

8	4	5	1	5
A	D	A	D	E
D	A	E	E	D
E	E	D	A	A

.

Now A wins by the plurality method, so the recursive plurality method ranks A third. Removing A from the ballots, we get

8	4	5	1	5
D	D	E	D	E
E	E	D	E	D

,

or, more compactly,

13	10
D	E
E	D

.

So D is ranked fourth, and E is ranked last. We summarize the ranking as follows:

$$C \succ B \succ A \succ D \succ E.$$

The notation "$X \succ Y$" (or "$Y \prec X$") will be used for "X is ranked above Y" in this chapter. (We use the symbol "$\succ$" because it looks a little bit like "$>$". Of course it isn't the same: X and Y are not numbers, so "$X > Y$" would make no sense.) ∎

The recursive runoff method, the recursive elimination method, recursive Borda count, the recursive method of pairwise comparison, and the recursive beatpath method are defined similarly. All of these methods generate, from a set of preference ballots, a *societal preference ranking*. We generally call a ranking method *recursive* if it is (or can be) derived by recursive application of a winner selection method.

The societal preference ranking, in general, will include ties, since winner selection methods don't always select a *sole* winner. (We still don't allow ties in individual preference ballots—individuals must always express strict preferences.)

Example 11.2. Assume that the preference schedule is

5	3	8	5	2	7	5
A	A	B	C	C	D	E
B	C	A	E	E	B	D
C	D	D	A	B	A	C
D	E	C	B	D	E	B
E	B	E	D	A	C	A

(11.2)

and that we use the recursive plurality method to rank the candidates. Then A and B are ranked equally in the first round. We therefore start the societal ranking with

$$A \sim B \succ \cdots,$$

using the symbol "$\sim$" to indicate a tie in societal preferences. Then we remove both A and B from the preference schedule:

5	3	8	5	2	7	5
C	C	D	C	C	D	E
D	D	C	E	E	E	D
E	E	E	D	D	C	C

,

or, more compactly,

8	8	7	7	5
C	D	C	D	E
D	C	E	E	D
E	E	D	C	C

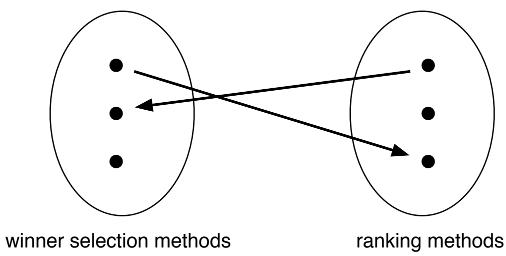
.

By plurality, C and D now tie for victory. Therefore the societal ranking is

$$A \sim B \succ C \sim D \succ E$$

according to the recursive plurality method. ∎

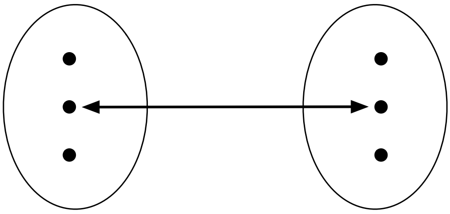
Since every winner selection method gives rise to a ranking method, we are tempted to say that winner selection methods are a special kind of ranking method. But is that true, or should we say it the other way around? We can certainly turn any ranking method into a winner selection method as well: We simply rank all candidates, then pick the top-ranked one(s) to be the winner(s). So each winner selection method gives rise to a ranking method, and each ranking method gives rise to a winner selection method. We can summarize this in the following abstract picture:



Each dot on the left-hand side denotes a winner selection method. There are many of them, infinitely many actually, but I can't draw infinitely many dots, so I drew three. Similarly, on the right-hand side, each dot denotes a ranking method. The arrow pointing from left to right indicates that the winner selection method represented by the uppermost dot on the left-hand side gives rise to the ranking method represented by the lowermost dot on the right-hand side. Similarly, the arrow pointing from right to left indicates that the ranking method represented by the uppermost dot on the right-hand side gives rise to the winner selection method represented by the middle dot on the left-hand side. From each dot on the left-hand side originates an arrow that goes to the right-hand side; I drew only one example. Similarly, from each dot on the right-hand side originates an arrow that goes to the left-hand side; again, I drew only one example.

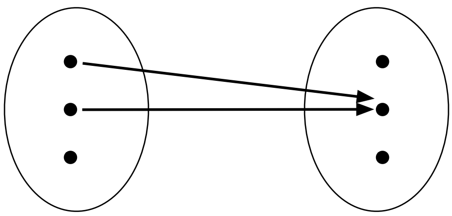
Using mathematical language, we say that we have here a "mapping" from the set of winner selection methods to the set of ranking methods, and another "mapping" the other way around. A *mapping* from a set S into a set T is in general a rule that assigns to each element of S exactly one element of T.

Suppose we start out with a winner selection method, and derive from it, as described above, a ranking method. Then we take that ranking method and derive from it, again as described above, a winner selection method. It is not hard to see that we will end up with the same winner selection method with which we started (see Exercise 11.5). Here is an abstract picture representing this statement:



A consequence is that two different winner selection methods can never give rise to the same ranking method.

Using mathematical language, we say that the mapping from the set of winner selection methods to the set of ranking methods is "injective." In general, a mapping from a set S into a set T is called *injective* if it never happens that the same element of T is assigned to two or more different elements of S. That is, this does not happen:



Recall that we call a ranking method *recursive* if it is derived by recursive application of a winner selection method. Are there any ranking methods that are not recursive? You may first think "Well, of course!", since you could define ranking methods in all sorts of ways not involving recursive application of a winner selection method. For instance, *one-shot Borda count* is a ranking method defined as follows: Compute the Borda scores for each candidate, and rank the candidates accordingly. This definition does not involve recursive application of any winner selection method. But how can we be sure that we can't give an alternative, equivalent definition of the same method that does involve recursive application of some underlying winner selection method?

If there were such an alternative definition of one-shot Borda count, the picture that we drew above tells us what the underlying winner selection method would have to be. It would have to be the winner selection method derived from one-shot Borda count. That winner selection method is Borda count (see Exercise 11.6). So if one-shot Borda count were a recursive method, it would have to be the same as recursive Borda count.

We will now give an example demonstrating that one-shot Borda count and recursive Borda count are not the same.

Example 11.3. Consider the preference schedule

$$\begin{array}{|c|c|c|}\hline 2 & 4 & 3 \\ \hline\hline C & B & A \\ \hline A & C & B \\ \hline B & A & C \\ \hline\end{array}\,. \tag{11.3}$$

Here A gets 17 Borda points, B gets 20 Borda points, and C gets 17 Borda points. So one-shot Borda count produces the following societal ranking:

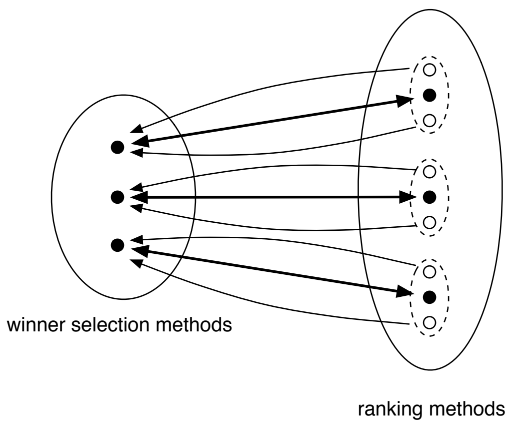
$$B \succ A \sim C.$$

On the other hand, the societal ranking produced by recursive Borda count is (see Exercise 11.7)

$$B \succ C \succ A.$$

So indeed one-shot Borda count and recursive Borda count do not produce the same ranking in this example. ∎

We can now complete our abstract diagram:



The solid dots on the right represent recursive ranking methods, and the open dots represent nonrecursive ones, such as one-shot Borda count.

In summary, it is natural to think of winner selection methods as a special class of ranking methods, but not the other way around. There are ranking methods that are *not* (derived from) winner selection methods; one-shot Borda count is an example.

Exercises

11.1. Suppose that the preference schedule is

4	4	3	2	1
A	E	C	B	E
E	C	A	E	A
C	B	E	A	B
D	D	D	C	D
B	A	B	D	C

.

How are the five candidates ranked by the recursive plurality method?

11.2. Suppose that the preference schedule is

3	2	3	1
C	B	A	B
B	C	C	C
A	A	B	A

.

How are the three candidates ranked by the recursive method of pairwise comparison?

11.3. For the preference schedule (11.2) in Example 11.2, prove that recursive application of any Condorcet-fair winner selection method generates the same ranking.

11.4. The *one-shot plurality method* ranks all candidates according to the number of first-place votes that they receive. (a) How would the one-shot plurality method rank the candidates if the preference schedule is as in Example 11.1? You will see from this example that the one-shot plurality method is not the same as the recursive plurality method. (b) Explain how you can be certain that the one-shot plurality method is not a recursive ranking method.

11.5. Explain in words: If we start out with a winner selection method, derive from it a ranking method, and then derive from that ranking method a winner selection method, we arrive at the same winner selection method that we started out with.

11.6. Explain in words: The winner selection method derived from one-shot Borda count is Borda count.

11.7. Verify that the ranking of candidates obtained, using recursive Borda count, for the preference schedule of Example 11.3 is $B \succ C \succ A$.

Chapter 12

Irrelevant Alternatives and Arrow's Theorem

In this chapter, we state and prove the most famous theorem in social choice theory, called Arrow's dictatorship theorem after Kenneth Arrow, who proved it in his dissertation, published in book form in 1951. It is a result for ranking methods quite similar to the Muller–Satterthwaite theorem of Chapter 9. Arrow won the Nobel Prize in Economics in 1972.

We return here to the (tacit) assumption of Chapters 9 and 10, temporarily suspended in Chapter 11, that both the number n of candidates and the number N of voters are fixed.

We can formulate fairness criteria for ranking methods just as we did for winner selection methods. I will give a few examples here, beginning with two that play roles in Arrow's theorem. The first is a very weak "majority rule" criterion.

Definition 12.1 (unanimity criterion for ranking methods). *A ranking method satisfies the unanimity criterion if it guarantees that $X \succ Y$ if X is preferred to Y by* every *voter.*

(Recall that the notation "$X \succ Y$" means that according to the societal ranking computed by the ranking method, X ranks above Y.) Most reasonable ranking methods satisfy the unanimity criterion.

Example 12.2. The following argument proves that the recursive plurality method satisfies the unanimity criterion. If every voter prefers X to Y, then as long as X has not been ranked yet, Y will not even get a single first-place vote. Only after X has been ranked, and therefore removed from the ballots, does Y have a chance to be ranked. Therefore X ranks higher than Y in the societal ranking. ■

There do exist examples of ranking methods that violate the unanimity criterion; an example is the recursive Smith method (see Exercise 12.1). Our second fairness criterion for ranking methods is a "no spoiler" criterion very similar to the IIC criterion of Chapter 10.

Definition 12.3 (IIA criterion). *A ranking method satisfies the independence-of-irrelevant-alternatives (IIA) criterion if the relative societal ranking of any two candidates X and Y depends only on the relative ranking of X and Y on each individual ballot—not on how voters rank other candidates.*

The criterion says that the relative ranking of X and Y should not depend on any candidate Z different from X and Y. For the purpose of comparing X and Y, Z should be irrelevant.

Example 12.4. If the preference schedule is

3	2	2
A	B	C
B	A	B
C	C	A

, (12.1)

the recursive plurality method ranks the candidates like this:

$$A \succ B \succ C.$$

If the preference schedule were

3	2	2
A	B	B
B	A	A
C	C	C

, (12.2)

the ranking would be

$$B \succ A \succ C.$$

On each *individual* ballot, the relative ranking of A vs. B is the same in (12.2) as in (12.1). Nevertheless, *society* prefers A to B if the preference schedule is (12.1), and it prefers B to A if the preference schedule is (12.2). Thus the recursive plurality method violates the IIA criterion. ∎

Arrow's theorem, the main result of this chapter, shows that it is in fact impossible to construct a reasonable ranking method that satisfies the IIA criterion. So here the conflict between "majority rule" and "absence of spoilers," which first became visible in Chapter 3 and was to some degree resolved in Chapter 5, manifests itself again.

Dictatorship of the kth voter, meaning that the kth voter's ranking of the candidates is taken to be society's ranking in all cases, does satisfy the IIA criterion. Unfortunately, nobody other than the kth voter likes this ranking method. For one thing, it violates the principle of one person, one vote, which we formulate for ranking methods in the same way as for winner selection methods (see Definition 1.2).

Definition 12.5 (principle of one person, one vote). *A ranking method satisfies the principle of one person, one vote if the societal ranking that it generates depends only on the reduced preference schedule.*

The principle of one person, one vote demands that all voters be treated equally. Similarly, the principle of independence of candidate names demands that all candidates be treated equally. It, too, is the same for ranking methods as for winner selection methods (see Definition 2.9).

Definition 12.6 (principle of independence of candidate names). *A ranking method satisfies the principle of independence of candidate names if the societal ranking does not change when the candidates are renamed or renumbered.*

We are now ready to state Arrow's theorem.

Theorem 12.7 (K. Arrow, 1951). *The only method for ranking $n \geq 3$ candidates that satisfies the unanimity and IIA criteria is dictatorship.*[7]

Suppose, for instance, that candidates A, B, and C have been interviewed for a job. Each member of the selection method ranks the three candidates. The committee uses some ranking method to derive from the individual rankings a joint ranking. Let us say that the joint ranking is

$$B \succ C \succ A.$$

A day after this ranking has been computed, it turns out that C faked his resume. He actually has no relevant experience at all, and will therefore not be considered further. He has turned into an "irrelevant alternative." Should B still be ranked above A by the committee? You may be tempted to answer "yes, of course!"—but if you did, you would assume independence of irrelevant alternatives, precisely the property which, according to Arrow's theorem, no reasonable ranking method can have.

Even without knowing Arrow's theorem, one can see quickly that a ranking method that satisfies the IIA criterion, and some other mild and reasonable conditions, must yield strange results in some cases. For instance, consider the preference schedules

1	1	1
A	B	C
B	C	A
C	A	B

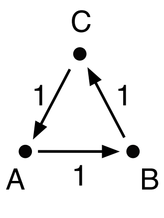
(12.3)

and

1	1	1
A	B	C
B	A	A
C	C	B

. (12.4)

The pairwise comparison graph for (12.3) looks as follows:

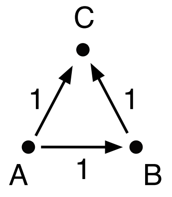


[7]As earlier, I use the singular *dictatorship* even though there are N different dictatorial ranking methods, one for each voter.

Let us suppose that we use a ranking method that satisfies the principles of independence of candidate names and one person, one vote. Such a method certainly has to rank A, B, and C equally in this example (see Exercise 12.4):

$$A \sim B \sim C.$$

On each ballot, the rankings of A and B, relative to each other, are the same in (12.4) as in (12.3). Therefore, if our ranking method satisfies the IIA criterion, the societal ranking of A and B, relative to each other, must be the same in (12.4) as in (12.3). Since it is $A \sim B$ for (12.3), it must be $A \sim B$ for (12.4). But the pairwise comparison graph for (12.4) is



and you will probably agree that the most reasonable ranking here would be

$$A \succ B \succ C.$$

We now begin to work towards the proof of Arrow's theorem. First, we prove that any ranking method that satisfies the assumptions of Arrow's theorem has a peculiar property. To explain this, we need two definitions.

Definition 12.8 (polarizing candidate). *Given a preference schedule, a candidate B is called polarizing if on each ballot, B appears either in the first position or in the last.*

For instance, if the preference schedule is

1	1	2	1
B	B	C	A
C	A	A	C
A	C	B	B

,

then B is polarizing. You might think of the example of a white supremacist. Most voters find his views revolting, and rank him last. Some voters are enthusiastic about his views, and rank him first. Hardly anybody would rank a white supremacist somewhere in the middle of the field of candidates.

Definition 12.9 (property P). *A ranking method is said to have property P if it ranks any polarizing candidate either strictly first or strictly last.*

If you think that this is a repetition of the definition of *polarizing candidate*, notice that Definition 12.9 refers to the *societal* ranking, not to individual rankings.

A ranking method that has property P might seem a bit suspect. For instance, think about this preference schedule:

50	50
A	B
C	D
D	E
E	F
F	C
G	G
B	A

.

You would probably agree that A and B do equally well here, and both arguably do much better than G. Nevertheless, a ranking method that has property P must rank either A strictly first and B strictly last, or vice versa.

On the other hand, recursive pairwise comparison, certainly a reasonable ranking method, comes very close to having property P. It fails to have property P only because of the possibility that a polarizing candidate might appear in first place on exactly half the ballots, and in last place on the other half of the ballots (see Exercise 12.5). Thus if the number of voters is odd, recursive pairwise comparison does have property P.

Lemma 12.10. *The assumptions of Arrow's theorem imply that the ranking method has property P.*

Proof. Suppose that the assumptions of Arrow's theorem were satisfied, and B were polarizing, but neither strictly first nor strictly last in society's ranking. There would then be candidates A and C, different from each other, so that

$$A \succeq B \succeq C.$$

(The notation $X \succeq Y$ indicates that $X \succ Y$ or $X \sim Y$—in society's ranking, X is either ranked higher than Y or equal with Y.) But note that on every single ballot, the ranking of A vs. B is the same as the ranking of C vs. B. Therefore, by the IIA criterion, if $A \succ B$ then $C \succ B$, in contradiction with $B \succeq C$. Therefore we must have $A \sim B$, and then, again by the IIA criterion, $C \sim B$ as well:

$$A \sim B \sim C.$$

We can now change the ballots so that on every single ballot, C ranks above A, while the position of B is not changed on any of the ballots. Again by the IIA criterion, this change must preserve society's ranking of A vs. B and B vs. C, so still

$$A \sim B \sim C, \tag{12.5}$$

but on the other hand, now

$$C \succ A \tag{12.6}$$

by the unanimity criterion. The contradiction between (12.5) and (12.6) implies that our initial assumptions were wrong, and thereby proves the lemma. □

Now we are ready to prove Arrow's theorem, Theorem 12.7.[8]

Proof. We begin by picking an arbitrary candidate A. (The argument can be done with *any* choice of A, and this point will be important later.) We start with some preference schedule in which A is placed last on every single ballot:

1	2	3	...	$N-1$	N
.	.	.	...	.	.
.	.	.	...	.	.
...	...	...	...	...	...
A	A	A	...	A	A

(This is a *detailed* preference schedule, not a reduced one. The numbers at the top of the columns simply label voters.) By the unanimity criterion, society must then place A last.

The strategy is now to switch A from last place to first place in the first, then the second, then the third ballot, and so on. We will prove that A remains in last place in the societal ranking, until suddenly, when A is moved from last to first place in the kth ballot (for some k between 1 and N), A goes into first place in the societal ranking. Voter k is, in this situation at least, pivotal. We will then prove that voter k is in fact a dictator—his ranking is society's ranking.

Let us begin by swapping A from last place to first place in the first ballot, leaving all else unchanged:

1	2	3	...	$N-1$	N
A	.	.	...	.	.
.	.	.	...	.	.
...	...	...	...	...	...
.	A	A	...	A	A

Since A is still polarizing, society places A either last or first, by Lemma 12.10. If A is still last in the societal ranking, we swap A from last place to first in the second ballot, again leaving all else unchanged:

1	2	3	...	$N-1$	N
A	A	.	...	.	.
.	.	.	...	.	.
...	...	...	...	...	...
.	.	A	...	A	A

We continue in this way until we reach a ballot in which swapping A from last position to first changes A's standing in society's ranking from last to first. This must happen eventually, since A will certainly be ranked first by society once A is in first place on every single ballot, by the unanimity criterion.

[8] I reproduce here the proof in Philip J. Reny, "Arrow's theorem and the Gibbard–Satterthwaite theorem: A unified approach," *Economics Letters* 70, issue 1, pages 99–105 (2001). It is very similar to the proof of the Muller–Satterthwaite theorem given in Chapter 9; this similarity is the central point of Reny's paper.

So there is a preference schedule of the form

1	2	...	$k-1$	k	$k+1$	...	$N-1$	N
A	A	...	A	.	.	...	.	.
.	.	...	.	.	.	...	.	.
.	.	...	.	.	.	...	.	.
...	...	...	...	...	...	...	...	...
.	.	...	.	A	A	...	A	A

(12.7)

for which A is ranked last by society, but if A is moved from last place to first place in ballot k, with no other changes in the preference schedule,

1	2	...	$k-1$	k	$k+1$	...	$N-1$	N
A	A	...	A	A	.	...	.	.
.	.	...	.	.	.	...	.	.
.	.	...	.	.	.	...	.	.
...	...	...	...	...	...	...	...	...
.	.	...	.	.	A	...	A	A

, (12.8)

then A is placed first by society. Voter k certainly has a pivotal influence in this example: As soon as he changes his mind and places A first instead of last, society follows suit.

We will now prove, in a sequence of steps, that voter k is a dictator. First, let B and C be any two candidates, different from each other and different from A. Assume that voter k ranks B over C. We will prove that society ranks B over C; thus society adopts, at least for two candidates that are different from A, voter k's rankings.

Without changing the relative ranking of B vs. C on any ballot, and therefore—by the IIA criterion—without changing society's ranking of B vs. C, we can change the ballots in such a way that A is first on ballots 1 through $k-1$, in between B and C on ballot k, and last on ballots $k+1$ through N:

1	2	...	$k-1$	k	$k+1$	...	$N-1$	N
A	A	...	A	.	.	...	.	.
.	.	...	.	B	.	...	.	.
...	...	...	...	...	...	...	...	...
.	.	...	.	A	.	...	.	.
...	...	...	...	...	...	...	...	...
.	.	...	.	C	.	...	.	.
...	...	...	...	...	...	...	...	...
.	.	...	.	.	A	...	A	A

. (12.9)

The rankings of B vs. A in (12.9) are just as in (12.7), so society's ranking of B vs. A is the same in (12.9) as in (12.7): $B \succ A$. (Recall that A is ranked strictly last by society when the preference schedule is (12.7).) But also, the rankings of A vs. C in (12.9) are just as in (12.8), so society's ranking of A vs. C is the same in (12.9) as in (12.8): $A \succ C$. (Recall that A is ranked strictly first by society when the preference schedule is (12.8).) Therefore

$$B \succ A \succ C,$$

and therefore

$$B \succ C.$$

We have proved that voter k calls the shots when it comes to comparing B and C—or in general, when it comes to comparing any two candidates different from A. His relative ranking of B vs. C is society's relative ranking. We say that voter k is an A^**-dictator*—brief terminology for a voter whose ranking decides the societal ranking between any two candidates *different from* A. Nobody other than voter k, the A^*-dictator, *ever* has any influence on how society ranks candidates B and C (both different from A) relative to each other.

All that is left to prove is that voter k in fact calls the shots even when it comes to comparing any candidate with A. So let X be any candidate other than A. We want to prove that voter k dictates the societal relative ranking of X vs. A. Let Y be a third candidate, neither X nor A. We emphasized from the outset that nothing was special about our choice of A. Thus the argument that we have given proves that there is a Y^*-dictator. No voter other than the Y^*-dictator can *ever* influence the relative societal ranking of A vs. X. However, preference schedules (12.7) and (12.8) show that there is at least one situation in which voter k influences the comparison of A vs. X: Depending on how voter k ranks A, X is either ranked above A (preference schedule (12.7)), or below A (preference schedule (12.8)). Therefore voter k must be the Y^*-dictator! So indeed voter k dictates the relative societal ranking of A vs. X as well.

In summary, voter k is the dictator: Society's ranking equals voter k's ranking in all cases. □

Exercises

12.1. Give an example that proves that the recursive Smith method violates the unanimity criterion.

12.2. Give an example that proves that recursive Borda count violates the IIA criterion.

12.3. Verify that dictatorship of the kth voter satisfies the unanimity and IIA criteria. (See also Appendix B, Example B.2.)

12.4. Explain why any ranking method that satisfies the principles of independence of candidate names and one person, one vote must produce the ranking

$$A \sim B \sim C$$

for the preference schedule (12.3).

12.5. Explain: (a) If a majority of voters places B strictly first, then the method of recursive pairwise comparison ranks B strictly first. (b) If a majority of voters places B strictly last, then the method of recursive pairwise comparison ranks B strictly last. (c) The method of recursive pairwise comparison has property P if the number N of voters is odd.

12.6. Suppose that a ranking method has Property P, and the preference schedule is

3	3
A	C
B	B
C	A

.

There are only two possible rankings that the ranking method might produce in this case. What are they?

12.7. Give an example that proves that Borda count does not have Property P.

12.8. Explain why Arrow's theorem does not hold for only two candidates.

Chapter 13

Fairness and Envy-Freeness

Suppose that several people jointly own an object that cannot be divided, such as a house, a painting, or a parrot. Suppose that they would like to transfer ownership of the object to one among them, who will then pay monetary compensation to the others. One might think of a divorce or an inheritance, for instance. Who should get the object, and what should the compensation payments be?

Of course there is no mathematical answer to this question. However, one can formulate guiding principles that one might want to use in making decisions of this kind, and study their mathematical consequences. This is what we will do in Chapters 13–15. We begin in this chapter with the fundamental notion of fairness, and a refinement called envy-freeness.

Example 13.1. Amelia and Bill (A and B) get divorced. They jointly own a house. One of them will get the house, but will have to compensate the other. How should they decide who will get the house and what the compensation amount will be? The simplest idea might be to flip a coin to decide who gets the house, consult a real estate agent about its market value, and have the person who gets the house pay half the market value to the other. Obviously this is not the best thing to do. We ought to take into account the central fact in all compensation and division problems:

The same object has different values to different people.

For instance, suppose that the real estate agent estimates that the house is worth \$440,000. B has done his own market research, and he agrees the house could fetch \$440,000, but only with patience. For a quick sale, it would need to be priced at \$400,000. A quick sale is important to him. He has no time to deal with this matter, and does not want the house—it evokes unpleasant memories of the three intolerably ill-behaved, good-for-nothing children whom the couple raised in the house. If he got it, he would immediately turn around and sell it for \$400,000.

On the other hand, A has fond memories of the three lovely, accomplished children whom the couple raised in the house. The house has sentimental value to her; it is worth \$500,000 to her.

As bad luck would have it, the coin toss results in the outcome that B gets the house. He has to pay \$220,000 to A in compensation—half of what the real estate agent said the house was worth, but more than half of what the house is worth to him. On the other hand, A does not get the house, and is paid only \$220,000—less than half of what the house is worth to her. Both A and B feel that they did not get their fair share.

To find a better solution, both A and B should reveal to a referee how much the house is worth to them. To avoid manipulative, dishonest bidding, it is best if neither knows how the other values the house. One therefore refers to this process as *sealed bidding*. The referee's task is to decide which of the two should get the house, and how much compensation that person should pay to the other. Our question is how the referee should think about this decision. ∎

Example 13.2. Three siblings, A, B, and C, inherit a house. They decide that one of them should get the house and compensate the others. To decide who should get the house, and how that person should compensate the others, they reveal to a referee (but not to each other) their estimates of the value of the house: A considers it worth \$300,000, while to B it is worth \$270,000, and to C \$210,000. What should the referee do? ∎

We first introduce some notation and terminology that we will use throughout this chapter and the following two. The number of people jointly owning the object is N. We use capital letters such as A, B, C, ... to denote the people jointly owning the object. Their *valuations* of the object are \$$a$, \$$b$, \$$c$, etc. ($A$ considers the object worth \$$a$, B considers it worth \$$b$, etc.) Another word for *valuation* is *bid*, and we will use these two words interchangeably.

A person whose bid is greatest is called a *highest bidder*. There can be several highest bidders; for instance, if $a = b = 400,000$ and $c = 350,000$, then A and B are highest bidders. The person to whom the referee assigns the object, for whichever reason, is called the *winning bidder*, and we use the notation \$$w$ for his bid, the *winning bid*. The *average bid*, that is, the average of all bids, is denoted by \$$m$. (The letter m stands for "mean" here.) The sum of all bids is denoted by S, so

$$m = \frac{S}{N}. \tag{13.1}$$

Each of the N people is entitled to one Nth of the object, and will therefore want to walk away with at least one Nth of what he or she thinks the object is worth. This amount is what we call the person's *fair share*. So A's fair share is \$$a/N$, B's fair share is \$$b/N$, etc. The average bid m is the sum of the fair shares.

If, say, A is the winning bidder, we denote the compensation amounts that A pays to B, C, etc., by \$$x_B$, \$$x_C$, etc., and define $x_A = a - x_B - x_C - \cdots$. What A obtains is worth \$$x_A$ in his eyes—the value of the object, which is \$a in his eyes, minus all the compensation amounts that A has to pay to the others. We call \$$x_A$, \$$x_B$, ... the *payouts*. A person's payout, in general, is the dollar value of what he or she walks away with, in his or her eyes. (The words "in his or her eyes" are needed only for the winning bidder, since for the others, there is nothing subjective about what they receive—they get cash.)

Example 13.3. In Example 13.2, $N = 3$, the bids (or valuations) are $a = 300,000$, $b = 270,000$, and $c = 210,000$ dollars, and the highest bidder is A. If the referee decides, for whichever reason, to assign ownership of the house to B, then B is the winning bidder, and $w = 270,000$. The average bid is

$$m = \frac{300,000 + 270,000 + 210,000}{3} = 260,000$$

dollars. The fair shares of A, B, and C are \$100,000, \$90,000, and \$70,000, respectively. If the referee requires B to pay \$100,000 to A and \$80,000 to C, then the payouts are $x_A = 100{,}000$, $x_B = 270{,}000 - 100{,}000 - 80{,}000 = 90{,}000$, and $x_C = 80{,}000$ dollars. ∎

Example 13.4. Three siblings, A, B, and C, have inherited a parrot from their aunt. Opinions are divided about its value; the bids are a, b, and c dollars. The referee assigns the bird to B, who will pay $\$x_A$ to A and $\$x_C$ to C for compensation. So B is the winning bidder, and the winning bid is $\$b$. Then

$$x_B = b - x_A - x_C.$$

This equation can be rewritten as

$$x_A + x_B + x_C = b.$$

Thus the sum of the payouts is the winning bid. ∎

It is clear that the conclusion of Example 13.4 is general, and we record it for future use.

Lemma 13.5. *In any compensation arrangement, the sum of all payouts equals the winning bid.*

We will now state the first of the two fundamental definitions of this chapter.

Definition 13.6 (fair compensation). *A compensation arrangement is called fair if the payout to each person is at least that person's fair share:*

$$x_A \geq \frac{a}{N}, \quad x_B \geq \frac{b}{N}, \quad etc.$$

Example 13.7. In Example 13.3,

$$\begin{aligned} x_A &= 100{,}000 = \frac{a}{3}, \\ x_B &= 90{,}000 = \frac{b}{3}, \\ x_C &= 80{,}000 > \frac{c}{3} = 70{,}000. \end{aligned}$$

In some sense, C is treated better than A and B: C gets more than her fair share, while the others get exactly their fair share, nothing more. Nonetheless, according to Definition 13.6, the arrangement is considered fair. ∎

Example 13.8. A, B, C, and D have inherited a house. They have different opinions on what the house is worth: $a = 380{,}000$, $b = 420{,}000$, $c = 440{,}000$, and $d = 360{,}000$.

The referee assigns the house to A, and asks him to pay \$110,000 to B, the same to C, and \$90,000 to D.

A's payout is $\$x_A = \$(380{,}000 - 2\times 110{,}000 - 90{,}000) = \$70{,}000$, less than A's fair share of \$95,000. The arrangement is therefore unfair, although B, C, and D are treated fairly: B's payout is $\$x_B = \$110{,}000$, greater than B's fair share of \$105,000. C's payout is $\$x_C = \$110{,}000$, equal to C's fair share. D's payout is $\$x_D = \$90{,}000$, equal to D's fair share. ∎

It will make sense in most cases for the referee to assign the object to a highest bidder; see also Chapter 14. However, there are situations in which it does make sense to award the object to somebody who is not a highest bidder.

Example 13.9. Three brothers, A, B, and C, have inherited a used car from their uncle. They decide to assign ownership of the car to one among them and have him compensate the others monetarily. Their valuations are $a = 6{,}000$, $b = 5{,}100$, and $c = 5{,}700$ dollars. Although A does not currently own a car, he has a history of reckless driving. Fearing for A's life (and the lives of other people), the referee awards the car to C, and asks him to pay \$2,000 to A and \$1,700 to B, A's and B's fair shares. Then C's payout is

$$\$x_C = \$(5{,}700 - 2{,}000 - 1{,}700) = \$2{,}000,$$

which is more than C's fair share, so this arrangement is fair. ∎

Is it always possible to find a fair compensation arrangement? The answer is yes, as long as the winning bid is high enough in comparison with the bids of the others.

Proposition 13.10. *It is possible to find fair compensation amounts if and only if the winning bid is at least as large as the average bid, i.e.,* $w \geq m$.

Proof. I will give the proof for an example only, but it will be clear that the argument works generally (see Exercise 13.3).

Suppose that $N = 4$ and that C is the winning bidder. Is it possible for C to compensate A, B, and D in such a way that the arrangement becomes fair? For fairness, C must pay at least $\$a/4$ to A, $\$b/4$ to B, and $\$d/4$ to D. The question is whether C can do that while retaining enough so that the arrangement is fair to him as well. The answer is yes if and only if

$$c - \frac{a}{4} - \frac{b}{4} - \frac{d}{4} \geq \frac{c}{4},$$

or

$$c \geq \frac{a+b+c+d}{4}, \tag{13.2}$$

that is, if and only if the winning bid c is at least as large as the average of all bids. □

We turn now to the second fundamental notion of this chapter, that of *envy-freeness*. First, an example.

Example 13.11. A, B, and C are college students sharing an apartment. They jointly own a cat. Eventually, their disagreements about what constitutes an appropriate level of cleanliness in the kitchen become irreconcilable. They decide to split up. One of them will get the cat. Of course, whoever gets the cat must compensate the others.

A loves the cat. She would give anything to keep her. Well, on second thought, perhaps not anything, but \$90 anyway. To B, the cat is a commodity. She takes a look at the *Weekly Town Shopper*, and finds that healthy adult cats go for about \$30, so that's what she bids. To C, the cat is a worthless burden. Of course, she doesn't tell her roommates that, but she discloses it to the referee, to make sure that she won't get saddled with the beast.

The referee decides to assign the cat to A, and to ask A to pay \$10 to B, and \$5 to C. It is easy to check that this is a fair arrangement in the sense of our definition (see Exercise 13.5).

However, C envies B, since B gets \$10 while C only gets \$5. (You might argue that that isn't a bad thing, since the cat is worthless to C anyway, but that is a different matter.) Furthermore, B envies A: In B's eyes, A gets the equivalent of \$30 (namely the cat, which in B's opinion is worth \$30), minus \$15 (the sum of the two compensation amounts). So in B's eyes, A gets the equivalent of \$15, whereas B only gets \$10. Although C envies B and B envies A, C does *not* envy A (envy is not "transitive"), for in C's eyes, the cat is worthless, so A simply suffers a loss of \$15 in C's eyes. ■

Definition 13.12 (envy-free compensation). *A compensation arrangement is called envy-free if nobody envies anybody else for what they receive.*

Example 13.13. A, B, C, and D have inherited a house. Their valuations of the house are a, b, c, and d dollars. A is assigned ownership of the house. To avoid envy among B, C, and D, it will be necessary for all of them to obtain the same compensation amount; let us call this amount \$$x$. To make sure that B does not envy A, we must make sure that

$$x \geq b - 3x.$$

The left-hand side of this inequality represents B's payout. The right-hand side represents the value of what A receives, in B's eyes. Solving the inequality for x, we find

$$x \geq \frac{b}{4}. \tag{13.3}$$

Similarly, to make sure that C and D do not envy A, we need

$$x \geq \frac{c}{4} \tag{13.4}$$

and

$$x \geq \frac{d}{4}. \tag{13.5}$$

To make sure that A does not envy any of the others, we need

$$a - 3x \geq x.$$

The left-hand side of this inequality represents A's payout. The right-hand side represents the cash amount that any of the others receives. Solving the inequality for x, we find

$$x \leq \frac{a}{4}. \tag{13.6}$$

Inequalities (13.3)–(13.6) can hold at the same time only if $a \geq b$, $a \geq c$, and $a \geq d$—in other words, if A is a highest bidder. We have thus concluded that a compensation arrangement is envy-free if and only if the winning bidder is a highest bidder, and all others receive equal compensation amounts that are fair to all of them (inequalities (13.3)–(13.5)) but do not exceed the fair share of the winning bidder (inequality (13.6)).

We observe now that inequality (13.6) is equivalent to

$$a - 3x \geq \frac{a}{4} \tag{13.7}$$

(see Exercise 13.6). Inequality (13.7) expresses that A is treated fairly: The left-hand side represents what A receives, in her own eyes, and the right-hand side represents her fair share. This leads us to another way of characterizing the envy-free compensation arrangements: A compensation arrangement is envy-free if and only all compensation amounts are equal, and the arrangement is fair. ■

Although this was just an example, it is clear that the arguments are general. We therefore arrive at the following conclusion.

Proposition 13.14. *For a compensation arrangement, the following three statements are equivalent to each other.*

1. *The arrangement is envy-free.*
2. *The winning bidder is a highest bidder, and all others receive equal compensation amounts not above the fair share of the winning bidder, but also not below the fair shares of any of the others.*
3. *All compensation amounts are equal to each other, and the arrangement is fair.*

Corollary 13.15. *For $N = 2$, fairness and envy-freeness are equivalent.*

Proof. This follows immediately from part 3 of Proposition 13.14. □

For $N > 2$, Proposition 13.14 shows that envy-freeness is a far more restrictive condition than fairness: It is equivalent to fairness with the extra requirement that all compensation amounts be equal. We will illustrate this point by a geometric picture.

Assume first that $N = 2$, $a > b$, and A is assigned ownership of the object. (Otherwise there would be no envy-free arrangement at all, by Proposition 13.14.) A pays $\$x_B$ to B as compensation. The arrangement is fair if and only if

$$a - x_B \geq \frac{a}{2} \tag{13.8}$$

and

$$x_B \geq \frac{b}{2}. \tag{13.9}$$

The left-hand sides of (13.8) and (13.9) represent payouts, and the right-hand sides fair shares. Inequalities (13.8) and (13.9) can be summarized as

$$\frac{b}{2} \leq x_B \leq \frac{a}{2}.$$

We call the interval from $b/2$ to $a/2$ the *fairness interval*. Any choice of x_B in this interval leads to a fair arrangement, while choices of x_B outside the interval lead to unfair arrangements. Envy-freeness is the same as fairness for $N = 2$ by Corollary 13.15, so the fairness interval is also the interval of envy-freeness.

Let us now assume that $N = 3$, $a > b > c$, and again A is assigned ownership of the object. (Otherwise there would be no envy-free arrangement, by Proposition 13.14.) A pays compensation amounts $\$x_B$ to B and $\$x_C$ to C. Here fairness means

$$a - x_B - x_C \geq \frac{a}{3}, \tag{13.10}$$

$$x_B \geq \frac{b}{3}, \tag{13.11}$$

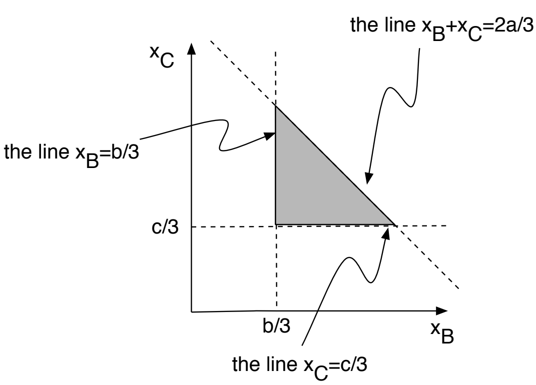
and

$$x_C \geq \frac{c}{3}. \tag{13.12}$$

The left-hand sides of these three inequalities represent payouts, and the right-hand sides fair shares. Inequality (13.10) can equivalently be written like this:

$$x_B + x_C \leq \frac{2}{3}a. \tag{13.13}$$

Inequalities (13.11)–(13.13) describe a triangle in the (x_B, x_C)-plane, which we refer to as the *fairness triangle*:



Any point in the interior or on the boundary of this triangle corresponds to a choice of x_B and x_C which yields a fair arrangement.

To understand why a pair (x_B, x_C) satisfies inequalities (13.11)–(13.13) if and only if it lies in the interior or on the boundary of the triangle, think about the geometric meanings of each of the inequalities individually. The points satisfying (13.11) lie on the vertical line $x_B = b/3$, or to its right. The points satisfying (13.12) lie on or above the horizontal line $x_C = c/3$. To understand the geometric meaning of inequality (13.13), first think about the equation

$$x_B + x_C = \frac{2}{3}a,$$

or, equivalently,

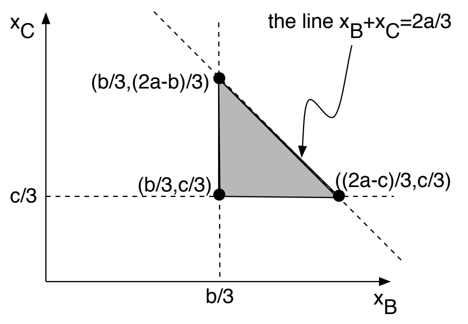
$$x_C = -x_B + \frac{2}{3}a. \tag{13.14}$$

This is the equation of a straight line in the (x_B, x_C)-plane with slope -1, intersecting the x_B-axis at $x_B = 2a/3$, and the x_C-axis at $x_C = 2a/3$. (If the notation confuses you, write "x" for x_B and "y" for x_C.) The *inequality* (13.13) is equivalent to

$$x_C \le -x_B + \frac{2}{3}a.$$

Therefore the pairs (x_B, x_C) satisfying inequality (13.13) correspond to points *on or below* the straight line described by (13.14).

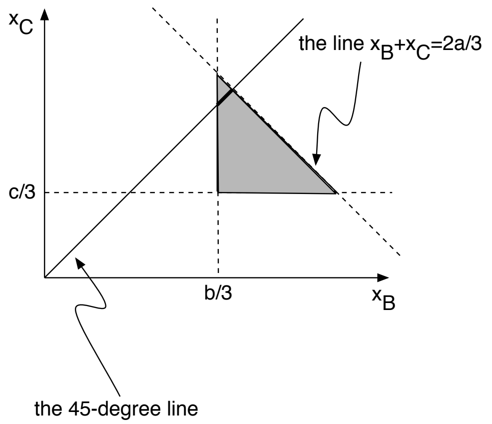
The coordinates of the vertices of the fairness triangle are easy to calculate (see Exercise 13.7):



The left lower vertex is $(b/3, c/3)$. The second coordinate of this point, $c/3$, is smaller than the first coordinate, $b/3$. The left lower vertex therefore lies *below* the 45-degree line. The left upper vertex is $(b/3, (2a-b)/3)$. The second coordinate of this point, $(2a-b)/3$, is greater than the first coordinate, $b/3$:

$$\frac{2a-b}{3} > \frac{b}{3}, \quad \text{since} \quad \frac{2a}{3} > \frac{2b}{3}, \quad \text{since} \quad a > b.$$

The left upper vertex therefore lies *above* the 45-degree line. Thus the 45-degree lines cuts through the fairness triangle like this:



The segment of the 45-degree line that lies inside the fairness triangle—indicated in bold in the figure—represents the fair arrangements with $x_B = x_C$, that is, the envy-free arrangements (Proposition 13.14).

> If $N = 4$, $a > b > c > d$, and the object is assigned to A, so that an envy-free arrangement is possible, the set of choices x_B, x_C, and x_D yielding a fair arrangement is a tetrahedron in three-dimensional space, which we call the *fairness tetrahedron*. The choices yielding envy-freeness correspond to the line segment inside this tetrahedron defined by $x_B = x_C = x_D$.

Exercises

13.1. Return to Example 13.1. Describe all fair arrangements in this example.

13.2. A, B, C, and D have inherited a house. Their valuations of the house are $a = 400{,}000$, $b = 360{,}000$, $c = 340{,}000$, and $d = 320{,}000$ dollars. (a) Propose a fair compensation arrangement in which A receives ownership of the house. (b) Propose a fair compensation arrangement in which B receives ownership of the house. (c) Explain why there is no fair compensation arrangement in which C receives ownership of the house. (d) Propose two different envy-free compensation arrangements.

13.3. Write out the proof of Proposition 13.10 for a case when $N = 5$ and A is the winning bidder.

13.4. (a) Prove that inequality (13.2) is equivalent to

$$c \geq \frac{a+b+d}{3}.$$

(b) Could the average bid m be replaced by the *average over all bids other than the winning bid* in Proposition 13.10?

13.5. Convince yourself that the arrangement that the referee decides upon in Example 13.11 is fair.

13.6. Explain why inequalities (13.6) and (13.7) are equivalent.

13.7. Explain why the vertices of the fairness triangle are $(b/3,(2a-b)/3)$, $(b/3,c/3)$, and $((2a-c)/3,c/3)$.

Chapter 14

Pareto-Optimality and Equitability

There are usually infinitely many possible fair compensation arrangements, and even infinitely many envy-free ones. Among these, which are best? It is, of course, not even clear what "best" should mean here. Best for whom? In this chapter, we discuss a notion of optimality which does make sense in this context, called Pareto-optimality after the Italian economist Vilfredo Pareto (1848–1923). We also think about the notion of equitability, one idea that can be used to derive specific fair compensation arrangements.

Definition 14.1 (Pareto-optimal compensation arrangement). *An arrangement is called Pareto-optimal if there is no alternative arrangement that raises at least one payout without lowering any of the others.*

We call an alternative arrangement raising at least one of the payouts without lowering any of the others an *objective improvement* of the original arrangement. Thus Pareto-optimal arrangements cannot be objectively improved upon. In later chapters, when we will think about fair division, Pareto-optimality will prove to be a subtle notion. Here, however, it is not at all subtle.

Proposition 14.2. *A compensation arrangement is Pareto-optimal if and only if the winning bidder is a highest bidder.*

Proof. To simplify the notation, we will give the proof for $N = 2$ only and assume $a > b$. It will be evident, however, that the argument works in general. So A and B jointly own an object, which we will refer to as a "house" to be concrete. They now get divorced, and therefore need to decide who should take ownership of the house, and how that person should compensate the other.

Let us first think about an arrangement in which A obtains the house and pays some amount \$$x$ to B. Could we objectively improve upon this arrangement? If we still assigned the house to A, but changed the payout amount x, then either x_A would be lowered (if x were raised), or x_B would be lowered (if x were lowered); therefore a change in x cannot yield an objective improvement.

If, on the other hand, we assigned the house to B, the sum of the payouts would decrease (Lemma 13.5), and therefore at least one of the payouts would decrease. Thus there is no objective improvement upon an arrangement in which A receives the house and pays \$$x$ to B; any such arrangement is Pareto-optimal.

Now let us think about an arrangement in which B obtains the house and pays some amount \$$x$ to A. Then $x_A = x$ and $x_B = b - x$. If instead we assigned the house to A and asked A to pay \$$(b-x)$ to B, then B would obtain the same payout as before, but A's payout would rise to $a-(b-x) = x + (a-b) > x$ dollars. Thus any arrangement in which B obtains the house can be objectively improved upon and is therefore not Pareto-optimal. □

For instance, the referee could simply assign the object to a highest bidder, without requiring any compensation to be paid at all. That would be Pareto-optimal, for anything one might do to improve the lots of the losing bidders would be to the detriment of the winning bidder; this is truly a "zero sum game." The word *optimal* is perhaps a bit misleading here: A Pareto-optimal division need not be, in any fair-minded person's eyes, a *good* division.

Now we turn to an idea which will allow us to calculate specific compensation amounts.

Definition 14.3 (equitable compensation). *A compensation arrangement is equitable if each person answers the question "Which fraction of the total value of the object does your payout represent?" the same way, that is,*

$$\frac{x_A}{a} = \frac{x_B}{b} = \cdots .$$

Example 14.4. Amelia (A) and Brad (B) own a beautiful grand piano. They both love the piano. Alas, they do not love each other, and therefore get divorced. They need to decide who should get the grand piano, and how that person should compensate the other. A values the piano at \$18,000, and B at \$24,000. To make the arrangement Pareto-optimal, B should get the piano and should compensate A. If B pays \$10,000 to A, is that equitable? In A's eyes, the fraction of the total value of the piano that A receives is

$$\frac{10,000}{18,000} = \frac{5}{9}.$$

This is greater than $1/2$, and therefore A is certainly treated fairly. In B's eyes, the fraction of the total value of the piano that B receives is

$$\frac{24,000 - 10,000}{24,000} = \frac{14,000}{24,000} = \frac{7}{12}.$$

This, too, is greater than $1/2$, so B is treated fairly as well. However,

$$\frac{5}{9} < \frac{7}{12}$$

(since $60 < 63$), and therefore the arrangement is not equitable. ■

We will illustrate a general method for calculating equitable compensation amounts with several examples.

Example 14.5. We return to Amelia, Brad, and their piano; see Example 14.4. B gets the piano, but pays \$$x_A$ to A. For equitability, we want

$$\frac{x_A}{18,000} = \frac{x_B}{24,000}.$$

Since $x_B = 24,000 - x_A$, we can write this equation in the form

$$\frac{x_A}{18,000} = \frac{24,000 - x_A}{24,000}. \tag{14.1}$$

This is an equation for x_A, and we could easily solve it right away. However, we will instead proceed in a slightly indirect manner. Our method for solving (14.1) is unnecessary when the number N of owners is just 2, as in this example, but it will prove useful when $N > 2$.

We denote the common value of the two quantities in (14.1) by q, so

$$\frac{x_A}{18,000} = q, \tag{14.2}$$

$$\frac{24,000 - x_A}{24,000} = q. \tag{14.3}$$

We seem to have made things more complicated, not simpler: We started out with a single equation in the unknown x_A, and now we have a coupled system of two equations in the two unknowns x_A and q. However, (14.2) and (14.3) are easy to solve if you observe that the sum of the numerators on the left-hand side is just 24,000. To take advantage of this observation, we isolate the numerators by multiplying the first equation by 18,000 and the second by 24,000:

$$x_A = 18,000q,$$
$$x_A - 24,000 = 24,000q,$$

where q is the same number in both equations. Now we add the two equations together. This yields a single equation for q:

$$24,000 = 42,000q,$$

or

$$q = \frac{24,000}{42,000} = \frac{4}{7}.$$

Therefore

$$x_A = \frac{4}{7} \times 18,000 \approx 10,285.71.$$

With this compensation amount, both A and B feel that they get $4/7$ of the value of the piano, about 57%. ∎

Example 14.6. Four pirates, A, B, C, and D, have stolen a bag of jewelry. They will assign the bag to one of them, who will have to compensate the others. They are in a rush, so they can't have the value of the jewelry assessed. Instead each guesses what the value of the jewelry might be. Their guesses are $a = 40,000$, $b = 36,000$, $c = 42,000$, and $d = 46,000$ dollars, respectively. The pirates' captain, A, decides he should get the bag. (He does not realize that this is not Pareto-optimal, and that all, including himself, could be better off if they gave the bag to D and asked D to compensate the others.)

Displaying an exceptional degree of sophistication by pirates' standards, they decide to choose compensation amounts x_B, x_C, and x_D that will make the arrangement equitable:

$$\frac{40,000 - x_B - x_C - x_D}{40,000} = q, \tag{14.4}$$

$$\frac{x_B}{36,000} = q, \tag{14.5}$$

$$\frac{x_C}{42,000} = q, \tag{14.6}$$

$$\frac{x_D}{46,000} = q, \tag{14.7}$$

where q is the same number for all four equations. This is a system of four coupled equations in the four unknowns x_B, x_C, x_D, and q. To solve these equations, we will use the same trick by which we solved (14.2) and (14.3). The key observation is that the sum of the numerators in (14.4)–(14.7) is just 40,000. Therefore we isolate the numerators, then sum the equations. Multiplying each equation by the denominator on the left-hand side, we get

$$\begin{aligned} 40,000 - x_B - x_C - x_D &= 40,000q, \\ x_B &= 36,000q, \\ x_C &= 42,000q, \\ x_D &= 46,000q. \end{aligned}$$

Summing these four equations, we find

$$40,000 = (40,000 + 36,000 + 42,000 + 46,000)q,$$

or

$$\begin{aligned} q &= \frac{40,000}{40,000 + 36,000 + 42,000 + 46,000} \\ &= \frac{40}{40 + 36 + 42 + 46} = \frac{40}{164} = \frac{10}{41}. \end{aligned} \tag{14.8}$$

Thus the compensation amounts will be

$$\begin{aligned} x_B &= \frac{10}{41} \times 36,000 \approx 8,780.49, \\ x_C &= \frac{10}{41} \times 42,000 \approx 10,243.90, \\ x_D &= \frac{10}{41} \times 46,000 \approx 11,219.51 \end{aligned}$$

dollars. After doing this calculation, unfortunately the pirates realize two problems. First,

$$q = \frac{10}{41} < \frac{1}{4},$$

so none of them gets a fair share. As we will explain shortly, this is a consequence of the captain's decision to take ownership of the jewelry even though he is a fairly low bidder. Second, B envies C and C envies D. ∎

Example 14.7. Now we do a very similar calculation with letters rather than numbers. Suppose that A, B, and C have inherited a house, which they value at a, b, and c dollars, respectively. Let us assume that the house is assigned to A, and A pays \$$x_B$ to B, and \$$x_C$ to C. (We do not necessarily assume that A is a highest bidder.) Equitability means

$$\frac{x_A}{a} = \frac{x_B}{b} = \frac{x_C}{c}.$$

Since $x_A = a - x_B - x_C$ this condition becomes

$$\frac{a - x_B - x_C}{a} = \frac{x_B}{b} = \frac{x_C}{c}. \tag{14.9}$$

We denote by the letter q the common value of the three quantities in (14.9). Thus

$$\frac{a - x_B - x_C}{a} = q, \tag{14.10}$$

$$\frac{x_B}{b} = q, \tag{14.11}$$

$$\frac{x_C}{c} = q. \tag{14.12}$$

Equations (14.10)–(14.12) should be thought of as a system of three equations in three unknowns, x_B, x_C, and q. (The valuations, a, b, and c, should be thought of as given here.) The key to solving this system is to observe that the sum of the numerators in (14.10)–(14.12) is just a. We therefore isolate the numerators, then sum the equations. To isolate the numerators, we multiply each equation by the denominator of its left-hand side:

$$a - x_B - x_C = qa, \tag{14.13}$$

$$x_B = qb, \tag{14.14}$$

$$x_C = qc. \tag{14.15}$$

Summing (14.13)–(14.15), we find

$$a = q(a+b+c),$$

or

$$q = \frac{a}{a+b+c}.$$

Inserting this formula into (14.14) and (14.15), we find

$$x_B = \frac{ab}{a+b+c} \tag{14.16}$$

and

$$x_C = \frac{ac}{a+b+c}. \tag{14.17}$$

■

After reading Examples 14.5 through 14.7, you will understand that in general the equitable division can be described as follows: Each person's payout is w/S times his or her bid. (See, for instance, (14.8), or (14.16) and (14.17).) Recall that $\$w$ is the winning bid and $\$S$ the sum of all bids.) For a person whose bid is $\$a$, the payout is

$$\frac{w}{S}a,$$

which can also be written as

$$\frac{w}{S/N}\frac{a}{N} = \frac{w}{m}\frac{a}{N}.$$

Thus each person's payout is w/m times his or her fair share. Now that we know that, it is very simple to calculate equitable compensation arrangements.

Example 14.8. A, B, C, and D have inherited a house. It is worth \$300,000 to A, \$270,000 to B, \$330,000 to C, and \$240,000 to D. The siblings decide that A should have the house, but should equitably compensate B, C, and D. To compute what this means, we need to compute the average bid:

$$m = \frac{300,000+270,000+330,000+240,000}{4} = 285,000.$$

The compensation paid to B, C, and D should be their fair shares, multiplied by

$$\frac{w}{m} = \frac{300,000}{285,000} = \frac{300}{285} = \frac{60}{57} = \frac{20}{19}.$$

The fair shares of B, C, and D are \$67,500, \$82,500, and \$60,000, respectively. Thus

$$x_B = \frac{20}{19} \times 67,500 \approx 71,052.63,$$
$$x_C = \frac{20}{19} \times 82,500 \approx 86,842.11,$$
$$x_D = \frac{20}{19} \times 60,000 \approx 63,157.89.$$

It will automatically be true that also

$$x_A = \frac{20}{19} \times 75,000.$$

This arrangement is fair because $20/19 > 1$. ■

In general, since each person's payout is w/m times his or her fair share, the equitable arrangement is fair if and only if $w/m \geq 1$, i.e., $w \geq m$.

Proposition 14.9. *An equitable compensation arrangement is fair if and only if the winning bidder is at least an average bidder:* $w \geq m$.

For the equitable arrangement to be envy-free, the bids of all those who are compensated monetarily must be equal to each other, since they receive q times their bid as compensation. Equitable compensation arrangements therefore typically create envy. However, one might argue that this is not so bad; perhaps low bidders *should* in fact be compensated less than high bidders.

Exercises

14.1. A and B jointly own a car. When A moves to another part of the country, B does not want to follow her there, and therefore they must decide who should get the car, and how much that person should compensate the other. A values the car at \$8,000, whereas B values it at \$10,400. (a) We want the arrangement to be Pareto-optimal. What does that tell us about what should be done? For parts (b) and (c) assume that B, who values the car more highly, gets it, and pays \$$x$ to A. (b) We want the arrangement to be fair. What does that tell us about x? (c) We want the arrangement to be equitable. What should x be?

14.2. Two brothers, A and B, find an old musical manuscript in the attic of the house that they own jointly. It is a movement of a string quartet by Beethoven. They have the manuscript examined by an expert. It proves to be original, in Beethoven's own hand. The expert thinks it could fetch about \$200,000 at an auction. A and B now jointly own it. To A, a musician, the manuscript is worth anything in the world, or, on second thought, \$500,000 anyway. On the other hand, B barely knows who Beethoven was. To him, the manuscript is worth whatever it can fetch at the auction—\$200,000. They decide that A should have the manuscript, but compensate B monetarily. To make the arrangement equitable, what should be the compensation amount?

14.3. Four siblings, A, B, C, and D, have inherited a vacation cottage, which they value at $a = 80{,}000$, $b = 60{,}000$, $c = 100{,}000$, and $d = 80{,}000$ dollars, respectively. They wish to assign ownership to one among them and have that person compensate the others. (a) To make the arrangement Pareto-optimal, who should get the cottage? (b) If A gets the cottage, how should x_B, x_C, and x_D be chosen to make the arrangement equitable? Is this a fair arrangement?

14.4. Three siblings, A, B, and C, have inherited a house, which they value at $a = 330{,}000$, $b = 300{,}000$, and $c = 240{,}000$ dollars, respectively. They decide to assign ownership of the house to B, but ask B to compensate A and C equitably. (a) What are the compensation amounts? (b) Is this arrangement fair? (c) Describe an objective improvement of this arrangement.

14.5. Two people, A and B, jointly own a house. They value it at \$$a$ and \$$b$, respectively. (a) They want to transfer ownership of the house to A, who should then compensate B equitably. Prove that the compensation amount should be

$$x_B = \frac{ab}{a+b}$$

dollars. (b) They want to transfer ownership of the house to B, who should then compensate A equitably. Prove that the compensation amount should be

$$x_A = \frac{ab}{a+b}$$

dollars. Note that the compensation amounts are the same in both cases. Of course, the difference is that in the first case, A compensates B, and in the second B compensates A.

Chapter 15

Equality, Equitability, and Knaster's Procedure

In the previous chapter, we studied equitability, a guiding principle that leads to specific fair compensation arrangements. A very different guiding principle, equality, is discussed in this chapter.

Definition 15.1 (equal compensation). *A compensation arrangement is called equal if all payouts are equal to each other.*

Example 15.2. Arnold (A) and his brother Bert (B) inherit their grandfather's motorcycle, built in 1962 but in reasonably good shape. To A, the motorcycle is worth \$2000. On the other hand, to B, who has no place to park it other than his friend's living room, it is worth only \$600. The two brothers therefore decide that A should have it, but pay compensation to B. Any compensation amount between \$300 and \$1000 would be fair. If the compensation amount is \$$x$, then $x_A = 2000 - x$, and $x_B = x$. If these two are to be equal to each other, we need

$$2000 - x = x,$$

so

$$2000 = 2x$$

or $x = 1000$. ∎

Example 15.3. A, B, C, and D have inherited a house. Their valuations of the house are $a = 300{,}000$, $b = 320{,}000$, $c = 280{,}000$, and $d = 260{,}000$ dollars. In one possible equal arrangement, the house would be given to D, who would then pay \$65,000 to A, B, and C. This would be equal, but not fair: A, B, and C would not get fair shares. Another equal arrangement would be to give the house to B, and have her pay \$80,000 to A, C, and D. This would be equal and fair at the same time. ∎

In general, in an equal compensation arrangement the payouts are all equal to \$$w/N$. This implies the following proposition (see Exercise 15.1).

Proposition 15.4. *An equal compensation arrangement is envy-free if the winning bidder is a highest bidder, and is unfair otherwise.*

We refer to the smallest of the payouts x_A, x_B, ... as the *minimal payout*. We can then ask how we should design a compensation arrangement to make the minimal payout as large as possible. Here is the answer.

Proposition 15.5. *A compensation arrangement maximizes the minimal payout if and only if it is equal, and the winning bidder is a highest bidder.*

Proof. The minimal payout cannot be greater than $\$w/N$, otherwise the average payout would have to be greater than $\$w/N$, but by Lemma 13.5 it is equal to $\$w/N$. The equal compensation arrangement makes all payouts equal to $\$w/N$, and therefore yields the greatest possible minimal payout, given that the winning bid is w. The greatest possible minimal payout overall is achieved by the equal arrangement when w is as large as possible, that is, when the winning bidder is a highest bidder. $\square$

Proposition 15.5 offers an argument in favor of using equal compensation arrangements in which the winning bidder is a highest bidder, i.e., Pareto-optimal equal arrangements.

We will now describe another procedure for calculating specific fair compensation arrangements, called Knaster's procedure in honor of the Polish mathematician Bronislaw Knaster (1893–1990). Knaster belonged to a group of distinguished mathematicians who worked at the University of Lwów (then Poland, now Ukraine) during the years preceding World War II. Members of this group, including Bronislaw Knaster, Hugo Steinhaus (1887–1972), and Stefan Banach (1892–1945), were among the first mathematicians with an interest in fair compensation and division.

Knaster's fair compensation procedure can be viewed as a compromise between equitability and equality. To motivate it, we first return to the notion of equitability. As Example 14.8 illustrates, the calculation of equitable compensation arrangements is utterly simple once one realizes that each person's payout is w/m times his or her fair share. Nonetheless, we will now describe a different procedure for calculating equitable arrangements, which we will call the *equitability procedure*. This procedure does not simplify the task of calculating equitable compensation arrangements, a task that hardly needs simplification. However, it clarifies the relation between equitable arrangements and arrangements computed by Knaster's procedure. In fact, Knaster's procedure is just a very slight variation on the equitability procedure.

Equitability Procedure

STAGE 1: The referee decides—in whichever way she likes—who should be the winning bidder, the person who should receive ownership of the object. As discussed earlier, the winning bidder will usually, but not always, be a highest bidder. Since the winning bidder believes that the object is worth $\$w$, but is entitled only to $1/N$ of it, the referee then asks the winning bidder to pay, in exchange for the object,

$$\$\left(w - \frac{w}{N}\right)$$

to the referee. Out of this amount, the referee pays each of the others their fair shares. This concludes the first stage. Everybody, including the winning bidder, has now received precisely their fair share.

Let us calculate how much cash the referee has left after the first stage. The sum of all the payments that the referee has made is the sum of all fair shares, which is S/N, minus the fair share of the winning bidder, which is w/N. The referee is therefore left with

$$w - \frac{w}{N} - \left[\frac{S}{N} - \frac{w}{N}\right] \tag{15.1}$$

dollars. Remembering that $S/N = m$, (15.1) can be simplified:

$$w - \frac{w}{N} - \left[\frac{S}{N} - \frac{w}{N}\right] = w - \frac{w}{N} - m + \frac{w}{N} = w - m.$$

So the left-over amount of cash is the discrepancy between the winning bid and the average bid.

STAGE 2: The referee distributes the left-over cash in such a way that each person's share is proportional to the person's bid: A receives

$$\frac{a}{S}(w-m)$$

dollars in cash in the second stage, B receives

$$\frac{b}{S}(w-m)$$

dollars, and so on. This works because the sum of all of these amounts is $\$(w-m)$, the left-over amount of cash:

$$\frac{a}{S}(w-m) + \frac{b}{S}(w-m) + \cdots = \frac{a+b+\cdots}{S}(w-m) = \frac{S}{S}(w-m) = w-m\,.$$

If $w < m$, that is, if the winning bid is smaller than the average bid, then the referee spends more money than he gets from the winning bidder in the first stage, but recovers his losses by making "negative payments" in the second stage. (To make negative payments means to collect money instead of handing it out.) The method is fair if and only if $w \geq m$, so that the payments in the second stage are not negative.

Proposition 15.6. *The equitability procedure generates equitable arrangements.*

Proof. The payout of a person whose bid is $\$a$ equals

$$\frac{a}{N} + \frac{a}{S}(w-m) = \left(\frac{1}{N} + \frac{w-m}{S}\right)a = qa$$

dollars, with

$$q = \frac{1}{N} + \frac{w-m}{S}. \tag{15.2}$$

Since q is the same for all N people, this proves the assertion. $\square$

Knaster's Procedure

STAGE 1: As in the equitability procedure.

STAGE 2: The left-over amount of $\$(w-m)$ is distributed equally, with each person receiving $\$(w-m)/N$.

We refer to the compensation arrangement obtained using Knaster's procedure as *Knaster's arrangement*. It is a compromise between equitability and equality, in the sense that the first stage of the procedure is equitable, but in the second stage, the left-over cash is distributed equally.

Proposition 15.7. *Knaster's compensation arrangement is fair if and only if the winning bidder is at least an average bidder:* $w \geq m$.

Proof. Everybody's payout is his or her fair share plus $\$(w-m)/N$. Thus the arrangement is fair if and only if $(w-m)/N \geq 0$, which is equivalent to $w \geq m$. $\square$

Example 15.8. If we apply Knaster's procedure to Example 15.2, A pays \$1000 in the first stage, but B receives only \$300. The left-over amount of \$700 is distributed equally. Half of it is returned to A, and half of it goes to B. Thus altogether, B receives \$650 in compensation for the motorcycle. ■

Example 15.9. In Example 15.3, suppose that B, the highest bidder, is awarded the house. In the first stage, B must pay three-quarters of what he thinks the house is worth to the referee, that is, \$240,000. Out of this amount, the referee pays A, C, and D their fair shares: \$75,000 to A, \$70,000 to C, and \$65,000 to D. The referee is then left with

$$240{,}000 - 75{,}000 - 70{,}000 - 65{,}000 = 30{,}000$$

dollars, which are distributed equally among the four people. In effect, the compensation amounts are

$$\begin{aligned} x_A &= 75{,}000 + 30{,}000/4 = 82{,}500, \\ x_C &= 70{,}000 + 30{,}000/4 = 77{,}500, \\ x_D &= 65{,}000 + 30{,}000/4 = 72{,}500 \end{aligned}$$

dollars. ■

The procedures for computing the equitable arrangement and Knaster's arrangement differ only in the second stage, in which the left-over amount of $\$(w-m)$ is distributed. In the equitability procedure, a person whose bid is $\$a$ gets

$$\frac{a}{S}(w-m) \tag{15.3}$$

dollars in the second round, whereas in Knaster's procedure everybody gets

$$\frac{w-m}{N} \tag{15.4}$$

dollars in the second round. To exhibit the difference between (15.3) and (15.4), we re-write (15.3) as follows:

$$\frac{a}{S}(w-m) = \frac{a}{S/N}\,\frac{w-m}{N} = \frac{a}{m}\,\frac{w-m}{N}. \tag{15.5}$$

Comparing (15.5) and (15.4), one sees that they differ by the factor a/m. So in comparison with the equitable arrangement, Knaster's procedure favors low bidders and reduces the payouts of high bidders.

When all bids are close to each other, the equitable arrangement and Knaster's arrangement will be *very* close to each other. To see this, notice first that the two procedures differ only in how they distribute the left-over cash, $\$(w-m)$, in the second stage. When all bids are close to each other, the winning bid w will be close to the average bid m, so $w-m$ will be close to zero. But in addition, the factor a/m by which (15.5) and (15.4) differ will be close to 1 when all bids are close to each other. For another look at this point, see Exercise 15.5.

Example 15.10. We return to Example 15.3 one more time. If B gets the house,

$$\frac{w}{m} = \frac{320{,}000}{(300{,}000+320{,}000+280{,}000+260{,}000)/4} = \frac{32}{29},$$

so the equitable compensation amounts should be $32/29$ times the fair shares:

$$\begin{aligned}
x_A &= \frac{32}{29}\times 75{,}000 \approx 82{,}758.62,\\
x_C &= \frac{32}{29}\times 70{,}000 \approx 77{,}241.38,\\
x_D &= \frac{32}{29}\times 65{,}000 \approx 71{,}724.14.
\end{aligned}$$

These values are quite close to those obtained using Knaster's procedure (see Example 15.9). ∎

Exercises

15.1. Prove Proposition 15.4.

15.2. In Example 14.4, which compensation amounts does Knaster's procedure yield, assuming that B gets the piano?

15.3. In Example 14.6, which compensation amounts does Knaster's procedure yield, assuming that A gets the booty?

15.4. Convince yourself that q, as defined in (15.2), equals $\frac{1}{N}\frac{w}{m}$.

15.5. A and B jointly own an object. Their valuations of the object are $\$a$ and $\$b$, respectively. They decide to transfer ownership to one of them and have that person compensate the others. (a) They calculate the compensation amount using Knaster's procedure. Prove that the compensation amount is $\$(a+b)/4$. (b) They calculate an equitable compensation amount. Prove that that amount is \$$ab/(a+b)$. (c) Verify:

$$\frac{a+b}{4} - \frac{ab}{a+b} = \frac{(a-b)^2}{4(a+b)}.$$

This is the difference between the compensation amounts in parts a and b. When a and b are close to each other, then $(a-b)^2$ is very small. (The square of a small number is very small. For instance, the square of 0.1 is 0.01.)

Chapter 16

Envy-Free, Pareto-Optimal, and Equitable Cake Cutting

In the remaining chapters of the book, we will think about the problem of sharing a cake among N people. We assume that the cake can be divided into arbitrarily small pieces. This makes a cake different from many other objects, which cannot reasonably be divided into pieces—houses, cars, parrots, valuable paintings, and so on. For such objects, the monetary compensation methods of Chapters 13 and 15 may be useful, whereas "cake cutting" methods are not.

The subject would be utterly frivolous if it were really primarily about cake. But the word "cake" stands for a divisible resource, to be shared among several parties. Situations in which limited resources must be shared are, of course, ubiquitous. The "cake" might be a piece of land, for instance. The problem is deadly serious if the piece of land is, say, East Jerusalem; see Chapter 23 for a discussion of this example. The word "cake" might also stand for flight departure and arrival times at, say, JFK Airport, to be assigned to airlines,[9] or perhaps for classroom resources at a university, to be shared among departments, or for bandwidth (rate of data transfer) on the Internet,[10] or for medical resources.[11] All of these problems resemble, to some degree, the cake cutting problems discussed here. Of course the real problems have many complications from which we abstract here.

In kindergarten, children learn the simplest, although arguably not the best, way of fairly sharing a cake between $N = 2$ people: "I cut, you choose." The child who cuts has a strong incentive to generate two pieces that she considers of equal value, since she does not know which of the two pieces will become hers. The other child may not agree that what he sees are two pieces of equal value—but in that case, he is free to take the piece that he considers more valuable. Both children are guaranteed to get at least half the cake in their

---

[9] Michael Ball, Andrew Churchill, and David Lovell, "Schedule compression by fair allocation methods," Technical Report, National Center of Excellence for Aviation Operations Research (NEXTOR), November 2007.

[10] Rong Pan, Blaji Prabhakar, Lee Breslau, and Scott Shenker, "Approximate fair allocation of link bandwidth," *IEEE Micro* 23, issue 1, pages 36–43 (2003).

[11] American Thoracic Society Bioethics Taskforce, "Fair allocation of intensive care unit resources," *American Journal of Respiratory and Critical Care Medicine* 156, issue 4, pages 1282–1301 (1997). This paper is about principles guiding resource allocation, not about methods of resource allocation.

eyes. It is important to say "in their eyes," since tastes differ; what is worth half the cake to one may not be worth half the cake to the other. In general:

Definition 16.1 (fair cake division). *When $N \geq 2$ people share a cake, a fair division is one that assigns to everybody a share that is worth, in his or her eyes, at least the fraction $1/N$ of the total.*

In more sophisticated mathematical language: Identify the cake with a set $C \subseteq \mathbb{R}^3$, and assume that associated with the ith person is a probability density ρ_i on C so that any piece $P \subseteq C$ is thought by the ith person to represent the fraction $\int_P \rho_i(u)\,du$ of the entire cake. A fair division is then a partition $C = P_1 \cup P_2 \cup \cdots \cup P_N$ with $P_i \cap P_j = \{\,\}$ for $i \neq j$ so that

$$\text{for all } i \in \{1, \ldots, N\} \quad \int_{P_i} \rho_i(u)\,du \geq \frac{1}{N}\,.$$

We use this way of formalizing the cake cutting problem even though the additivity assumption underlying it is clearly not always correct. For instance, I would not enjoy receiving 100 slices of apple pie 100 times more than one slice.

We say that a person gets his *fair share* if he receives a share which represents, in his or her eyes, at least the fraction $1/N$ of the total value of the cake. When $N = 2$, "I cut, you choose" guarantees a fair division of the cake. We therefore call it a *fair division method.* It was proposed in the Bible:

> And Abram said unto Lot, let there be no strife, I pray thee, between me and thee, and between my herdmen and thy herdmen; for we be brethren. Is not the whole land before thee? Separate thyself, I pray thee, from me: if thou wilt take the left hand, then I will go to the right; or if thou depart to the right hand, then I will go to the left. (*Genesis* 13: 8–9, King James Version)

If A and B divide a cake fairly, A thinks that she gets *at least* half the cake, and therefore she must think that she gets at least as much as B. Likewise, B thinks that he gets at least half the cake, and therefore he thinks that he gets at least as much as A. So when two people share a cake, "fair" means the same as "envy-free." Here is the general definition of "envy-freeness."

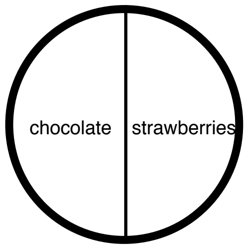
Definition 16.2 (envy-free cake division). *A cake division is called envy-free if nobody feels that he gets less than anybody else.*

In the language of the previous paragraph in small print, envy-freeness means

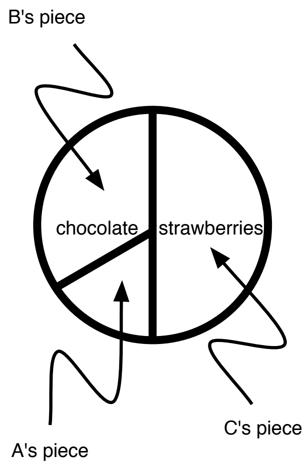
$$\text{for all } i, j \in \{1, \ldots, N\} \quad \int_{P_i} \rho_i(u)\,du \geq \int_{P_j} \rho_i(u)\,du.$$

Although fairness and envy-freeness are the same for $N = 2$, envy-freeness is a much stronger requirement than fairness when $N > 2$. Any envy-free division is fair (see Exercise 16.3), but a fair division need not be envy-free, as the following example shows.

Example 16.3. Suppose that the cake has a chocolate half and a strawberry half:



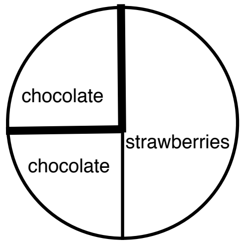
A, B, and C share this cake. A and B are allergic to strawberries; to them, the chocolate component is the *entire* cake. On the other hand, C does not like chocolate at all; to him, the strawberry component is the entire cake. Suppose we divide the cake like this:



Then A thinks that she gets one-third of the cake, by value—namely, one-third of the chocolate half, which to her is all that is of value. Similarly, B thinks that he gets two-thirds of the cake, and C thinks that she gets the entire cake. This is a fair division, since everybody thinks that he or she gets at least one-third of the cake. But it is not envy-free: A thinks that B gets two-thirds of the cake, more than her piece. In this example, this is the only envy that arises. For instance, A does not envy C: In A's eyes, C gets nothing (of value) at all! The example shows that for more than two people, a fair division may not be envy-free. ∎

A division *method* is called envy-free if it guarantees, in all cases, envy-free divisions. For $N = 2$, any fair method is envy-free. For $N > 2$, it is not easy to construct envy-free division methods; see Chapters 20 and 26.

Example 16.4. Let's return to the "I cut, you choose" method for two people. Let's think about the cake of Example 16.3. Suppose that two people A and B want to share this cake. A likes chocolate, but is allergic to strawberries—to her, the chocolate component is the entire cake. B likes strawberries, but does not eat chocolate—to him, the strawberry component is the entire cake. Unfortunately, they don't know this about each other. When A cuts the cake, she must cut it into two halves that are equal in her eyes, otherwise she runs the risk of getting a piece that she does not consider a fair share. If she is considerate, she will cut like this:



B will choose the piece that includes the entire strawberry component, and think that he gets the entire cake (all that is of value to him). A will be left with half the chocolate component, and will feel that that is half the cake. So in some sense, A has no reason to complain—in her eyes she gets her fair share, half the cake. But of course it would have been better to give her the *entire* chocolate component, since B will simply discard the piece of chocolate cake that he gets. ∎

Definition 16.5 (Pareto-optimal cake division). *We say that a cake division is Pareto-optimal if there is no alternative division that satisfies at least one person more without satisfying anybody less.*

We call an alternative division that satisfies at least one person more without satisfying anybody less an *objective improvement*. Thus a Pareto-optimal cake division is one that cannot be objectively improved upon. The case for Pareto-optimality is compelling: If it is possible to satisfy one person more without satisfying anybody less, why wouldn't one? A division *method* is called Pareto-optimal if it guarantees, in all cases, Pareto-optimal divisions. Example 16.4 shows that "I cut, you choose" is not Pareto-optimal.

Although it is clearly desirable to use a Pareto-optimal division, many divisions that you would find entirely unacceptable are Pareto-optimal. For instance, if to A, no part of the cake is entirely worthless, it is Pareto-optimal to give the entire cake to A, leaving the others empty-handed: To improve the lots of the empty-handed, you will have to take something away from A! This example shows that Pareto-optimality by itself does not imply fairness.

Lack of Pareto-optimality is not the only undesirable property of "I cut, you choose." The divider is sure to get exactly half the cake, in her eyes, but the chooser may get more than half the cake, in his eyes. So the chooser is better off than the divider. Abraham was generous when he volunteered to be the divider. We say that the division generated by "I cut, you choose" is (usually) not *equitable*. In general:

Definition 16.6 (equitable cake division). *A cake division among N people is called equitable if each of the N people gives the same answer to the question, "Which fraction of the total cake do you think you are getting?"*

A division *method* is called equitable if it guarantees, in all cases, equitable divisions. Example 16.4 shows that "I cut, you choose" is not equitable. We summarize the properties of "I cut, you choose" in the following proposition.

Proposition 16.7. *"I cut, you choose" is fair (and therefore, since $N = 2$, envy-free), but neither Pareto-optimal nor equitable.*

In Chapters 21–26, we will restrict consideration to cakes that consist of finitely many components that are (or can be thought of as) *homogeneous*, like the chocolate and strawberry components in Examples 16.3 and 16.4. We call such cakes *piecewise homogeneous*.

In the language introduced in the previous notes in small print, this assumption means that the ρ_i are step functions. Since any probability density can be approximated, in L^1, arbitrarily well by step functions, the assumption of piecewise homogeneity seems fairly minor.

Phrases like "one-third of the chocolate component" or "half the strawberry component" are defined *objectively*. You and I may disagree on whether or not coffee tastes good, but we can't very well disagree on the meaning of the phrase "half a cup of coffee" (assuming we agree on the definition of a "cup"). We say that the homogeneous components are *objectively divisible*. In general, we will denote the number of homogeneous (objectively divisible) components by n; in Examples 16.3 and 16.4, $n = 2$.

For a piecewise homogeneous cake, there is a very simple way of creating a division among N of people.

Definition 16.8 (equal division of a piecewise homogeneous cake). *For a piecewise homogeneous cake, an equal division is one in which each of the N people sharing the cake receives the fraction $1/N$ of each of the homogeneous components.*

However, it is easy to see that an equal division can be far from Pareto-optimal (see Exercise 16.4). In summary:

Proposition 16.9. *An equal division of a piecewise homogeneous cake is envy-free and equitable, but not Pareto-optimal.*

We have discussed three fundamental desirable properties of fair cake divisions: envy-freeness, Pareto-optimality, and equitability. There is a strong case for envy-freeness, and I think there is a compelling case for Pareto-optimality. Equitability seems to me to be the least important of these three properties. If I am one of the N people sharing the cake, I cannot tell whether the division is equitable unless I know the tastes of the other people. A nonequitable division will bother me only if it bothers me to see the other people greatly enjoy their shares, but why should it? For example, my neighbors have a gigantic dog named Percy, a Newfoundland. To me it would be intolerable to have Percy in my house. First, I am allergic, and second, I am also slightly afraid of dogs. But does it pain me to see my neighbors enjoy their dog, who is actually a very cute creature from a distance? Of course not, on the contrary.

Envy-freeness implies fairness. Pareto-optimality and equitability, together, also imply fairness.

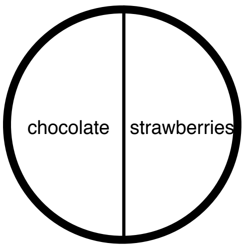
Proposition 16.10. *Any Pareto-optimal, equitable division of a piecewise homogeneous cake is fair.*

Proof. See Exercise 16.5. □

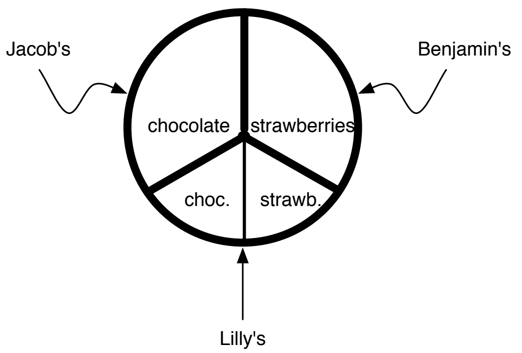
Pareto-optimality by itself does not imply fairness; for instance, as pointed out earlier, if A attaches some value to every part of the cake, then it is Pareto-optimal to give the whole cake to A, and give nothing to the others, but of course that is not fair. Equitability by itself does not imply fairness either (see Exercise 16.6).

Exercises

16.1. Benjamin, Jacob, and Lilly share a cake consisting of a chocolate half and a strawberry half:

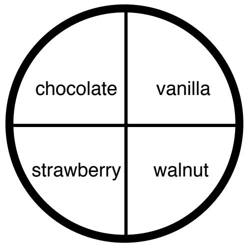


Benjamin likes strawberry cake twice as much as chocolate cake. Jacob likes chocolate cake twice as much as strawberry cake. Lilly goes for volume—she has no preference between the two halves. The cake is divided as follows:

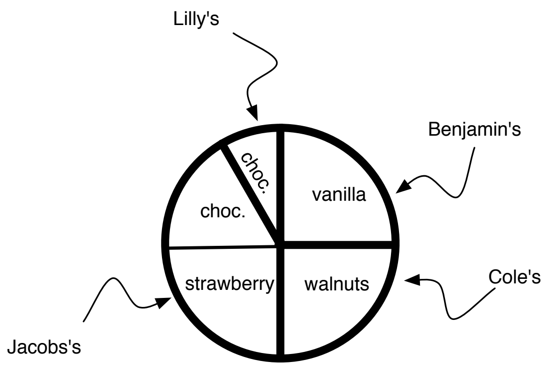


Jacob gets two-thirds of the chocolate half, Benjamin gets two-thirds of the strawberry half, and Lilly gets one-third of each half. (a) Is this fair? (b) Is it envy-free? (c) Is it Pareto-optimal? (d) Is it equitable?

16.2. Benjamin, Jacob, Lilly, and Cole share a cake consisting of four homogeneous components: chocolate, vanilla, strawberry, and walnut cake.



Benjamin likes only the vanilla piece. Jacob likes the chocolate and strawberry components equally, but does not like vanilla or walnuts at all. Lilly likes nothing other than the chocolate piece. Cole likes strawberries and walnuts equally, but does not like chocolate or vanilla at all. We give to Benjamin the entire vanilla component, to Jacob two-thirds of the chocolate and the entire strawberry component, to Lilly one-third of the chocolate component, and to Cole the entire walnut component:



(a) Is this fair? (b) Is it envy-free? If not, list all incidents of envy. (c) Is it equitable?

16.3. Explain why for any number N of people sharing a cake, any envy-free division is fair.

16.4. Give an example to prove that the equal cake division is not in general Pareto-optimal.

16.5. Prove: Any Pareto-optimal, equitable division of a piecewise homogeneous cake is fair. (Hint: Compare with the equal division.)

16.6. Give an example of an equitable division that is unfair.

Chapter 17

"I Cut, You Choose" for Three: Steinhaus's Method

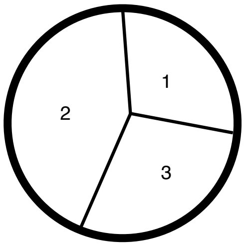
Hugo Steinhaus was perhaps the first mathematician who thought about the problem of generalizing "I cut, you choose" to more than two people. In the 1920s and 1930s, Steinhaus was among the leading members of a group of distinguished mathematicians working at the University of Lwów (then Poland, now Ukraine). In June of 1941, Lwów was occupied by Germany, and Hugo Steinhaus, who was Jewish, went into hiding. Although he suffered great hardship, he survived and lived until 1972. (He taught at the University of Notre Dame in Indiana for a brief time after the war, and spent time in England as well.) During the time when he was in hiding, in 1943, he invented the "lone divider method," the natural generalization of "I cut, you choose" to three people. A chapter on fair division can be found in his book *Mathematical Snapshots*, a wonderful book for nonmathematicians.

Here is how Steinhaus's lone divider method works. Of three people, one is designated the *divider*. You will see soon that the divider is at a disadvantage—she will get one-third of the cake, in her opinion, but no more, whereas the others can get more, in their views. The divider should therefore be designated by rolling a die. She cuts the cake into three pieces that she considers equally valuable, each worth one-third of the cake to her. Now the other two people, called the *choosers*, each designate the pieces that are worth at least one-third of the cake to them. Each chooser will be able to designate at least one of the three pieces (Exercise 17.4). We now distinguish two cases.

Case 1: A total of at least two pieces of cake are designated, either because at least one chooser designates two or three pieces or because each chooser designates one piece, but their choices are not the same. In this case, we can give each chooser a piece that they consider worth at least one-third of the cake, and give the remaining piece to the divider. Everybody gets his or her fair share that way—at least one-third of the cake in his or her eyes.

Case 2: Each chooser designates exactly one piece, and they designate the same piece. In that case, there is a conflict. To resolve it, we first give to the divider one of the two unchosen pieces, so the divider is satisfied. Then we combine the other unchosen piece with the piece that both choosers want. The combination must be worth more than two thirds of the cake to both choosers (Exercise 17.5). We use "I cut, you choose" to newly divide the combined piece among the two choosers. Again, everybody gets his or her fair share at least.

Example 17.1. A, B, and C share a cake, using Steinhaus's lone divider method. First, A cuts the cake into three pieces, not equal in volume, but equal in value in his opinion:



In B's opinion, only piece 2 represents a fair share. In C's opinion, pieces 2 and 3 represent fair shares. In this case, piece 3 should be assigned to C, piece 2 to B, and the remaining piece 1 to A. ∎

Example 17.2. A, B, and C have inherited several items from their great-aunt's house. There are a shelf full of books, a dining table, five dining chairs, and a 50-year-old parrot. They decide to determine who will get what using Steinhaus's lone divider method. The situation is not really ideal for this method, since the "cake" is not divisible into arbitrarily small pieces: We certainly don't want to divide the parrot. However, we will try to use the method anyway, and hope that it will leave the parrot intact. First, A divides the items into three groups of equal value to him:

1. the books,

2. the dining room furniture,

3. the parrot.

Both B and C are mostly interested in the books. The dining room furniture would not represent a fair share to them, nor would the parrot. According to Steinhaus's method, we should therefore assign to A either the dining room furniture, or the parrot. Let's give him the parrot. Then B and C should use "I cut, you choose" to decide how to split the books and the dining room furniture between them. ∎

Proposition 17.3. *The lone divider method is fair, but neither envy-free nor Pareto-optimal nor equitable.*

Proof. To see that it is neither envy-free nor equitable, think about a case in which both choosers have the same tastes, and they agree that one of the pieces that the divider has cut is worthless, one is worth 60% of the cake, and one is worth 40% of the cake. One of them will get the 60% piece, and the other will get the 40% piece. Although each chooser agrees that they got a fair share, the one who gets the 40% piece will envy the one who gets the 60% piece. Furthermore, the two choosers give different answers to the question, "Which fraction of the total cake do you think you got?" so the division is not equitable. Proving that the lone divider method is not Pareto-optimal is Exercise 17.6. □

It is natural to ask whether Steinhaus's idea could be extended to N people with $N > 3$. Let us try $N = 4$. One person is designated the divider and cuts the cake into four pieces that she considers equal to each other. Each of the three others (the "choosers") declares which of the pieces seem like fair shares to him or her. Each of the three choosers must be able to name at least one piece that he or she considers a fair share (see Exercise 17.4). Again, we distinguish two cases.

Case 1: It is possible to assign to each chooser one of the pieces that he or she considers a fair share.

Case 2: There is a conflict among the choosers; that is, it is not possible to assign to each chooser one of the pieces that he or she considers a fair share.

It is more complicated than before to understand precisely when we are in the desirable Case 1. It is customary to think about this issue in the following terms. Think of the people sharing the cake as girls, and the pieces of cake cut by the divider as boys. Each girl declares which boys she likes. Can we marry off the girls and boys in such a way that each girl gets married to a boy she likes? The answer is given by a mathematical theorem called the *marriage theorem*, which is the subject of the next chapter.

Exercises

17.1. A cake is to be divided among A, B, and C using Steinhaus's lone divider method. A cuts the cake into three pieces that he considers equal in value. The opinions of B and C about the values of the pieces cut by A are summarized in the following table:

	first piece	second piece	third piece
B	0.3	0.4	0.4
C	0.5	0.2	0.3

The table should be interpreted as follows. To B, the first piece is worth the fraction 0.3 of the cake (that is, 30%), the second piece is worth the fraction 0.4 (or 40%), and the third piece is worth the fraction 0.4 (40%) as well. To C, the three pieces are worth 50%, 20%, and 30% of the cake, respectively. How should the three people proceed, according to Steinhaus's method? (There are two possible ways of proceeding; describe one.)

17.2. Repeat Exercise 17.1 with this table:

	first piece	second piece	third piece
B	0.5	0.2	0.3
C	0.3	0.4	0.3

17.3. Repeat Exercise 17.1 with this table:

	first piece	second piece	third piece
B	0.5	0.2	0.3
C	0.4	0.3	0.3

17.4. Explain why, in the lone divider method, each chooser must be able to identify at least one piece that she considers a fair share for her.

17.5. Suppose that one of the three pieces created by the divider is considered a fair share by neither of the two choosers. Explain why it must then be true that the portion of the cake that remains after removal of this piece is considered worth more than two thirds of the cake by both choosers.

17.6. Give an example showing that the lone divider method is not Pareto-optimal.

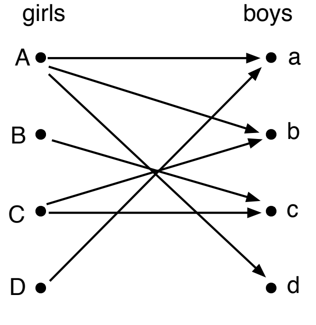
Chapter 18

Hall's Marriage Theorem

In this chapter, we will think about matchmaking among N boys and N girls (N is an integer, $N \geq 1$). Suppose that each of the girls declares which of the boys she likes. Is it possible to marry them all off in such a way that each girl marries a boy whom she likes? The last paragraph of Chapter 17 explains how we were lead to this problem while thinking about fair division.

If in fact our real concern were matchmaking, then maybe we should not only worry about which boys the girls like, but also about whether the attraction is mutual. But in the context of fair cake division, the "girls" are people, whereas the "boys" are slices of cake. We ask which slices of cake are considered fair shares by which people. In this context, it does not make sense to worry about the boys' feelings.

We will let dots on the page stand for girls and boys, and use arrows to indicate which girl likes which boy. Here is an example with $N = 4$. The girls are called A, B, C, and D here, and the boys a, b, c, and d.



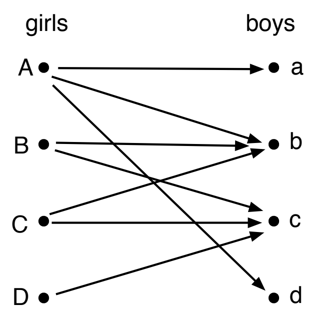
(If we did worry about the boys' feelings, we would simply leave out the arrows, and interpret a link between a girl and a boy as expressing the fact that they like each other.) In our example, if, for instance, A gets married to d and D gets married to a, then both A and D will be pleased. We say that "Ad" is an *acceptable* pair, and similarly Da is an acceptable pair. Also Bc and Cb are acceptable pairs. So there is a solution to the marriage problem in this example:

$$Ad, \quad Bc, \quad Cb, \quad Da.$$

In fact, you will see right away that this is the only solution.

It is easy to create examples in which the marriage problem is not solvable. For instance, suppose that one of the boys is Prince William, good-looking and in line for the throne of England. All girls claim that they love Prince William and nobody else in the world. Assuming that there are $N > 1$ girls, this conflict cannot be resolved, not at least without a dramatic modification of the laws regarding marriage prevailing in most of the world.

Here is another, less dramatic example in which the problem is not solvable:



A quick way of seeing that the marriage problem cannot be solved here is to observe that the three girls B, C, and D compete for only two boys, namely, b and c.

In general, suppose that $\mathcal{G}$ is a set of girls, and denote by $\mathcal{B}_{\mathcal{G}}$ the set of boys whom at least one of the girls in $\mathcal{G}$ likes. We denote by $\#\mathcal{G}$ the number of girls in $\mathcal{G}$, and by $\#\mathcal{B}_{\mathcal{G}}$ the number of boys in $\mathcal{B}_{\mathcal{G}}$. If there is a set $\mathcal{G}$ of girls for which

$$\#\mathcal{B}_{\mathcal{G}} < \#\mathcal{G}, \tag{18.1}$$

then the marriage problem is evidently not solvable—it is impossible to find suitable husbands for all the girls in $\mathcal{G}$, since they compete for too small a number of boys. In the example above, for $\mathcal{G} = \{B, C, D\}$, we obtain $\mathcal{B}_{\mathcal{G}} = \{b, c\}$, and therefore

$$\#\mathcal{B}_{\mathcal{G}} = 2 < \#\mathcal{G} = 3.$$

This proves that the marriage problem is unsolvable in this example. In the example of Prince William, if $\mathcal{G}$ is the set of *all* girls, then $\mathcal{B}_{\mathcal{G}} = \{\text{Prince William}\}$, and therefore

$$\#\mathcal{B}_{\mathcal{G}} = 1 < \#\mathcal{G} = N.$$

Now we turn this around and assume that we knew in fact that

$$\#\mathcal{B}_{\mathcal{G}} \geq \#\mathcal{G} \quad \text{for all nonempty sets } \mathcal{G} \text{ of girls.} \tag{18.2}$$

We call this condition the *marriage condition*. Is it true that there is a solution to the marriage problem whenever the marriage condition is satisfied? To put it differently, is (18.1) the *only* possible obstruction? The answer is yes. This is the content of Hall's marriage theorem, named in honor of the English mathematician Philip Hall (1904–1982), who proved it in 1935.

Theorem 18.1 (Hall's marriage theorem). *The marriage problem has a solution if and only if the marriage condition* (18.2) *holds.*

Proof. As we have already noted, the marriage problem *obviously* has no solution if the marriage condition is violated. Our task is to prove that the problem *does* have a solution if the marriage condition is satisfied. The proof uses the method of mathematical induction; if you are not familiar with this method, please study Appendix C before reading this proof.

To use mathematical induction, we must first prove that the marriage theorem is true when $N = 1$. When $N = 1$, there is only one girl and one boy. The only (nonempty) choice of $\mathcal{G}$ is the set that has the one girl as its element. If she likes the boy, then (18.2) holds ($\#\mathcal{G} = \#\mathcal{B}_{\mathcal{G}} = 1$), and the marriage problem has a solution. If she does not like the boy, then (18.2) is violated ($\#\mathcal{B}_{\mathcal{G}} = 0 < \#\mathcal{G} = 1$), and the marriage problem has no solution. So the marriage problem has a solution if and only if (18.2) holds.

The second step, using the *strong* principle of mathematical induction (see Appendix C), is to *assume* that we already know that the assertion is true when the number of girls and boys is $1, 2, 3, \ldots,$ or N, and *prove* that is must then also hold when there are $N + 1$ pairs. So we assume now that the marriage theorem holds when the number of girls and boys is $1, 2, 3, \ldots,$ or N. This is called the *inductive assumption.* We assume further that the number of girls and boys is $N + 1$ now, and that (18.2) holds. We want to prove that the marriage problem is solvable. We distinguish two cases.

Case 1: The marriage condition (18.2) holds "with room to spare," meaning that for any set $\mathcal{G}$ of girls with $\#\mathcal{G} \leq N$, $\#\mathcal{B}_{\mathcal{G}} > \#\mathcal{G}$, not just $\#\mathcal{B}_{\mathcal{G}} \geq \#\mathcal{G}$. (Note: If $\mathcal{G}$ is the set of all $N + 1$ girls, then $\#\mathcal{B}_{\mathcal{G}}$ could not possibly be greater than $\#\mathcal{G}$.) We then marry the first girl to any of the boys she likes. (She must like more boys than one, by our assumptions!) We remove the happy couple from the pool, and that leaves us with N girls and N boys. If you now take a subset $\mathcal{G}$ of the remaining girls, which of the boys who are still available do they like? They like the same ones as before, but some of them also like the boy who just got married. So we may no longer have $\#\mathcal{B}_{\mathcal{G}} > \#\mathcal{G}$, but we do still have $\#\mathcal{B}_{\mathcal{G}} \geq \#\mathcal{G}$. Therefore the marriage condition still holds for the smaller set of N girls and boys. Since we are allowed to *assume* the marriage theorem for N pairs, we know that we can match up the remaining N girls and boys. Therefore we are done in Case 1.

Case 2: The marriage condition holds "just barely." That is, there is some nonempty set of girls $\mathcal{G}$ with $\#\mathcal{G} \leq N$ so that $\#\mathcal{B}_{\mathcal{G}} = \#\mathcal{G}$. Let us write $k = \#\mathcal{G} = \#\mathcal{B}_{\mathcal{G}}$, and remember that $1 \leq k \leq N$. We set aside the k girls from this set $\mathcal{G}$, and the k boys whom at least one of them likes. For these k girls and boys, the marriage condition is satisfied. Since $k \leq N$, we are allowed to assume the marriage theorem for k pairs, so we can marry them off.

We return now to the remaining crowd of $N + 1 - k$ not-yet-married girls and boys. We claim that the marriage condition holds for them as well. To see this, take a set of s not-yet-married girls. If we put them together with the k married girls, then these

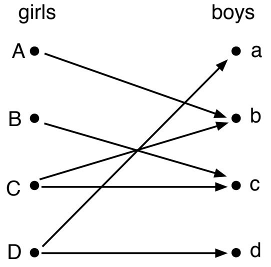
$s+k$ girls must like at least $s+k$ boys, since the marriage condition was assumed for the whole set of $N+1$ girls and boys. But none of the k already married girls likes any of the boys who are not married yet. Therefore the s not-yet-married girls must still have at least s not-yet-married boys left whom at least one of them likes.

So the not-yet-married crowd also satisfies the marriage condition, and since $N+1-k \le N$, we know, by inductive assumption, that we can marry them off. □

It is entirely unclear at first why Theorem 18.1 should be useful. When N is large, a direct exhaustive search, aimed at either finding a solution to the marriage problem or proving that there is none, certainly may be tedious; but it is far from clear why it should be any less tedious to check whether the marriage condition is satisfied or prove that it isn't. Theorem 18.1 will prove useful for *theoretical* purposes in Chapter 19. It will tell us that Steinhaus's lone divider method can be generalized to four or more people. However, it won't really help us find fair divisions in practice.

Exercises

18.1. The following diagram indicates which of the four girls A, B, C, and D likes which of the four boys a, b, c, and d.



Find a nonempty set $\mathcal{G}$ of girls that violates the marriage condition (18.2). In this example, are there other sets $\mathcal{G}$ that also violate (18.2)?

18.2. Four people want to share a cake. The first, called A, divides the cake into four pieces that are each worth one-fourth of the cake, in his opinion. We will call the four pieces a, b, c, and d. The others, called B, C, and D, have different opinions about the worth of the pieces. Their estimates of the worth of the pieces are summarized in the following table:

	a	b	c	d
A	0.25	0.25	0.25	0.25
B	0.25	0.10	0.40	0.25
C	0.10	0.10	0.70	0.10
D	0.20	0.40	0.20	0.20

.

For instance, B agrees that piece a is exactly 0.25 times the total value of the cake, or 25%. On the other hand, B thinks that piece b is 10% of the cake only.

(a) Represent A, B, C, D and a, b, c, d by dots, as we did in this chapter. Indicate by arrows which person considers which piece a fair share of the cake, i.e., at least 25% of the cake.

(b) Verify that the marriage condition (18.2) is satisfied here. (Do not use part (c) to do part (b), but instead verify (18.2) directly.)

(c) Give a solution of the marriage problem—that is, a fair assignment of pieces of cake.

(d) Explain why there are exactly two solutions of the marriage problem here.

(e) Verify that neither of the two solutions of the marriage problem yields an envy-free division of the cake.

Chapter 19

"I Cut, You Choose" for More than Three: Kuhn's Method

The natural generalization of Steinhaus's lone divider method to the case of $N > 3$ people was described and analyzed in 1967 by Harold W. Kuhn, a Professor of Mathematical Economics at Princeton University and a distinguished game theorist.[12]

We start out as in Chapter 17: There is a *lone divider* who cuts the cake into N pieces so that, in her view, each piece represents exactly $1/N$ of the cake. Then the other $N-1$ people, the *choosers*, specify which of the pieces represents, in their eyes, a fair share, i.e., at least the fraction $1/N$ of the cake. Each of them will at least consider one of the pieces a fair share (see Exercise 17.4).

To connect the situation with the marriage theorem, we think of the people sharing the cake as girls, and the pieces of cake as boys. We may then distinguish two cases.

Case 1: The marriage condition is satisfied. It is then possible to assign to each person a piece of cake that he or she likes, and therefore the problem can be solved. I have not told you how to actually *find* a satisfactory assignment. This is the *computational* aspect of the problem. In this chapter, we only describe the method *conceptually*. The computational aspect is too difficult to be included here. For small N, of course you can find the assignment by trial and error. For large N, that is still possible in theory, but in practice a more intelligent computational procedure is needed.

Case 2: The marriage condition is violated. This is the more difficult case, and we will think about it now. To say that the marriage condition is violated is to say that there is a set $\mathcal{C}$ of people for whom there are fewer than $\#\mathcal{C}$ pieces which at least one of them considers a fair share. (Recall that $\#\mathcal{C}$ denotes the number of elements in $\mathcal{C}$.) We call such a set *conflicted*. Certainly $\#\mathcal{C} < N$ (see Exercise 19.4). It is possible that $\#\mathcal{C} = N-1$ if $\mathcal{C}$ is equal to the set of choosers (see Exercise 19.4).

Let $\mathcal{C}_0$ denote a conflicted set of maximal size. That is, let $\mathcal{C}_0$ be a conflicted set containing as many people as possible. We have to phrase this in such a cautious way because there can be several conflicted sets of maximal size (see Exercise 19.5). We denote by k_0 the number of elements in $\mathcal{C}_0$:

$$k_0 = \#\mathcal{C}_0.$$

[12] H. W. Kuhn, "On games of fair division," in *Essays in Mathematical Economics in Honor of Oskar Morgenstern*, Martin Shubik, editor, Princeton University Press, 1967, pages 29–37.

Let us set aside the k_0 people in $\mathcal{C}_0$, together with the (fewer than k_0) pieces that at least one of them considers a fair share, and some extra pieces (chosen arbitrarily) to make it a total of exactly k_0 pieces. We will deal with $\mathcal{C}_0$ in the end, and first make sure that the other $N-k_0$ people get a fair share.

Lemma 19.1. *For the $N-k_0$ people who are outside $\mathcal{C}_0$ and the $N-k_0$ pieces of cake that have not been set aside, the marriage condition is satisfied.*

Proof. Suppose that among the $N-k_0$ people outside $\mathcal{C}_0$, there were a set $\mathcal{C}$ of people with the property that there were fewer than $\#\mathcal{C}$ pieces of cake that have not been set aside and are considered fair shares by at least one person in $\mathcal{C}$. Then the union $\mathcal{C}_0 \cup \mathcal{C}$ would be a conflicted set (see Exercise 19.7) larger than $\mathcal{C}_0$. This contradicts our assumption that $\mathcal{C}_0$ is a conflicted set of maximal size. $\square$

By Lemma 19.1 and by the marriage theorem, we can now distribute the pieces that have not been set aside to the people outside $\mathcal{C}_0$ in such a way that each person outside $\mathcal{C}_0$ receives a fair share.

What do we do with the k_0 people in $\mathcal{C}_0$? First notice that each of the pieces we have already given away was worth less than $1/N$ of the cake to each of those people—they did not consider those pieces fair shares, after all. So each of the people in $\mathcal{C}_0$ thinks that *more* than k_0/N of the cake has been set aside and is yet to be distributed. We recombine the pieces that have been set aside and do a fair division among the k_0 people in $\mathcal{C}_0$. Then each of the k_0 people in $\mathcal{C}_0$ receives at least $1/k_0$ of the part of the cake that has not yet been given away, which is worth more than $1/N$ of the original cake to them.

How do we create a fair division among the k_0 people in $\mathcal{C}_0$? The answer is, we use Kuhn's method! This appears to be circular—we were trying to *define* Kuhn's method, and here the definition suddenly calls for *using* Kuhn's method! However, remember that $k_0 < N$. If we knew already how Kuhn's method works when there are fewer than N people, we would now also know how it works when there are N people. In other words, we have reduced the size of our problem.

We can keep reducing the size of our problem in this way until we end up having to do fair division for only two people; this we do using "I cut, you choose."

For $N = 3$, this procedure is the same as that of Chapter 17, the lone divider method for three people proposed by Steinhaus (see Exercise 19.8). Let us therefore consider examples with $N = 4$.

Example 19.2. Four people want to share a cake. The first, called A, divides the cake into four pieces that are each worth one-fourth of the cake, in his opinion. We will call the four pieces a, b, c, and d. The others, called B, C, and D, have different opinions about the worth of the pieces. Their estimates of the worth of the pieces are summarized in the following table:

	a	b	c	d
A	0.25	0.25	0.25	0.25
B	0.20	0.40	0.10	0.30
C	0.15	0.45	0.10	0.30
D	0.20	0.45	0.05	0.30

For instance, B thinks that piece a is 0.20 times the total value of the cake, or 20%. This is why the number 0.20 appears in the second row, first column. On the other hand, B thinks that piece b is 40% of the cake. This is why the number 0.40 appears in the second row, second column.

The marriage condition is violated here: The three choosers compete for the same two pieces. In the notation used earlier, $\mathcal{C}_0 = \{B, C, D\}$ and $k_0 = 3$.

We assign to the divider one of the two pieces that neither of the three choosers considers a fair share—a or c. Arbitrarily, let us pick a. The remainder of the cake is worth 80% of the total to B, 85% of the total to C, and 80% of the total to D. We can recombine the three pieces and apply Steinhaus's lone divider method to divide it fairly among the three choosers. ∎

Example 19.3. We modify Example 19.2 as follows:

	a	b	c	d
A	0.25	0.25	0.25	0.25
B	0.20	0.50	0.10	0.20
C	0.15	0.45	0.20	0.20
D	0.30	0.20	0.30	0.20

Here $\{B, C\}$ is conflicted, since B and C compete for piece b only. On the other hand, $\{B, C, D\}$ is not conflicted, since there are three pieces (a, b, and c) that are fair shares for at least one of the three choosers B, C, and D. So here $\mathcal{C}_0 = \{B, C\}$, and $k_0 = 2$. We set aside B and C, together with piece b (the one that they both consider a fair share) and an arbitrarily chosen second piece, say a. To the people who have not been set aside (A and D), we assign pieces of cake that have not been set aside (c and d) in such a way that they get fair shares: D gets piece c, and A gets piece d. Then we recombine the pieces that have been set aside, a and b, and distribute them between B and C using "I cut, you choose." Note that both B and C consider the union of a and b to be more than half the cake: B considers it 70%, and C considers it 60% of the cake. Therefore each of them gets a fair share of the total cake if they fairly divide the union of a and b. ∎

This completes our description of the Steinhaus–Kuhn generalization of "I cut, you choose" to N people. The method always generates a fair division. However, as pointed out in Chapter 17, it is neither envy-free nor Pareto-optimal nor equitable.

Exercises

19.1. A, B, C, and D want to share a cake using Kuhn's method. The divider, A, cuts the cake into four pieces a, b, c, and d that are equal in her eyes. The following table describes how they view the value of each of those pieces:

	a	b	c	d
A	0.25	0.25	0.25	0.25
B	0.20	0.50	0.00	0.30
C	0.15	0.25	0.30	0.30
D	0.24	0.30	0.22	0.24

For instance, A considers each of the pieces worth one-quarter of the cake, B considers piece d worth 30% of the cake, etc. According to Kuhn's method, what should be done?

19.2. Repeat Exercise 19.1 with the table

	a	b	c	d
A	0.25	0.25	0.25	0.25
B	0.20	0.50	0.10	0.20
C	0.15	0.55	0.15	0.15
D	0.25	0.30	0.25	0.20

(There are several ways of proceeding according to Kuhn's method here; describe one.)

19.3. Repeat Exercise 19.1 with the table

	a	b	c	d
A	0.25	0.25	0.25	0.25
B	0.15	0.40	0.30	0.15
C	0.20	0.50	0.15	0.15
D	0.20	0.30	0.30	0.20

(There are several ways of proceeding according to Kuhn's method here; describe one.)

19.4. Explain why the divider cannot belong to any conflicted set.

19.5. Six people want to share a cake. The first, called A, divides the cake into six pieces that are each worth one-sixth of the cake, in her opinion. We will call the six pieces a, b, c, d, e, and f. The six people estimate the worth of the pieces as follows:

	a	b	c	d	e	f
A	1/6	1/6	1/6	1/6	1/6	1/6
B	0.1	0.3	0.3	0.1	0.1	0.1
C	0.1	0.1	0.1	0.5	0.1	0.1
D	0.1	0.1	0.1	0.5	0.1	0.1
E	0.1	0.1	0.1	0.5	0.1	0.1
F	0.1	0.1	0.1	0.1	0.3	0.3

Explain why in this case there are two different maximal conflicted sets $\mathcal{C}_0$.

19.6. In the example of Exercise 19.5, describe what should be done, according to Kuhn's method, up to the point when pieces are recombined and distributed among fewer than six people. (There are several different correct solutions to this problem.)

19.7. This exercise refers to a point in the proof of Lemma 19.1. Suppose that among the $N - k_0$ people outside $\mathcal{C}_0$, there were a set $\mathcal{C}$ of people with the property that there were fewer than $\#\mathcal{C}$ pieces of cake that have not been set aside and are considered fair shares by at least one person in $\mathcal{C}$. Explain why the union $\mathcal{C}_0 \cup \mathcal{C}$ would then be a conflicted set.

19.8. Explain why the procedure described in this chapter is that described in Chapter 17 when $N = 3$.

Chapter 20

The Method of Selfridge and Conway

In Chapter 19, we studied Kuhn's method, a general fair cake cutting procedure that works for any number N of participants. Unfortunately, this method is neither envy-free (unless $N = 2$) nor Pareto-optimal nor equitable. Here we describe an envy-free variation of the lone divider method for $N = 3$ people, proposed independently by John Selfridge, a mathematician at Northern Illinois University, and John Conway, a mathematician at Princeton University.

Let A, B, and C be the three people who want to divide the cake. The method has two rounds.

ROUND 1: The first step, as in Steinhaus's lone divider method, is to let one person—say A—divide the cake into three equal (from her point of view) pieces. We refer to A as the *divider*.

Next, B inspects the three pieces. If one piece is strictly larger than the other two in his opinion, he trims that piece, so that there is a tie for first place. The trimmings are set aside. For now, we will only divide the cake minus the trimmings in an envy-free way. The trimmings will be divided in the second round. We call B the *trimmer*. The trimmer might not trim anything at all from any piece, if he thinks that there is a tie for first place; in that case, there are no trimmings, and there will be no need for a second round.

Now C is asked to choose one of the three pieces. We call C the *chooser*. Of course, C will choose the piece that, in his opinion, is the largest—or, if there is a tie in his opinion, one of the pieces that are largest. So C will envy nobody in the first round.

Next, B gets to choose a piece. Recall that in his eyes, after the trimmings were removed there were two pieces that tied for first place. Perhaps C chose one of those two pieces, but even then B can still choose the other; so B has no reason to envy anybody in the first round. We do impose one extra constraint on B: If C did not choose the trimmed piece, then B must choose it. Since in his eyes the trimmed piece tied for first place, he has no reason to object to this rule.

Finally A takes the remaining piece. Note that the piece that remains for A is *not* the trimmed piece, for that one was taken either by C or by B. So what remains for A is one of the pieces that she originally cut.

It is convenient at this point to rename B and C. One of the two got the trimmed piece, while the other got one of the pieces originally cut by A. We will refer to the one who got the trimmed piece as X, and to the other as Y. So either $X = B$ and $Y = C$, or $X = C$ and $Y = B$. In any case, A does not envy Y—in her eyes, Y got precisely as large a share of the cake as she got. She certainly does not envy X—in her eyes, even if X were given the *entire* trimmings in addition to the piece that he already has, he would not be better off than she is. This is the key to the procedure, so keep it in mind.

This completes the first round. We have distributed the cake minus the trimmings without creating envy.

ROUND 2: Now we have to distribute the trimmings without introducing any envy. First we ask Y to cut the trimmings into three pieces that are equal in his eyes. Then we let X choose one of these three pieces. He gets to choose first, and therefore will not envy A or Y for their share of the trimmings—he just chooses the piece that he considers largest. A gets to choose one of the two remaining pieces of the trimmings, and Y takes the last piece. Since A chooses her share of the trimmings before Y gets his, she will not envy Y. But she also will not envy X, since she would not even envy X if he got *all* the trimmings, as we pointed out earlier. Since Y cut the trimmings, he envies neither A nor X for their share of the trimmings. Thus we have created an envy-free division of the entire cake.

Unfortunately, however:

Proposition 20.1. *The method of Selfridge and Conway is neither Pareto-optimal nor equitable.*

Proof. To prove that the method is not Pareto-optimal, think about what happens if the divider chooses the equal division. In the trimmer's (and anybody else's) eyes, all three pieces cut by the divider are of equal worth, each exactly one-third of the total cake. So there will be no trimmings. Each of the three people will simply get exactly one-third of the cake—in their eyes, and in anybody's eyes.

But this is not usually Pareto-optimal. We know that one can usually do much better. Consider, for instance, a cake consisting of three homogeneous components, say chocolate, strawberry, and rhubarb, and assume that A, B, and C value the three components as follows:

	chocolate	strawberry	rhubarb
A	1	0	0
B	0	1	0
C	0	0	1

In words: A considers the strawberry and rhubarb components worthless, B considers the chocolate and rhubarb components worthless, and C considers the chocolate and strawberry components worthless. Then each person can get, in his or her eyes, the *entire* cake!

To see that the method is not equitable either, think about the same cake, but this time assume these valuations:

	chocolate	strawberry	rhubarb
A	1/3	1/3	1/3
B	0	1/2	1/2
C	0	1/2	1/2

. (20.1)

If A cuts the cake into its three homogeneous components (in her eyes, they are each worth one-third of the cake), then B and C each get half the cake in their eyes, while A is left with one-third of the cake in her eyes (see Exercise 20.6). □

Example 20.2. We return to Example 17.2: A, B, and C have inherited from their great-aunt a shelf full of books, some dining room furniture, and the 50-year-old parrot. They will divide these items using the Selfridge–Conway method, hoping that the method will not call for cutting the parrot into pieces. First, A divides the items into three groups of equal value to him:

1. the books,
2. the dining room furniture,
3. the parrot.

B, the trimmer, thinks that the books are most valuable, the dining room furniture is second, and the parrot is least valuable. He therefore removes from the book collection the collected works of Charles Dickens, and 30 detective novels by Agatha Christie. Now he thinks that the remaining books and the dining room furniture are of equal value. C, the chooser, selects the dining room furniture. Next, B gets to choose his share, but since the trimmed piece (the book collection, minus the works of Dickens and the Agatha Christie novels) was not chosen by C, he must choose it. Finally, A gets the parrot. This completes the first round.

In the second round, C divides the "trimmings"—the collected works of Dickens, and the 30 detective novels. His interest in Dickens is limited, but he enjoys old-fashioned detective stories. So he divides the books as follows:

(i) 8 detective novels, and half the works of Dickens;

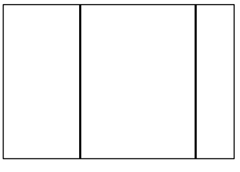
(ii) 8 detective novels, and half the works of Dickens;

(iii) 14 detective novels.

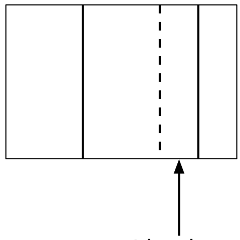
Next, B chooses 8 detective novels and half the works of Dickens. Then A chooses the 14 detective novels. So C gets the remaining 8 detective novels and the remaining half of the works of Dickens. ■

Exercises

20.1. A cake is to be divided among three people, A, B, and C, using the method of Selfridge and Conway. A cuts the cake into three slices that have, in her eyes, equal value:



B thinks that the middle piece is the most valuable, the left piece is second, and the right piece is third. He therefore trims the middle piece to make it as valuable as the third piece:



C decides that the rightmost piece is most valuable, and he takes it. (a) What should happen next? (b) After each of the three has taken a piece, who should cut the trimmings? (c) Who should be the first to choose a piece of the trimmings?

20.2. This exercise refers to Example 20.2. (a) Summarize what each of the three people get altogether. (b) Verify that A does not envy B. (c) Verify that A does not envy C.

20.3. After the first round of the Selfridge–Conway procedure, are any of the three people already guaranteed to have a fair share of the cake?

20.4. This exercise prepares Exercise 20.5. Suppose that three people, U, V, and W, share a cake. Suppose that W cuts the cake into three pieces that she considers equal in value, then U takes his share, then V takes his, and finally W takes his. Explain: U envies nobody, W envies nobody, and V does not envy W but may envy U.

20.5. In the second round of the Selfridge–Conway procedure, the trimmings are divided into three pieces by Y, the chooser who did not get the trimmed piece in the first round. Then the three people choose their shares of the trimmings in the order X, A, Y (first X, then A, then Y). Recall that X is either B or C, and X got the trimmed piece in the first round.

Here we will explore whether it has to be done that way or whether variations are possible. There are six orderings in which the three people may take their shares of the trimmings in the second round:

$$
\begin{array}{l}
(1)\ A, X, Y \\
(2)\ A, Y, X \\
(3)\ X, A, Y \\
(4)\ X, Y, A \\
(5)\ Y, A, X \\
(6)\ Y, X, A
\end{array}
$$

In each case, let us assume that the person listed last is the one who divides the trimmings into three pieces. It makes sense to let the person who divides the trimmings choose last, since he considers all three pieces of the trimmings equal in value, and will therefore envy nobody no matter which piece of the trimmings he gets. Selfridge and Conway proposed ordering (3), which guarantees envy-freeness overall. (a) For each of the possible orderings, list who might envy whom overall (that is, after both rounds are complete). (Hint: See Exercise 20.4.) Is there any ordering other than (3) which guarantees envy-freeness overall? (b) Explain in your own words why ordering (3) guarantees envy-freeness overall.

20.6. This exercise refers to a point in the proof of Proposition 20.1. Show: If the table (20.1) describes the valuations, and if A cuts the cake into its homogeneous components, then B and C each get half the cake in their eyes, while A is left with one-third of the cake in her eyes.

Chapter 21

The Geometry of Pareto-Optimal Division between Two People

Since Pareto-optimality is such a sensible requirement, it is natural to try to understand its meaning in an intuitive, geometric way. We will focus in this chapter on the simplest interesting case, that of two people sharing a piecewise homogeneous cake.[13] As before, we denote the number of homogeneous components by n. Let's imagine asking each of the two people to specify how they value each of the components.

Example 21.1. A and B divide a cake consisting of four homogeneous components. The following table indicates their valuations of the various components of the cake:

	1st component	2nd component	3rd component	4th component
A	0.60	0.10	0.20	0.10
B	0.20	0.05	0.10	0.65

For instance, A thinks that the first component of the cake represents 60% of the total cake, B thinks that the fourth component represents 65%, and so on. Our goal is to describe all Pareto-optimal divisions in a simple geometric way. ∎

In general, we denote A's and B's valuations of the jth component by a_j and b_j, respectively:

	1st component	2nd component	...	nth component
A	a_1	a_2	...	a_n
B	b_1	b_2	...	b_n

A thinks that the jth component of the cake represents a fraction a_j of the total cake, while B thinks that it represents a fraction b_j of the total cake. Of course,

$$a_1 + a_2 + \cdots + a_n = 1 \quad \text{and} \quad b_1 + b_2 + \cdots + b_n = 1.$$

(To both of them, the whole cake is the whole cake.)

[13]For a discussion of this subject in full generality and at a much higher mathematical level, see Julius Barbanel, *The Geometry of Efficient Fair Division*, Cambridge University Press, 2005, from which I derived the title of this chapter.

Let us assume that $a_j > 0$ and $b_j > 0$ for all j. If $a_j = 0$ and $b_j > 0$, for instance, for some j, then it is clear that to achieve Pareto-optimality, we must give the entire jth component to B and divide the rest of the cake Pareto-optimally. If $a_j = b_j = 0$, the jth component is worthless to both parties, and it does not matter what we do with it.

Recall that a division is Pareto-optimal if there is no alternative division that pleases one person more (that is, makes her think that she gets a larger share of the cake) without pleasing the other less. You will realize immediately that Pareto-optimality has to do with giving to A what A desires, and to B what B desires, for if instead we gave to A what B desires, and to B what A desires, then the two could both gain if they traded. So crucial quantities to think about here are the "A-to-B valuation ratios." By definition, the *A-to-B valuation ratio for the jth component* is a_j/b_j.

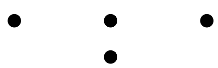
Let us imagine the components of the cake lined up in order of decreasing A-to-B valuation ratios:

$$\frac{a_1}{b_1} \geq \frac{a_2}{b_2} \geq \cdots \geq \frac{a_n}{b_n}.$$

(We simply assume that the components are numbered that way.) In the example at the beginning of the chapter, this was the case:

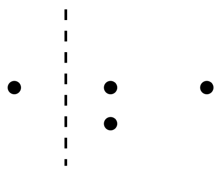
$$\frac{0.6}{0.2} > \frac{0.1}{0.05} = \frac{0.2}{0.1} > \frac{0.1}{0.65}.$$

So the components that are more valuable to A than to B are placed to the left of the components that are more valuable to B than to A. It is useful to symbolize each component with a dot, arranged on the page in such a way that dots further to the left correspond to components for which the A-to-B valuation ratio is larger. In our example, we would draw the dots like this:

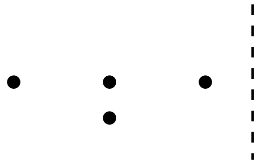


The second and third dots are vertically aligned to indicate that the A-to-B valuation ratios of the second and third components of the cake are equal.

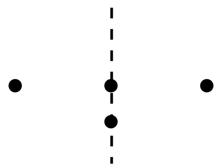
We can create a division by drawing a vertical line separating the dots into ones to the left of the line, and ones to the right of the line, assigning components corresponding to dots on the left to A, and assigning components corresponding to dots on the right to B. For instance, we could draw the line like this:



In this case, A would get the first component, and B would get the remaining three. Or we could even draw the line like this:



In this case, A would get the entire cake, and B would get nothing at all. The vertical line may pass straight through one of the points, or several vertically aligned points, for instance:



In such a case, the picture does not describe the division completely. A component corresponding to a point that the vertical line passes through is divided—some of it goes to A, some to B. The picture does not specify how it is divided. *All* of it may go to A, *all* of it may go to B, or the two may share it. If the vertical line passes through several vertically aligned points, as in our example, then the corresponding components need not all be divided in the same way.

The divisions that are created in this way are characterized by the existence of a *threshold value* r of the A-to-B valuation ratio: If $a_j/b_j > r$, the jth component is assigned to A, and if $a_j/b_j < r$, it is assigned to B. (If $a_j/b_j = r$, the jth component may be assigned to A, to B, or be split.) We will call a division of this sort a *threshold division*. In Example 21.1, the choice of threshold $r = 2.5$ leads to the division "Assign the first component to A, and the second, third, and fourth components to B." (The same division is obtained if one chooses r to be any number between 2 and 3.) If one takes r to be 2, one must specify how the second and third components are to be divided. "Assign the first and third components to A, and the second and fourth to B" is a threshold division with $r = 2$. "Assign the first, second, and third components to A, and the fourth to B" is another.

In a threshold division, all pieces that are split (that is, assigned partially to A and partially to B) have the same valuation ratio. For instance, in an equal division, *all* components of the cake are divided, with half assigned to A and the other half to B. Therefore an equal division is not a threshold division, unless all components have equal valuation ratios; see Exercises 21.2 and 21.3.

The central result of this chapter can now be stated as follows.

Theorem 21.2. *The threshold divisions are precisely the Pareto-optimal divisions.*

> ***Proof.*** The proof has two steps. First we assume that a division is *not* a threshold division, and prove that then it is not Pareto-optimal. Second we assume that a division is a threshold division, and prove that then it is Pareto-optimal. (If the logic of this argument is not clear to you, please read Sections B.1 through B.3 of Appendix B.)
>
> We begin with the first step, so we suppose that we are given a division, and that it is not a threshold division. There are then two components of the cake, say the jth and

the kth ones, so that

$$\frac{a_j}{b_j} > \frac{a_k}{b_k}, \tag{21.1}$$

A gets some of component k, and B gets some of component j. (Recall that we assume throughout that each person attaches positive value to each component of the cake. Therefore there are no issues concerning division by zero in (21.1).)

Let's think about what happens if A and B swap some of their holdings of components j and k, as follows: A gives a small fraction ϵ of component k to B, and B gives a small fraction δ of component j to A. (The Greek letters ϵ = epsilon and δ = delta are frequently used in mathematics to denote small positive numbers.) The fraction of the cake that A receives, in her eyes, changes by

$$C_A = -\epsilon a_k + \delta a_j. \tag{21.2}$$

A gains as a result of the swap if $C_A > 0$, and loses if $C_A < 0$. We will choose ϵ and δ so that $C_A = 0$, so A's satisfaction remains the same:

$$C_A = -\epsilon a_k + \delta a_j = 0,$$

that is,

$$\delta = \epsilon \, \frac{a_k}{a_j}. \tag{21.3}$$

(We are careful, of course, to choose ϵ so small that A in fact has the fraction ϵ of component k, and B in fact has the fraction $\delta = \epsilon \, a_k/a_j$ of component j, since otherwise the trade would be impossible.) After the swap, the fraction of the cake that B receives, in his eyes, changes by

$$C_B = \epsilon b_k - \delta b_j.$$

We will now prove that $C_B > 0$, so B benefits from the swap.

We use (21.3) to rewrite the formula for C_B:

$$C_B = \epsilon b_k - \delta b_j = \epsilon b_k - \epsilon \, \frac{a_k}{a_j} \, b_j = \epsilon \left(b_k - \frac{b_j}{a_j} \, a_k \right). \tag{21.4}$$

By assumption, $\epsilon > 0$. Inequality (21.1) implies that

$$b_k - \frac{b_j}{a_j} \, a_k > 0 \tag{21.5}$$

as well. Therefore, by (21.4), $C_B > 0$, so B is strictly more satisfied after the swap than before. Since A is as satisfied as before, the given division was not Pareto-optimal. This completes the first step of the proof.

We now turn to the second step of the proof. We assume that we are given a threshold division. We will prove that this division is Pareto-optimal. To simplify the notation, we will assume here that the cake consists of just three homogeneous components, with

$$\frac{a_1}{b_1} > \frac{a_2}{b_2} = r > \frac{a_3}{b_3}. \tag{21.6}$$

The first component is assigned entirely to A, and the third to B. The second component is divided. (This includes the possibility that A gets it in its entirety, or B gets it in its entirety.) We will prove that no alternative division could please one person more without pleasing the other less.

Any alternative division can be obtained from the given division by reassigning some of what A has to B and/or vice versa. So let us assume that a fraction q_1 of the first component (which is entirely A's in the given threshold division) is shifted from A to B; here q_1 is a number between 0 and 1. Similarly, let us assume that a fraction q_3 of the third component (which is entirely B's in the given division) is shifted from B to A, with $0 \le q_3 \le 1$. The second component is shared between A and B in the given threshold division. We can therefore shift a fraction of the second component from A to B, or from B to A. Let us assume that a fraction q_2 of the second component is shifted from A to B, with $-1 \le q_2 \le 1$. If $-1 \le q_2 < 0$, we take this to mean that the fraction $|q_2|$ of the second component is shifted from B to A. After the shifts are complete, the fraction of the total cake that A receives, in her opinion, has changed by

$$C_A = -q_1 a_1 - q_2 a_2 + q_3 a_3. \tag{21.7}$$

If $C_A > 0$, then A has gained, and if $C_A < 0$, she has lost. Similarly, the fraction of the total cake that B receives, in his view, has changed by

$$C_B = q_1 b_1 + q_2 b_2 - q_3 b_3. \tag{21.8}$$

From (21.6), we know that

$$a_1 > r b_1, \quad a_2 = r b_2, \quad \text{and} \quad a_3 < r b_3. \tag{21.9}$$

We will combine (21.7)–(21.9) to derive a relationship between C_A and C_B, in order to prove that $C_A > 0$ implies $C_B < 0$, and $C_B > 0$ implies $C_A < 0$.

Combining (21.7) and (21.9), we find

$$\begin{aligned} C_A &= -q_1 a_1 - q_2 a_2 + q_3 a_3 \\ &\le -q_1 r b_1 - q_2 r b_2 + q_3 r b_3. \end{aligned}$$

Factoring out r, then using (21.8), we get

$$\begin{aligned} &-q_1 r b_1 - q_2 r b_2 + q_3 r b_3 \\ &= -r(q_1 b_1 + q_2 b_2 - q_3 b_3) \\ &= -r C_B. \end{aligned}$$

In summary, these calculations prove that

$$C_A \le -r C_B. \tag{21.10}$$

Since $r > 0$, we conclude that C_A and C_B have opposite signs (unless both are zero). If $C_B > 0$, then $C_A < 0$; if B gains, then A loses. Similarly, if $C_A > 0$, then $C_B < 0$; if A gains, then B loses. We have concluded that it is impossible to devise an alternative division that pleases one person more without pleasing the other less. The threshold division that we started out with is therefore Pareto-optimal, as we wanted to prove.

Although we have carried out the second step of the proof only for a special case, it is not so hard to convince yourself that the same argument applies in the general case. □

Exercises

21.1. A and B want to share a cake consisting of four homogeneous components. Their valuations of the components are as follows:

	1st component	2nd component	3rd component	4th component
A	0.50	0.30	0.10	0.10
B	0.25	0.40	0.05	0.30

(a) Compute the A-to-B valuation ratios.

(b) Which of the following are threshold divisions? (Be careful: The components are *not* listed here in such a way that the A-to-B valuation ratio increases or decreases from left to right.)

(i) Give half of the first component to A, half of it to B, and give the rest to B.

(ii) Give half of the first component to A, half of it to B, and give the rest to A.

(iii) Split the first component 1:2 (1/3 of it goes to A, and 2/3 of it goes to B), give the second component to B, split the third component 3:1 (3/4 of it goes to A, and $1/4$ of it goes to B), and give the fourth component to B.

(iv) Give the first, second, and third components to A, and split the fourth component evenly.

(v) Give the first and second components to A, and the third and fourth to B.

(vi) Give the first component to A, and the second, third, and fourth components to B.

21.2. Explain why an equal division is not a threshold division, unless all components have equal valuation ratios. (See also Appendix B, Example B.3.)

21.3. Assume that all components of the cake have equal valuation ratios:

$$\frac{a_1}{b_1} = \frac{a_2}{b_2} = \cdots = \frac{a_n}{b_n}. \tag{21.11}$$

Prove that then A and B have the same valuations:

$$a_1 = b_1, \quad a_2 = b_2, \quad \ldots, \quad a_n = b_n. \tag{21.12}$$

Please note that you are not asked to prove that (21.12) implies (21.11)—that is obvious! You are asked to prove that, conversely, (21.11) implies (21.12). In other words, you are asked to prove that if all the valuation ratios are the same, then their common value is 1. (See also Appendix B, Example B.4.)

21.4. Verify: Inequality (21.1) implies inequality (21.5).

21.5. ($*$) (This problem has a star only because its purpose is to illustrate the first half of the proof of Theorem 21.2, which is in small print. The problem itself is not difficult.) Two boys named Andrew and Benjamin (A and B) have stolen a bag of nuts and a bag of chocolates. A likes nuts, so to him the bag of nuts is 2/3 of the booty, and the bag of chocolates is 1/3 of the booty. B likes chocolates, so to him the bag of nuts is 1/3 of the booty, and the bag of chocolates is 2/3 of the booty. They decide to split both the nuts and the chocolates evenly. (a) Explain why this is not a threshold division. (b) After the two boys have split both bags evenly, they decide to make a trade. A will give a fraction ϵ of the bag of chocolates to B, and B will give a fraction δ of the bag of nuts to A. Find $\epsilon > 0$ and $\delta > 0$ such that $\epsilon \leq 1/2$ and $\delta \leq 1/2$ (so that the trade will in fact be possible), and such that after the trade, A will be as satisfied as before, whereas B will be more satisfied than before.

Chapter 22

The Adjusted Winner Method of Brams and Taylor

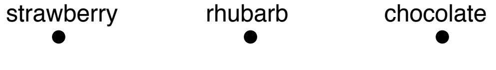
In Chapter 21, we described, in a geometric way, the Pareto-optimal divisions of a piecewise homogeneous cake between two people. There are many such divisions, ranging from giving the entire cake to A to giving the entire cake to B. These two extremes correspond to drawing the vertical line in Chapter 21 all the way to the right or all the way to the left. Neither of these two extremes is fair. The reasonable locations for the vertical line are in between. There is exactly one way of placing the vertical line that leads to an *equitable* division. To place the vertical line that way, and use the resulting equitable Pareto-optimal division, is called the *adjusted winner method*. The adjusted winner method was proposed in 1994 by Steven Brams, a Professor of Political Science at New York University, and Alan Taylor, a Professor of Mathematics at Union College. It was patented in 1999 (U.S. Patent 286031); the assignee of the patent is New York University.

Example 22.1. Suppose A and B want to divide a cake that consists of a strawberry component, a rhubarb component, and a chocolate component, and their valuations are as follows:

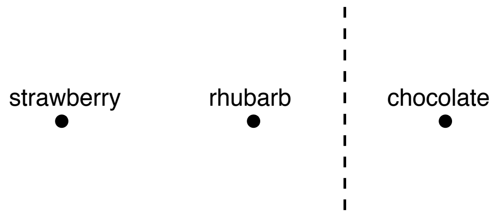
	strawberry	rhubarb	chocolate
A	0.5	0.3	0.2
B	0.3	0.2	0.5

.

The A-to-B valuation ratios are $5/3$ for strawberry, $3/2$ for rhubarb, and $2/5$ for chocolate. As discussed in Chapter 21, we symbolize the three components of the cake with three dots, placing dots corresponding to components with greater A-to-B valuation ratios further to the left:



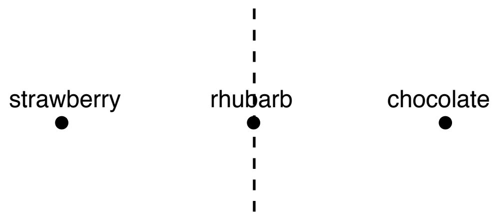
As shown in Chapter 21, to divide the cake in a Pareto-optimal way, we must first place a vertical division line. Brams and Taylor suggest to start by placing the line so that the dots with A-to-B valuation ratio greater than 1 are to the left of the line, and those with A-to-B valuation ratio less than 1 are to the right:



This corresponds to a division in which A gets the strawberry and rhubarb components, and B gets the chocolate component.

If this were what we did, then A would feel that she would get 80% of the cake, while B would feel that he would get 50%. The division is fair, but not equitable. It is fair by accident; in general, placing the vertical line so that it separates the components with valuation ratio > 1 from those with valuation < 1 need not result in a fair division.

To make the division equitable, we must move the vertical line. The process of moving the line is called the "adjustment," and that is where the name "adjusted winner" comes from. In this example, we must move the vertical line further to the left, to satisfy B more, at the expense of A. The smallest meaningful move to the left results in the following diagram:



We must now specify which fraction of the rhubarb component should go to A, and which should go to B. Let us denote by p the fraction of the rhubarb component that A retains; so p is a number between 0 and 1. Of course, B then receives the fraction $1-p$ of the rhubarb component. The fraction of the total cake that A receives, in her eyes, is

$$0.5+0.3p. \tag{22.1}$$

(0.5 for the strawberry component and $0.3p$ for a fraction of p of the rhubarb component.) The fraction of the cake that B receives, on the other hand, is

$$0.5+0.2(1-p). \tag{22.2}$$

(0.5 for the chocolate component and $0.2(1-p)$ for a fraction of $1-p$ of the rhubarb component.) To make the division equitable, we have to choose p so that (22.1) and (22.2) become equal:

$$0.5+0.3p = 0.5+0.2(1-p).$$

Subtracting 0.5 from both sides, we find

$$0.3p = 0.2(1-p),$$

or

$$0.3p = 0.2 - 0.2p.$$

Adding $0.2p$ to both sides:

$$0.5p = 0.2.$$

Dividing both sides by 0.5:

$$p = \frac{0.2}{0.5} = \frac{2}{5} = 0.4.$$

So if we give to A the entire strawberry component and 40% of the rhubarb component, and to B the remaining 60% of the rhubarb component and the entire chocolate component, then the division will be both Pareto-optimal and equitable. Let us verify that it is indeed equitable. The fraction of the total cake that A gets, in her view, is

$$0.5 + 0.3p = 0.5 + 0.12 = 0.62.$$

The fraction of the cake that B gets, in his view, is

$$0.5 + 0.2(1-p) = 0.5 + 0.12 = 0.62.$$

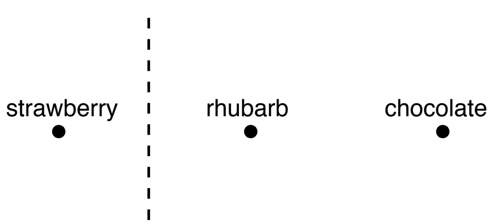
So both A and B get 62% of the cake, in their view. ∎

Example 22.2. Let us change the valuations of Example 22.1 as follows:

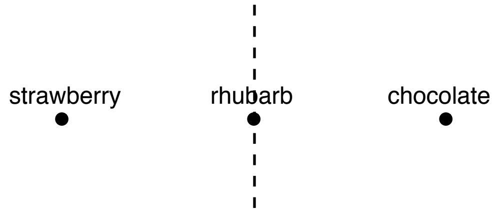
	strawberry	rhubarb	chocolate
A	0.40	0.10	0.30
B	0.05	0.15	0.80

.

Again we begin by drawing the line so that the dots corresponding to components with valuation ratios > 1 are to the left, and those corresponding to components with valuation ratios < 1 are to the right:



This corresponds to a division that assigns 40% of the cake to A, in her view (so it is not even fair!), and 95% of the cake to B, in his view. To make the division equitable, we must move the vertical line to the right. The smallest meaningful move to the right results in this diagram:



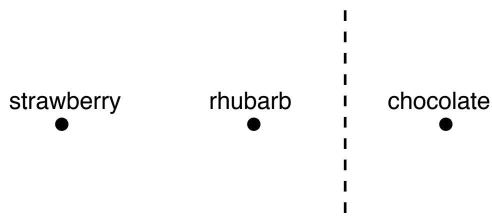
Let us now try to find a number p so that an equitable division will result if the fraction p of the rhubarb component goes to A, and the fraction $1-p$ goes to B. The equation for p is

$$0.4+0.1p=0.8+0.15(1-p).$$

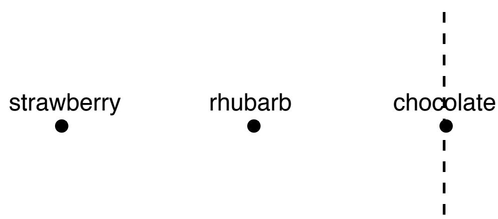
The left-hand side of this equation represents the value of A's share (the strawberry component and a fraction p of the rhubarb component), in A's eyes. The right-hand side represents the value of B's share (the chocolate component, and a fraction $1-p$ of the rhubarb component), in B's eyes. Solving the equation for p, we find

$$p=\frac{0.55}{0.25}=\frac{11}{5}.$$

The trouble is that this is greater than 1! This simply means that no equitable division can be created by giving the strawberry component to A, the chocolate component to C, and splitting the rhubarb component. To achieve equitability, we must draw the vertical line yet further to the right. Drawing it as follows won't help:



That is the same as drawing it through the rhubarb dot and choosing $p=1$. We already know that this is not sufficient to achieve equitability: We would need $p=11/5>1$ to achieve equitability. So we have to move the line so that it passes through the chocolate dot:



The corresponding division assigns the strawberry and rhubarb components to A, and splits the chocolate component, giving a fraction p of it to A and the fraction $1-p$ to B. To make the division equitable, we need

$$0.4+0.1+0.3p = 0.8(1-p). \tag{22.3}$$

Solving for p, we find

$$p = \frac{0.3}{1.1} = \frac{3}{11}\,.$$

We have found the equitable, Pareto-optimal division: Give the strawberry and rhubarb components to A, and also give $3/11$ of the chocolate component to A. Give the remaining $8/11$ of the chocolate component to B.

With this division, A thinks that she gets the fraction

$$0.40+0.10+\frac{3}{11}\times 0.30 \approx 0.58$$

of the cake, and B thinks that he gets the fraction

$$\frac{8}{11}\times 0.80 \approx 0.58.$$

The two fractions are indeed the same, as they should be by (22.3). Both A and B think that they get 58% of the cake. ∎

By design, the adjusted winner method is Pareto-optimal and equitable. From Exercise 16.5 we then know that it is fair as well—and therefore envy-free, since $N=2$.

It does not really matter where the initial vertical line is drawn. It makes sense, from a practical point of view, to start with a vertical line separating the items with valuation ratio >1 from those with valuation <1, since in most cases the final vertical line will be near that location. However, any other choice of the initial vertical line would lead to the same outcome.

Exercises

22.1. John and Jane divorce. They have to decide who gets custody of their 11-year-old, who gets their home, who gets the vacation house on Cape Cod, and who gets their substantial savings account. Notice that only one of these items is divisible in an objective way: It is quite clear what it means to split the savings account at a ratio of 2:1, for instance. Custody of the 11-year-old might, with a stretch, be thought of as divisible: He can spend alternate months with John and with Jane (that would be even splitting), or he can spend one day a week with John and the rest with Jane (that might be called splitting at a ratio of 1:6). The two houses are certainly not divisible. But we can still try out the adjusted winner method. If we are lucky, we will never *have* to answer the question how to split either of the two houses! The two are asked to value the items. This is what they say:

	custody	home	vacation house	cash
John	0.50	0.15	0.15	0.20
Jane	0.15	0.45	0.10	0.30

What should be done, according to the adjusted winner method?

22.2. (a) John and Jane split up. They jointly own a car. They need to decide who should get the car, and which compensation that person should pay to the other. John values the car at \$6,000, while Jane values it at \$4,000. They would like to use the adjusted winner method, but unfortunately that's impossible—the car is not divisible! But then they remember that they also own a joint bank account of \$20,000. They decide to distribute both items, car and bank account, simultaneously, using the adjusted winner method. What will be the outcome? (Luckily, it turns out that a division of the car does not become necessary.)

(b) John and Jane do *not*, in fact, own a joint bank account, but they are so fond of the adjusted winner method that they open one, just to be able to use the method. Each pays \$10,000 into the account. Then they distribute both car and account using the adjusted winner method. This will *amount* to giving the car to John, but asking him to pay a certain net compensation amount to Jane. What is that amount, after taking into account that each of them paid \$10,000 into the joint account, but does not necessarily get \$10,000 back? Of course, John and Jane don't actually have to go to the bank and open the account. They can just *imagine* they had opened it, then compute what would happen using the adjusted winner method.

(c) ($*$) Explain: If part (b) is repeated with a *very large* bank account, the result is Knaster's arrangement.

22.3. Think about a cake consisting of five homogeneous (objectively divisible) pieces. Suppose that A values the pieces as follows:

$$a_1 \;\; a_2 \;\; a_3 \;\; a_4 \;\; a_5.$$

(The sum $a_1+a_2+a_3+a_4+a_5$ equals 1.) Similarly, B values the pieces as follows:

$$b_1 \;\; b_2 \;\; b_3 \;\; b_4 \;\; b_5.$$

(The sum $b_1+b_2+b_3+b_4+b_5$ equals 1 as well.) Assume that

$$\frac{a_1}{b_1} > \frac{a_2}{b_2} > \frac{a_3}{b_3} > \frac{a_4}{b_4} > \frac{a_5}{b_5}.$$

We consider two divisions of the cake.

Division 1: We assign pieces 1 and 2 to A, and pieces 3, 4, and 5 to B.

Division 2: Starting with Division 1, we introduce the following modifications. We choose numbers p_1, p_2, p_3, p_4, and p_5 between 0 and 1, and shift a fraction p_1 of the first piece from A to B, a fraction p_2 of the second piece from A to B, a fraction p_3 of the third piece from B to A, a fraction p_4 of the fourth piece from B to A, and a fraction p_5 of the fifth piece from B to A.

(a) Under Division 1, A thinks that she receives $a_1 + a_2$ percent of the cake. Under Division 2, she thinks that she receives $a_1 + a_2 + F$ percent of the cake, where F is some number. Write down a formula for F.

(b) Under Division 1, B thinks that he receives $b_3 + b_4 + b_5$ percent of the cake. Under Division 2, he thinks that he receives $b_3 + b_4 + b_5 + G$ percent of the cake. Write down a formula for G.

(c) Prove by calculation: If $G > 0$, then $F < 0$.

(d) Prove by exploiting the fact that threshold divisions are Pareto-optimal: If $G > 0$, then $F < 0$.

Chapter 23

Conflict Resolution Using the Adjusted Winner Method

Example 23.1. My friends Lisa and Martin bought a house together. The house was beautiful, but in serious need of renovation. They started making plans for the renovation, but found quickly that they disagreed on several important issues. I suggested to them that they should use the adjusted winner method to negotiate about these issues, of course with me serving as the referee. Be it that they are not mathematically inclined, or that they don't trust me as a referee—they rejected my proposal. Here is their punishment: I am putting them into this book.

Below are five items on which they disagreed (taking poetic license):

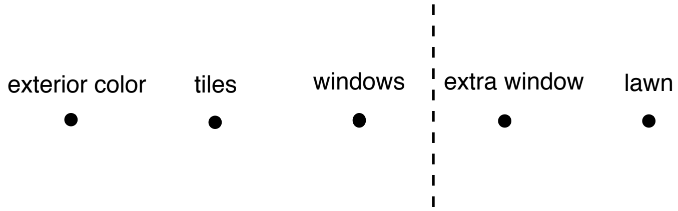
1. Lisa wanted to paint the house a light blue. Martin wanted to paint it a dark green.
2. Lisa wanted the kitchen floor tiles to be a uniform reddish brown. Martin wanted them sand-colored with irregular patterns.
3. Lisa wanted the existing old windows restored. Martin wanted them replaced.
4. Lisa wanted to break an extra window into the exterior wall of the kitchen. Martin wanted to place an extra shelf there, and therefore wanted the solid wall preserved.
5. Lisa wanted a good-sized lawn. Martin, recognizing that he would be the one to cut that lawn, wanted fruit trees instead.

If Lisa and Martin had been asked to indicate the relative importance of these issues to them, the outcome might have looked like this:

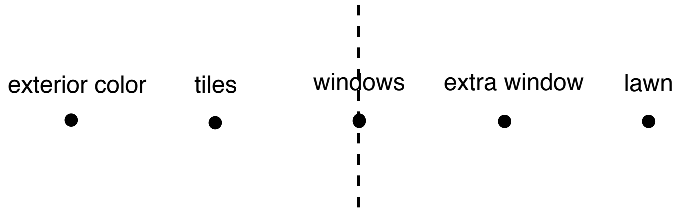
	exterior color	tiles	window replacement	extra window	lawn
Lisa	0.30	0.20	0.25	0.10	0.15
Martin	0.10	0.13	0.20	0.17	0.40

.

In the table, the issues are already ordered in such a way that the Lisa-to-Martin valuation ratio decreases from left to right. We begin by drawing a vertical line that separates the items with a Lisa-to-Martin valuation ratio > 1 from those with a Lisa-to-Martin valuation ratio < 1:



With this division, Lisa thinks that she gets 75% of the "cake," while Martin thinks that he gets 57%. To make the division equitable, we must move the vertical line to the left:



Now the issue of the replacement windows is to be split—perhaps some are to be restored, the others replaced? Suppose that Lisa retains control over a fraction p of this issue, in some sense, so Martin gets control over the fraction $1-p$. Then Lisa gets the fraction $0.50+0.25p$ of the "cake," and Martin gets the fraction $0.57+0.20(1-p)$. We want to choose p such that

$$0.50+0.25p = 0.57+0.20(1-p).$$

Solving this equation for p, we find

$$p = \frac{0.27}{0.45} = 0.6.$$

So the adjusted winner method suggests that the lawn should be converted to fruit trees, no extra window should be broken into the kitchen wall, and 40% of the windows should be replaced, while 60% of them should be restored.[14] Lisa should get to select the kitchen tiles, and she should get to select the color of the house. ∎

Example 23.2. This example is taken from a manuscript titled, "Adjusted Winner and the Future Negotiations over East Jerusalem" by Moshe Hirsch of Hebrew University in Jerusalem, presented at the Annual Meeting of the Public Choice Society, New Orleans, Louisiana, March 2005 (cited and presented here with the author's permission). Hirsch proposed to look at the Israeli–Palestinian conflict from the point of view of the adjusted winner method. Specifically he focused on the future of East Jerusalem. To do this, one has to first identify the items that are contentious between the two sides. Hirsch suggested the following list:

[14] That would probably look rather odd. Maybe Lisa and Martin were wise not to use the adjusted winner method.

1. The Palestinian neighborhoods outside the Old City.

2. The Muslim and Christian Quarters in the Old City.

3. Al-Haram al-Sharif, or Temple Mount. This site is sacred to both Jews and Muslims. Although all of Jerusalem is under Israeli control and considered part of Israel by the Israeli government, the al-Haram al-Sharif is administered by the Muslim authorities, and the Chief Rabbinate has ruled that Jews should not enter the site.

4. The Jewish and Armenian Quarters in the Old City.

5. The Western Wall—a site of utmost significance to Jews.

6. The Israeli neighborhoods outside the Old City.

None of these items can be considered objectively divisible like a cup of coffee or a pile of nuts. So what will the adjusted winner method tell us, if anything? The answer might be this: It will suggest which issues should be resolved wholly in one or the other party's favor, and on which issues they must compromise and share in some ways.

Hirsch interviewed experts on both sides, analyzed the positions taken by the two sides during negotiations, and arrived at the following estimates of the two sides' valuations.

Palestinian valuations:

1. Palestinian neighborhoods outside the Old City: 19% of the total.

2. Muslim and Christian Quarters in the Old City: 22%.

3. Al-Haram al-Sharif/Temple Mount: 48%.

4. Jewish and Armenian Quarters in the Old City: 6%.

5. Western Wall: 4%.

6. Israeli neighborhoods outside the Old City: 1%.

Israeli valuations:

1. Palestinian neighborhoods outside the Old City: 0%.

2. Muslim and Christian Quarters in the Old City: 9.5%.

3. Al-Haram al-Sharif/Temple Mount: 22%.

4. Jewish and Armenian Quarters in the Old City: 18%.

5. Western Wall: 31%.

6. Israeli neighborhoods outside the Old City: 19.5%.

Note that these items are listed in the order of decreasing Palestinian/Israeli valuation ratio. The valuation ratio is greater than 1 for items 1–3, less than 1 for items 4–6. If the Palestinians got items 1–3, they would think that they got 89% of East Jerusalem. If the Israelis got items 4–6, they would think that they got 68.5% of East Jerusalem. To make the division equitable, it is necessary to compromise on the issue of the Temple Mount. If the Israelis get, in addition to items 4–6, a fraction $1-p$ of the Temple Mount, and the Palestinians retain the fraction p, the Israelis will think that they get $68.5+22(1-p)$ percent of the total, and the Palestinians will think that they bet $41+48p$ percent of the total. We calculate p from the equation

$$68.5+22(1-p)=41+48p,$$

so

$$p=\frac{49.5}{70}\approx 0.71.$$

In fact,

$$68.5+22\times(1-0.71)\approx 75$$

and

$$41+48\times 0.71\approx 75,$$

so if the Israelis get 29% of the Temple Mount, and the Palestinians 71%, then both sides will think that they get about 75% of East Jerusalem.

The adjusted winner method suggests that Israel should get 29% of the Temple Mount, but what could this possibly mean? Hirsch makes several suggestions: Palestinian sovereignty with some official functions for Israel, for instance. Or Israel might cede the Temple Mount to the Palestinians in return for some sort of financial compensation. Another option is to bring in more issues—West Bank settlements, the question of Palestinian refugees, normalization of diplomatic relations, etc. Then Israel might cede the Temple Mount in return for Palestinian concessions on other issues.

Of course the adjusted winner method does not seriously resolve the dispute over East Jerusalem. However, it seems to me that it may in fact be a helpful way of thinking about the problem. It suggests that the central issue is the Temple Mount. If there were a way of giving, in some sense, 29% of it to the Israelis and 71% of it to the Palestinians, then both sides could get three-quarters of East Jerusalem; that is an interesting conclusion. ∎

Chapter 24

The Effect of Dishonesty on the Adjusted Winner Method

The adjusted winner method has one rather serious flaw: For a dishonest person, it is very easy to manipulate the method to his or her advantage. We will not analyze this in generality, but illustrate it with an example.

Suppose that A and B share a cake that consists of two homogeneous components, and the valuations look like this:

	first component	second component
A	0.6	0.4
B	0.2	0.8

.

The adjusted winner method assigns to B a fraction p of the second component, and to A the remaining fraction $1-p$ of the second component as well as the entire first component, with

$$0.6+0.4(1-p)=0.8p,$$

so

$$1.2p=1,$$

or $p=5/6$. So B gets the fraction $5/6$ of the second component. This arrangement gives both people

$$60+\frac{40}{6}=\frac{5}{6}\times 80=66.66\ldots$$

percent, or two-thirds of the cake.

Suppose now that B knows A's valuations in advance. She decides to announce not her own honest valuations, but instead valuations that maximize her share, given that A values the pieces 60:40. So the valuations are now as follows:

	first component	second component
A	0.6	0.4
B	x	$1-x$

.

B will try to choose x to maximize the share that she will receive in her *honest* view.

We denote the fraction of the cake that B receives, in her honest view, by $S(x)$. If B were honest, she would pick $x=0.2$, and get two-thirds of the cake, so $S(0.2)=2/3$.

To understand how B should choose x to maximize $S(x)$, we will work out a general formula for $S(x)$. In fact, there are different formulas for $S(x)$, depending on the range that x lies in.

Case 1: First we consider values of x for which the valuation ratio for the first component is greater than 1, and that for the second component is less than 1:

$$x < 0.6.$$

In the initial round of the adjusted winner method, the first component is assigned to A, giving him the fraction 0.6 of the cake, and the second to B, *seemingly* giving her the fraction $1-x$ of the cake.

Case 1.1: Within Case 1, we begin with values of x for which B *appears* (based on her announced valuations) to get the worse share of the two in the initial round: $1-x < 0.6$, or $x > 0.4$. Recalling that we are assuming $x < 0.6$, we see that the range of values of x to be considered in Case 1.1 is

$$0.4 < x < 0.6.$$

In this case, A must hand a fraction $1-p$ of the first component to B in the adjustment round, retaining a fraction p, with

$$0.6p = 1-x+(1-p)x,$$

so

$$p = \frac{1}{x+0.6}.$$

B obtains the second component and the fraction $1-p$ of the first component. In her *honest* view, she gets the fraction

$$S(x) = 0.8+\left(1-\frac{1}{x+0.6}\right)\times 0.2 = 1-\frac{0.2}{x+0.6} \tag{24.1}$$

of the cake.

Case 1.2: Still within Case 1, we now consider values of x for which B *appears* (based on her announced valuations) to get the better share of the two in the initial round: $1-x > 0.6$, or

$$x < 0.4.$$

In the adjustment round, B gets to keep only a fraction p of the second component, with

$$(1-x)p = 0.6+0.4(1-p),$$

or

$$p = \frac{1}{1.4-x}.$$

In her *honest* opinion, the fraction of the cake that B gets is then

$$S(x) = \frac{0.8}{1.4-x}. \tag{24.2}$$

Case 2: We now proceed to the case when B inflates her valuation of the first component so much that she exceeds A's valuation of it:

$$x > 0.6.$$

In this case, the valuation ratio for the first component is smaller than 1, and that for the second component is greater than 1. So the initial division assigns the second component to A, and the first to B. Thereby A receives the fraction 0.4 of the cake, and B seemingly receives a much greater fraction of the cake, namely, x. Therefore B gets to retain only a fraction p of the first component, with

$$px = 0.4 + 0.6(1 - p),$$

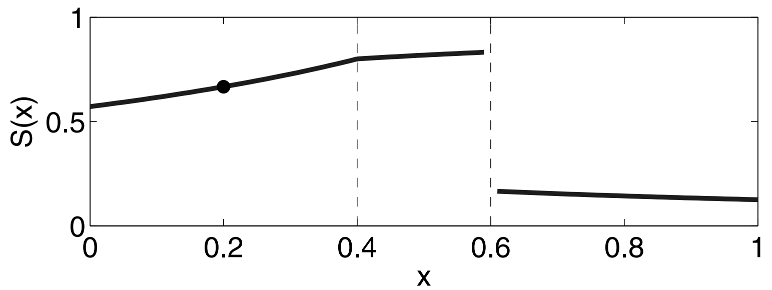
or

$$p = \frac{1}{x + 0.6}.$$

In her *honest* view, the fraction of the cake that she gets is

$$S(x) = \frac{0.2}{x + 0.6}. \tag{24.3}$$

We now plot $S(x)$ as a function of x, using (24.1), (24.2), and (24.3):



You can see that the graph consists of three pieces, corresponding to the three cases $x < 0.4$, $0.4 < x < 0.6$, and $x > 0.6$. If B is honest, the relevant value is $S(0.2) = 2/3$, indicated as a dot.

To maximize $S(x)$, B should choose x very slightly below 0.6. She will then, in her eyes, obtain nearly the fraction

$$1 - \frac{0.2}{1.2} = \frac{5}{6}$$

of the cake (I used (24.1) with $x = 0.6$ here), a marked improvement over the fraction $2/3 = 4/6$ that she would have got if she had honestly declared $x = 0.2$.

It is intuitively clear why B can improve her share by choosing a value of x slightly below 0.6: In this way, she will still be assigned the second component (which truly is her favorite piece), but she will downplay how much the second component is worth to her, in fact so much that she will get a fraction of the first component as well.

But B must lie carefully: If she chooses x slightly above 0.6, then her share plummets. She then receives little more than the fraction

$$\frac{0.2}{1.2} = \frac{1}{6}$$

of the cake, in her honest estimate. (Here I used formula (24.3) with $x = 0.6$.)

We conclude with an interesting observation. Suppose that A recognized that B will announce a valuation for the first piece very slightly below 0.6, and a valuation for the second component very slightly above 0.4. Suppose now that we gave A a second chance: He may announce new valuations. Which valuations should he announce to maximize *his* share? The answer is: He should value the first component very slightly above B's valuation of it, and the second very slightly below B's valuation of it—but this means that he should announce his *honest* valuations!

This is an example of what is called a *noncooperative equilibrium* or a *Nash equilibrium*: Each person announces the valuations that are most favorable for him or her, *given* the valuations announced by the other person. Nash equilibria are named after the mathematician John Nash, winner of the 1994 Nobel Prize in Economics for his work in game theory, and subject of the book *A Beautiful Mind* by Sylvia Nasar, which was turned into a movie.

Exercises

24.1. In the plot of $S(x)$, you see that the piece of the graph with $x < 0.4$ and the piece of the graph with $0.4 < x < 0.6$ fit together at $x = 0.4$. Explain from formulas (24.1) and (24.2) why this is so.

24.2. This exercise refers to the example discussed in this chapter. When A announces his honest valuations, and B dishonestly announces a valuation x very slightly below 0.6 for the first component and $1 - x$ for the second component, approximately which percentage of the cake does A receive in his view?

24.3. (a) Using formula (24.1), explain why $S(x)$ increases as x increases from 0.4 to 0.6.
(b) Using formula (24.2), explain why $S(x)$ increases as x increases from 0 to 0.4.
(c) Using formula (24.3), explain why $S(x)$ decreases as x increases if $x > 0.6$.

24.4. A and B share a cake that consists of two homogeneous components.

(a) Their honest valuations of the pieces are as follows:

	first component	second component
A	0.70	0.30
B	0.30	0.70

Based on these valuations, which percentage of the cake will each of them get, in his or her eyes, when the adjusted winner method is used?

(b) B expects that A's valuations will be 0.70 for the first piece and 0.30 for the second. He therefore decides to pretend that the first piece is *almost* 70% of the total to him as well. So the announced valuations are

	first component	second component
A	0.70	0.30
B	0.69999	0.30001

.

If the adjusted winner method is used based on these dishonest valuations, (i) which percentage of the cake will A get, in her view, and (ii) which percentage of the cake will B get, in his *honest* view? (Assume that the valuations in part (a) still reflect his honest view.)

(c) When A sees B's valuations, she smells a rat: They are just suspiciously close to her own. She therefore says, "Since our valuations are very nearly identical, let us just cut each of the two components in half—you get one half, and I get the other." (The referee has put down the rule that each of the two is allowed to insist on equal division of the two pieces at any stage of the process.) Now each of the two gets, in his or her honest view, exactly 50% of the cake. Explain: (i) Both are much worse off than in part (a). (ii) A is about as well off as in part (b), whereas B is much worse off than in part (b).

Thus by insisting on equal division, A did not help herself, but she punished B for his dishonesty, which she suspected based on the similarity between her valuations and those announced by B.

(d) Suppose again that B expects that A's valuations will be 0.70 for the first piece and 0.30 for the second. From his course on Mathematics of Social Choice, he remembers that he can gain an advantage by announcing similar valuations. Unfortunately, he does not quite remember the details. He announces that the first component is a little bit more than 70% of the cake to him, and the second component a little bit less than 30%. So the announced valuations are

	first component	second component
A	0.70	0.30
B	0.70001	0.29999

.

If the adjusted winner method is used based on these dishonest valuations, (i) which percentage of the cake will A get, in her view, and (ii) which percentage of the cake will B get, in his *honest* view? (Assume that the valuations in part (a) still reflect his honest view.)

24.5. Explain: In the adjusted winner method, an honest person will always get at least a fair share, whereas a dishonest person may get less than a fair share. (Hint for the second part: See Exercise 24.4(d).)

Chapter 25

Proportional Allocation

Again we think about two people sharing a piecewise homogeneous cake. We will now think about an alternative to the adjusted winner method called the *method of proportional allocation*.

Example 25.1. John and Jane share a fruit tart composed of apple, strawberry, and rhubarb components. Their valuations are as summarized in the following table:

	apple	strawberry	rhubarb
John	0.1	0.5	0.4
Jane	0.5	0.2	0.3

.

It may seem natural to give 5 times more apple tart to Jane than to John, reflecting the fact that Jane attaches 5 times more value to this component than John; thus we would give $1/6$ of the apple component to John, and $5/6$ to Jane. Similarly, we might want to divide the strawberry component at a ratio of 5:2, giving John $5/7$ of it, and Jane $2/7$, to reflect the fact that John considers it 50% of the entire tart, while Jane only considers it 20% of the entire tart. By the same reasoning, we would give $4/7$ of the rhubarb component to John and $3/7$ to Jane. This is what is called proportional allocation.

Let us calculate which percentage of the tart John thinks he gets in this example:

$$\frac{1}{6} \times 0.1 + \frac{5}{7} \times 0.5 + \frac{4}{7} \times 0.4 = \frac{253}{420} \approx 0.602.$$

When we calculate which percent of the tart Jane thinks she gets, we discover, surprisingly, that the division is equitable:

$$\frac{5}{6} \times 0.5 + \frac{2}{7} \times 0.2 + \frac{3}{7} \times 0.3 = \frac{253}{420} \approx 0.602.$$

The method of proportional allocation assigns each of the two about 60% of the tart. The adjusted winner method would have assigned each of the two about 67% of the tart (see Exercise 25.2). ∎

We will prove now that the equitability of the proportional allocation method in Example 25.1 is not accidental.

Proposition 25.2. *When two people share a cake using proportional allocation, the outcome is always equitable and fair.*

Proof. To keep the notation as simple as possible, we consider the case of A and B sharing a cake consisting of $n = 2$ homogeneous components. However, all arguments apply to $n > 2$ without important changes.

Suppose that the valuations are as shown in the following table:

	first component	second component
A	a_1	a_2
B	b_1	b_2

.

Here

$$0 \le a_1, a_2, b_1, b_2 \le 1$$

and

$$a_1 + a_2 = b_1 + b_2 = 1. \tag{25.1}$$

When the method of proportional allocation is used, A and B receive the fractions

$$\frac{a_j}{a_j + b_j} \quad \text{and} \quad \frac{b_j}{a_j + b_j} \tag{25.2}$$

of the jth component of the cake, respectively ($j = 1$ or 2). The sum of the two expressions in (25.2) equals 1, as it ought to.

A thinks that her share of the jth piece represents the fraction

$$\frac{a_j}{a_j + b_j}\, a_j = \frac{a_j^2}{a_j + b_j}$$

of the total cake. Therefore A's answer to the question "Which fraction of the cake did you get, in your view?" would be

$$\frac{a_1^2}{a_1 + b_1} + \frac{a_2^2}{a_2 + b_2}. \tag{25.3}$$

Similarly B's answer to the same question would be

$$\frac{b_1^2}{a_1 + b_1} + \frac{b_2^2}{a_2 + b_2}. \tag{25.4}$$

It is not immediately obvious, but true, that the two expressions (25.3) and (25.4) are equal to each other. To prove this, we prove that the difference (25.3) − (25.4) equals zero:

$$\left(\frac{a_1^2}{a_1 + b_1} + \frac{a_2^2}{a_2 + b_2}\right) - \left(\frac{b_1^2}{a_1 + b_1} + \frac{b_2^2}{a_2 + b_2}\right) = \frac{a_1^2 - b_1^2}{a_1 + b_1} + \frac{a_2^2 - b_2^2}{a_2 + b_2}$$

$$= \frac{(a_1 - b_1)(a_1 + b_1)}{a_1 + b_1} + \frac{(a_2 - b_2)(a_2 + b_2)}{a_2 + b_2}$$
$$= (a_1 - b_1) + (a_2 - b_2) = (a_1 + a_2) - (b_1 + b_2) = 1 - 1 = 0.$$

We have shown that the method of proportional allocation is equitable when $n = 2$—A and B will give the same answer to the question "Which fraction of the cake did you get in your opinion?"

We want to prove now that the common answer that A and B will give to this question is always at least 1/2. In other words, we want to prove that proportional allocation is a fair method. Remembering that for two people fairness and envy-freeness are the same thing, we might as well aim at proving envy-freeness instead of fairness, that is, at showing that what A receives in A's opinion is at least as much as what B receives in A's opinion:

$$\frac{a_1^2}{a_1 + b_1} + \frac{a_2^2}{a_2 + b_2} \geq \frac{a_1 b_1}{a_1 + b_1} + \frac{a_2 b_2}{a_2 + b_2}, \tag{25.5}$$

so A does not envy B. Of course, it then also follows that B does not envy A, since we can just swap the names of A and B.

To prove inequality (25.5), we use a trick. Inequality (25.5) has an aesthetically unappealing feature: The right-hand side remains the same if we swap "a" and "b", but the left-hand side does not. We will remove this lack of symmetry, and it will turn out that the resulting, symmetric form of (25.5) is easy to prove.

The left-hand side in inequality (25.5) is (25.3), which we have shown to be the same as (25.4). Since (25.3) and (25.4) are equal, their common value also equals their average. (In general, if $X = Y$, then both X and Y also equal the average $(X+Y)/2$. Considering how obvious an observation this is, it is surprising how useful it can be!) So inequality (25.5) can equivalently be written as

$$\frac{1}{2}\left[\left(\frac{a_1^2}{a_1 + b_1} + \frac{a_2^2}{a_2 + b_2}\right) + \left(\frac{b_1^2}{a_1 + b_1} + \frac{b_2^2}{a_2 + b_2}\right)\right]$$
$$\geq \frac{a_1 b_1}{a_1 + b_1} + \frac{a_2 b_2}{a_2 + b_2}. \tag{25.6}$$

In this symmetric form, the inequality is easy to prove. To do this, multiply both sides of (25.6) by 2 first:

$$\frac{a_1^2}{a_1 + b_1} + \frac{a_2^2}{a_2 + b_2} + \frac{b_1^2}{a_1 + b_1} + \frac{b_2^2}{a_2 + b_2} \geq \frac{2a_1 b_1}{a_1 + b_1} + \frac{2a_2 b_2}{a_2 + b_2}.$$

Then prove that the left-hand side minus the right-hand side is ≥ 0:

$$\frac{a_1^2}{a_1 + b_1} + \frac{a_2^2}{a_2 + b_2} + \frac{b_1^2}{a_1 + b_1} + \frac{b_2^2}{a_2 + b_2} - \left(\frac{2a_1 b_1}{a_1 + b_1} + \frac{2a_2 b_2}{a_2 + b_2}\right)$$
$$= \frac{a_1^2 - 2a_1 b_1 + b_1^2}{a_1 + b_1} + \frac{a_2^2 - 2a_2 b_2 + b_2^2}{a_2 + b_2} = \frac{(a_1 - b_1)^2}{a_1 + b_1} + \frac{(a_2 - b_2)^2}{a_2 + b_2} \geq 0. \qquad \square$$

Proportional allocation gives each of the two people the same percentage of the cake (in his or her view). This common percentage is at least half the cake, but it is usually less than what the adjusted winner method would yield. Why then should we bother to talk about proportional allocation at all? The answer is: Proportional allocation is less vulnerable to manipulation by a dishonest person than the adjusted winner method. If the two people do not trust each other, they might therefore prefer proportional allocation.

I will illustrate that proportional allocation is less manipulable than the adjusted winner method using the example of Chapter 24. So suppose, as in Chapter 24, that the honest valuations are as follows:

	first component	second component
A	0.6	0.4
B	0.2	0.8

. (25.7)

Proportional allocation assigns 58.3% of the cake to A, and the same percentage to B (see Exercise 25.3). Now suppose that B announces dishonest valuations:

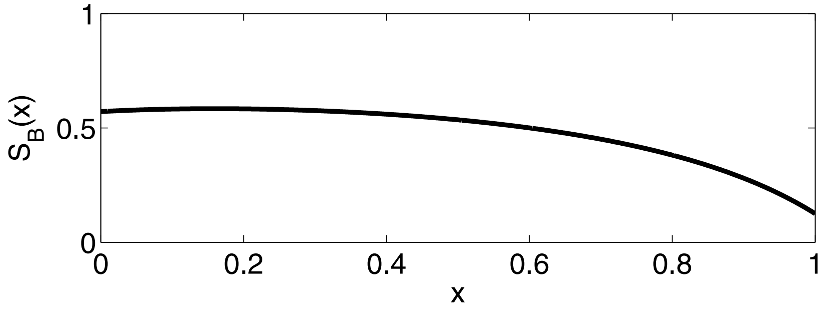
	first component	second component
A	0.6	0.4
B	x	$1-x$

.

Suppose further that B chooses x to maximize her share of the cake. The share of the cake that she will now get, in her honest view, is

$$S_B(x) = \frac{0.2x}{x+0.6} + \frac{0.8(1-x)}{1.4-x} \tag{25.8}$$

(see Exercise 25.4). The problem of choosing x in such a way that $S_B(x)$ is maximized is not quite so simple here. But we can use a computer or graphing calculator to plot $S_B(x)$ as a function of x, and the result looks like this:

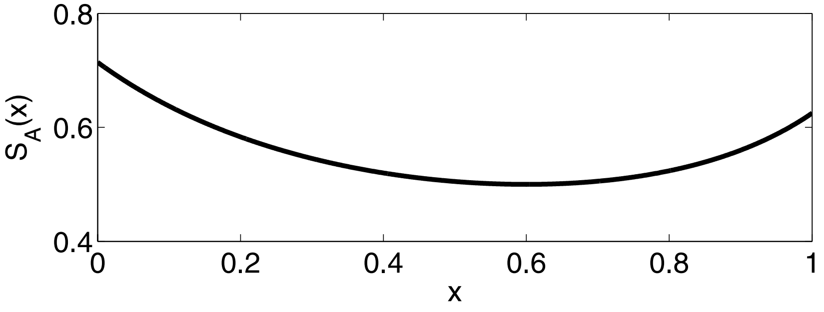


Over a wide range of values of x, around $x = 0.2$, $S_B(x)$ is astonishingly flat. If B chose a value of x greater than 0.5, her share would in fact decline. The calculator tells us that the optimal choice is $x \approx 0.16$. This choice of x gives B a total of about 58.4% of the cake—hardly more than the 58.3% that honesty would have gotten her! Dishonesty yields almost no advantage at all for B. Interestingly, the value of x that is optimal for B is *smaller* than 0.2. This means that to optimize her share of the cake, B should *exaggerate*, not downplay, how much she likes the second component of the cake, in order to get more of it.

It is also interesting to study A's share of the cake as a function of x:

$$S_A(x) = \frac{0.6^2}{x+0.6} + \frac{0.4^2}{1.4-x} \tag{25.9}$$

(see Exercise 25.4). Again using a graphing calculator, we see that this is what the graph of $S_A(x)$ as a function of x looks like:



Although B cannot help herself much by being dishonest, she can hurt or help A substantially by being dishonest. For instance, if she chooses $x = 0.16$ to maximize her share of the cake, that also raises A's share of the cake, from 58.3% to 60.3%! (This confirms that proportional allocation is not Pareto-optimal.)

Exercises

25.1. A and B share a cake consisting of two homogeneous components. Their valuations are as follows:

	first component	second component
A	0.6	0.4
B	0.4	0.6

.

(a) What is the division generated by proportional allocation? Which percentage of the cake do A and B think they get with proportional allocation?

(b) Propose an alternative equitable division that pleases both people more than the one generated by proportional allocation.

25.2. Verify that in Example 25.1 the adjusted winner method would have assigned to each of the two people about 67% of the cake.

25.3. For the valuation table (25.7), verify that the proportional allocation method assigns approximately 58.3% of the cake to each person.

25.4. (a) Explain equation (25.8). (b) Explain equation (25.9).

Chapter 26

Dividing a Piecewise Homogeneous Cake among $N > 2$ People

In this chapter, we explore to which extent the idea of the adjusted winner method can be extended to the case when the number N of people sharing the cake is greater than 2. For $N = 2$, we gave a simple, complete characterization of the Pareto-optimal divisions in Chapter 21. For $N > 2$, such a characterization is much more difficult, and that is a central reason why division of a piecewise homogeneous cake is more difficult for $N > 2$ than for $N = 2$. However, the following lemma is a step towards a better understanding of Pareto-optimality for general N.

Lemma 26.1. *Suppose that $N \geq 2$ people, including two people called A and B, share a cake consisting of $n \geq 1$ homogeneous components. For any j and k between 1 and n, let a_j, a_k, b_j, and b_k denote the fractions of the cake that components j and k represent in the opinion of A and B, respectively. Assume*

$$a_j > 0, \quad b_k > 0, \tag{26.1}$$

and

$$\frac{a_j}{a_k} > \frac{b_j}{b_k}. \tag{26.2}$$

Then any division in which A gets some of component k and B gets some of component j is not Pareto-optimal.

For motivation, first think about the case when $a_j \gg a_k$ (that is, a_j is *much* larger than a_k) and $b_k \gg b_j$. Then of course (26.2) holds. A strongly prefers component j over component k, and B strongly prefers k over j. If A has some of component k and B has some of component j, A and B can achieve an objective improvement (a change that benefits both and does not adversely affect anybody else) by trading. But in fact the preferences need not be strong, and it is not even necessary for A or B to prefer j over k, or vice versa, at all. Inequality (26.2) just states that A judges component j more favorably, in comparison with component k, than B.

The number a_k is allowed to be zero. In that case, the left-hand side of inequality (26.2) is infinity. You might hesitate here, wondering how reasonable it is to say that when $a_j = 0$; but fortunately we assumed in (26.1) that $a_j > 0$. If the left-hand side of inequality

(26.2) is infinity, then the inequality holds, since the right-hand side is finite; this is assured because we also assumed in (26.1) that $b_k > 0$.

Here is the proof of Lemma 26.1.

Proof. Assume (26.1) and (26.2), and suppose we are given a division in which A gets some of component k, and B gets some of component j. We want to show that both A and B can benefit from a trade in which A hands a small fraction of component k to B, and B hands a small fraction of component j to A.

Let $\epsilon > 0$ and $\delta > 0$ denote small numbers. Let us ask whether A and B will both benefit if A hands the fraction ϵ of component k to B, and B hands the fraction δ of component j to A. In this trade, A benefits if

$$\delta a_j > \epsilon a_k, \tag{26.3}$$

and B benefits if

$$\epsilon b_k > \delta b_j. \tag{26.4}$$

Inequalities (26.3) and (26.4) can be summarized like this:

$$\frac{a_j}{a_k} > \frac{\epsilon}{\delta} > \frac{b_j}{b_k}. \tag{26.5}$$

Can we choose ϵ and δ in such a way that (26.5) holds? The answer is yes if and only if

$$\frac{a_j}{a_k} > \frac{b_j}{b_k}. \tag{26.2}$$

If this inequality holds, we simply pick some number, say r, with

$$\frac{a_j}{a_k} > r > \frac{b_j}{b_k},$$

then pick a small $\delta > 0$ and set $\epsilon = r\delta$. Of course we must be careful to pick δ so small that B in fact has the fraction δ of component j, and A has the fraction $r\delta = \epsilon$ of component k, prior to the trade. $\square$

The *method of proportional allocation* discussed for $N = 2$ in Chapter 25 has a simple and natural extension to the case $N > 2$, as illustrated by the following example.

Example 26.2. A, B, and C share a cake consisting of two homogeneous components. Their valuations are as shown in the following table:

	first component	second component
A	a_1	a_2
B	b_1	b_2
C	c_1	c_2

.

When using the method of proportional allocation, one assigns to A, B, and C the fractions

$$\frac{a_1}{a_1+b_1+c_1}, \quad \frac{b_1}{a_1+b_1+c_1}, \quad \text{and} \quad \frac{c_1}{a_1+b_1+c_1}$$

of the first component of the cake, respectively. The second component is divided analogously. ∎

Proportional allocation is not Pareto-optimal for $N = 2$ (see Chapter 25). Might it be Pareto-optimal for $N > 2$, by some miracle? Of course not. In fact, consider any example in which each person attaches positive value to each component of the cake. Then proportional allocation gives *some* of each component to each person. Lemma 26.1 therefore implies that proportional allocation is almost never Pareto-optimal, as illustrated by the following example.

Example 26.3. A, B, and C share a cake consisting of two components. Their valuations are as follows:

	first component	second component
A	0.5	0.5
B	0.3	0.7
C	0.4	0.6

The proportional allocation method does not yield a Pareto-optimal division. This follows from Lemma 26.1, since

$$\frac{0.5}{0.5} > \frac{0.3}{0.7},$$

so A judges the first component more favorably, in comparison with the second component, than B, yet A gets some of the second component, and B gets some of the first. ∎

In fact, if N people share a cake using proportional allocation, and if each of them attaches positive value to each component of the cake, then proportional allocation is *never* Pareto-optimal, unless they all have precisely the same taste (see Exercise 26.4). In addition, proportional allocation is neither equitable nor envy-free for $N > 2$. For $N = 3$, this is shown by the following example.

Example 26.4. A, B, and C share a cake consisting of two homogeneous components. Their valuations are as shown in the following table:

	first component	second component
A	2/3	1/3
B	8/9	1/9
C	1/9	8/9

The proportional allocation method assigns to A the fraction

$$\frac{2/3}{2/3+8/9+1/9} = \frac{2}{5}$$

of the first component, and the fraction

$$\frac{1/3}{1/3+1/9+8/9} = \frac{1}{4}$$

of the second. Similarly, B receives the fraction

$$\frac{8/9}{2/3+8/9+1/9} = \frac{8}{15}$$

of the first component, and the fraction

$$\frac{1/9}{1/3+1/9+8/9} = \frac{1}{12}$$

of the second. In A's view, A receives the fraction

$$\frac{2}{5} \times \frac{2}{3} + \frac{1}{4} \times \frac{1}{3} = 0.35$$

of the cake. In B's view, B receives the fraction

$$\frac{8}{15} \times \frac{8}{9} + \frac{1}{12} \times \frac{1}{9} = 0.4833\ldots$$

of the cake. Since $0.35 \neq 0.4833\ldots$, the division is not equitable. Furthermore, in A's view, B receives the fraction

$$\frac{8}{15} \times \frac{2}{3} + \frac{1}{12} \times \frac{1}{3} = 0.3833\ldots$$

of the cake. Since $0.3833\ldots > 0.35$, A envies B. ∎

We summarize the properties of proportional allocation for $N > 2$ as follows.

Proposition 26.5. *For $N > 2$, proportional allocation is fair, but neither envy-free nor Pareto-optimal nor equitable.*

Proof. We have proved all of these statements already, with one exception: We have not shown yet that proportional allocation is fair for $N > 2$. The proof of that is below, in small print.

To simplify the notation, I will only discuss the case of $N = 3$ people, A, B, and C. It will be clear, however, that the argument is general. I will prove that A receives a fair share. By symmetry, then of course B and C also receive fair shares.

As always, let us denote the number of homogeneous pieces by n. The fraction of the cake that A receives, in her eyes, is

$$\sum_{j=1}^{n} \frac{a_j^2}{a_j + b_j + c_j}.$$

Thus we want to show that

$$\sum_{j=1}^{n} \frac{a_j^2}{a_j + b_j + c_j} \geq \frac{1}{3}, \tag{26.6}$$

assuming of course that

$$0 \leq a_j \leq 1, \quad 0 \leq b_j \leq 1, \quad 0 \leq c_j \leq 1 \quad \text{for } j = 1, 2, \ldots, n,$$

and

$$\sum_{j=1}^{n} a_j = \sum_{j=1}^{n} b_j = \sum_{j=1}^{n} c_j = 1. \tag{26.7}$$

We will write the difference between the two sides in inequality (26.6) as a sum of positive terms:

$$\begin{aligned}
\sum_{j=1}^{n} \frac{a_j^2}{a_j + b_j + c_j} - \frac{1}{3} &= \sum_{j=1}^{n} \frac{a_j^2}{a_j + b_j + c_j} - \sum_{j=1}^{n} \frac{a_j}{3} \\
&= \sum_{j=1}^{n} \frac{a_j^2 - a_j(a_j + b_j + c_j)/3}{a_j + b_j + c_j} \\
&= \sum_{j=1}^{n} a_j \, \frac{a_j - (a_j + b_j + c_j)/3}{a_j + b_j + c_j}.
\end{aligned} \tag{26.8}$$

Now observe that

$$\sum_{j=1}^{n} (a_j + b_j + c_j)/3 \times \frac{a_j - (a_j + b_j + c_j)/3}{a_j + b_j + c_j} \tag{26.9}$$

$$= \frac{1}{3} \sum_{j=1}^{n} \left(a_j - \frac{a_j + b_j + c_j}{3} \right) = 0$$

by (26.7). Subtracting the expression (26.9), which after all is just zero, from (26.8), we obtain

$$\sum_{j=1}^{n} \frac{\left(a_j - (a_j + b_j + c_j)/3\right)^2}{a_j + b_j + c_j}, \tag{26.10}$$

a sum of positive terms as desired. For aesthetic reasons, we rewrite (26.10) like this:

$$\frac{1}{3} \sum_{j=1}^{n} m_j \left(\frac{a_j - m_j}{m_j} \right)^2 \tag{26.11}$$

with

$$m_j = \frac{a_j + b_j + c_j}{3}.$$

Expression (26.11) is the amount by which A exceeds her fair share. Note that m_j is the average valuation of component j by the three people, and the expression

$$\frac{a_j - m_j}{m_j}$$

appearing in (26.11) is the deviation of a_j from m_j, relative to m_j itself. Thus (26.11) shows that with the method of proportional allocation, A exceeds her fair share by a greater amount if her valuations deviate more strongly from the average valuations. □

We know that for $N = 2$ there is precisely one envy-free, Pareto-optimal, and equitable division procedure, namely, the adjusted winner method. What if $N > 2$? Is there still a division procedure that guarantees envy-freeness, Pareto-optimality, and equitability? The answer is no. Here is a simple example for which we will prove, with some effort, that there is no envy-free, Pareto-optimal, and equitable division.

Example 26.6. A cake consisting of three homogeneous components—chocolate, strawberry, and rhubarb—is to be divided among three people, A, B, and C. The following table shows which fraction of the total cake each of the three components represents to each of the three people:[15]

	chocolate	strawberry	rhubarb
A	0.4	0.5	0.1
B	0.3	0.4	0.3
C	0.3	0.3	0.4

.

For instance, A thinks that the strawberry component is worth half the cake, C thinks that the chocolate component is worth 30% of the cake, etc.

There is a simple fair and equitable division here: Give the entire chocolate component to A, the entire strawberry component to B, and the entire rhubarb component to C. I will call this "division $\mathcal{D}_0$." Under division $\mathcal{D}_0$, each of the three people receives 40% of the cake, in his or her view. ■

We will now prove the following lemma.

Lemma 26.7. *In Example* 26.6*, if $\mathcal{D}$ is a Pareto-optimal, equitable division of the cake, then $\mathcal{D} = \mathcal{D}_0$.*

Before proving this lemma, we address a subtlety that is of very little importance here but which may puzzle an attentive reader: Lemma 26.7 leaves open the possibility that there is no Pareto-optimal and equitable division at all in Example 26.6, that is, that $\mathcal{D}_0$ is not Pareto-optimal. However, Theorem 26.12 will imply that $\mathcal{D}_0$ is indeed Pareto-optimal.

Now we turn to the proof of Lemma 26.7.

Proof. Suppose $\mathcal{D}$ is a Pareto-optimal and equitable division.

Step 1: Under $\mathcal{D}$, each person must get at least 40% of the cake in his or her eyes. Proof: If each person got less than 40% (they all get the same percentage in their eyes, since $\mathcal{D}$ is equitable), then $\mathcal{D}_0$ would be an objective improvement over $\mathcal{D}$, but that is impossible since $\mathcal{D}$ is Pareto-optimal.

[15]See S. Brams and A. Taylor, *Fair Division: From Cake-Cutting to Dispute Resolution*, Cambridge University Press, 1995. Brams and Taylor attribute this example to J. H. Reijnierse and J. A. M. Potters.

Step 2: Under $\mathcal{D}$, A gets no rhubarb cake. Proof: If A got any of the rhubarb cake, then B could get neither chocolate nor strawberry cake, by Lemma 26.1. But then B would get less than 30% of the cake, in his view, which is impossible by Step 1.

Step 3: Under $\mathcal{D}$, B gets no rhubarb cake either. Proof: If B got any rhubarb cake, then C could get neither chocolate nor strawberry cake, again by Lemma 26.1. Thus C would then get rhubarb cake only—but not the entire rhubarb component, since B would get some of it. So C's share of the cake, in her eyes, would be less than 40% of the cake. By Step 1, this is impossible.

Steps 2 and 3 imply that C gets the entire rhubarb component of the cake.

Step 4: Under $\mathcal{D}$, A gets no strawberry cake. Proof: If A got strawberry cake, then B would not get chocolate cake, by Lemma 26.1, since $0.4/0.5 > 0.3/0.4$. But then B would get only strawberry cake (by Step 3, he gets no rhubarb cake), and not all of it, since A would get some of it, so in his eyes, he would get less than 40% of the cake. By Step 1, this is impossible.

We have now concluded that A gets chocolate cake only. Since by Step 1 she must, in her view, get at least 40% of the cake, she must get the *entire* chocolate component. She gets nothing else, so she thinks that she is getting exactly 40% of the cake. Therefore each person must think that he or she gets 40% of the cake, since $\mathcal{D}$ is equitable. Since C gets the entire rhubarb component, which is worth 40% of the cake to her, she cannot get anything beyond the rhubarb component. What remains for B is the strawberry component. So $\mathcal{D} = \mathcal{D}_0$. □

Lemma 26.7 implies the following.

Theorem 26.8. *For $N = 3$, there is no general fair division method for piecewise homogeneous cakes that is envy-free, Pareto-optimal, and equitable.*

Proof. In Example 26.6, if there were an envy-free, Pareto-optimal, and equitable division, that division would have to be $\mathcal{D}_0$ by Lemma 26.7. But $\mathcal{D}_0$ is not envy-free: A envies B for the strawberry component. □

You may object that the preceding discussion is too complicated: There seem to be much simpler examples demonstrating that, in general, Pareto-optimality and equitability alone, without even imposing the additional requirement of envy-freeness, can be incompatible with each other when $N = 3$, as the following example shows.

Example 26.9. Suppose that A, B, and C share a cake consisting of two homogeneous components—chocolate and strawberry. Suppose that their valuations are described by the following table.

	chocolate	strawberry
A	1	0
B	0	1
C	0	1

In words, to A the strawberry component is worthless, and to B and C the chocolate component is worthless. Then any Pareto-optimal division evidently requires that A be given the entire chocolate half of the cake, which to him is the entire cake. But equitability would then require that B and C *each* get the entire strawberry component, which is impossible. Thus there is no division that is both Pareto-optimal and equitable in this example. ∎

Example 26.9 is correct, but it feels a little bit like cheating. In this example, B and C are only interested in the strawberry component of the cake, and A is only interested in the chocolate component. Therefore the real problem is not how to divide the entire cake among A, B, and C, but how to divide the strawberry component among B and C—and that, of course, can be done in a Pareto-optimal, envy-free, equitable way. This is why we gave Example 26.6, in which each person attaches positive value to each component of the cake, so there are no zeros in the table of valuations.

We will return to this point, but first we take up the seemingly unrelated idea of *maximizing the minimal share*, similar to the idea of maximizing the minimal payout discussed in Chapter 15.

Example 26.10. A, B, and C share a cake. In a first proposed division, A thinks he gets 40% of the cake, B thinks she gets 70%, and C thinks he gets 65%. In a second proposed division, A, B, and C each think that they get 45% of the cake. The first division does not maximize the minimal share, for in the first division, the minimal share is 40% of the cake, whereas in the second division, it is 45% of the cake. From the information given, we don't know whether the second division maximizes the minimal share. ∎

Proposition 26.11 (S. Willson, 1995).[16] *Assume that each of the people sharing the cake attaches positive value to each of the homogeneous components of the cake. Then a cake division maximizes the minimal share if and only if it is Pareto-optimal and equitable.*

Proof. Let us first suppose that we are given a cake division that maximizes the minimal share. We want to prove that such a division is Pareto-optimal and equitable.

To see that the division is equitable, suppose it were not. Then we could increase the minimal share by taking away a small amount of cake from one of those with maximal share, and distribute it among those with minimal share. For this argument to work, it is important that those with minimal share attach positive value to each of the homogeneous components of the cake, so that they are certain to benefit from the extra cake they receive.

To see that the division is Pareto-optimal, suppose it were not. We could then find an alternative, objectively improved division which would raise the share of some without lowering anybody's share. Since the given division maximizes the minimal share, the improved one certainly would as well. But according to what we have already proved, it would then be equitable, and therefore it would improve *everybody's* share over the original division. This contradicts our assumption that the original division maximizes the minimal share.

This completes the proof that a cake division that maximizes the minimal share is equitable and Pareto-optimal. Now suppose conversely that we are given a cake division

[16] I learned Proposition 26.11, Theorem 26.12, and their proofs from an unpublished 1995 manuscript titled "Fair division using linear programming" by Stephen J. Willson (Iowa State University, Department of Mathematics), cited here with the author's permission. My presentation of the results and their proofs differs substantially from Willson's, but the ideas are his.

that is equitable and Pareto-optimal. We want to prove that such a division maximizes the minimal share.

If it did not maximize the minimal share, there would be an alternative division with a greater minimal share. Since the given division is equitable, such an alternative division would be an objective improvement. This is impossible because the given division is Pareto-optimal. □

We can now say more precisely in which sense Example 26.9 is cheating: That example is made possible only by the fact that some people consider some components of the cake entirely worthless. There could be no such example otherwise.

Theorem 26.12 (S. Willson, 1995). *Assume that each of the people sharing the cake attaches positive value to each of the homogeneous components of the cake. Then there exists a Pareto-optimal, equitable division. Furthermore, any such division is fair.*

Proof. By Proposition 26.11, it suffices to prove that there is a division that maximizes the minimal share. You may think, "Of course there is such a division. After all, the minimal share can't be infinitely large; in fact it can't be greater than 1. Therefore there must be some division with the largest possible minimal share."

The conclusion is correct. The argument is flawed, but it does not seem very important to me to convince you of that if you don't see it already. Just in case you don't see the flaw and would like to see it, here is a much simpler example of similarly flawed mathematical reasoning that may help in understanding the point: Let S denote the set of all numbers x with $x < 1/2$. Does S have a largest element? You may think, "Obviously yes. After all, elements of S can't be infinitely large; in fact, they can't be greater than $1/2$. Therefore there must be some largest element of S." If you thought this way, you would surely believe that the largest element of S is $x = 1/2$. The trouble, of course, is that $1/2$ is not an element of S, since $1/2$ is not smaller than $1/2$. In fact S does not have any largest element.

I will now give a correct completion of the proof of Theorem 26.12. I will put it in small print since it requires some mathematical background that the reader of this book is not assumed to have.

As always, denote by n the number of homogeneous components that the cake consists of. To specify a cake division, we must specify the fraction x_{ij} of the jth component that the ith person is to receive, $1 \le i \le N$, $1 \le j \le n$. There are two sets of constraints on the x_{ij}:

$$0 \le x_{ij} \le 1 \quad \text{for all } i \text{ and } j, \tag{26.12}$$

$$\sum_{i=1}^{N} x_{ij} = 1 \quad \text{for all } j. \tag{26.13}$$

Denote by a_{ij} the fraction of the cake that the jth component represents to person i. The fraction of the cake received by the ith person, in his or her opinion, is then

$$\sum_{j=1}^{n} x_{ij} a_{ij},$$

and the minimal share is

$$\min_{i=1,2,\ldots,N} \sum_{j=1}^{n} x_{ij} a_{ij}. \tag{26.14}$$

The constraints (26.12) and (26.13) define a compact subset of $\mathbb{R}^{Nn}$, and (26.14) defines a function on this set that is continuous and must therefore have a maximum.

Alternative argument: (26.12)–(26.14) define a linear programming problem that is feasible and bounded and therefore has a solution that can be computed using, for instance, the simplex method. Thus Willson's reasoning does not only prove that there exists a Pareto-optimal, equitable division, but also yields a way of computing one.

Any equitable and Pareto-optimal cake division is fair (see Exercise 16.5). $\square$

We can now show that division $\mathcal{D}_0$ in Example 26.6 is Pareto-optimal: There is an equitable, Pareto-optimal division by Theorem 26.12, and by Lemma 26.7 any such division must be equal to $\mathcal{D}_0$.

We have proved that for $N \geq 3$, one cannot always find an envy-free, Pareto-optimal, and equitable division, but one *can* find one that is Pareto-optimal and equitable, assuming each person attaches positive value to each component of the cake. Dropping the requirement of envy-freeness makes the problem solvable.

In fact, the division problem becomes solvable if any one of the three requirements—envy-freeness, Pareto-optimality, or equitability—is dropped. For instance, one can always find a division that is envy-free and equitable: The equal division is an example. It gives each of the N people exactly $1/N$ of the cake, in everybody's point of view. It is not usually the best possible envy-free and equitable division. See, for instance, Exercise 26.3, where an envy-free, equitable division gives each person 80% of the cake, far more than the 33.33...% that the equal division would give them.

Using the notation introduced in the small-print part of the proof of Theorem 26.12, the constraint of envy-freeness is

$$\sum_{j=1}^{n} x_{ij}a_{ij} \geq \sum_{j=1}^{n} x_{kj}a_{ij} \quad \text{for all } i \text{ and } k \text{ with } 1 \leq i,k \leq N, \tag{26.15}$$

and the constraint of equitability is

$$\sum_{j=1}^{n} x_{ij}a_{ij} = \sum_{j=1}^{n} x_{kj}a_{kj} \quad \text{for all } i \text{ and } k \text{ with } 1 \leq i,k \leq N. \tag{26.16}$$

The linear programming problem of maximizing

$$\sum_{j=1}^{n} x_{1j}a_{1j}$$

under the constraints (26.12), (26.13), (26.15), and (26.16) is feasible (since the equal division satisfies all constraints) and bounded. Therefore an optimal envy-free, equitable division can be computed using, for instance, the simplex method. In this context, "optimal" means "best among envy-free, equitable divisions," not necessarily "Pareto-optimal." For instance, in Example 26.6, there is an envy-free, equitable division that is best among all envy-free, equitable divisions, but such a division cannot be Pareto-optimal. There is no objective improvement that is envy-free and equitable, but there must be an objective improvement that either fails to be envy-free or fails to be equitable, since in Example 26.6, there is no division that is envy-free, Pareto-optimal, and equitable.

A general procedure that is envy-free and Pareto-optimal (but not in general equitable) was created by J. H. Reijnierse and J. A. M. Potters.[17] A complete description of this procedure is beyond the scope of this book. The idea is to imagine that all people sharing the cake are given the same amount of money and bid on the various homogeneous components. The price of a given component rises when the demand for it is too high to be met, and falls when the demand for it is too low for all of it to be sold. A "market" of this sort, appropriately formalized, will produce a division that is envy-free (if X envied Y for Y's share, X could have bought it himself, after all he had the same amount of money as Y) and Pareto-optimal (if there were an objectively better division, the market would have found it). Neither of these statements is obvious; they were proved by Reijnierse and Potters.

Which of the three requirements—envy-freeness, Pareto-optimality, and equitability—should one drop? This is of course a philosophical question, not a mathematical one. It seems to me that one should surely not dispense with Pareto-optimality: If there is an objectively better division, why not use it? I can see an argument for dropping the requirement of equitability: Why should I care about how other people like what they get? I made this argument before, in Chapter 16, using my neighbors' dog Percy as an example. On the other hand, I also see an argument for retaining the requirement of equitability (and therefore dropping the requirement of envy-freeness): A Pareto-optimal, equitable division is at least fair, even though not envy-free in general, and in addition it maximizes the minimal share, by Proposition 26.11.

Exercises

26.1. In Example 26.6, which division would the method of proportional allocation yield? For each of the three people, compute the percentage of the cake that they would think they would get. What is the minimal share?

26.2. A cake consisting of two components is to be shared among three people, A, B, and C. Their valuations are

	first component	second component
A	0.3	0.7
B	0.4	0.6
C	0.8	0.2

.

(a) Which division does proportional allocation yield? (b) For each of the three people, compute how much of the cake he or she gets in his or her eyes. Is the division fair? (c) Is it equitable? (d) Is it Pareto-optimal?

[17] J. H. Reijnierse and J. A. M. Potters, "On finding an envy-free Pareto-optimal division," *Mathematical Programming* 83, pages 291–311 (1998).

26.3. Three people, A, B, and C, share a cake consisting of three components. The valuations are like this:

	first component	second component	third component
A	0.8	0.1	0.1
B	0.1	0.8	0.1
C	0.1	0.1	0.8

.

Suppose that we give the first component to A, the second to B, and the third to C. (a) Show that this is an envy-free division. (b) Show that it is equitable. (c) ($*$) Show that it is Pareto-optimal. (d) Why do the results of parts (a) through (c) not contradict Theorem 26.8?

26.4. ($*$) A cake consisting of two components is to be shared among three people, A, B, and C. Their valuations are like this:

	first component	second component
A	a_1	a_2
B	b_1	b_2
C	c_1	c_2

.

Assume that all entries in this table are positive, that is, nobody considers any of the components of the cake entirely worthless. Show that proportional allocation is Pareto-optimal if and only if all three people have the same valuations: $a_1 = b_1 = c_1$ and $a_2 = b_2 = c_2$.

26.5. Give an example illustrating that the statement in Exercise 26.4 need not hold when there are zeros in the table, that is, when some people consider some of the cake components worthless.

Appendix A

Sets

A *set* S is a collection of objects. If X is one of the objects that belong to the set S, we say that X is an *element* of S. The notation for this is $X \in S$. If X is *not* an element of S, we write $X \notin S$.

For instance, suppose that I is the set of all presidents in U.S. history who have ever been impeached by the House of Representatives. You may be surprised to learn that

$$\text{Richard Nixon} \notin I.$$

(Nixon escaped impeachment by resigning.) In fact,

$$I = \{\text{Andrew Johnson, Bill Clinton}\}.$$

(Both were acquitted by the Senate.)

As this example illustrates, we often indicate a set S as a list of its elements, surrounded by curly brackets. Some sets contain just a single element. For example, the set of all U.S. presidents who did not speak English at home is

$$\{\text{Martin Van Buren}\}.$$

(Van Buren, U.S. president from 1837 to 1841, spoke Dutch to his wife.) A set can even have no element at all. We then call it the *empty set*, and denote it like this:

$$\{\,\}.$$

For instance, at this time the set of all female U.S. presidents is still the empty set. The notion of the empty set may seem a bit strange, but it is sometimes convenient in mathematics and in logic.

Given two sets, say S_1 and S_2, we can form a new set by forming their *union*. The objects in the union are the objects that belong to S_1, to S_2, or to both. We write

$$S = S_1 \cup S_2 = \text{set of all objects } X \text{ with } X \in S_1 \text{ or } X \in S_2. \tag{A.1}$$

On the right-hand side of (A.1), we have used the convention that "A or B" always means that A is true, or B is true, or both are true. In other words, "or" is *inclusive*, not *exclusive*. In mathematics, the word "or" is always used in the inclusive sense.

For example, let

$$B = \text{set of all U.S. presidents who were born on July 4th}$$

and

$$D = \text{set of all U.S. presidents who died on July 4th.}$$

Then

$$B \cup D = \text{set of all U.S. presidents who were born or died on July 4th.}$$

The "or" is inclusive, i.e., a president who was born *and* died on July 4th would be contained in $B \cup D$. However, there has not been any such president. In fact,

$$B = \{\text{Calvin Coolidge}\}$$

(Coolidge was born on July 4, 1872), and

$$D = \{\text{John Adams, Thomas Jefferson, James Monroe}\}.$$

(Adams and Jefferson both died on July 4, 1826, the 50th anniversary of the Declaration of Independence—one in Massachusetts and the other in Virginia. Monroe died exactly 5 years later, on July 4, 1831).

Given two sets S_1 and S_2, we can also form a new set by forming their *intersection*. The elements of the intersection are the objects that belong to both S_1 and S_2. We write

$$S = S_1 \cap S_2 = \text{set of all objects } X \text{ with } X \in S_1 \text{ and } X \in S_2.$$

For instance, let

$$M = \text{set of all U.S. presidents who were born in Massachusetts}$$

and

$$R = \text{set of all U.S. presidents who were re-elected for a second term.}$$

The set M has four elements:

$$M = \{\text{John Adams, John Quincy Adams, John F. Kennedy, George H. W. Bush}\}.$$

The set R, of course, is fairly large. However,

$$M \cap R = \{\ \}.$$

(Of the four presidents born in Massachusetts, none spent more than four years in office.)

Given a set S, we can *remove* an element X from S. The notation for this is $S - \{X\}$. So

$$S - \{X\} = \text{set of all elements of } S \text{ that are not } X.$$

If $X \notin S$, then $S - \{X\} = S$. More generally, if S_1 and S_2 are sets,

$$S_1 - S_2 = \text{set of all elements of } S_1 \text{ that are not in } S_2.$$

For instance, suppose that

$$P = \text{set of U.S. presidents born prior to 1776}$$

and

$$F = \text{set of U.S. presidents who signed the Declaration of Independence.}$$

The set P has 8 elements:

$$P = \{\text{George Washington, John Adams, Thomas Jefferson, James Madison,}$$

$$\text{James Monroe, John Quincy Adams, Andrew Jackson, William Harrison}\}.$$

(The first 7 of these were in fact the first 7 U.S. presidents. Andrew Jackson was succeeded by Martin Van Buren, who was born in 1782. Van Buren was succeeded by William Harrison, who was born in 1773, and was president for precisely one month, from his inauguration on March 4, 1841, to his death on April 4, 1841.) The set F only has 2 elements:

$$F = \{\text{John Adams, Thomas Jefferson}\}.$$

So

$$P - F = \{\text{George Washington, James Madison, James Monroe,}$$

$$\text{John Quincy Adams, Andrew Jackson, William Harrison}\}.$$

If S_1 and S_2 are sets, the notation

$$S_1 \subseteq S_2$$

indicates that S_1 is a *subset* of S_2, i.e., that every element of S_1 is also an element of S_2. (The empty set counts as a subset of any set.) For example, suppose that

$$T = \text{set of U.S. presidents elected in a year ending with a 0 between 1840 and 1960}$$

and

$$O = \text{set of U.S. presidents who died in office.}$$

Then

$$T = \{\text{William Harrison, Abraham Lincoln, James Garfield, William McKinley,}$$

$$\text{Warren Harding, Franklin D. Roosevelt, John F. Kennedy}\},$$

and surprisingly,

$$T \subseteq O.$$

William Harrison died a month after taking office in 1840. Abraham Lincoln, who took office in 1860 and was assassinated soon after winning a second term. James Garfield was elected in 1880 and was fatally shot on July 2, 1881. William McKinley was first elected in 1896, re-elected in 1900, and shot in 1901. Warren Harding, elected in 1920, died in 1923, apparently of natural causes. Franklin D. Roosevelt, re-elected for a third term in 1940, died in office in 1945, shortly after having won a fourth term. John F. Kennedy, elected in 1960, was assassinated in 1963. The 20-year-curse was broken by Ronald Reagan, elected in 1980, who very narrowly escaped an assassination attempt in 1981. George W. Bush completed his second term in good health—a fact which is surely entirely unrelated to some people's belief that he was not rightfully elected in 2000.

If $S_1 \subseteq S_2$, the two sets may in fact be equal. (If $S_1 = S_2$, then of course every element of S_1 is also an element of S_2!)

Exercises

A.1. Suppose that $S_1 = \{3,4,5\}$ and $S_2 = \{4,5,6\}$. Find (a) $S_1 - S_2$, (b) $S_2 - S_1$, (c) $S_1 \cup S_2$, (d) $S_1 \cap S_2$.

A.2. Let $\mathbb{N} = \{1,2,3,4,\ldots\}$, and $E = \{2,4,6,\ldots\}$. (Notice that the symbol $\mathbb{N}$ is not the letter N. Typically N denotes a number, whereas $\mathbb{N}$ always denotes the set $\{1,2,3,4,\ldots\}$.) True or false? (a) $4 \in E$, (b) $8 \subseteq \mathbb{N}$, (c) $\mathbb{N} \subseteq E$, (d) $E \in \mathbb{N}$, (e) $9 \notin E$, (f) $E \subseteq \mathbb{N}$, (g) $E - \mathbb{N} = \{\ \}$.

Appendix B

Logic

In this appendix, I will discuss a few points of logic that have caused confusion for some of my students when I have taught courses from this text. One can read this appendix before reading the main text, as long as one omits those examples that refer to specific points from the main text.

B.1 $A \Rightarrow B$, $B \Rightarrow A$, and $A \Leftrightarrow B$

We assume that "A" and "B" are statements, i.e., sentences that are either true or false. "The moon rises in the east and sets in the west" is a statement in this sense, while "What a mess!" is not.

If "A" and "B" are statements, then so is "$A \Rightarrow B$," or in words, "If A, then B," or "A implies B." This statement is considered true *unless* A is true and B is false. This convention, common in mathematics and logic, can sometimes be a bit strange. For instance, mathematicians would consider the sentence "If $2+2=5$, then the Earth does not revolve around its axis" a true statement. If you think that "nonsensical" would be a better word, I certainly sympathize with your point of view. But the convention of mathematicians and logicians is that $A \Rightarrow B$ is considered true unless A is true and B is false, and therefore $A \Rightarrow B$ is true if A is false, no matter what B may be.

The *converse* of "$A \Rightarrow B$" is "$B \Rightarrow A$," which can also be written as "$A \Leftarrow B$," or expressed in words as "If B, then A", or "B implies A." Everybody knows that $A \Rightarrow B$ and $B \Rightarrow A$ are not the same. If A implies B, then B may or may not imply A. "If N is divisible by 6, then N is divisible by 2" is true, but "If N is divisible by 2, then N is divisible by 6" is false. Nonetheless one can sometimes confuse $A \Rightarrow B$ with $B \Rightarrow A$. Here are a few examples.

Example B.1. Think about the following conversation. "All Democrats support universal health care." "That is not true! My cousin supports universal health care, and he certainly is not a Democrat." Why is this an example of confusing $A \Rightarrow B$ with $B \Rightarrow A$? (This is Exercise B.1.) ■

Example B.2. If your solution to Exercise 12.3 is "This is true by Arrow's theorem," you confused "The method satisfies the unanimity and IIA criteria (and $n \geq 3$) $\Rightarrow$ The method is dictatorship" (Arrow's theorem) with "The method is dictatorship $\Rightarrow$ The method satisfies the unanimity and IIA criteria." ∎

Example B.3. Here is a proposed "solution" to Exercise 21.2: "If all components have equal valuation ratios, all dots in the diagram are vertically aligned. The equal division, in which all components are divided equally, is therefore a threshold division." To understand the logical error here, let "A" denote the statement "The equal division is a threshold division," and "B" the statement "all components have equal valuation ratios." Exercise 21.2 asks you to prove $A \Rightarrow B$. The proposed "solution" proves instead $B \Rightarrow A$. ∎

Example B.4. Many students "solve" Exercise 21.3 by saying, "Well, *of course*, if equations (21.12) holds, then equations (21.11) hold: The ratios are all equal to 1!" But this only proves the utterly obvious statement (21.12)$\Rightarrow$(21.11), not (21.11)$\Rightarrow$(21.12), which is what the exercise is about. ∎

"$A \Leftrightarrow B$," or in words, "A is equivalent to B" or "A if and only if B" means that $A \Rightarrow B$ and $A \Leftarrow B$ are both true; see also Exercise B.2.

B.2 Negation

If "A" is a statement, then so is "not A." One calls "not A" the *negation* of A. It is true if A is false, and false if A is true. In mathematics and logic, one uses the short notation "$\neg A$" for "not A." I have not found it necessary anywhere in the main text of this book to use this notation, but in this appendix it is useful.

B.3 The Contrapositive

The statements

$$A \Rightarrow B$$

and

$$\neg B \Rightarrow \neg A$$

are equivalent (see Exercise B.3). This means: If $A \Rightarrow B$ is true, then so is $\neg B \Rightarrow \neg A$, and if $\neg B \Rightarrow \neg A$ is true, then so is $A \Rightarrow B$. In other words, $A \Rightarrow B$ if and only if $\neg B \Rightarrow \neg A$. The statement $\neg B \Rightarrow \neg A$ is called the *contrapositive* of $A \Rightarrow B$. The word "contrapositive" sounds a lot like "converse," but the two words have entirely different meanings. The *contrapositive* of $A \Rightarrow B$ is just another way of saying $A \Rightarrow B$. The *converse* of $A \Rightarrow B$ is the statement $B \Rightarrow A$, which may be true or may be false if $A \Rightarrow B$ is true.

Example B.5. Mr. Kaltschauer is in the habit of taking a cold shower every morning. He has not had a cold in years, and he attributes his good health to his hardy lifestyle. He says to his neighbor, who has colds frequently, "You get colds because you don't take cold showers."

To analyze Mr. Kaltschauer's comment, let us denote by "A" the statement "You take cold showers," and by "B" the statement "You get colds." Personal experience suggests to Mr. Kaltschauer that $A \Rightarrow \neg B$. However, what he says to his neighbor is $\neg A \Rightarrow B$, which is the same as (namely, the contrapositive of) $\neg B \Rightarrow A$, and thus the converse of $A \Rightarrow \neg B$. ∎

Example B.6. Consider the statement "A cake division that is Pareto-optimal and equitable is fair." (This is true; see Exercise 16.5.) You might think about it the following way. Let us assume that we are given a Pareto-optimal cake division. Statement A is "The division is equitable," and statement B is "The division is fair." Our assertion is then $A \Rightarrow B$. Its contrapositive is $\neg B \Rightarrow \neg A$. In words: "If the division is not fair, it is not equitable." Or, including explicitly the assumption of Pareto-optimality underlying our discussion: "A cake division that is Pareto-optimal but unfair is not equitable." ∎

B.4 Indirect Proof

In many places in this book, we use indirect proofs. They are typically of the following form. We want to prove a statement of the form $A \Rightarrow B$. We assume that B were false, and prove that in that case, we can deduce something that either contradicts A or is just patently false. Since any false statement implies $\neg A$ (and, for that matter, anything at all; see Section B.1), in either case we prove $\neg B \Rightarrow \neg A$, the contrapositive of $A \Rightarrow B$.

B.5 Quantifiers

Suppose that S is a set, and for any $x \in S$, "$A(x)$" is a statement, which depends on x. For example, S could be the set of all real numbers, and "$A(x)$" could be the statement "$x^5 - x^3 - 2 = 0$." Or S could be the set of all preference schedules for which there is a Condorcet candidate, and "$A(x)$" could be the statement "The beatpath method makes the Condorcet candidate in preference schedule x the only winner."

Let us think about the statements

$$\text{For all } x \in S,\ A(x) \text{ is true} \tag{B.1}$$

and

$$\text{There is an } x \in S \text{ for which } A(x) \text{ is true.} \tag{B.2}$$

The phrases "for all" and "there is" are called *quantifiers* in logic. They quantify for how many x the statement $A(x)$ is true—for all, or for at least one. (By "there is," we always mean "there is at least one" here.) Quantifiers are so common in mathematics, in logic, and in everyday reasoning that it is useful to think, in a formal way, about the negation of a statement involving a quantifier. The negation of (B.1) is

$$\text{There is an } x \in S \text{ for which } \neg A(x) \text{ is true.} \tag{B.3}$$

The negation of (B.2) is

$$\text{For all } x \in S,\ \neg A(x) \text{ is true.} \tag{B.4}$$

Thus when one negates a statement beginning with a quantifier, two things change: First, the quantifier changes: "For all" becomes "there is," and vice versa. Second, the statement that is being quantified is negated: "$A(x)$" turns into "$\neg A(x)$."

Example B.7. "Every child in first grade must keep a journal during winter vacation" is a "for all" statement: S is the set of children in first grade, and for a child $x \in S$, $A(x)$ is the statement "x must keep a journal during winter break." The negation is *not* "Children in first grade need not keep journals during winter break." That would be "For all $x \in S$, $\neg A(x)$ is true." The correct negation of "For all $x \in S$, $A(x)$ is true" is "There is an $x \in S$ for which $\neg A(x)$ is true," that is, "There is at least one child in first grade who need not keep a journal during winter break." ∎

Example B.8. There are many "for all" and "there is" statements in disguise in this book. For instance, let us take the statement that Borda count violates the majority criterion (Chapter 2). You might say, "Not necessarily. Think about this preference schedule:

1	2	1
A	A	B
B	C	A
C	B	C

Here there is a majority candidate, namely, A, and indeed A wins by Borda count. Therefore your statement is not necessarily true." But that would reflect a misunderstanding of the statement that Borda count violates the majority criterion. By convention, this is a "there is" statement: We say that Borda count violates the majority criterion because *there is* at least one example exhibiting a violation of the majority criterion. We don't say (although we reasonably *could*) that Borda count violates the majority criterion for some preference schedules, and satisfies it for others. ∎

In mathematics and in logic, it is customary to make the convention that the statement

$$\text{For all } x \in \{\},\ A(x) \text{ is true.}$$

is considered true, no matter what $A(x)$ may be. (The symbol $\{\}$ denotes the empty set.) For instance, the statement

My six-year-old son is taller than all those of his first grade classmates who are more than 65 years old.

is true because he has no classmates who are more than 65 years old. (This is another example of a statement that you might want to call "nonsensical" instead of "true," but mathematicians and logicians make the convention of calling it true.)

B.6 When Is Proof by Example Valid?

To prove a statement of the form (B.2), one need not do anything other than produce a single $x \in S$ for which $A(x)$ is true. So here proof by example is perfectly valid.

Similarly, if one wants to prove that (B.1) is *false*, a single counterexample suffices. To prove that (B.1) is false is the same as proving that (B.3) is true, which is a "there is" statement, and for "there is" statements, it is enough to produce a single example—in this case, an example of an $x \in S$ for which $\neg A(x)$ is a true statement.

On the other hand, to prove (B.1), it is *not* enough to check whether $A(x)$ is true for some $x \in S$ that one chooses. Statement (B.1) is that $A(x)$ is true *for all* $x \in S$, and to prove that, one would have to check *all* $x \in S$—or come up with some general argument that proves that it is true for all $x \in S$ without checking all $x \in S$ explicitly.

Example B.9. Proposition 2.10 makes two statements. The first statement is that sequential comparison satisfies the Condorcet criterion. This means that *for all* preference schedules for which there is a Condorcet candidate, sequential comparison makes the Condorcet candidate the sole winner. This is a "for all" statement, and to prove it requires a general argument.

The second statement of Proposition 2.10 is that sequential comparison violates the unanimity criterion. This is, in disguise, a "there is" statement: It is the statement that *there is* an example of a preference schedule in which a candidate X is unanimously, by all voters, ranked above Y, and nonetheless Y is among the winners according to sequential comparison. In the proof of Proposition 2.10, this is proved with a single example, preference schedule (2.4). ∎

Example B.10. In Exercise 4.4, is it enough to give an example in which there is a Condorcet candidate, and to show that in that example, $S = \{X\}$? The answer is no. The exercise asks you to prove the following: If X is a Condorcet candidate, then $S = \{X\}$. This is a "for all" statement in thin disguise: For all preference schedules for which there is a Condorcet candidate X, $S = \{X\}$. Therefore a single example is not sufficient, you have to find a general argument. ∎

Exercises

B.1. Explain the logical problem in Example B.1.

B.2. The words "A if and only if B" mean the same as "$A \Rightarrow B$ and $A \Leftarrow B$." Does the "if" correspond to "$\Rightarrow$" and the "only if" to "$\Leftarrow$", or is it the other way around?

B.3. Explain why $\neg B \Rightarrow \neg A$ is just another way of saying $A \Rightarrow B$.

B.4. Explain: To prove that a winner selection method, or a ranking method, or a fair division method, *satisfies* a certain criterion, a single example is never enough. To prove that a method *violates* the criterion, a single example is always enough.

Appendix C

Mathematical Induction

Carl Friedrich Gauss (1777–1855) was among the most influential mathematicians in history, perhaps *the* most influential. People tell the following story about him.[18] His primary school teacher wanted to occupy the class for a while, and asked them to add the natural numbers from 1 to 100:

$$1+2+3+\cdots+99+100 = \ ?.$$

Within seconds, the young Gauss said "5050," the correct answer. How did he do it—or in any case, how could he have done it? He observed that the sum can be computed like this:

$$(1+100)+(2+99)+(3+98)+\cdots+(49+52)+(50+51)$$

$$= 101+101+101+\cdots+101+101 = 50\times 101 = 5050.$$

Proposition C.1. *If $N \geq 1$ is any natural number, then*

$$1+2+\cdots+N = \frac{N(N+1)}{2}. \tag{C.1}$$

Try it out: For $N=2$, the left-hand side is $1+2$, and the right-hand side is $(2\times 3)/2 = 3$. For $N=3$, the left-hand side is $1+2+3$, and the right-hand side is $(3\times 4)/2 = 6$. For $N=4$, the left-hand side is $1+2+3+4$, and the right-hand side is $(4\times 5)/2 = 10$. For $N=100$, the right-hand side of (C.1) is $(100\times 101)/2 = 5050$, the answer given by Gauss. For $N=1$, we must first pause to think about what the left-hand side of (C.1) is supposed to mean. It is the sum of the first N natural numbers in general, so for $N=1$, it is the "sum of the first 1 natural numbers"—and although that is a bit confusing and even ungrammatical, we take it to mean 1. Then (C.1) holds for $N=1$, since the right-hand side is $(1\times 2)/2 = 1$.

How can we be sure that (C.1) holds for *all* natural numbers N? One possibility is to imitate the argument of the six-year-old Gauss, with 100 replaced by N (Exercise B.1). We will instead give a proof based on the following principle.

[18] It is not known whether this story is true.

Principle of mathematical induction. Suppose that $A(N)$ is a statement about natural numbers N. To prove that $A(N)$ is true *for all* natural numbers N, it suffices to prove the following:

(i) $A(1)$ is true, and

(ii) if $A(N)$ is true, then $A(N+1)$ is true as well.

If we know that $A(N)$ implies $A(N+1)$, and if $A(1)$ holds, then $A(2)$ holds, and therefore $A(3)$ holds, and therefore $A(4)$ holds, and so on. You might think of an infinite line of dominoes, one corresponding to each natural number. If the first domino falls, and if each domino that falls pushes down the next one in the line, all will fall eventually.

We now use the principle of mathematical induction to prove Proposition C.1.

Proof. Here $A(N)$ is the statement that (C.1) holds. $A(1)$ is the statement that it holds for $N=1$, and we already discussed why it does. To prove (C.1) for all N, we must prove that if $A(N)$ holds, so does $A(N+1)$. It is important to understand that we do not assume that $A(N)$ holds for all N; we merely prove that *if* it holds for a given N, then $A(N+1)$ holds as well.

If $A(N)$ holds, then the sum of the first N natural numbers equals $N(N+1)/2$. We want to prove that $A(N+1)$ must then hold as well, which means that the sum of the first $N+1$ natural numbers equals $(N+1)(N+2)/2$—namely, the expression that you get when you replace "N" by "$N+1$" on the right-hand side of (C.1).

If the sum of the first N natural numbers equals $N(N+1)/2$, then

$$\text{sum of first } N+1 \text{ natural numbers} = \text{sum of first } N \text{ natural numbers} + N + 1$$

$$= \frac{N(N+1)}{2} + N + 1 = \frac{N(N+1)}{2} + \frac{2(N+1)}{2} = \frac{N(N+1)+2(N+1)}{2}$$

$$= \frac{(N+2)(N+1)}{2} = \frac{(N+1)(N+2)}{2}.$$

Indeed this is the desired expression. □

Gauss's teacher might have made the task a little harder by asking for the sum of the *odd* natural numbers below 100:

$$1+3+5+\cdots+99 = \ ?.$$

Is there still a formula that applies to this problem? Let us first try some examples:

$$1+3=4, \tag{C.2}$$
$$1+3+5=9, \tag{C.3}$$
$$1+3+5+7=16, \tag{C.4}$$
$$1+3+5+7+9=25, \tag{C.5}$$
$$1+3+5+7+9+11=36. \tag{C.6}$$

The numbers on the right-hand side are perfect squares: 2^2, 3^2, 4^2, 5^2, 6^2. This leads us to the following guess.

Proposition C.2. *For any natural number N,*

$$1+3+5+\cdots+(2N-1)=N^2. \tag{C.7}$$

(Notice that $2N-1$ is an odd number whenever N is a natural number.) Substituting $N=2$, 3, 4, 5, and 6 into formula (C.7), we get the examples (C.2) through (C.6). Formula (C.7) is easy to prove by mathematical induction.

Proof. To prove (C.7) by induction, we must first prove that it is true when $N=1$. What is the left-hand side when $N=1$? In general, the left-hand side is the sum of the odd numbers between 1 and $2N-1$. For $N=1$, it is "the sum of the odd numbers between 1 and 1." Again, this is a bit confusing perhaps, and even ungrammatical, but we take it to be equal to 1. The right-hand side is $1^2=1$. So for $N=1$, (C.7) is true.

Now we *assume* that (C.7) is true for a given N, and we *prove* that it is then true for $N+1$ as well, that is,

$$1+3+5+\cdots+(2N-1)+(2(N+1)-1)=(N+1)^2. \tag{C.8}$$

Note that $2(N+1)-1=2N+1$, so we can write (C.8) more simply like this:

$$1+3+5+\cdots+(2N-1)+(2N+1)=(N+1)^2. \tag{C.9}$$

Thee left-hand side of (C.9) is

$$[1+3+5+\cdots+(2N-1)]+2N+1.$$

Using (C.7), this equals

$$N^2+2N+1,$$

and this equals $(N+1)^2$. So indeed (C.9) holds. $\square$

Gauss's teacher could have made the task *much* harder by asking for the sum of the first 100 *squares*:

$$1^2+2^2+3^2+4^2+\cdots+99^2+100^2 = ?.$$

There is a formula just like (C.1) that is applicable here.

Proposition C.3. *For any natural number N,*

$$1^2+2^2+3^2+4^2+\cdots+(N-1)^2+N^2=\frac{N(N+1)(2N+1)}{6}. \tag{C.10}$$

For instance, for $N=2$,

$$1^2+2^2=\frac{2\times3\times5}{6},$$

since both sides of the equation are equal to 5. For $N=100$, we find

$$1^2+2^2+3^2+4^2+\cdots+99^2+100^2=\frac{100\times101\times201}{6}=338,350.$$

Formula (C.10) is harder to find than formula (C.1), but once somebody *tells* you the formula, it is not so hard to *prove* that it holds using mathematical induction.

Proof. $A(N)$ is now the statement that (C.10) holds, that is, the statement that the sum of the first N squares equals $N(N+1)(2N+1)/6$. The first step in the proof is to verify that $A(1)$ holds. The left-hand side of (C.10) is $1^2 = 1$ when $N = 1$. The right-hand side is $(1 \times 2 \times 3)/6 = 1$. So indeed $A(1)$ is true.

To complete the proof, we must prove the following: If $A(N)$ holds, then so does $A(N+1)$. Indeed, if the sum of the first N squares equals $N(N+1)(2N+1)/6$, then the sum of the first $N+1$ squares equals

$$\begin{aligned} \frac{N(N+1)(2N+1)}{6} + (N+1)^2 &= \frac{N(N+1)(2N+1)}{6} + \frac{6(N+1)^2}{6} \\ = \frac{N(N+1)(2N+1)+6(N+1)^2}{6} &= \frac{(N(2N+1)+6(N+1))(N+1)}{6} \\ = \frac{(2N^2+7N+6)(N+1)}{6} &= \frac{(N+2)(2N+3)(N+1)}{6} \\ &= \frac{(N+1)(N+2)(2N+3)}{6}, \end{aligned} \tag{C.11}$$

which is the right-hand side of (C.10) with "N" replaced by "$N+1$." ◻

The principle of mathematical induction is used only once in this course, in Chapter 18. There, it is used not in the exact form that we have presented, but in a slightly stronger form:

Strong principle of mathematical induction. Suppose that $A(N)$ is a statement about natural numbers N. To prove that $A(N)$ is true for *all* natural numbers N, it suffices to prove the following:

(i) $A(1)$ is true, and

(ii) if $A(1)$, $A(2)$, $A(3)$, ..., and $A(N)$ are true, then so is $A(N+1)$.

If we know (i) and (ii), then we know $A(1)$ holds, and therefore $A(2)$ holds as well, and therefore $A(3)$ holds, therefore $A(4)$ holds, and so on. I will give an example illustrating this principle. First we need a definition.

Definition C.4. *A natural number $N \geq 2$ is called a composite number if it can be written as a product of two natural numbers between* 2 *and $N-1$. It is called a prime number if it is not a composite number.*

The ten smallest prime numbers are 2, 3, 5, 7, 11, 13, 17, 19, 23, and 29.

Proposition C.5. *Let N be a natural number. Then N is either equal to* 1*, or a prime number, or a product of two or more prime numbers.*

You almost certainly know this, but you may not know how to prove it. Here is a proof using the strong principle of mathematical induction.

Proof. Here $A(N)$ is the statement "N is either equal to 1, or a prime number, or a product of two or more prime numbers." Clearly $A(1)$ is true! To complete the proof, using the strong principle of mathematical induction, we must now prove the following: If $A(1)$, $A(2),\ldots,A(N)$ are true, then $A(N+1)$ is also true. So suppose that $A(1)$, $A(2),\ldots,A(N)$ are true. To prove that $A(N+1)$ is true, we will prove that either $N+1$ is a prime number or $N+1$ is a product of prime numbers. If $N+1$ is a prime number, then we are done. But if it is not a prime number, then it is a composite number, so it can be written as a product,

$$N+1 = N_1\,N_2,$$

where N_1 and N_2 are natural numbers between 2 and N. Remember that we are assuming that $A(1)$, $A(2)$, $A(3),\ldots,A(N)$ are true. In particular, we are assuming that $A(N_1)$ and $A(N_2)$ are true. Since both N_1 and N_2 are ≥ 2, this means that N_1 and N_2 are either prime numbers or products of prime numbers. In either case, $N+1 = N_1\,N_2$ is also a product of prime numbers. $\square$

Exercises

C.1. Derive the formula $1+2+3+\cdots+N = N(N+1)/2$ using an argument like the one ascribed to the six-year-old Gauss for $N = 100$. Start like this:

$$1+2+3+\cdots+N = (1+N)+(2+(N-1))+(3+(N-2))+\cdots.$$

You will need to think about the case of even N and odd N separately.

C.2. Verify that (C.11) is the right-hand side of (C.10) with "N" replaced by "$N+1$."

C.3. Use mathematical induction to prove the following: For all natural numbers N,

$$1+3+3^2+\cdots+3^{N-1} = \frac{3^N-1}{2}.$$

C.4. Use mathematical induction to prove the following: For all natural numbers N and all numbers $r \neq 1$,

$$1+r+r^2+\cdots+r^{N-1} = \frac{r^N-1}{r-1}.$$

C.5. Approximately what is

$$1+\frac{1}{5}+\frac{1}{5^2}+\frac{1}{5^3}+\frac{1}{5^4}+\cdots+\frac{1}{5^{100}}\ ?$$

(Hint: Use the result of Exercise C.4. Plug in $r = 1/5$, $N = 101$, and notice that $1/5^{101}$ is very close to zero.)

Appendix D

Solutions to Selected Exercises

1.3. When there are 4 candidates, each voter confers 4 Borda points on her favorite candidate, 3 on her second favorite, etc. So she confers $4+3+2+1=10$ Borda points altogether. When there are 18 voters, the total number of Borda points is therefore $18 \times 10 = 180$. Similarly, when there are 20 voters and 3 candidates, the total number of Borda points is $20 \times (3+2+1) = 120$.

1.4. Each head-to-head competition gives rise to one pairwise comparison point, regardless of whether one of the two candidates wins or they tie. Therefore there are half as many pairwise comparison points in total as there are *ordered* pairs (X,Y) of candidates with $X \neq Y$ (half as many because the head-to-head competition between X and Y is the same as the head-to-head competition between Y and X). There are $n(n-1)$ ordered pairs of candidates (each candidate can be paired with $n-1$ others), and therefore $n(n-1)/2$ pairwise comparison points in total. For $n=4$, $n(n-1)/2 = 4 \times 3/2 = 6$.

1.5. Suppose that j is the number of voters preferring X to Y, and k is the number of voters preferring Y to X. Then $j+k=N$. If $j>k$, then the margin of victory is $j-k = j+k-2k = N-2k$, which is even if and only if N is even. If $k>j$, then similarly the margin of victory is $k-j=k+j-2j=N-2j$, which again is even if and only if N is even.

1.7. A majority candidate X is in first place on more than half of the ballots. In any head-to-head competition between X and another candidate Y, all those voters who rank X first would vote for X, and therefore X would win.

1.8. Let us do our best to try to make B the winner, even though A is the first choice of 3 out of the 4 voters. We have to put A first in 3 ballots, but let us put B second in those 3 ballots, and first in the 4th ballot. Also, let us try to put A last in the 4th ballot, to make it hard for A to win:

3	1
A	B
B	C
C	A

.

Even though we have clearly done the best we possibly could have done to help B win, in fact A has 10 Borda points, and B has 9, so A wins.

1.9. As in Exercise 1.8, we do our best to help B win, given the constraint imposed by the problem that A should be the first choice of three out of five voters. This means placing B second on the three ballots on which A is first, placing B first on the remaining two ballots, and placing A last on those ballots:

3	2
A	B
B	C
C	A

Here indeed A has 11 Borda points, and B has 12, so B wins.

2.1. (a) True. (b) False. (c) False: Borda count satisfies the unanimity criterion, but not the majority criterion. (d) False: Sequential comparison satisfies the majority criterion, but not the unanimity criterion.

2.4. (a) Choose any candidate X. Since X is not a Condorcet candidate (there is no Condorcet candidate, by assumption), there is a candidate, say Y, who beats X in head-to-head competition. Since Y is not a Condorcet candidate either, there is a candidate, say Z, who beats Y in head-to-head competition. Continuing like this, we can construct a chain of candidates in which each member of the chain is beaten by the next member in head-to-head competition. Eventually, a candidate must appear for a second time in this chain, since there are only finitely many candidates. Thus the circle closes. (b) If the preference schedule is

1	1	1
D	D	D
A	B	C
B	C	A
C	A	B

,

then a majority prefer A to B, B to C, C to A (so there is a circle in societal preferences), but everybody prefers D to the rest, so D is a Condorcet candidate.

2.5.

1	1
Z	X
X	Y
Y	Z

.

2.9. (a) A is eliminated first, since A has the fewest first-place votes. Then C beats B, so C is the overall winner. (b) C is eliminated first, since C has the most last-place votes. Then A beats B, so A is the overall winner. (c) C is the majority candidate, with 8 out of 15 first-place votes. (d) C is a majority candidate, but the Coombs method does not make C the winner.

2.10. If X ranks above Y on every ballot, then X derives more Borda points than Y from every voter, and therefore has more Borda points overall. Therefore Y is not among the winners.

3.1. (a) B wins by the plurality method. However, if C is retroactively disqualified, then A wins. (b) A wins by the runoff method. However, if B is retroactively disqualified, then C wins. (c) The elimination method and the runoff method are the same when $n = 3$. (d) B wins by the Coombs method. However, if C is retroactively disqualified, then A wins. (e) B has 20 Borda points, while A and C only have 17 each. So B wins by Borda count. However, if C is retroactively disqualified, then A wins. (f) In head-to-head competition, A beats B, B beats C, and C beats A. Thus $W = \{A, B, C\}$ by pairwise comparison. However, when A is retroactively disqualified, then $W = \{B\}$. (g) C wins by sequential comparison if the candidates are ordered alphabetically. However, if A is retroactively disqualified, then B wins.

3.2. The spoilers in the solution to Exercise 3.1 are all losing spoilers, except for part (f), where A is a winning spoiler, and there is no losing spoiler.

3.3. The simplest example is

1	1
A	B
C	A
B	C

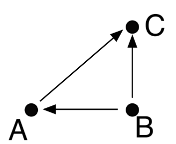
.

In head-to-head competition, A ties with B, A beats C, and B ties with C. Therefore $W = \{A\}$ by the method of pairwise comparison. If, however, C is retroactively disqualified, then $W = \{A, B\}$. Thus C is a losing spoiler.

3.4. (a) In head-to-head competition, A beats both B and C, 7:6. (b) By the plurality method, C wins. When A is retroactively disqualified, however, then B wins.

3.5. Suppose that a majority of voters considers A the best candidate and assigns a check to him, but some voters don't consider A desirable at all and assign no check to him. Suppose that B, on the other hand, is considered acceptable by everybody, so every voter assigns a check to B. Then B beats A. (Of course, quite arguably B *should* beat A in such a situation, so it is not clear that lack of majority-fairness in the sense explained here should be considered a flaw of approval voting.)

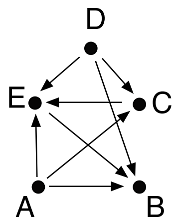
4.1. It is best to draw the pairwise comparison graph first:



(There is no need to note margins of victory, since they play no role here.) This shows that B is a Condorcet candidate. Therefore $S = \{B\}$, and the method of pairwise comparison must make B the sole winner. The dominating sets are

$$\mathcal{D}_A = \{A,B\}, \quad \mathcal{D}_B = \{B\}, \quad \mathcal{D}_C = \{A,B,C\}.$$

4.2. Again it is best to draw the pairwise comparison graph first:



To find, for instance, $\mathcal{D}_A$, we first note that A itself belongs to $\mathcal{D}_A$. But if A belongs to $\mathcal{D}_A$, so does D since A does not beat D in head-to-head competition. A and D beat the three others in head-to-head competition, and therefore

$$\mathcal{D}_A = \{A,D\}.$$

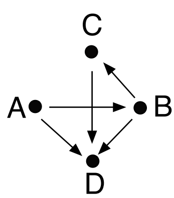
Similarly, we find

$$\mathcal{D}_B = \{A,B,C,D,E\}, \quad \mathcal{D}_C = \{A,B,C,D,E\},$$

$$\mathcal{D}_D = \{A,D\}, \quad \mathcal{D}_E = \{A,B,C,D,E\}.$$

Thus there are two dominating sets in this example: $\{A,D\}$ and $\{A,B,C,D,E\}$. The Smith set is $\{A,D\}$. In the pairwise comparison tournament, A and D get 3.5 points each, C gets 1.5 points, E gets 1 point, and B gets 0.5 points. Thus A and D are also the winners by pairwise comparison.

4.5. Both the set of winners by pairwise comparison and the Smith set depend only on the pairwise comparison graph without margins of victory. The following pairwise comparison graph yields $W = \{A\}$ by the pairwise comparison method, but $S = \{A,B,C\}$:



Here is a preference schedule that produces this pairwise comparison graph:

2	1	1
A	C	B
B	A	C
C	B	A
D	D	D

.

4.8. Suppose we are given an example that demonstrates that the method of sequential comparison violates the unanimity criterion, that is, an example in which every voter prefers candidate X over candidate Y, yet Y wins by the method of sequential comparison. Then Y is a Smith candidate by Proposition 4.10, and therefore the same example demonstrates that the Smith method violates the unanimity criterion.

5.1. (a) The Smith candidates are A, B, and C. Among these, A has one first-place vote, and B and C have two each. So B and C win according to the a posteriori Smith-fair plurality method. (b) After D is removed from the preference schedule, B has four first-place votes, C has three, and A has two. Therefore B is the sole winner according to the a priori Smith-fair plurality method.

5.2. (a) C wins by the elimination method, regardless of whether or not D is on the ballot, so D is not a weak spoiler. (b) A wins by Borda count, regardless of whether or not D is on the ballot, so D is not a weak spoiler. (c) Since pairwise comparison is a priori Smith-fair (see Exercise 5.8), it does not permit weak spoilers, so D is not a weak spoiler. In fact, C wins by pairwise comparison, regardless of whether or not D is on the ballot.

5.4. Consider this example:

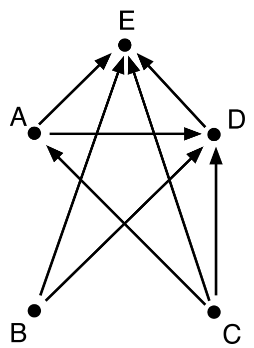
2	2	3	3
A	A	B	C
B	C	C	B
C	B	A	A

.

In head-to-head competition, B and C tie. They both beat A. Therefore $S = \{B, C\}$, but by the plurality method, A is the winner.

5.5. In head-to-head competition, a Smith candidate X must not only beat every candidate outside the Smith set S (for $n-k$ points), but also tie with at least one other Smith candidate (this is where the extra half point comes from), otherwise $S - \{X\}$ would be a dominating set smaller than S.

5.6. The pairwise comparison graph looks as follows:



So A, B, and C each beat both D and E in head-to-head competition. This implies that $\{A,B,C\}$ is a dominating set. It is clear that it does not have any dominating subset, therefore $S = \{A,B,C\}$.

5.8. We must prove that the outcome of the pairwise comparison tournament can never depend on the presence or absence of non-Smith candidates. Suppose that some number l of non-Smith candidates were removed a priori. This would affect the scores of all Smith candidates in the pairwise comparison tournament in the same way: Each of them would lose l points. Therefore the outcome of the tournament would remain unchanged.

5.10. (a) Consider this example:

2	2	3	3
A	A	B	C
B	C	C	B
C	B	A	A

.

In head-to-head competition, B and C tie. They both beat A, so $S = \{B,C\}$. If we use the plurality method, when A is the sole winner. So A is a non-Smith candidate who affects the outcome of the election, thus a weak spoiler. But A is not a spoiler in the sense of Chapter 3; in fact, a sole winner can never be a spoiler in the sense of Chapter 3. (b) Consider this preference schedule:

2	1	1
A	C	B
B	A	C
C	B	A

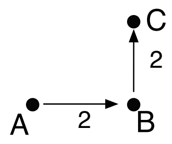
.

(This is Example 3.3.) Here $S = \{A,B,C\}$, so there is no weak spoiler, since there is no non-Smith candidate. If we use the pairwise comparison method, $W = \{A\}$, but if B (a loser) is retroactively disqualified, then $W = \{A,C\}$. Thus B is a losing spoiler in the sense of Exercise 3.2.

6.1. The unmatched beatpath $C \to A$, of strength 15, eliminates A. The unmatched beatpath $A \to B$, of strength 15, eliminates B. The unmatched beatpath $B \to D$, of strength 11, eliminates D. Thus there cannot be an unmatched beatpath against C, otherwise nobody would be left standing! At first sight, you might think that $D \to C$ is an unmatched beatpath, since $C \to A \to D$ is only of strength 3. But $C \to A \to B \to D$ is of strength 11.

6.2. A, B, and C form a circle, and they each beat D, E, and F in head-to-head competition. Therefore $S = \{A,B,C\}$. Since the beatpath method is a priori Smith-fair, one may disregard non-Smith candidates a priori, and only consider A, B, and C and the arrows among them. C is not a winner because the beatpath $B \to C$ is unmatched. There are no unmatched beatpaths against A and B, so $W = \{A,B\}$.

6.3. The pairwise comparison graph looks like this:

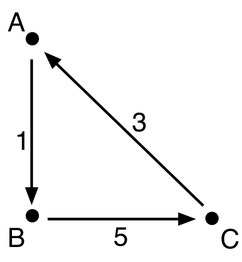


Here A has a beatpath against B of strength 2 which B cannot match, and B has a beatpath against C of strength 2 which C cannot match. So $W = \{A\}$ by the beatpath method.

6.4. If 3 voters prefer A to B and 3 voters prefer B to C, then at least 2 voters must prefer both A to B and B to C, since there are only a total of 4 voters altogether. If also 3 voters prefer C to A, then among the 2 voters who prefer both A to B and B to C, there must be at least one who also prefers C to A—and therefore has circular preferences. This contradiction proves the assertion.

6.5. The key to both statements is that no beatpath ever enters the Smith set from outside, since no non-Smith candidate beats any Smith candidate in head-to-head competition. For part (a), observe that if $X \in S$ and $Y \notin S$, then X beats Y in head-to-head competition, so there is a direct beatpath from X to Y. This beatpath is unmatched, since a matching beatpath would have to enter S from outside. For part (b), note that a beatpath against a Smith candidate that would pass through any non-Smith candidate would have to enter the Smith set S from outside.

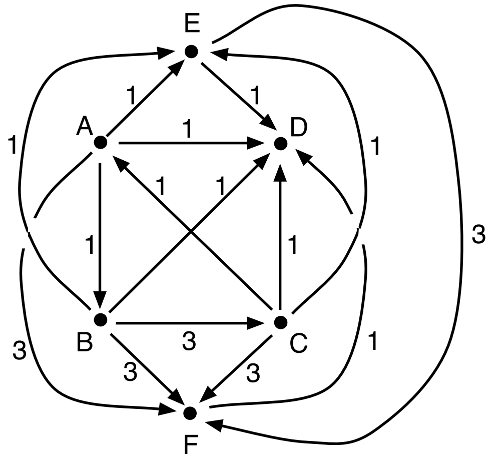
6.6. For the preference schedule of Exercise 3.1, the pairwise comparison graph looks like this:



B has a beatpath of strength 5 against C which C cannot match, and C has a beatpath of strength 3 against A which A cannot match. Therefore B is the sole winner. Since A would win in the absence of C, C is a losing spoiler.

7.1. (a) B is eliminated first, since B has the fewest first-place votes. Then A beats C, 16:8, so A wins. (b) Now C is eliminated first, since now C has the fewest first-place votes. Then B beats A, 13:11, so B wins.

7.2. The pairwise comparison graph is



The Smith set is

$$S = \{A, B, C\}.$$

Removing the non-Smith candidates D, E, and F from the preference schedule, we obtain

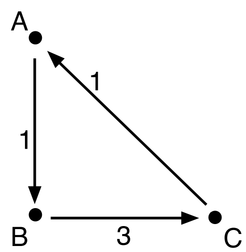
3	2	1	1
A	B	C	C
B	C	A	B
C	A	B	A

.

(a) The Borda scores, based on the preference schedule with D, E, and F removed, are

$$A: \ 14, \quad B: \ 15, \quad C: 13.$$

So $W = \{B\}$ by a priori Smith-fair Borda count. (b) Since the beatpath method is a priori Smith-fair, we are allowed to focus only on the piece of the pairwise comparison graph involving Smith candidates:



Here B has a beatpath of strength 3 against C, and C cannot match that beatpath. There are no other unmatched beatpaths. Therefore $W = \{A, B\}$ by the beatpath method. (Note that neither A nor B even have a single first-place vote!)

7.3. When X is moved upwards in a ballot, X cannot lose any first-place vote, and nobody other than X can gain any first-place vote. If X has a plurality of first-place votes prior to the change, the same therefore holds after the change.

7.4. Suppose that swapping X with the candidate Y immediately above X on a single ballot turns X from a winner into a loser, according to the a priori Smith-fair plurality method. Then the swap must change the Smith set; otherwise moving X upwards in a ballot could not possibly spoil the outcome for X. Denoting by S the Smith set prior to the swap, and by S' the Smith set after the swap, we know from Lemma 7.11 that $S' \subseteq S$, $S' \neq S$, $X \in S'$, $Y \in S - S'$. The number of candidates in S' cannot be one; otherwise X would be the only Smith candidate after the change, and thus a Condorcet candidate, and therefore X would not be turned into a loser by the swap. The number of candidates in S' cannot be two either; otherwise S' would contain X and one other candidate with whom X would have to tie in head-to-head competition, and both would win according to the a priori Smith-fair plurality method after the swap. So S' must contain at least three candidates. In addition, $S - S'$ contains at least one candidate (namely, Y), so S must contain at least four candidates.

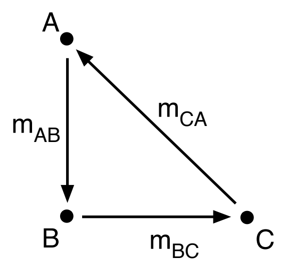
7.5. In Example 7.1, the Smith set is the set of all candidates, both before and after the change in the ballots, and a priori Smith-fair runoff is therefore the same as plain runoff.

7.6. This follows from the fact that the number of Borda points that a candidate X derives from a given ballot equals the number of candidates above whom X is ranked on that ballot plus 1.

8.1. (a) A has a beatpath against B of strength 13, which B cannot match (her beatpath against A is of strength 11). Therefore B does not win. Similarly, B has a beatpath against C of strength 15, and C cannot match that beatpath. Therefore C does not win. The sole winner is A. (b) B has a beatpath against C of strength 15, which C cannot match. Therefore C is not a winner. Neither A nor B have unmatched beatpaths against them, so $W = \{A, B\}$. (c) B has a beatpath of strength 15 against C, and C cannot match that. (The beatpath from C to B is only of strength 13.) Therefore C is not a winner. Similarly, C has a beatpath of strength 15 against A, and A cannot match that. Therefore A is not a winner. The sole winner is B.

8.2. Since N is odd, no head-to-head competition can lead to a tie. Assume there is a tie in pairwise comparison. Then A cannot lose against both B and C: Otherwise either B or C would be a Condorcet candidate (depending on which of them beats the other in head-to-head competition) and therefore the sole winner by pairwise comparison. Similarly, A cannot win against both B and C: Otherwise A would be a Condorcet candidate and therefore the sole winner by pairwise comparison. So A must win one head-to-head competition, and lose another, if there is to be a tie in pairwise comparison. The same is then, of course, true for B and C as well: Each candidate wins one competition, and loses the other. If A beats B, B must beat C, and C must beat A. If A beats C, C must beat B, and B must beat A.

8.3. If there is no circle in societal preferences, then one candidate must win two head-to-head competitions, and therefore be a Condorcet candidate and the sole winner according to the beatpath method, which after all is Condorcet-fair. This shows that there can be a tie according to the beatpath method only if there is a circle in societal preferences. Suppose now that there is a circle, for instance,



with margins $m_{AB} > 0$, $m_{BC} > 0$, and $m_{CA} > 0$. We say that m_{AB} is the strength of the *direct beatpath* against B, m_{BC} is the strength of the *direct beatpath* against C, and m_{CA} is the strength of the *direct beatpath* against A. A candidate wins, according to the beatpath method, if and only if the direct beatpath against the candidate is of minimal strength (not stronger than the direct beatpath against any other candidate); see Exercise 8.1 for examples. Thus there is a tie if and only if there are at least two different links of minimal strength along the circle.

8.5. (a) First, we call the candidate whom the first voter puts first A, the one whom she puts second B, and the one whom she puts last C, so the first column of the preference schedule is

A
B
C

.

If there is to be no Condorcet candidate, of the two remaining voters, one must rank B first, and the other C. (In any other case, one candidate would have two first-place votes, and therefore be a majority candidate.) If necessary, we renumber the voters so that the second voter is the one who puts B first, so the preference schedule takes this form:

A	B	C
B	?	?
C	?	?

.

B is certain to beat C in head-to-head competition, 2:1. If B is not to be a Condorcet candidate, B must lose against A in head-to-head competition, and therefore the third voter must rank A above B:

A	B	C
B	?	A
C	?	B

.

If A is not to be a Condorcet candidate, then the second voter must now place C above A:

A	B	C
B	C	A
C	A	B

.

(b) When there is a Condorcet candidate, all Condorcet-fair methods make that candidate the sole winner. When there is no Condorcet candidate, then up to renaming candidates and renumbering voters, the preference schedule is

A	B	C
B	C	A
C	A	B

.

Any winner selection method that obeys the principles of one person, one vote and independence of candidate names must yield a three-way tie in this case: $W = \{A, B, C\}$.

9.1. (a) When the preference schedule is (9.1), $W = \{A, B, C\}$. When the preference schedule is changed to (9.2), B becomes the Condorcet candidate, and therefore the sole winner. So A no longer wins, even though no voter has changed his or her ranking of A vs. B or A vs. C. (b) When the preference schedule is (9.1), A is the sole winner by the beatpath method. But when the preference schedule is changed to (9.2), B becomes the Condorcet candidate, and therefore the sole winner, even though no voter has changed his or her ranking of A vs. B or A vs. C.

9.2. Suppose that a winner selection method satisfies the unanimity criterion. Suppose that X is ranked first by each voter. Then for each other candidate Y, X is preferred to Y by every voter, and therefore Y is not a winner. Thus X is the sole winner.

9.3. Suppose that a winner selection method is Condorcet-fair, and that a candidate X ranks first on every ballot. Then X is a Condorcet-candidate, and therefore the sole winner. Therefore the method is Pareto-efficient.

9.5. When there are only two candidates, then knowing how each voter ranks X in comparison with all other candidates—that is, with the other candidate, since there is only one other candidate—means knowing the detailed preference schedule. So when there are only two candidates, the IIC criterion does not mean anything more than that one can determine the winners if one knows the detailed preference schedule, and any winner selection method satisfies that condition. An example of a single-winner method that is also Pareto-efficient and monotonic is the plurality method, with ties broken by alphabetic order. So this is an example of a method that satisfies all assumptions of the Muller–Satterthwaite theorem (for only two candidates), yet is not dictatorial.

10.1. Without any dishonest votes, the Borda scores of A, B, and C are 16, 15, and 17, respectively. The winner according to Borda count is therefore C. If one of the two voters whose ranking is

B
A
C

dishonestly votes

A
B
C

instead, A will get an extra Borda point at the expense of B, so now the Borda scores of A, B, and C are 17, 14, and 17, respectively. There is a tie for victory between A and C, and since ties are resolved alphabetically, A now wins. This is an example of strategic voting because the dishonest voter prefers A to C.

10.2. We have to explain why nobody can change the election outcome in a way that he or she likes by casting a dishonest ballot. Those voters who are not the dictator cannot affect the election outcome at all. The dictator can only affect the election outcome by casting a ballot on which the first-ranked candidate is not the one whom he or she truly likes best. Therefore the dictator cannot change the election outcome *in a way that he or she likes* by casting a dishonest ballot.

10.3. When there are only two candidates, the plurality method (with ties resolved alphabetically) is an example of a Pareto-efficient, strategy-proof single-winner method that is not dictatorship.

11.1. The candidate with the largest number of first-place votes is E, so E is ranked first. After E is removed, the preference schedule becomes

4	4	3	2	1
A	C	C	B	A
C	B	A	A	B
D	D	D	C	D
B	A	B	D	C

.

Now C has the most first-place votes, so C is ranked second. After removal of C, the preference schedule becomes

4	4	3	2	1
A	B	A	B	A
D	D	D	A	B
B	A	B	D	D

,

or in compact form,

7	4	2	1
A	B	B	A
D	D	A	B
B	A	D	D

.

Now A has the most first-place votes, so A is ranked third. After removal of A, the preference schedule becomes

7	4	2	1
D	B	B	B
B	D	D	D

,

or, compactly,

7	7
D	B
B	D

.

So D and B are ranked equally, at the bottom. The overall societal ranking is

$$E \succ C \succ A \succ B \sim D.$$

11.2. C is a Condorcet candidate, and therefore C is ranked first. After removal of C, the preference schedule is

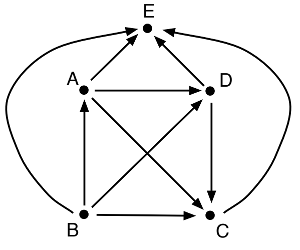
6	3
B	A
A	B

.

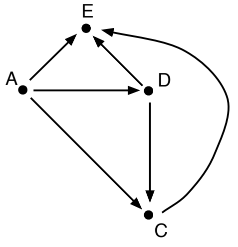
So B is ranked second, and A third:

$$C \succ B \succ A.$$

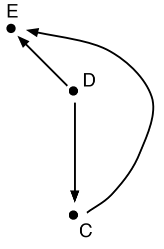
11.3. The pairwise comparison graph looks like this:



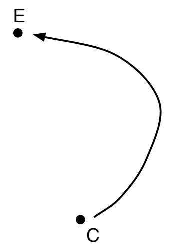
B is a Condorcet candidate, so B wins by any Condorcet-fair method and is ranked first. After B is removed from the preference schedule, the pairwise comparison graph looks like this:



Now A is a Condorcet candidate, and therefore wins by any Condorcet-fair method. So A is ranked second. Then A is removed, and the remaining pairwise comparison graph looks like this:



So now D is a Condorcet candidate, wins by any Condorcet-fair method, and is ranked third. After removal of D, the remaining pairwise comparison graph is



C is ranked fourth, and E last:

$$B \succ A \succ D \succ C \succ E.$$

11.4. (a) $C \succ A \sim E \succ B \succ D$. This is not the ranking computed, using the recursive plurality method, in Example 11.1. (b) If one-shot plurality ranking were a recursive ranking method, it would have to be equivalent to recursive application of the plurality (winner selection) method, since the plurality method is the winner selection method derived from one-shot plurality ranking. But in part (a) we have an example demonstrating that one-shot plurality ranking is *not* the same as recursive application of the plurality method.

11.5. We start out with some winner selection method, which we call M, and denote by M' the ranking method derived by recursive application of M. From the ranking method M', we can derive a winner selection method, and we call that method M''. We want to prove that M'' and M are the same. To determine the winner(s) using M'', one ranks the candidates using M', then makes the first-ranked candidate(s) the winner(s). But the first-ranked candidate(s) according to M' is (are) the winner(s) according to M. Thus M'' and M select the same winner(s) and are therefore the same method.

11.6. The winner selection method derived from one-shot Borda count declares the candidate(s) ranked first by one-shot Borda count the winner(s). One-shot Borda count ranks the candidate(s) with the largest Borda score first. So the winner selection method derived from one-shot Borda count declares the candidate(s) with the largest Borda score the winner(s), and thus it is Borda count.

12.1. The example in the proof of Proposition 4.14 works here:

1	1	1
A	C	B
B	A	D
D	B	C
C	D	A

.

$S = \{A, B, C, D\}$, and therefore the recursive Smith method ranks all four candidates equally, even though every voter prefers B to D.

12.2. Suppose that the preference schedule is

3	2	2
A	B	C
B	A	B
C	C	A

.

Recursive Borda count ranks the candidates like this:

$$B \succ A \succ C.$$

If the preference schedule is altered like this:

3	2	2
A	B	C
C	A	B
B	C	A

,

the ranking becomes

$$A \succ C \succ B.$$

The relative societal ranking of A vs. B has changed—before, B was preferred to A, and now A is preferred to B by society—even though all voters compare A with B the same way after the change as before.

12.6. Either A is strictly first and C is strictly last, or vice versa, so either $A \succ B \succ C$ or $C \succ B \succ A$.

12.8. For two candidates, ranking by the plurality method satisfies the unanimity criterion and the IIA criterion, and it is not dictatorship.

13.1. The house should be assigned to A. The compensation amount that A pays to B should lie between \$200,000 and \$250,000.

13.2. (a) For example, A could get ownership of the house and pay \$100,000 to each of the others. (This is the *equal* arrangement; see Chapter 15.) (b) For example, B could get ownership of the house and pay each of the others their fair shares. This will leave B with a payout of \$95,000, above his fair share of \$90,000. (c) C's bid is below the average bid of \$355,000. (d) For an envy-free arrangement, A must get the house, and B, C, and D must all receive the same compensation, not smaller than \$90,000 (so B, the second-highest bidder, is treated fairly), but not larger than \$100,000 (so A is treated fairly).

13.3. Suppose that $N = 5$ and that A is the winning bidder. Is it possible for A to compensate B, C, D, and E in such a way that the arrangement becomes fair? For fairness, A must pay at least \$$b/5$ to B, \$$c/5$ to C, \$$d/5$ to D, and \$$e/5$ to E. The question is whether A can do that while retaining enough so that the arrangement is fair to her as well. The answer is yes if and only if

$$a - \frac{b}{5} - \frac{c}{5} - \frac{d}{5} - \frac{e}{5} \geq \frac{a}{5},$$

or

$$a \geq \frac{a+b+c+d+e}{5},$$

that is, if and only if the winning bid a is at least as large as the average bid.

13.4. (a) Subtract $c/4$ from both sides of inequality (13.2):

$$\frac{3}{4}c \geq \frac{a+b+d}{4}.$$

Then multiply both sides by $4/3$. (b) Yes, since w lies above the average of all bids if and only if w lies above the average of all bids except the winning one. This is analogous to saying that your score on an exam is above the class average if and only if it is above your classmates' average. Whether or not your own score is included in the average can certainly affect the numerical value of the average, but has no bearing on the question whether or not your score is above average.

13.5. In this example, $x_A = 75 > a/3 = 30$, $x_B = 10 = b/3$, and $x_C = 5 > c/3 = 0$.

13.6. In inequality (13.7), subtract $a/4$ from both sides and add $3x$ to both sides:

$$\frac{3}{4}a \geq 3x.$$

Then divide both sides by 3:

$$\frac{a}{4} \geq x.$$

This is inequality (13.6).

13.7. The left upper vertex lies on the lines $x_B = b/3$ and $x_B + x_C = 2a/3$, so $x_C = 2a/3 - x_B = 2a/3 - b/3 = (2a-b)/3$. The left lower vertex lies on the lines $x_B = b/3$ and $x_C = c/3$. The right vertex lies on the lines $x_C = c/3$ and $x_B + x_C = 2a/3$, so $x_B = 2a/3 - x_C = 2a/3 - c/3 = (2a-c)/3$.

14.1. (a) B should get the car. (b) $4{,}000 \leq x \leq 5{,}200$. (c) The winning bid is $w = 10{,}400$ dollars, and the average bid is $m = 9{,}200$ dollars. Therefore the compensation should be $w/m = 10{,}400/9{,}200 = 104/92 = 26/23$ times A's fair share:

$$x = \frac{26}{23} \times 4{,}000 \approx 4{,}521.74.$$

14.2. The winning bid is $w = 500{,}000$ dollars. The average bid is $m = 350{,}000$ dollars. Therefore the compensation amount should be

$$\frac{w}{m} = \frac{500{,}000}{350{,}000} = \frac{50}{35} = \frac{10}{7}$$

times B's fair share, that is,

$$x_B = \frac{10}{7} \times 100{,}000 \approx 142{,}857.14.$$

14.3. (a) The highest bidder, C. (b) The winning bid is $w = 80,000$ dollars, and the average bid is
$$m = \frac{80,000 + 60,000 + 100,000 + 80,000}{4} = 80,000$$
dollars as well. Therefore the compensation amounts should be the fair shares: $x_B = 15,000$, $x_C = 25,000$, and $x_D = 20,000$. This is a fair arrangement since $w \geq m$—namely, $w = m$.

14.4. (a) The winning bid is $w = 300,000$ dollars. The average bid is $m = 290,000$ dollars. Therefore the compensation amounts should be 30/29 times the fair shares:
$$x_A = \frac{30}{29} \times 110,000 \approx 113,793.10,$$
$$x_C = \frac{30}{29} \times 80,000 \approx 82,758.62.$$
(b) Yes, since $30/29 > 1$. (c) If the house is assigned to A, then $w/m = 33/29 > 30/29$, and therefore the equitable arrangement yields a higher payout for all three siblings.

14.5. (a) The winning bid is $w = a$ dollars. The average bid is $m = (a+b)/2$ dollars. Therefore the compensation amount should be
$$x_B = \frac{w}{m}\frac{b}{2} = \frac{a}{(a+b)/2}\frac{b}{2} = \frac{ab}{a+b}$$
dollars. (b) The winning bid is $w = b$ dollars. The average bid is $m = (a+b)/2$ dollars. Therefore the compensation amount should be dollars.
$$x_A = \frac{w}{m}\frac{a}{2} = \frac{b}{(a+b)/2}\frac{a}{2} = \frac{ab}{a+b}$$

15.1. Suppose that A receives the object and pays \$$x$ to all others. To make the arrangement equal, A's payout must be \$$x$ as well, that is,
$$a - (N-1)x = x,$$
or
$$x = \frac{a}{N}.$$
Thus the compensation amount in an equal arrangement must be the winning bidder's fair share. If the winning bidder is not a highest bidder, this is unfair to those with higher bids. If, on the other hand, the winning bidder is a highest bidder, it is fair to everybody, and since all compensation amounts are equal to each other, it is also envy-free, by Proposition 13.14.

15.2. B first pays \$12,000 to the referee in exchange for receiving the piano. The referee pays \$9,000 to A and is left with \$3,000, which are split evenly: B gets \$1,500 back, and A gets another \$1,500, bringing her total compensation to $x_A = 10,500$ dollars.

15.3. First, A gets the booty and pays the referee \$30,000. The referee pays \$9,000 to B, \$10,500 to C, and \$11,500 to D, spending a total of \$31,000. The referee has paid \$1,000 of his own money at this point, so he charges each of the four pirates \$250. (These are very civilized pirates, so they go along with all this.) The overall result is $x_B = 8,750$, $x_C = 10,250$, and $x_D = 11,250$. Each pirate gets \$250 less than his fair share—a consequence of A's unwise decision to take the booty even though he is a relatively low bidder.

15.4.

$$\frac{1}{N} + \frac{w-m}{S} = \frac{1}{N}\left(1 + N\frac{w-m}{S}\right) = \frac{1}{N}\left(1 + \frac{w-m}{S/N}\right)$$
$$= \frac{1}{N}\left(1 + \frac{w-m}{m}\right) = \frac{1}{N}\frac{m+w-m}{m} = \frac{1}{N}\frac{w}{m}.$$

15.5. Suppose, for instance, that A is the winning bidder. (For the following calculations, it does not matter whether $a \geq b$ or $a < b$.) (a) In the first stage of Knaster's procedure, A receives the object and pays $\$a/2$ to the referee. The referee pays $\$b/2$ to B and is left with $\$(a-b)/2$. (If $a < b$, this is negative.) In the second stage, the referee distributes this amount equally: $\$(a-b)/4$ to A, and $\$(a-b)/4$ to B. (If $a < b$, then A and B pay the referee $\$(b-a)/4$ each.) The effective compensation amount is

$$x_A = \frac{b}{2} + \frac{a-b}{4} = \frac{2b+a-b}{4} = \frac{a+b}{4}$$

dollars. (b) The winning bid is $w = a$, and the average bid is $m = (a+b)/2$. The compensation amount therefore is

$$\frac{w}{m}\frac{b}{2} = \frac{a}{(a+b)/2}\frac{b}{2} = \frac{ab}{a+b}.$$

If B is the winning bidder, the answers to parts (a) and (b) are the same, except that a and b need to be swapped. However, the formulas $(a+b)/4$ and $ab/(a+b)$ remain the same when a and b are swapped, so in fact the answers to parts (a) and (b) are the same no matter who is the winning bidder. (c)

$$\frac{a+b}{4} - \frac{ab}{a+b} = \frac{(a+b)^2 - 4ab}{4(a+b)} = \frac{a^2 + 2ab + b^2 - 4ab}{4(a+b)} = \frac{(a-b)^2}{4(a+b)}.$$

16.1. (a) Benjamin gets two-thirds of the strawberry component of the cake. In his view, the strawberry component is two-thirds of the cake, by value. So the fraction of the cake that he thinks he gets, by value, is $2/3 \times 2/3 = 4/9 > 1/3$. Similarly, Jacob thinks he gets 4/9 of the cake, by value. Lilly gets one-third of the cake in her eyes, since she simply goes by volume. So all three get a fair share, and therefore the division is fair. (b) Benjamin prefers his share to Jacob's or Lilly's, since he prefers strawberry to chocolate. Jacob prefers his share to Benjamin's and Lilly, since he prefers chocolate to strawberry. In Lilly's eyes, all three shares are equal. Therefore the division is envy-free. (c) Yes, the division is Pareto-optimal. To see this, notice first that any objective improvement of the proposed division could not assign less than one-third of the total volume to Benjamin, since the third of the volume that he gets in the proposed division is maximally desirable to him. For similar reasons, an objective improvement could not assign less than one-third of the volume to Jacob. It could also not assign less than one-third of the total cake volume to Lilly, since she determines the value of her portion by volume, and gets one-third in the proposed division. So none of the three children could get less than one-third of the volume of the cake in an objectively improved division. But one would have to get more than one-third for it to be an improvement. This is impossible. (d) The division is not equitable, since Benjamin and Jacob each think they get 4/9 of the cake, by value, whereas Lilly thinks she gets one-third, by volume or by value.

16.2. (a) By value, Benjamin thinks he gets the entire cake, Jacob thinks he gets 5/6 of the cake, Lilly thinks she gets one-third of the cake, and Cole thinks he gets half the cake. These fractions are all greater than 1/4, so the division is fair. (b) The division is not envy-free: Lilly envies Jacob. This is the only incident of envy. (c) The division is not equitable, since different children think that they get different fractions of the cake, by value.

16.3. Any person who thinks that his or her fair share is worth no less than the shares of any of the other people must think that his or her fair share is worth at least $1/N$ of the cake.

16.5. An equitable division that is unfair is not Pareto-optimal: The equal division is an objective improvement. (For more on the logic of this argument, see Example B.6 in Appendix B.)

16.6. Suppose that A and B share a cake consisting of a strawberry half and a chocolate half, and that A likes strawberries but does not each chocolate, whereas B likes chocolate but does not each strawberries. If we assign the chocolate half to A, and the strawberry half to B, each thinks that they get none of the cake, by value. This is equitable—zero equals zero—but not fair.

17.1. Each chooser is assigned a piece that he or she considers fair. C must get the first piece, and B can get either the second or the third. A gets the remaining piece.

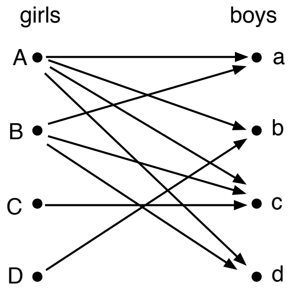
17.2. B should get the first piece, C the second, and A the third.

17.3. There is a conflict: B and C consider the first piece—but none of the others—fair. The first piece, together with one of the others, should therefore be set aside. The remaining piece should be assigned to A. Then B and C should recombine the two pieces that have been set aside and divide them among each other using “I cut, you choose.”

17.6. Consider a cake that consists of three homogeneous components, and suppose that A likes only the first component, B likes only the second, and C likes only the third. In the lone divider method, the divider’s share will be worth to her exactly one-third of the cake. However, in this example, there is obviously a division in which each of the three people gets a piece that is worth the entire cake to him or her.

18.1. The three girls A, B, and C compete for the two boys b and c, so $\mathcal{G} = \{A, B, C\}$ violates the condition. There is no other example of a set $\mathcal{G}$ violating the marriage condition here.

18.2. (a) Diagram:



(b) This problem is so small that we can easily discuss all nonempty sets $\mathscr{G}$ of "girls" systematically. The ones with only one element are easy: Since each "girl" likes at least one "boy" (namely, each of A, B, C, and D considers at least one piece of cake a fair share), $\#\mathscr{B}_{\mathscr{G}} \geq \#\mathscr{G}$ if $\#\mathscr{G} = 1$. Second, let us think about sets $\mathscr{G}$ with $\#\mathscr{G} = 2$. Since A, B, and D consider at least two pieces of cake fair shares, $\#\mathscr{B}_{\mathscr{G}} \geq 2 = \#\mathscr{G}$ if A, B, or D belong to $\mathscr{G}$; but one of them must belong to $\mathscr{G}$ if $\mathscr{G}$ has two elements. Third, let us think about sets $\mathscr{G}$ with $\#\mathscr{G} = 3$. If A is included in $\mathscr{G}$, then $\mathscr{B}_{\mathscr{G}} = \{a,b,c,d\}$, so $\#\mathscr{B}_{\mathscr{G}} \geq \#\mathscr{G}$. If A is not included in $\mathscr{G}$, then $\mathscr{G} = \{B,C,D\}$ and $\mathscr{B}_{\mathscr{G}} = \{a,b,c,d\}$. Finally, if $\mathscr{G} = \{A,B,C,D\}$, then of course $\#\mathscr{B}_{\mathscr{G}} = \#\mathscr{G} = 4$.

(c) and (d) C must get c, and D must get b. Then B can get either a or d, and A gets the remaining piece.

(e) B will envy C for piece c, which B considers more valuable than any of the others, but can't get, since it is the only piece that C considers a fair share.

19.1. The marriage condition is satisfied here. There is exactly one way of assigning the pieces to the four people so that everybody gets a fair share: b to D, d to B, c to C, and a to A.

19.2. $\{B,C\}$ is a conflicted set, while $\{B,C,D\}$ is not. So $\{B,C\}$ is a maximal conflicted set. We set aside B and C, together with piece b (which they both consider a fair share) and—arbitrarily—piece c (which neither B nor C considers fair, but we must set aside as many pieces as people). The remaining people are A and D, and we can assign the remaining pieces to them so that both A and D get fair shares: A gets d and D gets a. Pieces b and c are recombined, and B and C share them using "I cut, you choose."

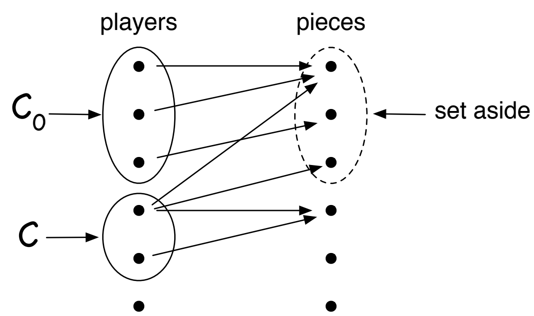
19.3. $\{B,C,D\}$ is a (in this case, *the*) maximal conflicted set. We therefore set aside B, C, and D, together with the pieces b, c, and one extra piece, say a. Then A gets piece d. Pieces a, b, and c are recombined, and B, C, and D share them using Steinhaus's lone divider method.

19.4. The divider considers each of the N pieces a fair share—that's how she cuts the pieces. So if a set $\mathcal{S}$ of people includes the divider, the total number of pieces that at least one person in $\mathcal{S}$ considers a fair share must be N.

19.5. Here $\{B,C,D,E\}$ and $\{C,D,E,F\}$ are conflicted, and they are both maximal conflicted sets, since $\{B,C,D,E,F\}$ is not conflicted, and no set of people that includes A is conflicted.

19.6. We set aside one of the maximal conflicted sets, for instance, $\{B,C,D,E\}$, together with those pieces that are considered a fair share by any of the people in this set, i.e., b,c, and d, and one extra piece, let's say a. The remaining pieces, e and f, are distributed among the remaining people in such a way that they each get a fair share. We know from Lemma 19.1 that this must be possible, and indeed it is: A gets piece e, and F gets piece f (or vice versa). Then the pieces that had been set aside, a, b, c, and d, are recombined and divided fairly among the people who had been set aside, B, C, D, and E.

19.7. Here is a picture that explains this point:



20.1. (a) B should take the trimmed piece, since C did not take it. (b) C. (c) B.

20.2. (a)

A: parrot, 14 detective novels
B: all books, minus half the works of Dickens, and minus 22 detective novels
C: dining room furniture, half the works of Dickens, 8 detective novels.

(b) A gets the parrot and some books, while B gets books only, and not even all of them. Even if B got *all* the books and A got *only* the parrot, A would consider their shares equal. (c) A considers the parrot and the dining room furniture equally valuable. He had a chance to choose half the works of Dickens and 8 detective novels in the second round, but chose 14 detective novels instead, so he must consider those at least equally valuable.

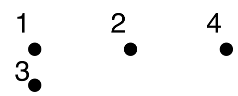
20.3. The divider is guaranteed to have a fair share after the first round, since he gets one of the original pieces. The trimmer is not: In his view, only one of the original three pieces may be fair, and he is forced to trim that piece to make it tie with the second-most valuable piece, which may not be fair in his eyes. The chooser is not guaranteed a fair share in the first round either: The only piece that he considers fair may be the one that the trimmer trims, and after trimming it may no longer be fair in the chooser's eyes.

20.4. U envies nobody because he gets to choose first. W envies nobody because he considers all three pieces equal. V does not envy W because he gets to choose prior to W.

20.5. (a) (1) X may envy A. (2) Y may envy A. (3) Nobody envies anybody. The only *possible* envy is that A might envy X. Indeed, A might envy X for his share of the trimmings in the second round. But overall A will never envy X, since A would not even envy X if X got *all* the trimmings in the second round: That would just give X one of the three original pieces overall, and A considers all three of those pieces equal in value. (4) Y may envy X. (5) A may envy Y. (6) X may envy Y. (b) See the explanation given under (3) above.

20.6. B, the trimmer, sees two pieces that are worth half the cake each, and one worthless piece. Therefore B sees no need to do any trimming at all; two pieces already tie for first place. Then C takes one of the two pieces that both B and C consider worth half the cake (the strawberry and rhubarb components of the cake), B takes the other, and A gets the chocolate component, which she considers worth one-third of the cake.

21.1. (a) The valuation ratios are 2, 3/4, 2, and 1/3. The components should therefore be ordered like this:



(b) (i) Threshold division. The vertical line goes through the first and third components.

(ii) Not a threshold division. Since the first component is split, the vertical line would have to go through the first and third components. But then the second and fourth components would have to go to B, not A, in a threshold division.

(iii) Threshold division. The vertical line goes through the first and third components.

(iv) Threshold division. The vertical line goes through the fourth component.

(v) Not a threshold division. Since the second component goes to A, the third component, which has a greater A-to-B valuation ratio than the second component, would have to go to A as well in a threshold division.

(vi) Threshold division. The vertical line goes through the first and third components.

21.2. Each component is split. Therefore the vertical line must run through each component. This is possible only if the dots representing the components are all aligned vertically, i.e., if they all have equal A-to-B valuation ratios.

21.3. Assume (21.11). Suppose that r is the common value of all the A-to-B valuation ratios:

$$a_j = rb_j$$

for all j. Then

$$1 = a_1 + a_2 + \cdots + a_n = r(b_1 + b_2 + \cdots + b_n) = r,$$

so $r = 1$, and therefore (21.12) holds.

21.5. (a) The A-to-B valuation ratio of the nuts is 2, while that of the chocolates is 1/2. In the even division, both are shared. In a threshold division, items that are shared must have the same A-to-B valuation ratio. Thus the even division is not a threshold division. (b) For A to be as satisfied as before, the value of what he loses, in his eyes, must equal the value of what he gains. The fraction ϵ of the bag of chocolates that he gives up is the fraction $\epsilon/3$ of the booty to him. The fraction δ of the bag of nuts is the fraction $2\delta/3$ of the booty to him. So we need to set $\epsilon = 2\delta$. For instance, $\epsilon = 1/2$ and $\delta = 1/4$ will work: A gives all the chocolates he has (half a bag) to B, and in return B gives half the nuts he has (a quarter of a bag) to A. Seen out of B's eyes, he gains the fraction $2\epsilon/3$ of the booty and loses the fraction $\delta/3$. With $\epsilon = 1/2$ and $\delta = 1/4$, his total gain is the fraction

$$\frac{1}{2} \times \frac{2}{3} - \frac{1}{4} \times \frac{1}{3} = \frac{1}{3} - \frac{1}{12} = \frac{1}{4}$$

of the booty.

22.1. In order of decreasing A-to-B valuation ratio, the items are

custody, vacation house, cash, home.

The initial division is to give the custody and the vacation house to John, and the cash and the home to Jane. With this division, John thinks that he gets 65% of the total, and Jane thinks that she gets 75%. To make the division equitable, we must take something away from Jane and give it to John. Since the cash is, among the two items that Jane receives, the one for which the John-to-Jane valuation ratio is larger (that ratio is 2/3 for the cash, and 1/3 for the home), we should redistribute the cash. Jane retains a fraction p of the cash, and John gets the fraction $1 - p$ of it. In Jane's eyes, she gets the fraction

$$0.45 + 0.30p$$

of the total, and John gets, in his eyes, the fraction

$$0.50 + 0.15 + 0.20(1 - p).$$

Setting these two expressions equal to each other, we get

$$0.45 + 0.30p = 0.65 + 0.20(1 - p).$$

Subtracting 0.45 from both sides and adding 0.20p to both sides, we find

$$0.5p = 0.4,$$

so

$$p = \frac{4}{5}.$$

Thus Jane should get the home and 80% (namely, 4/5) of the cash, and John should get the custody, the vacation home, and 20% of the cash. With this arrangement, both will feel that they got 69% of the total.

22.2. (a) John thinks that the car represents the fraction $6{,}000/26{,}000 = 3/13$ of their total joint belongings, whereas the cash represents the fraction $20{,}000/26{,}000 = 10/13$. Jane thinks that the car represents the fraction $4{,}000/24{,}000 = 1/6$ of their total belongings, whereas the cash represents the fraction $20{,}000/24{,}000 = 5/6$. The following table summarizes this:

	car	cash
John	3/13	10/13
Jane	1/6	5/6

.

The John-to-Jane valuation ratios are $18/13 > 1$ for the car, and $12/13 < 1$ for the cash, so at first John should get the car, and Jane the cash. John then thinks that he gets $3/13$ of the total, whereas Jane thinks she gets $5/6$. To make the division equitable, some of the cash has to be redistributed. Denoting by p the fraction of the cash that Jane retains, so that John receives the fraction $1-p$ of the cash, John now thinks that he receives this fraction of the total:

$$\frac{3}{13}+(1-p)\,\frac{10}{13}=1-\frac{10}{13}\,p\,.$$

Jane thinks she gets this fraction:

$$\frac{5}{6}\,p\,.$$

We set these two fractions equal to each other and solve for p:

$$\begin{aligned}1-\frac{10}{13}\,p&=\frac{5}{6}\,p,\\ 1&=p\left(\frac{10}{13}+\frac{5}{6}\right),\\ p&=\frac{78}{125}.\end{aligned}$$

So Jane should get

$$p\times 20{,}000=\frac{78}{125}\times 20{,}000=12{,}480$$

dollars in cash, while John gets the rest of the cash and the car.

(b) Jane pays \$10,000 into the joint account and gets \$12,480 back. John pays \$10,000 as well, but only gets back \$7,520 (namely, the amount left over after Jane gets her \$12,480). In effect, John has paid \$2,480 in compensation to Jane.

(c) Repeat the calculations of parts (a) and (b) with a bank account containing not \$20,000 but $v\times \$1{,}000$, where v is a large number. Don't use any specific number for v, just leave the letter v. I am omitting the calculation, which is lengthy but entirely analogous to the calculations in parts (a) and (b). The conclusion from the calculation is that, in effect, Jane gets

$$\frac{5+24/v}{2+10/v}\times \$1{,}000$$

in compensation for the car. (You can convince yourself that for $v=20$, the case discussed in parts (a) and (b), you get \$2,480 from this formula, the answer found in part (b).) For very large v, this is very close to

$$\frac{5+0}{2+0}\times \$1{,}000=\$2{,}500,$$

the compensation amount that Knaster's procedure would yield (see Exercise 15.5).

22.3. (a) $F = -p_1a_1 - p_2a_2 + p_3a_3 + p_4b_4 + p_5b_5$.
(b) $G = p_1b_1 + p_2b_2 - p_3b_3 - p_4b_4 - p_5b_5$.
(c) Suppose that $G > 0$, that is,

$$p_1b_1 + p_2b_2 > p_3b_3 + p_4b_4 + p_5b_5. \tag{D.1}$$

Now recall that

$$\frac{a_1}{b_1} > \frac{a_2}{b_2} > \frac{a_3}{b_3} > \frac{a_4}{b_4} > \frac{a_5}{b_5}.$$

Let r be a number between a_2/b_2 and a_3/b_3:

$$\frac{a_1}{b_1} > \frac{a_2}{b_2} > r > \frac{a_3}{b_3} > \frac{a_4}{b_4} > \frac{a_5}{b_5}.$$

Then

$$a_1 > rb_1, a_2 > rb_2, \qquad \text{but } a_3 < rb_3, a_4 < rb_4, \text{ and } a_5 < rb_5. \tag{D.2}$$

Using (D.2) in (D.1), we obtain

$$p_1\frac{a_1}{r} + p_2\frac{a_2}{r} > p_3\frac{a_3}{r} + p_4\frac{a_4}{r} + p_5\frac{a_5}{r}\,.$$

Multiply both sides of this inequalities by r:

$$p_1a_1 + p_2a_2 > p_3a_3 + p_4a_4 + p_5a_5\,.$$

This means $F < 0$.

(d) If G were positive and F were not negative, the altered division would be an objective improvement of the threshold division that we started out with.

24.1. If you evaluate either formula at $x = 0.4$, you get $S(0.4) = 0.8$.

24.2. You see the answer by looking at Case 1.1 and setting $x = 0.6$. This yields $p = 1/1.2 = 5/6$. This is the fraction of component 1 that A retains, so in his eyes, he gets the fraction

$$\frac{5}{6} \times 0.6 = 0.5$$

of the cake.

24.3. (a) As x increases,

$$\frac{0.2}{x+0.6}$$

decreases, and therefore

$$1 - \frac{0.2}{x+0.6}$$

increases.
(b) As x increases, $1.4 - x$ decreases, and therefore

$$\frac{0.8}{1.4-x}$$

increases.

(c) As x increases, so does $x+0.6$, and therefore

$$\frac{0.2}{x+0.6}$$

decreases.

24.4. (a) In the initial round, A gets the first component and B the second. Each gets, in their eyes, 70% of the cake, so there is no adjustment round.

(b) In the initial round, A gets the first component, and B the second. Since it appears as though B, in his view, got only 30% of the cake, A must give him some of her share of the first component. The percentage p of the first component that she retains is determined by

$$0.7p = 0.3 + (1-p)0.7,$$

approximating 0.69999 by 0.7 and 0.30001 by 0.3. So $p = 1/1.4 = 5/7$. Therefore (i) A, in her view, gets half the cake, and (ii) B, in his *honest* view, gets the fraction

$$0.7 + \frac{2}{7} \times 0.3 \approx 0.79$$

of the cake. His dishonesty raised his share from 70% to 79% and lowered A's share from 70% to 50%.

(c) Now both get half the cake, which is much worse than in part (a) for both of them, and much worse than in part (b) for B.

(d) Now in the initial round, B gets the first component, and A the second. Furthermore, based on the announced valuations, it appears as though B were more satisfied than A. He must therefore abandon some of his share of the first component to B, retaining only a fraction p, with

$$0.7p = 0.3 + (1-p)0.7,$$

where I have approximated 0.70001 by 0.7 and 0.29999 by 0.3. So $p = 5/7$. In her view, A receives the fraction

$$\frac{2}{7} \times 0.7 + 0.3 = 0.5$$

of the cake, whereas B, in his *honest* view, only gets the fraction

$$\frac{5}{7} \times 0.3 \approx 0.21$$

of the cake. His dishonesty has brought him down to 21% (far less than a fair share!), whereas he would have gotten 70% if he had been honest.

24.5. The solution to part (d) of Exercise 24.4 demonstrates that a dishonest person can get much less than a fair share. On the other hand, to an honest person, the dishonesty of the other is not transparent. (The honest person does not know the dishonest person's true valuations.) Since she is guaranteed a fair share if both people are honest, she must be guaranteed a fair share even if the other person is dishonest.

25.1. (a) A gets 60% of the first component and 40% of the second component. B gets the rest. Each thinks that he or she gets the fraction

$$0.6 \times 0.6 + 0.4 \times 0.4 = 0.52$$

of the cake, that is, 52%. (b) The adjusted winner method would assign the first component to A and the second to B. Each would then think that he or she gets 60% of the cake.

25.2. In the initial division, John gets the strawberry and rhubarb components, and Jane gets the apple component. Thereby John thinks that he gets 90% of the cake, while Jane thinks she gets 50%. To make the division equitable, John must give up some fraction of the rhubarb component. The fraction that he retains is denoted by p. The value of p is determined from the equation

$$0.5 + 0.4p = 0.5 + 0.3(1 - p),$$

so $p = 3/7$. So John gets the entire strawberry component and 3/7 of the rhubarb component, which he considers worth

$$\frac{1}{2} + \frac{3}{7} \times \frac{2}{5} = \frac{47}{70} \approx 0.67$$

of the cake. Jane gets the entire apple component and 4/7 of the rhubarb component, which she considers

$$\frac{1}{2} + \frac{4}{7} \times \frac{3}{10} = \frac{47}{70}$$

of the cake.

25.3. A receives 3/4 of the first component and 1/3 of the second component, and therefore thinks that he gets the fraction

$$\frac{3}{4} \times 0.6 + \frac{1}{3} \times 0.4 = \frac{7}{12} \approx 0.583$$

of the cake. Similarly, B thinks she gets the fraction

$$\frac{1}{4} \times 0.2 + \frac{2}{3} \times 0.8 = \frac{7}{12} \approx 0.583$$

of the cake.

25.4. (a) The fraction of the first component that B receives is

$$\frac{x}{x + 0.6}.$$

In her honest view, the first component is worth the fraction 0.2 of the whole cake. Therefore her share of the first component is worth the fraction

$$\frac{0.2x}{x + 0.6}$$

of the whole cake to her. This is the first of the two summands in equation (25.8). Similarly, the fraction of the second component that she receives is

$$\frac{1-x}{1-x+0.4} = \frac{1-x}{1.4-x}.$$

In her honest view, the second component is worth the fraction 0.8 of the cake, and therefore her share of the second component is, to her, worth the fraction

$$\frac{0.8(1-x)}{1.4-x}$$

of the cake. This is the second summand in equation (25.8). (b) Similar to part (a).

26.1. A receives the fraction

$$\frac{0.4}{0.4+0.3+0.3} = \frac{2}{5}$$

of the chocolate component, the fraction

$$\frac{0.5}{0.5+0.4+0.3} = \frac{5}{12}$$

of the strawberry component, and the fraction

$$\frac{0.1}{0.1+0.3+0.4} = \frac{1}{8}$$

of the rhubarb component. Therefore A's share is

$$\frac{2}{5} \times 0.4 + \frac{5}{12} \times 0.5 + \frac{1}{8} \times 0.1 = 0.380833\ldots.$$

Similarly, B receives the fractions $3/10$, $1/3$, and $3/8$ of the chocolate, strawberry, and rhubarb components, respectively. His share is

$$\frac{3}{10} \times 0.3 + \frac{1}{3} \times 0.4 + \frac{3}{8} \times 0.3 = 0.335833\ldots.$$

C receives the fractions $3/10$, $1/4$, and $1/2$. Her share is

$$\frac{3}{10} \times 0.3 + \frac{1}{4} \times 0.3 + \frac{1}{2} \times 0.4 = 0.365\,.$$

The minimal share is B's, which is just barely more than a fair share.

26.2. (a) A gets 1/5 of the first component and 7/15 of the second component. B gets 4/15 of the first component and 2/5 of the second component. C gets $8/15$ of the first component and $2/15$ of the second. (b) In her eyes, A gets the fraction $1/5 \times 0.3 + 7/15 \times 0.7 = 0.3866\ldots$ of the cake. In his eyes, B gets the fraction $4/15 \times 0.4 + 2/5 \times 0.6 = 0.3466\ldots$. In her eyes, C gets the fraction $8/15 \times 0.8 + 2/15 \times 0.2 = 0.4533\ldots$. Each gets, in his or her eyes, more than one-third of the cake, so the division is fair. (c) No. (d) No, by Lemma 26.1.

For example, C judges the first component more favorably, in comparison with the second, than B. Therefore for the division to be Pareto-optimal, either B should get none of the first component or C should get none of the second. But proportional allocation assigns some of each component to each person in this example.

26.3. (a) Each person thinks that he or she gets 80% of the cake, while the two others get 10% only. (b) Each person thinks that he or she gets 80% of the cake. (c) It is sufficient to prove that the *average* share of the three people cannot exceed 80% of the cake. This will then imply that the proposed division, in which each person's share is 80% of the cake, is Pareto-optimal; any objective improvement would raise the average share above 80%.

Let us denote the fraction of the first, second, and third components that A receives by x_1, x_2, and x_3. A's share of the cake is then the fraction

$$0.8x_1 + 0.1x_2 + 0.1x_3. \tag{D.3}$$

Similarly, B's share is the fraction

$$0.1y_1 + 0.8y_2 + 0.1y_3, \tag{D.4}$$

and C's share is the fraction

$$0.1z_1 + 0.8z_2 + 0.1z_3. \tag{D.5}$$

The *sum* of (D.3) through (D.5) is

$$\begin{aligned}
&(0.8x_1 + 0.1x_2 + 0.1x_3) + (0.1y_1 + 0.8y_2 + 0.1y_3) + (0.1z_1 + 0.1z_2 + 0.8z_3)\\
&= (0.8x_1 + 0.1y_1 + 0.1z_1) + (0.1x_2 + 0.8y_2 + 0.1z_2) + (0.1x_3 + 0.1y_3 + 0.8z_3)\\
&\le (0.8x_1 + 0.8y_1 + 0.8z_1) + (0.8x_2 + 0.8y_2 + 0.8z_2) + (0.8x_3 + 0.8y_3 + 0.8z_3)\\
&= 0.8(x_1 + y_1 + z_1) + 0.8(x_2 + y_2 + z_2) + 0.8(x_3 + y_3 + z_3) = 0.8 + 0.8 + 0.8 = 2.4.
\end{aligned}$$

So the *average* share, obtained by averaging (D.3) through (D.5), is not larger than the fraction $2.4/3 = 0.8$ of the cake—80% of the cake. (d) The proof of Theorem 26.8 shows that *there are examples* in which there is no envy-free, Pareto-optimal, equitable division. This does not mean that there is *never* an envy-free, Pareto-optimal, equitable division.

26.4. Under proportional allocation, each person gets *some* of each of the two components. By Lemma 26.1, the division is therefore *not* Pareto-optimal if

$$\frac{a_1}{a_2} > \frac{b_1}{b_2}.$$

But we can reverse the roles of A and B and then find that the division is also not Pareto-optimal if

$$\frac{a_1}{a_2} < \frac{b_1}{b_2}.$$

Therefore Pareto-optimality requires

$$\frac{a_1}{a_2} = \frac{b_1}{b_2},$$

and similarly

$$\frac{a_1}{a_2} = \frac{c_1}{c_2}.$$

Thus, if the division is Pareto-optimal, the three ratios a_1/a_2, b_1/b_2, and c_1/c_2 have the same value. Let us denote this common value by r:

$$a_1 = ra_2, \quad b_1 = rb_2, \quad c_1 = rc_2. \tag{D.6}$$

Summing these three equations, we find

$$a_1 + b_1 + c_1 = r(a_2 + b_2 + c_2). \tag{D.7}$$

But $a_1 + b_1 + c_1 = a_2 + b_2 + c_2 = 1$. Therefore (D.7) implies $r = 1$, and therefore, by (D.6),

$$a_1 = a_2, \quad b_1 = b_2, \quad c_1 = c_2.$$

In words, A, B, and C have the same valuations.

26.5. Suppose that a cake consisting of two components is to be shared among two people, and the table of valuations is

	first component	second component
A	1	0.5
B	0	0.5

.

The method of proportional allocation yields the result that A should get the entire first component, and A and B should equally share the second component. There is no objective improvement of this division.

A.1. (a) $\{3\}$. (b) $\{6\}$. (c) $\{3,4,5,6\}$. (d) $\{4,5\}$.

A.2. (a) True. (b) False (8 is an element, not a subset of $\mathbb{N}$). (c) False. (d) False (E is not an element of $\mathbb{N}$, but a subset). (e) True. (f) True. (g) True.

B.1. The first person says, "Being a Democrat implies supporting universal health care" ($A \Rightarrow B$). The second person answers, "No, supporting universal health care does not imply being a Democrat" ($B \not\Rightarrow A$). But $B \not\Rightarrow A$ does not contradict $A \Rightarrow B$.

B.2. "A if B" means $B \Rightarrow A$. "A only if B" means $A \Rightarrow B$.

B.3. Suppose that $A \Rightarrow B$. Let us prove that $\neg B \Rightarrow \neg A$. So suppose that $\neg B$. Then A cannot be true, for if it were, B would be true, since we are assuming $A \Rightarrow B$. So $A \Rightarrow B$ implies $\neg B \Rightarrow \neg A$. That conversely $\neg B \Rightarrow \neg A$ implies $A \Rightarrow B$ is proved in the same way.

B.4. To say that a method *satisfies* a criterion is a statement about *all* preference schedules or cake division problems. It is a "for all" statement. To say that a method *violates* a criterion is to say that there is at least one example exhibiting the violation. It is a "there is" statement.

C.1. We break the sum $1+2+3+\cdots+N$ into many sums of pairs of integers. Each of the sums of pairs of integers equals $N+1$. If N is even, then this looks like this:

$$1+2+3+\cdots+N = (1+N)+(2+(N-1))+(3+(N-2))+\cdots$$

$$+((N/2-1)+(N/2+2))+(N/2+(N/2+1)).$$

There are $N/2$ pairs. So the total is $N/2\times(N+1) = N(N+1)/2$, as we wanted to show. If N is odd, it looks like this:

$$1+2+3+\cdots+N = (1+N)+(2+(N-1))+(3+(N-2))+\cdots$$

$$+((N-1)/2+(N+3)/2)+(N+1)/2.$$

Note that here there are $(N-1)/2$ pairs, and then the number $(N+1)/2$, which does not belong to a pair. The total is therefore $(N-1)/2\times(N+1)+(N+1)/2 = N(N+1)/2$, as desired.

C.3. $A(N)$ is the statement

$$1+3+3^2+\cdots+3^{N-1} = \frac{3^N-1}{2}.$$

For $N=1$, the left-hand side has only one term, namely, 1. (By convention, the "sum" is then 1.) The right-hand side is $(3^1-1)/2 = 1$ for $N=1$. Now assume that $A(N)$ holds for some N. We have to prove that then $A(N+1)$ holds, that is, that the sum of the first N powers of 3 equals $(3^{N+1}-1)/2$. The sum of the first N powers of 3 equals

$$1+3+3^2+\cdots+3^{N-1}+3^N,$$

and since we assume $A(N)$, we may write this as

$$\frac{3^N-1}{2}+3^N = \frac{3^N-1+2\times 3^N}{2} = \frac{3\times 3^N-1}{2} = \frac{3^{N+1}-1}{2},$$

as desired.

C.4. For $N=1$, the left-hand side is a sum with a single summand, namely, 1, and by convention, that is 1. The right-hand side equals 1 for $N=1$. Now assume that for a given N,

$$1+r+r^2+\cdots+r^{N-1} = \frac{r^N-1}{r-1}. \tag{D.8}$$

We want to prove that

$$1+r+r^2+\cdots+r^N = \frac{r^{N+1}-1}{r-1}. \tag{D.9}$$

The left-hand side of (D.9) can be written as

$$1+r+r^2+\cdots+r^{N-1}+r^N = \frac{r^N-1}{r-1}+r^N, \tag{D.10}$$

by (D.8). The right-hand side of (D.10) is

$$\frac{r^N - 1}{r - 1} + \frac{r^N(r-1)}{r-1} = \frac{r^{N+1} - 1}{r - 1}\,.$$

This completes the proof.

C.5. The sum is

$$\frac{(1/5)^{101} - 1}{1/5 - 1} \approx \frac{-1}{1/5 - 1} = \frac{5}{4}\,.$$

Index