NIJMEGEN INSTITUTE FOR
COGNITION RESEARCH AND
INFORMATION TECHNOLOGY

COALITION FORMATION AND SOCIAL CHOICE

nici

AD VAN DEEMEN

Coalition Formation and Social Choice

ISBN 90 373 0091 X/CIP/NUGI 659

Coalition Formation and Social Choice

een wetenschappelijke proeve
op het gebied van de sociale wetenschappen,
in het bijzonder de politicologie

Proefschrift

ter verkrijging van de graad van doctor
aan de Katholieke Universiteit van Nijmegen,
volgens besluit van het college van decanen
in het openbaar te verdedigen op
maandag 22 april 1991, des namiddags te 3.30 uur

door

Adrianus Marinus Arnoldus van Deemen
geboren op 13 september 1952 te Nijmegen

NICI Technical Report 91-01
Nijmegen Institute for Cognition Research
and Information Technology

Promotores:
Prof. dr. Th.G.G. Bezembinder
Prof. dr. H.C.M. de Swart, Katholieke Universiteit Brabant

Contents

Acknowledgements

This book is the result of a research project at the Department of Political Science of the Catholic University of Nijmegen. This project, coded as SW/3/86, started in January 1986. I am indebted to the people who made this project possible.

Many other people have made this book possible. In particular, I must thank the participants of the Social Choice Theory colloquia. My debt to my two promotores is great. They taught me too much to mention. I am grateful for their numerous comments, corrections and improvements. Above all, I am grateful for their stimulating support at the unavoidable critical moments in writing this book.

My thanks to my wife and my children can neither be expressed in plain language, nor in mathematical formulae. Without them, this book would not have been written.

List of Symbols and Abbreviations

The following list contains the most important symbols.
$:=$: definition, is by definition.
$\square$: end of proof.
X: a social choice problem.
Ω: the set of finite subsets of X with at least three elements.
A: agenda, an element of Ω.
N: the set of players, agents, individuals, voters, actors.
S: a coalition, a subset of N.
R: preference on X.
P: Strict preference on X.
I: Indifference on X.
$\beta(A, R)$: the set of R-best elements of A.
$\mu(A, R)$: the set of maximal choices of A.
C: choice function.
p: a preference profile, i.e. an element of Π.
Π: the set of preference profiles.
R_i^p: the preference of agent i in profile p.
$B(X)$: the set of complete and reflexive relations on X.
$A(X)$: the set of reflexive, complete and acyclic relations on X.
$Q(X)$: the set of reflexive, complete and quasi-transitive relations on X.
$O(X)$: the set of reflexive, complete and transitive relations on X.
$L(X)$: the set of antisymmetric, complete and transitive relations on X.
F: a social choice rule.
Φ: the collection of social choice rules.
Φ_{SDF}: the set of social decision functions.

Φ_{QS}: the set of quasi-transitive social choice rules.
Φ_{SWF}: the set of social welfare function.
$Con(A, p)$: the Condorcet set for A under p.
$b_i(x, p)$: the Borda score of x in R_i^p.
$\sigma(A, R)$: the generalized optimal choice set of A.
$G = (N, W)$: simple game.
W: the nonempty set of winning coalitions.
W^{size}: the set of minimum size coalitions.
W^{min}: the set of minimal winning coalitions.
Θ: weak order of policy positions.
W^{MC}: the set of minimal and closed coalitions.
ρ: preference on coalitions.
π: strict preference on coalitions.
ι: indifference on coalitions.
$h(R, R')$: the Hamming distance between R and R'.

The following list contains the logical symbols:
$\Leftrightarrow$: equivalence, if and only if.
$\Rightarrow$: implication, if ..., then ....
$\wedge$: conjunction, and.
$\vee$: inclusive disjunction, or.
$\neg$: negation.
$\forall$: universal quantifier, for all.
$\exists$: existential quantifier, there exists.

The following list explains the most important abbriviations:
BCS: Black-Condorcet System,
BR: Borda Rule,
GESTS: Generalized Stable Set,
GOCS: Generalized Optimal Choice Set,
IIA: Independence of Irrelevant Alternatives,
PC: Pareto Condition,
PR: Plurality Rule,
ND: Nondictatorship,
QS: Quasi-transitive Social choice rule,

SCR: Social Choice Rule,
SDF: Social Decision Function,
SMD: System of Majority Decision,
SWF: Social Welfare Function,
UD: Unrestricted Domain.

Chapter 1

Introduction

1.1 The Relevance of Political Coalition Formation

Essential in politics is winning. What counts is the enforcement of a decision, the passing of a bill, winning an election, carrying a policy through, the formation of a majority coalition, etc. However, in politics it is impossible to win by staying alone. Coalitions must be formed in order to enforce the victory. Even a dictator cannot rule by staying alone. A dictator too must form alliances with some societal elites in order to enforce his or her preferences. Thus, coalition formation is at the heart of political life.

This view on politics is not new. Especially the American political scientist W. Riker opts for this view in several places of his work. In his *The Theory of Political Coalitions* and in his *Introduction to Positive Political Theory*, which Riker wrote in cooperation with P. Ordeshook, he introduces without hesitation the behavioral assumption that the purpose of each political actor is to form winning coalitions (see for example Riker and Ordeshook 1973: 179). In his more recent work on the place of political science in public choice, Riker argues that "[m]ost things that people want they cannot get by themselves" (Riker 1988: 249). He then constructs a continuum of possible forms of cooperation among people. At one extreme of this continuum

"... there is the ideal type of a team effort against nature,

> guided by a Platonic rule of justice. There is a joint product which all participants desire with equal intensity; they agree about the appropriate means to obtain it; all members know and accept a role in the division of labor, and they always choose the action prescribed by the role, so there is never a private interest apart from the shared interest in the success of the group endeavor". (Riker ibid: 249).

This ideal type of cooperation is called harmonious by Riker. At the other side of the continuum there is

> " ... the ideal type of cooperation in which goals are only partially shared. .... In this kind of cooperation, then, a winning coalition excercises the authority of the entire group to support outcomes that, while perhaps benefiting the whole body, still benefit especially the members of the winning coalition. The winning coalition is able to do this because the losers are, in some way, constrained to continue to participate and thus to recognize the obligation, even the validity, of the winners' advantage". (Riker ibid: 249).

This type of cooperation is called exploitative and, according to Riker, political science is "the study of cooperation in deciding on social actions, policies, and norms, where the cooperation is closer to the exploitative ideal than to the harmonious" (Riker ibid: 251).

In Riker's view, the winners take all and the losers are obliged to accept the consequences of the actions of the winners. This last aspect is interesting. The outcomes of the execution of power of a winning coalition concerns the whole population, hence also the losers. The losers are obliged to take these outcomes for granted, whether these outcomes are beneficial for them or not. If the outcomes are beneficial, then the losers have "luck". In the other case they have "bad luck". In either case, the losers are non-excludable. In this respect, the outcomes of the execution of power of a winning coalition have the character of a public good (cf. Barry 1980, Holler and Packel 1983).

1.2 The Game-Theoretical Study of Political Coalition Formation

Coalition formation being so relevant in political life, it is no surprise that political scientists have spent considerable research efforts and resources to investigate this phenomenon. In particular, there is a well-established tradition in political science of studying coalition formation from a game-theoretical point of view. For convenience, we will call it the game-theoretical tradition or approach. It is not the aim of this section to give an overview of this tradition. We only sketch the 'historical moments' that in our view are crucial for the development of this tradition.

The starting shot is given in 1962 in the already mentioned seminal work *The Theory of Political Coalitions* of William Riker. The basic statement of Riker's theory is the well-known *size principle* (Riker (1962: 33, 47):

> "In social situations similar to n-person, zero-sum games with side-payments, participants create coalitions just as large as they believe will ensure winning and no larger".

Why formulating this principle? After all, a great part of game theory formulated in Von Neumann and Morgenstern (1953) and made accessible for social scientists in the lucid work of Luce and Raiffa (1957), deals with coalitions. This part is called n-person cooperative game theory. Hence what is the rationale for the size principle and, more generally, for coalition formation theories? According to Riker (1962, esp. chapter 2), n-person cooperative game theory is oriented towards payoff structures. It tries to investigate what a player[1] gains when joining a specific coalition. More technically, it tries to establish what payoff vectors may emerge. As such, it says nothing about the coalitions that may be formed. Game theory is payoff-oriented and is not designed to predict coalitions[2]. In order to predict coalitions, other theories must be constructed in which additional concepts and assumptions with respect to the standard apparatus

[1] Alternatively, we speak of 'player', 'actor', 'agent' and 'individual'. In all cases, we mean the same thing, namely a separable decision-making unit.

[2] Also cf. Ordeshook (1986: 408).

of cooperative game theory must be used.

After the starting shot fired by Riker, the next historical event to memorize in the game-theoretical tradition of coalition formation theories is the appearance of De Swaan's *Coalition Theories and Cabinet Formation* (1973). In this work, a number of existing theories of coalition formation with a root in game theory are critically investigated. De Swaan also presents a new theory of coalition formation, namely *policy distance theory*. In this theory, the notions of policy position and of policy distance are introduced. As the name of the theory suggests, especially the notion of policy distance is of crucial importance. The basic idea, here, is that each actor strives to form a winning coalition with a policy position that is as close as possible to his own policy position.

The important new aspect of policy distance theory is that it allows the construction of a preference profile[3] - one preference relation for each actor - concerning the set of winning coalitions. With the aid of this preference profile a dominance relation is constructed over the set of winning coalitions. If dominance is defined, then also undominance can be defined. The set of undominated coalitions, given the constructed dominance relation, is called the core. The prediction is that only coalitions from the core will be formed.

In his theory, De Swaan uses standard game theoretical concepts besides plausible behavioral and political assumptions. Moreover, De Swaan's work indicates a research possibility, namely the investigation of actors' preferences for coalitions, that hitherto has not been explored. It is precisely this research line we wish to explore in this work.

De Swaan's work is also important from another point of view: it confronts the investigated theories with data about cabinet formation in nine West-European multi-party systems, thus combining the purely theoretical approach with an empirical approach. Together with *Government Coalitions in Western Europe* of Taylor and Laver (1973), De Swaan's work is the starting shot of an empirical research tradition that confronts the game-theoretical coalition theories with data about cabinet formation in multi-party systems.

The works of De Swaan and Taylor and Laver show an impressive

[3] De Swaan speaks of a matrix of preferences.

empirical support of Axelrod's conflict of interest theory (Axelrod 1970). In a nutshell, Axelrod's theory says that only winning coalitions will be formed that

1. are minimal in "the sense that they contain no more members than is necessary" to win (De Swaan 1973: 75) and

2. are closed in the sense that they contain only members that are adjacent on a one-dimensional policy scale (Axelrod 1970: 169, De Swaan 1973).

Winning coalitions satisfying these two properties are called *minimal closed*. In Axelrod's view, these coalitions have minimal conflict of interest and will therefore be formed.

De Swaan was so impressed by the empirical success of Axelrod's theory that he decided to employ it as "a basis for the classification of parties and party systems" (De Swaan 1982: 237). Depending on the number of feasible minimal closed coalitions and given the outcome of elections and the parties' policy positions, he calls a party system *monolemmatic*, *dilemmatic* or *polylemmatic*, a classification, however, that did not find much acceptance in political science.

Nowadays the relevance of the empirical results of De Swaan, Taylor and Laver and others working within the empirical tradition is highly questioned (Browne and Franklin 1986, Laver 1986, Nolte 1988, Tops 1989). For an account of this topic, consider Tops (1989).

Another important moment in the history of the game-theoretical tradition is the work of the mathematician and game theorist Peleg. He presents a coalition formation theory in which a specific actor plays an essential role (Peleg 1980, 1981). This actor is called dominant. More specifically, a dominant actor is an actor with two characterising properties. First, it is an actor who has the ability to control completely the internal opposition of a winning coalition (cf. Van Deemen 1989). Second, it is an actor with more opportunities to form winning coalitions than any nondominant actor. Thus, a dominant actor has more threat potentials than any nondominant actor participating in the coalition formation process (Van Deemen 1989). The essence of Peleg's theory is that this dominant actor will have enough power to determine the process of coalition formation in the way he likes (Peleg 1981, Van Deemen 1989).

Peleg's work introduces the *actor-oriented approach* in the game-theoretical study of political coalition formation. Having remained unnoticed for a remarkable time, Peleg's theory was put under the attention of political scientists in Van Deemen (1987, 1989). These works also present new results and a variation (the center-player approach) to Peleg's actor-oriented approach. We will extensively discuss our contributions to Peleg's coalition theory in chapter 5.

1.3 Theoretical Problems in the Game-Theoretical Study of Political Coalition Formation

1.3.1 The Absence of an Actor-Oriented Policy Theory

Since the celebrated work of De Swaan, it is customary within the game-theoretical tradition to classify coalition formation theories into two classes (Grofman 1984, Van Deemen 1989). The theories in the first class predict coalitions by using information about the power positions of the relevant political actors. Since no notion of policy is used in these theories, these coalition formation theories are called policy-blind. A well-known representative of this class is the already mentioned and very frequently discussed minimum size theory (Riker 1962, Riker and Ordeshook 1973). Also Peleg's dominant actor theory is policy-blind.

The second class of coalition formation theories contains the so-called policy theories. These theories all deal with coalition formation processes in simple games in which the policy position of a player in a one-dimensional or multi-dimensional scale is relevant. Thus, besides the power positions also the policy position of a political actor in a uni- or multi-dimensional space is used as an important variable for predicting coalitions. Examples of one-dimensional theories are the already mentioned policy distance theory of De Swaan (1973) and Axelrod's conflict of interest theory. An example of a multi-dimensional theory is the competitive coalition theory of McKelvey, Ordeshook and Winer (1978).

This classification is widely accepted. For example, it can be found in the bundle *Coalitions and Collective Action* edited by M. Holler (1984).

It also allows to detect the 'blind spots' in the theoretical developments within the game-theoretical tradition. So this classification makes clear that there is no actor-oriented coalition formation theory that belongs to the class of policy theories. That is, there is no theory in the style of Peleg's in which besides the power also the policy position of a player is crucial. This observation leads to the formulation of our first problem:

Problem 1 *Is it possible to formulate an actor-oriented coalition formation theory within the game-theoretical tradition in which policy positions (one for each player) play a decisive role.*

Solving this problem would provide a competitive theoretical alternative for Peleg's theory of dominant players that, in contrast to Peleg's theory, belongs to the class of policy theories. The problem may also have empirical relevance. Applying the coalition formation theories to cabinet formation in multi-party systems, De Swaan (1973), Taylor and Laver (1973) and Browne and Dreijmanis (1982) show that policy positions of political parties in multi-party systems are important in the formation of government coalitions.

1.3.2 Players' Preferences for Coalitions

Coalition formation theories formulated within the game-theoretical tradition typically ignore the players' preferences with respect to coalitions. There is only one exception: policy distance theory. As already has been noticed, this theory leads to the construction of a set of preferences, one preference for each player, with respect to coalitions. The thus resulting preference profile is used to derive a prediction set.

Since game theory starts from the assumption of rational agents, i.e. agents who choose alternatives which are best according to their preferences, this ignorance is, indeed, remarkable and difficult to explain. Perhaps it has to do with the fact that the theories may become complex when besides players' preferences for payoff structures, also players' preferences for coalitions are introduced. The theories then must deal with two types of individual preferences that may, in addition, interact with each other.

Another possible cause of this ignorance may be the game-theoretical origin of the coalition theories itself. In game theory, it is assumed that players have preferences with respect to payoffs and not with respect to coalitions that may be formed in order to get as much payoff as possible. Shifting the attention away from payoff towards coalition formation does not, by itself, lead to efforts of constructing theories in which the formation of coalitional preferences is essential. The relevance of payoff preferences does not imply the relevance of coalition preferences.

In our view, to see the relevance of players' coalition preferences requires another view on coalition formation processes. It is then necessary to see coalition formation itself as a choice process. Preferences become relevant in situations in which a choice must be made from a set of alternatives (a choice problem). In particular, coalition preferences become relevant in situations in which a selection must be made from a set of possible coalitions.

This view can be made more complicated by seeing a player in a game as an agent with two preferences. One preference is concerned with the possible outcomes of the game and the other preference with the possible coalitions in the game. Since the formation of a coalition is a method to realize an outcome, it is reasonable to assume that the player's preference for coalitions is, in some way or another, determined by the player's preference concerning the outcomes.

The main theme of this work is to explore the idea that coalition formation is a choice process that is guided by the players' preferences with respect to coalitions. More specifically, the above discussion leads to the following problem:

Problem 2 *1. Consider a nonempty finite set of agents and a nonempty finite set of possible winning coalitions. How will these agents form preferences with respect to this set of possible winning coalitions and how will the agents use these preferences concerning coalitions in order to form a winning coalition?*

2. Consider a nonempty finite set of agents who have to make collectively a choice from a nonempty set of mutually exclusive alternatives (outcomes, policies, payoffs etc.). Also consider a nonempty finite set of possible winning coalitions. Suppose each agent has a preference on the set of alternatives. Also suppose that agents only can choose an alternative by forming a winning coalition. How will an agent *form a preference concerning the set of winning coalitions? How can the players' preferences concerning coalitions be explained in terms of their preferences concerning the set of alternatives from which a choice collectively has to be made?*

In order to solve the first part of this problem, theories must be constructed that describe and explain the formation of players' preferences for coalitions and that use the players' preferences for coalitions in order to predict a set of coalitions. The solution of the second part of the problem requires the formulation of theories that explain the formation of players' preferences concerning coalitions in terms of preferences concerning alternatives (outcomes, policies, payoff). Of course, also these latter theories must explain how the players' preferences concerning coalitions may be used in order to predict a set of coalitions.

The formulated problem is theoretically relevant. Solving it would give us theories that describe and explain the formation and use of preferences concerning coalitions in an explicit way. Knowing the preferences concerning the coalitions would enable us to predict the coalitions that may be formed. Theories for the second problem even can explain the formation of coalitions in terms of other preferences. Knowing the preferences of each player with respect to a set of outcomes would be sufficient then to predict the coalitions that may be formed. Since the complexity is raised by introducing coalition preferences and aggregation of these preferences, the theories may be better approximations of political reality than the already existing ones. But, of course, on this last point, only empirical research can decide.

1.4 Simple Games and Social Choice

In order to solve the two formulated problems, we will use both simple game theory and social choice theory.

1.4.1 Simple Game Theory

Our exploration of coalition formation as a choice process will be within the game-theoretical tradition. Especially we will use n person simple game theory where $n \geq 3$. This theory was first defined and studied in Von Neumann and Morgenstern (1953 Ch. X). It has been further refined in Shapley (1962, 1967, 1981).

Simple games are cooperative games in which two types of coalitions play a role, namely winning ones and losing ones. Essential of simple games is that the winners - the members of a winning coalition - take all. The losers - members of a losing coalition - have no power to control the game. The decisions of the winners concern the whole set of players and the losers are obliged to take these decisions for granted, whether the effects of the winners' decisions are favourable for them or not. An example is the majority voting game. In this game only a majority coalition of voters can win, i.e. determine a winning alternative.

It must be stressed that simple game theory deals with winning or losing in general, that is, without referring explicitly to the rules that determine winning or losing. The advantage of this abstract approach is that classes of simple games can be studied without referring to particular rules as majority rule, unanimity rule, bargaining rules, etc.

Since winning and losing is essential for politics, simple games are extremely useful to model political situations and processes, especially coalition formation[4]. With the choice of simple game theory as the game-theoretical framework for our theories, we keep in line with tradition[5].

1.4.2 Social Choice Theory

Social choice theory is about choice processes in which two or more agents are involved. The theory has its roots in the Enlightenment of the eighteenth century. Especially Borda and Condorcet have contributed a great deal to the development of this theory in that time[6]. The development of the theory is speeded up in this century by the works of Arrow (1963), Black (1957), Fishburn (1973) and Sen (1970). Advanced treatments of the theory are given in Kelly (1978, 1988), Sen (1977, 1986), Storcken (1989) and Suzumura (1983).

[4]Cf. Shapley (1962, 1967, 1981). In Shapley (1981) the terms 'simple game' and 'political structure' even are used interchangeably.

[5]Cf. Von Neumann and Morgenstern (1953 Ch. X); De Swaan (1973, esp. Ch.4); Van Deemen (1987, 1989, 1990, 1990a), Peleg (1981), Van Roozendaal (1990).

[6]Cf. Black (1957, Part II); also see chapter 3.

In essence, social choice theory deals with the aggregation of individual preferences into a social preference that, in its turn, can be used to determine a social choice[7]. The term 'social' refers here to the fact that at least three agents are involved in the choice process and, hence, at least three indiviudal preferences are to be aggregated.

Since coalition formation is considered as a choice process involving more than two persons, social choice theory will be important in this work. Coalition formation will be seen as a process in which the aggregation of players' coalitional preferences into a binary relation (social preference) on a set of winning coalitions is essential. The resulting aggregated relation determines a prediction set (a set of social choices).

However, with respect to determining social choices with the aid of a social preference, there is a problem that first must be solved before using social choice theory to study coalition formation processes. Social choice theory mainly deals with the determination of best social choices. An alternative is socially best if it is socially preferred to every other alternative. Of course, if a socially best alternative exists, then it must be a social choice. The problem, however, is what the social choice must be when a socially best alternative does not exist. This problem occurs when the social preference yielded by an aggregation process is cyclic. When the social preference is reflexive, complete and acyclic, the problem will not occur[8] (Sen 1970, Suzumura 1983). However, it is not guaranteed apriori that aggregation of coalitional preferences always leads to a reflexive, complete and acyclic relation. Thus, socially best coalitions need not exist. Therefore, it is necessary to develop theories that allow the selection of alternatives even when best alternatives do not exist. These theories may then be used to predict sets of coalitions even when socially best coalitions do not exist. This leads to the formulation of our third

---

[7] We are referring here to the relational approach in social choice theory. See chapter 2 of this work. There is also a socalled functional approach in which social choices are determined without referring to social preferences. Again cf. chapter 2, especially section 2.9.

[8] The notions of cycle, acyclicity, completeness, reflexivity and others will be precisely defined and studied in chapter 2. To give a handhold, a relation R satisfies reflexivity if for all $x : xRx$, completeness if for all $x, y : xRy$ or yRx. R is cyclic if it contains a cycle of the form $x_1Px_2P \dots Px_nPx_1$ where P is strict social preference (if xRy, then not yRx). R satisfies acyclicity if it it is not cyclic.

problem:

Problem 3 *Let A be a nonempty and finite set of alternatives (winning coalitions). Let R be a relation on A that is obtained by aggregating individual preferences concerning A. What alternatives must be selected from A when R is cyclic?*

Analogously to game theory, theories that can deal with cyclic aggregated relations are called solution theories. The general problem of finding such theories belongs to the domain of social choice theory. This problem is as old as social choice theory itself. Indeed, it is not an easy problem to solve. However, it is a highly relevant problem since solution theories not only allow us to predict coalitions when best coalitions are not guaranteed. They also may tell us what to do when we meet e.g. cyclic majorities. As is well known, this latter problem was a great obsession for Condorcet[9] and is called the Condorcet paradox. Thus, theories that can handle cyclic social preferences may be able to dissolve the Condorcet paradox[10].

1.5 Summary of Purposes

The purpose of this work is to solve the above mentioned problems. Thus, the first purpose is to fill up the 'blind spot' in the classification of coalition formation theories (see section 1.3.1). We will try to construct a theory that deals with coalition formation in games in which a typical player rules and in which policy plays a major role. Since we will define a player that is in the center of a policy order we will call this coalition theory the center player theory. It will be the counterpart of Peleg's policy-blind dominant player theory.

However, we have a more ambitious purpose. Our main purpose is to work out the idea that coalition formation is a social choice process in which the choice behavior of each player is guided by his preference for coalitions. According to this idea, the players' coalitional preferences must be aggregated in such a way that an aggregation relation - a social preference - on a set of coalitions results that can be used as a basis for deriving a prediction set - a social choice set - of coalitions.

[9]Cf. Black (1957, Part II) and chapter 3 of this work.

[10]Cf. chapter 3, section 3.1.3).

In order to realize this, it is necessary to realize another purpose first, namely the construction of a solution theory that is useful in generating prediction sets of coalitions from a possibly cyclic binary relation, obtained by aggregating players' preferences for coalitions. The fulfillment of this purpose is mainly important for the realization of our more ambitious purpose. However, it also has value as such. It may help to solve problems that already exist a long time, especially the problem of the existence of cyclic majorities (the Condorcet paradox).

We refer to the distinction between theoretical and experimental physics. In theoretical physics, theories are invented, mathematized, elaborated or refined. In experimental physics (aspects of) the theories are tested. We wish to make this distinction also for political science. Since in political science research the experiment is seldom used and observations often arise from political life, we prefer to speak of quasi-experimental research. So we distinguish purely theoretical political science from quasi-experimental political research. Theoretical political science already has some history. Pioneering works are Riker and Ordeshook (1973) and Ordershook (1986). In these works it is also called mathematical political science and positive political theory.

Our work will be an exercise in theoretical or mathematical political science. We proceed in a purely theoretical way and want to leave the empirical tests of the presented and elaborated theories to political researchers belonging to the quasi-experimental research tradition. In order to illustrate the working of the theories, we will use examples that mostly are taken from reality. However, these examples are only used for illustration purposes.

1.6 Overview

We summarize the subsequent chapters.

Chapter 2. In order to study coalition formation processes as aggregation processes of coalitional preferences, we first have to study social choice theory. This will be done by making a clear distinction between the descriptive part and the solution part of the theory.

The basic concepts of the descriptive part will be presented in chapter 2, sections 2 to 7. The concepts are social state (alternative, outcome etc.), agenda, individual (player, agent, actor), coalition, preference and choice, rational choice and social choice rule.

In section 7, we make a start with the discussion of the solution part. There, the main concept is collective rationality. We will see that collective rationality implies the selection of socially best alternatives. We also discuss conditions that guarantee the existence of socially best alternatives.

An important subject in social choice theory is the investigation of a set of conditions deemed desirable for an aggregation to satisfy. This will be dealt with in section 8. Famous in this respect is Arrow's theorem (Arrow 1963). Arrow formulates a number of - in his view - reasonable conditions. Surprisingly enough, he shows that no aggregation process can satisfy simultaneously this set of reasonable conditions, that is, he shows that it is impossible that an aggregation process satisfies these conditions. We will discuss the Arrow conditions and Arrow's interesting theorem extensively. We also present a new proof of Arrow's theorem.

In the last section of this chapter, we study two alternative conceptual frameworks in social choice theory, namely, the Kelly-Suzumura framework and the Fishburn framework. The essence of both frameworks is that social choices are determined without using social preferences. They start just from the opposite position by trying to determine a social preference by observing social choice sets.

Chapter 3. The study of social choice processes as presented in chapter 2 is rather abstract. It deals with aggregation rules and processes in general, not with particular aggregation rules. We will study particular rules in chapter 3. We also will pay a lot of attention to the history of social choice theory, especially to pieces of work of Marquis de Condorcet that are, are far as we know, not cited in the standard literature of social choice.

Section 1 of this chapter is devoted to the system of majority decision. We formulate an exact definition of it and discuss its relevance.

We present May's theorem that states necessary and sufficient conditions for an aggregation rule to be the system of majority decision. The main problem with the system of majority decision is the possibility of the Condorcet paradox. This paradox exists if the aggregation of individual preferences by the system of majority decision leads to a cyclic majority relation. We will spend a lot of energy to this paradox that is named after its discoverer Marquis de Condorcet (1743-1794). First we discuss the the solution of Condorcet to the paradox. Some documents written by Condorcet give another view on Condorcet's solution than suggested by Black (1957: 175). Then we present another solution to the paradox. We propose a theory called the theory of stable majorities that is based on the notion of majority dominance. According to this theory, alternative x majority dominates alternative y if there is a path $x\bar{a}Mx_1\bar{a}Mx_2\ldots\bar{a}My$ from x to y, where $\bar{a}M$ denotes strict majority[11]. A stable majority solution is a set in which no two alternatives majority-dominate each other. Further, for each alternative y outside a stable majority solution there is an alternative x in that majority solution such that x majority-dominates y. We are able to prove, by using a theorem in chapter 4, that a stable majority solution is nonempty. The theory of stable majority solutions is an instance of a more general solution theory of generalized stable sets that will be presented and studied in chapter 4.

In section 2 the plurality rule is studied. Again we will present material of Condorcet on this subject.

In section 3 the Borda rule is presented and discussed. We formulate a definition that can be used to aggregate preferences with indifference. Also some interesting results of Condorcet concerning this rule will be presented among which its violation of the condition of independence of irrelevant alternatives. Condorcet also has discoverded the valuable insight that different rules may lead to different social choices for the same situations. He shows this by presenting several situations for which the majority choice disagrees with the Borda choice or the plurality choice, or for which the Borda choice agrees

[11]That is, $x\bar{a}My$ if the number of individuals that prefer x to y is strictly greater than the number of people that prefer y to x.

with majority choice but disagrees with plurality choice. We will present a number of these situations designed by Condorcet.

In section 4 we apply the majority idea as studied in section 1 to the Dutch electoral system. We show, by constructing a voting situation, that it is possible that a party A preferred by a majority of voters to party B still gets less seats in parliament than does B. We propose a design principle in order to avoid this curiosity.

Chapter 4. In chapter 2 we made a start with discussing the solution part of social choice theory. As we have seen in that chapter, if a social preference is cyclic, then a best social choice cannot exists. The aim of this chapter is to formulate theories that specify the existence of social choices when a best social choice does not exist. In order to be able to do this, they all must be able to handle cyclic social preferences in some way or another. This chapter is important since the theories presented here are needed to find prediction sets for coalition theories that will be developed in later chapters. That is, this chapter must solve our problem 3.

This chapter also will have a more mathematical character. We will use methods from digraph theory in order to formulate and prove results. The necessary mathematical concepts and techniques are presented in section 1.

In section 2, we present the first theory that is able to deal with cyclic social preferences. This theory says that an alternative x must be a social choice if there is no alternative y such that y is strictly socially preferred to x. An alternative x with this property is called maximal and the theory therefore is called the theory of maximal choices. This theory does not work for every cyclic preference. We will give conditions for the existence of nonempty sets of maximal choices.

Since the theory of maximal choices may fail in yielding a social choice, we study in section 2 another theory which is called the theory of generalized optimal choice sets. The origin of this theory is Schwartz (1972, 1986). It appears to be equivalent with the theory of admissible sets of Kalai and Schmeidler (1976) and Kalai et.al. (1977) and with the theory of dynamic solutions of Shenoy

(1979, 1980). The concept of generalized optimal choice set is a generalization of the concept of set of maximal choices.

Crucial in this theory is the notion of minimal undominated set. A set of alternatives is undominated if for no x in this set there is an alternative y outside this set such that y is socially strictly preferable to x. An undominated set is minimal if none of its proper subsets has this property. The generalized optimal choice set of a set A is the union of all minimal undominated subsets of A. The theory says that only choices from this set will be or must be a social choice. We give results of Schwartz that show that this theory can handle any cyclic social preference. Hence, it is able to yield a social choice set for every social preference obtained by aggregation of individual preferences. We also will present some new results for this theory among which the most important is the contraction theorem.

A shortcoming of the theory of generalized optimal choice sets is that inside generalized optimal choice sets there may be dominance in the sense that one alternative in that set may be socially strictly preferred to another alternative in that set. That is, a generalized optimal choice set may not be internally stable. To avoid this shortcoming, a version of stable set theory is formulated in section 5 of this chapter. The origin of this theory is Von Neumann and Morgenstern (1953). A set of alternatives S is stable if for no x and y in S we have that x is strictly socially preferred to y. It is externally stable if for each y outside S there is an x inside S such that x is socially strictly preferred to y. In this section we will discuss the relevance of stable set theory. We also discuss the nice sociological interpretation of these sets in terms of 'standards of behavior' as is done in Von Neumann and Morgenstern (1953). Stable sets do not exist for every cyclic social preferences. We formulate some results concerning the nonemptiness of these sets.

The properties of internal and external stability of stable sets are appealing. However, the possible emptiness of these sets is a serious shortcoming. To dissolve this shortcoming we formulate in section 6 a new theory that will be called the theory of generalized stable sets. We give an informal description. A set of alternatives V is generalized internally stable if for no x and y there is a path

$xPx_1Px_2...x_nPy$ from x to y, where P is the asymmetric part of a social preference R. V is called generalized externally stable if for each y not in V there is an x in V such that there is a path $xPx_1Px_2...x_nPy$ from x to y where P is the asymmetric part of a social preference R. A set satisfying generalized external and generalized internal stability is called a generalized stable set. We prove an existence theorem that shows that generalized stable sets are always nonempty. Hence this theory can handle any cyclic social preference and thus is able to yield a social choice set for every social preference obtained by aggregating individual preferences. Its advantage to the theory of generalized optimal choice sets is that generalized stable sets satisfy generalized internal stability (and hence internal stability).

In section 7 of this chapter we summarize the most important differences and commonalities of the presented solution theories. With this section we stop our efforts to solve problem 3. The theory of sets of maximal choices and the theory of generalized stable sets will be used in the subsequent chapters, especially in chapter 6 and 7. The theory of generalized stable sets also is used to formulate the theory of stable majority solutions discussed in chapter 3.

Chapter 5. In this chapter we solve our problem 1. In section 2 the basic concepts of simple game theory are presented. In this section we also discuss Riker's minimum size theory in more detail. In section 3 we study particular classes of simple games, namely weak games, oligarchic games and dictatorial games.

In section 4, we briefly discuss the formulation of Arrow's theorem and some of its variations in terms of simple game theory. We want to accentuate that we do not have the aim to study how to represent social choice rules by simple games and, conversely, how to associate simple games with social choice rules. For this exercise we refer to Peleg (1983, 1984).

In section 5 we present Peleg's theory of dominant players in detail. We will present an additional result and also a stronger hypothesis than formulated in Peleg (1981). In order to illustrate the working of this theory, we give a computation example. In this example we use the game representation of the parliamentary system of Germany in 1987.

Then, as the counterpart of Peleg's theory, we present the theory of center players. This will be done in section 6. In this theory, a particular player called center player plays a decisive role in the formation of political coalitions. This player owes his power to his position in a policy order. It is a player who can form winning coalitions with players on the left of him in that order, with players on the right of him in that order or with players on either sides of him. We give a sufficient condition for the existence of a center player in a simple game. We also formulate a result about the uniqueness of this player. With this theory we have made a start to solve problem 1.

In section 7, we formulate Axelrod's conflict of interest. We prove a result that shows a connection between this theory and the theory of center players.

In section 8 we refine the theory of center players by introducing the notion of balance of coalitions. This section will solve problem 1.

To illustrate the theory of center players including the theory of balanced coalitions we provide a computation example. This is done in section 9. In this example, the game representation of the Dutch parliament since its election of 6 September 1989 is used. It appears that the Dutch political party Christian -Democratic Appel is the center party in this game.

Chapter 6 In this chapter we will solve the first part of problem 2. We present two theories in which the formation of coalitional preferences and their use in coalition formation processes are described and explained.

As noticed in section 1.2 of this introductory chapter, the first theory that uses coalition preferences in order to predict sets of coalitions is De Swaan's policy distance theory. In section 2 of chapter 6, we will present a re-examination of both the open and closed version of this theory. According to both theories, a player will prefer a coalition

S to a coalition T if the distance between his policy position and the expected policy position of S is less than the distance between his policy position and the expected policy position of T. The behavioral assumption is that each player tries to be a member of a coalition with an expected policy position that is as close as possible to his own. It appears that both the open and closed version of policy distance theory are inconsistent.

In section 3 we present a new theory called the power excess theory of coalition formation. Crucial in this theory is the notion of power excess. The power excess of a player i in a coalition S is the difference between the power of i and the power total of $S - \{i\}$. Power can be measured in several ways, for example by using one of the standard power indices or by using, simply, the weights of the players in the case of a weighted majority game. A player's preference concerning the possible winning coalitions accords with the power excess of that player in the several coalitions. If i has a larger power excess in coalition S than in coalition T, then he will prefer S to T. The behavioral assumption is that each player strives to form a coalition in which he has maximal power excess. We formulate a policy-blind version and a policy version of this theory. It appears that the policy-blind version is closely related with Riker's minimum size theory and that the policy version is related to an adjusted version of minimum size theory in which policy is used as an explaining variable.

In this chapter we have solved the first part of problem 2. We have examined the merits of an already existing theory (policy distance theory) and we have formulated a new theory in which players'preferences concerning coalitions are essential. These players' preferences are used to determine prediction sets of coalitions. However, we did not yet solve the second part of problem 2.

Chapter 7. In this chapter we solve part 2 of problem 2. We first define the notion of social choice game. This will be done in section 2. A social choice game is a simple game in which each player has a preference concerning a set of alternatives. More specifically, a social choice game is a simple game with the folllowing parameters:

1. a set of alternatives,
2. a set of players,
3. a set of preferences, one for each player, on the set of alternatives, and
4. a set of winning coalitions.

The problem is to construct individual preferences on the set of winning coalitions by using the individual preferences on the set of alternatives.

In order to do this we introduce in section 3 the notion of Hamming distance function. By using the operation of set difference between two preferences and by determining subsequently the cardinality of the thus obtained set difference, we obtain a measure of the dissimilarity of two preferences. The function that assigns to each ordered pair of preferences the cardinality of the set difference of the preferences is called the Hamming function. We prove that this function is a metric. In this section we also introduce the notion of betweenness of preferences and the notion of linear profile. We will prove that if the set of individual preferences constitutes a linear profile, a majority choice will exist.

In section 4 we present the first coalition formation theory of this chapter. This theory, called conflict minimization theory, is based on the idea that each player strives to form a coalition with minimal conflict. In order to determine the size of conflict in a coalition a conflict index is introduced, specifically, a conflict index that uses information about the Hamming distances of the pairs of preferences at hand. This is the so-called Hamming conflict index. We then develop the descriptive part and the solution part of this theory. It appears that conflict minimization theory leads to aggregated relations for which the set of maximal choices (the core) is not empty. In order to demonstrate the working of the theory, we present a computation example.

In section 5 we present the second theory to solve our problem. This theory is called preference distance theory and is inspired by De Swaan's policy distance theory. Starting point of this theory is that in each winning coalition a social preference will be produced concerning the set of alternatives. The production of the social preference in a coalition of course depends on the social choice rule used in that coalition. If each player knows the social choice rule to be used in a coalition, then each player can calculate the Hamming distance between his preference and the social preference of each coalition. He then is able to determine his preference concerning coalitions. That is, he will prefer a coalition with a social preference that is close to his own preference to a coalition with a social preference that is far away. Moreover, he will strive towards a coalition with a social preference on the set of alternatives that is as close as possible to his own preference on the set of alternatives. It appears that this theory not automatically leads to a relation for which the set of maximal choices is nonempty (for which a nonempty core exists). We formulate a sufficient condition for the existence of a nonempty core of coalitions. However, in order to guarantee a nonempty prediction set of coalitions for any situation, we link this theory to the theory of generalized stable sets as developed in chapter 4. Finally, we present a computation example that shows the working of the preference distance theory of coalition formation

By presenting two theories that both use individual preferences on a set of alternatives as determinants of individual preferences on a set of winning coalitions and by using these players' preferences on coalitions to predict sets of coalitions, we have solved our last problem. So we have reached the end of our journey in that chapter.

Chapter 2

Fundamentals of Social Choice Theory

2.1 Introduction

In this chapter some basic concepts and results of social choice theory will be presented. These concepts and results will be used in the following chapters to study coalition formation processes. Our presentation of social choice theory follows that of Arrow (1963), Blair and Pollack (1982), and Sen (1970, 1977, 1986). Sen (1977, 1986) calls this format the relational approach and distinguishes it from, what he calls, the functional approach (see Kelly 1978, 1988; Fishburn 1973, Suzumura 1983).

The basic concepts and related assumptions which will be discussed in this chapter are: social state, agenda, individual, preference, choice, rationality, preference profile, social choice system and collective rationality. Further, we present a formulation of one of the most important results in social choice theory, namely, Arrow's theorem. We provide a new and easy proof of this theorem. This proof will give a good illustration of the working and the range of Arrow's theorem.

The relational approach as discussed by Sen (1970, 1977, 1986) has a clear structure. As it will be presented in this work, it consists of two parts, namely, a *descriptive part* and a *solution part*. The descriptive part starts with the notion of social choice problem. A social choice problem is a set of alternatives from which collectively a choice must be made. Further, this part deals with the preferences and choice behavior of the concerned individuals, the nature of social choice rules and the social structuring of social choice problems as produced by these rules. The social structuring of a social choice problem is usually called a social preference. The second part, the part on solution, shows how a social choice may or will be produced. It explains or evaluates how a social choice problem may or will be solved using the information as produced in the descriptive part. Both parts of the theory will be treated in this chapter. Sections 2 through 7 of this chapter deal with the descriptive part. In section 7 a start is made of the discussion of the solution part. We turn back to the solution part in chapter 4 where it will be studied more extensively. In section 8 the descriptive part and the solution part are brought together by presenting Arrow's theorem. First the relevant conditions in this theorem are described and discussed. Subsequently, our proof of the theorem is presented. Also the meaning of this theorem is shortly discussed. In the final section we discuss shortly the functional approach and compare it with the relational approach.

2.2 Social States

Starting point for the theory of social choice is a nonempty set of choice objects, called *social states*. This set will be denoted by X. In the sequel, X is called a *social choice problem*. The power set of X, that is, the set of all subsets of X, is denoted by $\mathcal{P}(X)$. Social states will be denoted by x, y. A social state is a primitive term of the theory, that is, a term that is not explicated by the theory. The theory takes it, so to say, for granted.

The term 'social state' was introduced by Arrow in his path-breaking work *Social Choice and Individual Values*. According to Arrow, if the term had to be defined, then

> "the most precise definition of a social state would be a complete description of the amount of each type of commodity in the hands of each individual, the amount of labor to be supplied by each individual, the amount of each productive resource invested in each type of productive activity, and the amounts of various types of collective activity, such as municipal services, diplomacy and its continuation by other means,

and the erection of statues for famous men." (Arrow 1963: 17).

In short, a definition of a social state would entail a complete description of a societal vector whose components are values of significant sociological, economical or political variables that characterize a societal situation. In this view, distinct social states correspond to distinct societal vectors. This distinction may be threefold: societal vectors may differ in the values of the components, they may have different components, or both.

The view of social states as societal vectors leads to a handy and beautiful structuring of the social choice problem under scrutiny, that is, of X. The societal vectors can, then, be arranged into a matrix (see figure 2.1).

The social states are presented by the rows in this matrix. The columns

$$\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & & & \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{pmatrix}$$

Figure 2.1: Matrix representation of a social choice problem

contain the relevant variables that characterize a societal situation. The entries show the values of the respective variables in the corresponding social state. This societal decision matrix would be a nice mathematical description of X. In order to resolve the social choice problem a selection of a row from this matrix must be made. Several basic definitions and assumptions of the theory would be expressible in simple matrix theoretical terms.

Unfortunately, the view on social states as societal vectors is not exhaustive. Many social choice problems cannot be described in terms of societal decision matrices. To pin the theory only to such matrices would therefore mean an apriori reduction of the set of possible fields of application of the theory and thus of its relevance. Besides that, it neither fits the aim of the theory. As Arrow formulated it, the problem of social choice is how, under which plausible assumptions, a choice can be made (Arrow

1963: 103 e.f.). This means that the attention is on the social choice mechanism itself, not on the nature of the input of the mechanism, that is, on the nature of the social choice problem. The apriori matrix view is, therefore, not only too thin, it is also unnecessary.

A social state can best be seen as anything that might be conceived of as a solution of a social choice problem and therefore as a choice object. This view does not refer to any particular model of a social state. It allows all the degrees of freedom needed to make an interpretation. To avoid an annoying use of the term 'social state', we also make use of the equivalent terms 'alternative' or 'option'

Actually, not every conceivable social state is feasible. Some social states will be politically unattainable, while others are economically too costly to consider. Still others do not fit into the cultural climate in which the decision making process takes place. Another number of social states will not be considered as solutions of decision-making or policy problems because they are, technically or ethically, not implementable. In fact, there are numerous reasons for a conceivable social state of being not feasible.

Any finite collection of at least three feasible social states will be called an *agenda*. An agenda will be denoted by A. The set of possible agendas is denoted by Ω. Of course, $\Omega \subset \mathcal{P}(X)$.

The order of the alternatives on the agenda may be very relevant in producing a social choice. Different orders for the same agenda may produce different social choices. See for excellent illustrations of this fact, Straffin (1980). This fact is also well-known in political practice. Therefore, agenda formation processes are not only very interesting but also highly relevant. However, in this work we leave this theme aside, be it with some regret. A treatment of this theme would not change our basic results on coalition formation and is therefore superfluous. We confine ourselves to the assumption that nonempty agendas with at least three but in any case a finite number of alternatives will be formed.

2.3 Individuals and Coalitions

Another basic set for the theory is the set of *individuals*. Again, an 'individual' is a primitive term. It refers to a decision-making unit that is clearly distinguishable from other units. Usually, it stands for human beings. However, it may refer to any living being capable of intentional action. We think hereby of dolphins and chimpansees. For a study of dolphin cognition see Schustermann et.al. (1986), especially the challenging hypotheses of Jerison about the dolphins'world in this work. For a study of the formation of coalition preferences in chimpansee colonies see De Waal (1981). A path-breaking and daring work on animal thinking, which is in our view a necessary condition for intentional choice, is Griffin (1984). The term 'individual' also may refer to a collective of individuals that together can be treated as one behavioral unit or one social actor without loss of relevant information. We think hereby of interest groups, political parties, etc.

The set of individuals will be denoted by N. The individuals in this set are denoted by i, j, etc. It is assumed that N has at least three members and is finite unless stated otherwise. Further, it will always be assumed that N is given apriori, that is, it is fixed in advance. We do not work with variable populations in this monograph. The reason for this convention is not because the notion of variable populations is unimportant but because it does not touch upon our central research problem.

A nonempty subset of N is called a *coalition*. A coalition will be denoted by S. The notion of coalition does not assume anything particular about cooperative behavior. A coalition is, as stated, a set of individuals, nothing more and nothing less. It is quite possible that two ore more individuals are considered as a coalition while they do not know each other, let alone that they have the ability to communicate face-to-face with each other. This general approach allows us to collect individuals together into a set that have some characteristic in common. It allows us, for example, to form the maximal set of individuals that have the same preference vis-a-vis a particular pair of social states.

This approach differs in a fundamental aspect from the game-theoretical oriented coalition theories in political science (Van Deemen 1989, De Swaan 1973, 1985, Ordeshook 1986, Riker 1962, Riker and Ordeshook 1973). In this tradition, a coalition is seen as "an agreement among two or more persons to control their actions (choices or strategies)" (Ordeshook 1986: 302). In contrast to social choice theory, this approach starts from the assumption that cooperation is the essence of coalitions.

The coalitions we want to study in this work are formed on a cooperative basis and are, as such, the result of rational choices of intentionally behaving individuals. In the subsequent chapters we will always consider coalitions from this cooperative point of view. Only in this chapter we keep in mind the more general notion of coalition as used in social choice theory.

2.4 Preference and Choice

Consider a social choice problem X. It contains options that might be conceived of as possible candidates for the solution of this problem. Each individual will appreciate these options according to his own value system. Some options will be more valued than others. Still others will be judged to be of the same value. This relative valuation is the essence of the notion of *individual preference*. In fact, it is a binary relation on the set X with some nice properties. These nice properties will be discussed below.

Given a set of individual preferences, is it possible to derive a relation over the set of social states that may function, in a way of speaking, as a social value? In other words, is there a value that reflects in a more or less reasonable way the individual preferences and that has the same reasonable properties as these individual preferences? This is the basic question posed in Arrow's *Social Choice and Individual Values*. In the same work Arrow also provides an answer in the form of a theorem, the so-called Arrow's impossibility theorem. This answer, which will be discussed later in this chapter, is very disturbing if not shocking. It belongs to the hard core of social choice theory. Just as in the individual case, these social values can be thought of as binary relations on X. We will call them *social preferences* after Arrow (1962) and Sen (1970, 1977, 1986).

To formulate a general frame for preferences, we study in this section preferences without looking at their individual or collective status. Later, these concepts will be used both on the micro-level and on the macro-level. We do not study the notion of utility and its relation to social choice. For this see Bezembinder (1987).

Formally, a preference on X is a binary relation on X, denoted by R. As is standard, we write xRy instead of $(x, y) \in R$.

Definition 2.1 *Let R be a preference on X.*

- $xPy := xRy$ *and* $\neg(yRx)$*;*

- $xIy := xRy$ *and* yRx.

P is the asymmetric part of R and is called *strict preference*. I is the symmetric part of R and is called *indifference*. Clearly, $P \cap I = \emptyset$.

A preference may or may not satisfy any of the following properties:

Definition 2.2 *A preference R on X is*

1. reflexive := $\forall x \in X : xRx$;
2. asymmetric := $\forall x, y \in X$: *if* xRy, *then* $\neg(yRx)$
3. antisymmetric := $\forall x, y \in X$: *if* xRy *and* yRx, *then* $x = y$.
4. symmetric := $\forall x, y \in X$: *if* xRy, *then* yRx;
5. complete := $\forall x, y \in X$: xRy *or* yRx;
6. transitive := $\forall x, y, z \in X$: *if* xRy *and* yRz, *then* xRz;
7. *quasi-transitive* := $\forall x, y, z \in X$: *if* xPy *and* yPz, *then* xPz;
8. cyclic := *there is an R-cycle in X where an R*-cycle *is a finite sequence* $x_1, \ldots, x_m$ *in X such that* $x_1Px_2Px_3 \ldots x_{m-1}Px_mPx_1$.
9. acyclic := *R contains no R-cycles.*

In fact, preferences can satisfy several conditions simultaneously. The most used packages of conditions are given in the next definition:

Definition 2.3 *A preference R on X is a*

1. partial order := *R satisfies reflexivity, anti-symmetry and transitivity;*
2. weak order := *R satisfies reflexivity, completeness and transitivity;*
3. linear order := *R satisfies reflexivity, anti-symmetry, transitivity and completeness.*

The following sets of preferences are important:

1. $B(X)$ is the set of reflexive and complete preferences on X,

2. $A(X)$ is the set of reflexive, complete and acyclic preferences on X,

3. $Q(X)$ is the set of reflexive, complete and quasi-transitive preferences on X,

4. $O(X)$ is the set of weak orders on X, and

5. $L(X)$ is the set of linear orders on X.

It is evident that $L(X) \subset O(X) \subset Q(X) \subset A(X) \subset B(X)$. A nicely elaborated classification system for preferences satisfying specific packages of conditions can be found in Storcken (1989).

Definition 2.4 *Let $R \in B(X)$, $A \in \Omega$ and $x \in A$.*

1. *x is* R-best *for $A := \forall y \in A : xRy$;*

2. *x is* R-maximal *for $A := \neg\exists y \in A : yPx$.*

The set of R-best social states for an agenda A will be denoted by $\beta(A, R)$. The set of R-maximal social states for A is denoted by $\mu(A, R)$.

Theorem 2.1 1. *$\beta(A, R) \subseteq \mu(A, R)$ for every $A \in \Omega$ and every $R \in B(X)$.*

2. *$\beta(A, R) = \mu(A, R)$ if R is complete.*

Since the proof of this proposition is easy, we leave it to the reader. Also cf. Sen (1970) or Pattanaik (1971).

The following fundamental result gives necessary and sufficient conditions for the existence of a best social state. For a proof of this well-known result, see Sen (1970: 16) or Suzumura (1983: 32).

Lemma 2.1 *Let R be a preference on X. $\beta(A, R) \neq \emptyset$ for every $A \in \Omega$ if and only if R is reflexive, complete and acyclic.*

According to this lemma, if R is cyclic or if R is not complete, then for some agenda a best alternative does not exist.

2.4.1 Choice

A choice is the outcome of a behavioral act. This behavioral act consists of selecting alternatives from an agenda. A static description of choice behavior is provided by the notion of a choice function.

Definition 2.5 *A* choice function *is a function* $C : \Omega \rightarrow \mathcal{P}(X)$ *satisfying*

- $C(A) \subseteq A$ *for every* $A \in \Omega$;

- $C(A) \neq \emptyset$ *for every* $A \in \Omega$.

In short, a choice function is a rule that describes how to each agenda a nonempty subset of that agenda is assigned. If C is a choice function and if A is an agenda, then $C(A)$ is called the *choice set of* A.

The definition of a choice function does not imply that choices are produced on the base of information on preferences. The definition does not preclude the production of choices on the basis of, for example, chance mechanisms, religious codes, oracles or fortune tellers.

2.4.2 Rational Choice

Rationality is the selection of a best alternative (see Plott 1973, Schwartz 1986, Suzumura 1983). This simple view on rationality consists of two components. The first is the value component. This component has to do with determining what is best and is captured by means of the notion of preference. The second is the behavioral component. This component has to do with selecting an alternative and is summarized by the notion of a choice function. Thus, the notion of rationality is in essence a combination of the notion of preference and the notion of a choice function.

Definition 2.6 *A choice function* $C : \Omega \rightarrow \mathcal{P}(X)$ *is* rational := *there is a preference* R *on* X *such that* $C(A) = \beta(A, R)$ *for every* $A \in \Omega$.

If so, we say that R rationalizes C. The following result is a corollary of lemma 2.1:

Theorem 2.2 *A preference* R *rationalizes a choice function* $C : \Omega \rightarrow \mathcal{P}(X)$ *if* R *is reflexive, complete and acyclic.*

Since transitivity implies quasi-transitivity which implies, in its turn, acyclicity, a preference R also rationalizes a choice function $C : \Omega \to \mathcal{P}(X)$ if R is reflexive, complete and transitive or if R is reflexive, complete and quasi-transitive. In the first case we will speak of *transitive rationality*, in short: *T-rationality*. In the second case we speak of *quasi-transitive rationality*, in short: *Q-rationality*.

Essential for the notion of rationality as presented so far is the existence of a best social state. Hence the importance of lemma 2.1. If there is no best social state, it is not possible to speak of rationality. A best social state will not exist if the relevant preference is incomplete or cyclic. The problem of how an individual might choose in this case is not solved by social choice theory or by any other rational choice theory. That is, the choice behavior of individuals having cyclic preferences is beyond the scope of any rational choice theory sofar. In contrast, the problem of how to choose *collectively* in the case of cyclic social preferences is actually one of the oldest problems in social choice theory. Already Marquis de Condorcet (1785, 1789, 1791) and Dodsgon (1867) have studied this problem. We return to the work of Marquis de Condorcet in chapter 3.

2.5 Preference Profiles

In the sequel it is assumed that each individual $i \in N$ has a preference on X that is complete and transitive. This assumption implies that each individual is T-rational.

An n-tuple of preferences, one and only one for each individual, is called a *preference profile*. Formally:

Definition 2.7 *A* preference profile *is a mapping* $p : N \to O(X)$.

A preference $p(i) \in O(X)$ in a preference profile p will be denoted by R_i^p. This can be read as 'the preference of individual i in profile p'. If the context is clear, we only write R_i. The asymmetric and symmetric parts of an individual preference can be read in a similar way. The set of all possible preference profiles will be denoted by Π. Thus

$$\Pi := (O(X))^N.$$

In representing preference profiles, we often will use a compact notation. In this notation xy is written for xPy and (xy) for xIy. Before each preference we put the frequency of individuals having that preference. We give an example:

Example 2.5.1

$$\begin{array}{rl} 3: & xyzw \\ 2: & z(wy)x \\ 1: & (xyzw) \end{array}$$

In this profile three individuals have the linear order $xyzw$. Two individuals have the weak order $z(wy)x$. These two individuals are indifferent with respect to w and y. One individual has the preference $(xyzw)$. This individual is completely indifferent.

We abstract away from the moral aspects and the possibly strategic use of individual preferences. What counts are the preferences as reported by the individuals, not whether they are good or immoral, insincere or sophisticated. A pioneering work on the strategic use of preferences is Pattanaik (1978). Also consider Moulin (1983). For a more elaborated game-theoretical approach to this theme, consider the work of Peleg (1984).

In social choice theory it is usually assumed that individual preferences as reflected in a preference profile are exogeneously determined. Information about the way they are formed is not incorporated into the theory. This fact is critized by Elster (1983). He calls it the thin view on rationality and he argues that the model should also incorporate information about preference *formation* processes. This critique is taken seriously in this work and it will be met partially. In the subsequent chapters we will present coalition theories that show how actors form their preferences with respect to coalitions. In Elster (1986) a number of other failures and limitations is given.

Hindess (1988) also critizes the model of individual rational choice that is presented here as a starting point for social choice theory. His main point is that this model does not take into account the fact that any individual is imbedded in a social structure. Each individual is seen as an atom. Hindess' critique is very interesting. In our opinion, it is a serious attack on the rational choice model. The position and functioning of an individual within a social structure is, probably, relevant with respect to the formation of his preference. A theory of preference and choice is, therefore, only adequate if it can explain the relation between the positional and the structural characteristics of a social network on the one side and the properties of individual preference on the other. The trouble with this point of view is that we do not know how such an adequate theory will look like.

A third critique has a more psychological flavour. It says that the individual rational choice model does not reckon with the personal history of the concerned individual. An individual preference and with that an individual choice must, in some way or another, be related to the individual's life history. This view implies that another life history of an individual may have led to another preference and another choice behavior. Therefore, to be adequate, a choice theory must deal with the history of an individual in explaining individual preference and choice. This critique deals with the static character of the rational choice model. To meet it means to construct dynamic models that explain the changing of tastes and values in terms of, say, psychological variables.

2.6 Social Choice Rules

The process of selecting an alternative (solution, social choice) from a social choice problem will be reconstructed in this work as a process consisting of two stages. In the first stage, a rule or device assigns a preference to a preference profile. In the second stage, the assigned preference is used to produce a social choice. In this section we study the first stage. In the next section the second stage is studied.

The preference assigned to a profile by a rule can be interpreted as a social value of the group concerned. It will therefore be called a *social preference*. The purpose of a social preference is to structure the social choice problem in order to make a social choice. In general, the nicer the properties of a social preference, the better the problem is structured and the easier it is to select a social choice. The preference-assigning rule or device is called a *social choice rule*, in short, SCR.

Definition 2.8 *1. F is a* social choice rule *(henceforward SCR) := F is a mapping from Π into $B(X)$.*

2. *The image $F(p)$ of a preference profile $p \in \Pi$ under an SCR F is called a* social preference.

The collection of all possible SCRs is denoted by Φ. That is,

$$\Phi := (B(X))^{\Pi}.$$

The definition of an SCR is very general in its description of the properties of social preferences. It just says that a social preference, as the image of a preference profile under an SCR, is reflexive. A more restrictive SCR is obtained by using the set of reflexive, complete and acyclic relations on X as its range.

Definition 2.9 *F is a* social decision function *(henceforth SDF) := F is a mapping from Π into $A(X)$.*

The set of all social decision functions is denoted by Φ_{SDF}. That is,

$$\Phi_{SDF} := (A(X))^{\Pi}.$$

The name 'social decision function' is taken from Sen (1970; also see Sen 1977, 1986). However, Sen defines an SDF in a rather difficult way (cf. 1970: 52). Our definition is equivalent to his but more simple. Clearly, an SDF is an SCS.

A further restriction is obtained by demanding quasi-transitivity instead of acyclicity of social preferences.

Definition 2.10 *F is a* quasi-transitive social choice rule *(henceforth QS) := F is a mapping from Π into $Q(X)$.*

The collection of all QS is denoted by Φ_{QS}. Formally:

$$\Phi_{QS} := (Q(X))^{\Pi}.$$

It is easy to verify that every QS is an SDF. However, not every SDF is a QS.

In his original work, Arrow demands, besides the conditions of completeness and reflexivity, the condition of transitivity of social preferences. This leads to the notion of a social welfare function.

Definition 2.11 *F is a* social welfare function *(henceforward SWF) := F is a mapping from Π into $O(X)$.*

The set of all possible SWF is denoted by Φ_{SWF}. Formally:

$$\Phi_{SWF} := (O(X))^{\Pi}.$$

The name 'social welfare function' chosen by Arrow is unfortunate, since it might be confused with the notion of Bergson-Samuelson social welfare function as used in the so-called Paretian Welfare Economics (de Graaff 1957, Samuelson 1967, 1977, Sen 1970, 1986). In Paretian Welfare Economics, social welfare is considered as an increasing function of individual utility indices. An individual utility index U_i for an individual i is a real-valued mapping with as domain a set of factors that are relevant for the welfare of i. A Bergson-Samuelson social welfare function is a real-valued function of the form

$$W = W(U_1, U_2, \ldots, U_n),$$

where U_i is the utility index of individual i and $\partial W/\partial U_i > 0$ for all i. As such, a Bergson-Samuelson social welfare function is a representation of a relation on a set of social states that is based on a value judgment. It is not a reconstruction of a mechanism that produces the value on which such a social welfare function must be based. And this is, according to Sen (1970, 1986), the fundamental difference with an Arrowian social welfare function. An Arrowian SWF is, as we have accentuated, a rule of producing social values, it is not a social value itself. It is a mechanism that produces reflexive, complete and transitive social preferences. In contrast, a Paretian social welfare function is a social preference, a representation of a social value that might, eventually, be produced by an Arrowian SWF. So, an Arrowian SWF is a rule, a Bergson-Samuelson social welfare function is an output produced by a rule.

In another work, Arrow (1967: 68) calls an SCR a *constitution*. He already proposes this term in his famous work (Arrow 1963). However, there a constitution is only another term for a social welfare function (see Arrow o.c: 105). Sen (1970, 1977, 1986) and Suzumura (1983) use the term *collective choice procedure*. Fishburn (1973) calls an SCR a social choice function while Kelly (1988) uses the same term as we do. Clearly, there is a diversity in nomenclature.

Since $O(X) \subset Q(X)$, we have $\Phi_{SWF} \subset \Phi_{QS}$.
Since $Q(X) \subset A(X)$, we have $\Phi_{QS} \subset \Phi_{SDF}$.
Since $A(X) \subset B(X)$, we have $\Phi_{SDF} \subset \Phi$.
Hence, because of transitivity of proper inclusion, we have

$$\Phi_{SWF} \subset \Phi_{QS} \subset \Phi_{SDF} \subset \Phi.$$

This collection of nested sets of social choice rules will be called the *Hierarchical Class of Social Choice Rules.*

In the mathematical systems theory of Mesarovic and Takahara (1975), a system is defined as a relation between an input set and an output set. If an input-output relation has the characteristics of a function, then Mesarovic and Takahara call it a *functional system.* So, in this sense, an SCR is a functional system. The input set is the set of all possible preference profiles Π, the output set is the set of all reflexive and complete binary relations $B(X)$. Seeing an SCR as a system reminds us that we must clearly distinguish between preference profiles (input), social preferences (output) and rules. Rules, in order to be systems, must be constant, i.e. may not vary in assigning output to input. Consequently, rules must be evaluated independently from the content of social states and from the nature of individual preferences.

In this study we preclude systems which use chance procedures. We do not wish to study choice properties of roulette wheels or dice. However, systems which involve magic, oracles, fortune telling, religious codes or something like that are allowed.

Contrary to what is often thought, the interpretation of an SCR need not be limited to voting procedures. According to Nakamura (1975, 1979) and Plott (1975), an SCR may be seen as a rule to play a game. Further, an SCR may represent a representative system, an electoral system or a democratic system in the direct sense (see Fishburn 1973, Murakami 1968, Mueller 1989). It also may represent a decision-making system in which violence and suppression are the ruling factors. See Fishburn (1973: 210-212) for a study of suppression in a social choice theoretical context. This broad set of interpretation possibilities is the result of the abstraction level of the social choice theoretical concepts and is one of the attractive properties of the theory. It allows both the general study of classes of social choice rules and the study of particular systems, e.g. the system of majority decision.

2.7 Collective Rationality and Social Choice

In the previous section we have discussed the first stage of a social choice process. We now turn to the second stage of a social choice process. In this stage, a choice set is specified for some agenda given the social preference information as produced in the first stage. That is, if F is an SCR, $p \in \Pi$ a preference profile and A an agenda, then $F(p)$ will be used in this stage to generate a choice set $C(A)$ where C is a choice function. This choice set can be thought of as a solution for the social choice problem under scrutiny. (Hence the name 'solution part') and will be called *social choice set*. An element in a social choice set will be called a *social choice*.

The question what kind of social choice sets should result from a social preference is at the hard core of social choice theory. Our answer to this question is closely related with our view on what a good solution is for a social choice problem. The first obvious candidate is the set of $F(p)$-best options for an agenda A. This leads to the classical concept of *collective rationality*.

Definition 2.12 *An SCR F satisfies* collective rationality *(henceforward CR)* $:= \forall p \in \Pi \forall A \in \Omega[\beta(A, F(p)) \neq \emptyset]$.

Consequently, if F satisfies collective rationality, then $C : \Omega \rightarrow \mathcal{P}(X)$ defined by $C(A) = \beta(A, F(p))$ is a choice function. The following result is a consequence of lemma 2.1.

Theorem 2.3 *Let F be a SCR. F satisfies collective rationality if and only if* $F \in \Phi_{SDF}$.

Thus, an SCR satisfies CR if and only if it is a social decision function. Clearly, also Q-systems satisfy collective rationality. Since a SWF is a QS[1], also a SWF satisfies collective rationality. Keeping in line with

[1]Consider the hierarchical class of social choice rules.

section 4.2 of this chapter, we call a QS *collective Q-rational* and a SWF *collective T-rational.*

Collective rationality is an important condition. It guarantees the existence of a best social choice and, hence, the possibility of selecting a social optimum. However, it has some serious shortcomings. First, this condition does not indicate what to choose in the case of cyclic social preferences. It indicates that there is no solution for a social choice problem in this case. It stops, so to say at optimality. If there is no best social choice, we cannot look any further with the notion of rationality. To cure this myopia, some alternatives have been formulated in the course of time (see Miller 1980, 1983; Banks 1985; Schwartz 1986; McKelvey 1986). In chapter 4 we present two extensions of the notion of collective rationality that both have a root in n-person game theory. There we also present our own contribution. These extensions will show what to select when a best alternative does not exist.

Another drawback of the condition of collective rationality is that it is inconsistent with some sets of conditions that all appear to be rather reasonable. This is expressed in a number of socalled impossibility theorems (cf. a.o. Kelly 1978, Schwartz 1986, Sen 1970, 1977, 1986, Storcken 1989, Suzumura 1983). The most celebrated of these is Arrow's theorem. This theorem will be studied in the next section.

In the previous section we noticed that an SCR is a system that must be clearly distinguished from its input set and output set. Consequently, the evaluation of a social choice must be clearly distinguished from the evaluation of the SCR producing that social choice. Neither the quality or content of social choices yielded by a SCR nor the quality of the individual preferences used as input for that SCR can be used to qualify that SCR. To be sure, it is quite possible that a very reasonable and fair SCR yields a rather criminal social choice. If, for example, a majority wants fascism, then the system of majority decision[2] will generate fascism. This does not mean that for this reason the system of majority decision is a bad system. 'Only' the social choice, then, is alarming, 'only' the preferences of a majority, then, are pathological.

[2]See for a study of this rule the next chapter

2.8 Arrow's Impossibility Theorem

An important task of social choice theory is to study relevant conditions that social choice systems may or should satisfy. The methodological status of these conditions is twofold. Firstly, these conditions can be interpreted as ethical constraints imposed on social choice systems. The conditions, then, are used in a normative manner. Adherents of this normative view are, among others, Arrow (1963, 1977) and Kelly (1978). Secondly, the conditions can be considered as positive laws that govern the behavior of social choice systems. Pronounced followers of this behavioral view are, among others, Plott (1976) and Schwartz (1986).

It is not necessary to choose already for a particular view. This is the advantage of the formal character of social choice theory. Sometimes it is useful to work with both views. For example, if we want to evaluate a real-life social choice system, then we must first investigate the conditions this system satisfies. After this we must compare the package of prevailing conditions with a package of conditions we like a social choice system to satisfy. Then we can make an evaluation and, eventually, criticize the existing system. To give another example, if decision-makers want another social choice procedure, then we may ask them for the conditions that the new system should satisfy. If the package of desired conditions is not internally inconsistent, we may design a system satisfying these conditions. Once installized, the conditions, then, serve as laws that regulate the behavior of the system.

Unrestricted Domain

The first condition we deal with is *Unrestricted Domain.*

Definition 2.13 (Unrestricted Domain) *An SCR F satisfies* unrestricted domain *(henceforward UD) := the domain of F includes all logically possible individual preferences.*

Since we have taken Π as the domain of an SCR, this condition is mathematically redundant. However, in social choice theory, it is tradition to take it up and to discuss it.

Arrow presents this condition in the second edition of his *Social Choice and Individual Values* (Arrow, 1963: 96). However, there is an important difference between our formulation and Arrow's one. Arrow formulates UD only for SWF's. Our formulation is concerned with social choice rules in general.

This condition forbids that a social choice rule only works for specific preference profiles. It says that there are no restrictions for the input set of a social choice rule and hence no constraints to form an individual preference, provided, of course, all individual preferences are reflexive, complete and transitive. In this sense, this condition has to do with individual freedom to form a preference.

Sometimes it is argued that UD is too robust. Preferences do obey some regularities and it is possible that just these regularities allow one to avoid the difficulties typically encountered in social choice processes (i.e. the difficulties as appearing in impossibility theorems). Thus, it may be useful to restrict the domain of a social choice rule. Within this view, a frequently studied domain restriction is *Single-Peakedness* which roughly means that the social states "can be ordered along a line in such a way that, as we pass from left to right along the line, each individual's preference increases up to a peak or to an indifference plateau, and then decreases thereafter." (Fishburn 1973: 101). For a review of this and other domain restrictions, consider Sen (1986).

Pareto Condition

Another important condition is the Pareto condition.

Definition 2.14 (Pareto Condition) *An SCR F satisfies the* Pareto condition *(henceforward PC) := for every preference profile $p \in \Pi$ and for every $x, y \in X$: if for all $i \in N : xP_i^p y$, then $xF(p)y$ and not $yF(p)x$.*

The Pareto condition says that if everyone strictly prefers x to y, then x also is or should be socially strictly prefered to y. This condition seems reasonable. To deny it would imply that there are possible situations in which everyone strictly prefers x to y while y is socially preferred. If the agenda is $\{x, y\}$, this would mean that, if collective rationality is satisfied, y would be the social choice while everyone finds x better. Another argument for the Pareto condition is that it avoids needless concessions (cf. Coombs and Avrunin, 1987).

An argument against the Pareto condition can be given by using Sen's socalled liberal paradox (Sen 1970: Ch.6 and 6*, Sen 1974). This fascinating paradox shows that the Pareto condition is inconsistent with a weak condition of liberalism. This, what Sen calls, condition of minimal liberalism requires "that each individual is entirely decisive in the social choice over at least one pair of alternatives ...." (Sen 1970: 79). Sen's liberal paradox implies that the Pareto condition must be given up when requiring minimal liberalism. This and other related problems have been studied extensively in Wriglesworth (1985). A game-theoretical approach to the liberal paradox is given in Gardenfors (1985).

Binary Independence of Irrelevant Alternatives

The most controversial of all conditions in the theory of social choice is *Independence of Irrelevant Alternatives.* The binary version of this condition says that if the individual preferences with respect to some pair of alternatives remain the same, then the social preference with respect to this pair of alternatives must remain the same, irrespective of the changes of the individual preferences with respect to other pairs of alternatives.

Definition 2.15 (Binary Independence of Irrelevant Alternatives) *Let F be a SCR. F satisfies* Binary Independence of Irrelevant Alternatives *(henceforth IIA) := for all $p, q \in \Pi$ and for all $x, y \in X$, if for all $i \in N$,*

$$R_i^p \cap (\{x,y\} \times \{x,y\}) = R_i^q \cap (\{x,y\} \times \{x,y\}),$$

then

$$F(p) \cap (\{x,y\} \times \{x,y\}) = F(q) \cap (\{x,y\} \times \{x,y\}).$$

This is the binary version of IIA. It is also possible to use an agenda with three or more elements. We then get an n-ary version. This version says that if individual preferences with respect to the alternatives on this agenda remain the same, then the social preference with respect to the alternatives on this agenda must remain the same, irrespective of the changes of the individual preferences with respect to the alternatives outside this agenda. In following Blau (1972), Kelly (1978) proves that the binary version of

IIA is equivalent to the n-ary version whenever F is an SDF. See Kelly (1978: 31-32). Also see Sen (1986: 1096-1097).

According to Plott (1976), IIA is a universal law. In his view there are neither real-life systems nor conceivable systems that violate this principle. It is obeyed by all our "current societal models" (Plott 1976: 535). According to Schwartz (1986: 33),

> "Independence of Irrelevant Alternatives and therewith Binary Independence are eminently reasonable assumptions to make in a realistic study of collective choice. I know of no real-world collective-choice process that violates either condition."

To see why it is very difficult to escape from IIA, consider the negation of it. Then there are choice situations p and q and alternatives $x, y \in X$ such that $R_i^p \cap (\{x,y\} \times \{x,y\}) = R_i^q \cap (\{x,y\} \times \{x,y\})$ for all $i \in N$, while $F(p) \cap (\{x,y\} \times \{x,y\}) \neq F(q) \cap (\{x,y\} \times \{x,y\})$. But how to explain this difference? What alternatives other than x and y might have caused this difference in the respective social preferences? According to Fishburn (1973: 7), the idea of allowing irrelevant alternatives "to influence the social choice introduces a potential ambiguity into the choice process that can at least be alleviated if not removed by insisting on the independence condition." However, the violation of the principle implies more than a "potential ambiguity". If IIA is not required, then not only preferences with respect to irrelevant alternatives but also every other variable operative in this world and everything else we can conceive of might influence the production of a social preference and with that the outcome of a social choice process. In this case, it is very difficult to say anything meaningful about social choice processes.

Nondictatorship

The essence of decision-making is power. How must power be distributed among the individuals and coalitions? This problem is not easy. If power is concentrated in the hands of a few, then it is likely that some problems will occur with respect to the acceptability of the social choices and with that with the implementation of these. If power is dissipated too much, then decision deadlocks may result already by a low degree of divergence of individual preferences.

The most concentrated form of decision-making power is dictatorship. A dictator has the power to dictate the social preference, irrespective of the preferences of the other individuals. Clearly, dictatorship is in contradiction with any notion about democracy. The condition that forbids the existence of a dictator is called the Nondictatorship condition. In chapter 5 of this work we study other concentrations of decision-making power.

Definition 2.16 (Nondictatorship) *A SCR F satisfies* Nondictatorship *(henceforward ND) := there is no $i \in N$ such that for every $p \in \Pi$ and for every $x, y \in X$: if $xP_i^p y$, then $xF(p)y$ and not $yF(p)x$.*

We are now ready to formulate the first fundamental law of political science.

Theorem 2.4 (Arrow's Impossibility Theorem) *There is no $F \in \Phi_{SWF}$ satisfying UD, IIA, PC and ND.*

We give a simple proof that illustrates the working of the conditions in an illuminating way[3]. Alternative proofs can be found, among others, in Arrow (1963), Sen (1970), Fishburn (1973), Hansson (1975), Kelly (1978), Schwartz (1986) and Storcken (1989).

Proof of Arrow's theorem *We first prove the theorem to be true for three alternatives and two individuals who both have linear orders as preferences with respect to these three alternatives. We also restrict the range of social welfare functions to the set of possible linear orders on this set of three alternatives. We denote the three alternatives with a, b, c. The two individuals are, say, Romeo and Julia. Because of UD we must consider all* $3! \times 3! = 36$ *possible preference profiles. For this consider the following 6 by 6 matrix.*

[3]The inspiration of this proof is found in the second proof method of the Gibbard-Satterthwaite theorem as presented in the excellent work of Schmeidler and Sonnenschein (1978).

		Julia					
		a b c	b a c	c b a	a c b	b c a	c a b
	a b c	a b c	ac (PC) bc (PC) bac (1)	ba (2) cb (10) cba (T)	ac (PC) cb (10) acb (T)	bc (PC) ca (8) bca (T)	ab (PC) ca (8) cab (9)
	b a c	bc (PC) ab (12) abc(T)	b a c	ba (PC) cb (10) cba (T)	ac (PC) cb (10) acb (11)	bc (PC) ca (8) bca (T)	ca (8) ab (12) cab (T)
	c b a	bc (4) ab (12) abc (T)	ba (PC) ac (6) bac (T)	c b a	ac (6) cb (PC) acb (T)	ca (PC) bc (4) bca (T)	ca (PC) ab (12) cab (T)
Romeo	a c b	ab (PC) bc (4) abc (T)	ba (2) ac (PC) bac (3)	ba (2) cb (PC) cba (7)	a c b	bc (4) ca (8) bca (T)	ab (PC) ca (8) cab (T)
	b c a	bc (PC) ab (12) abc (T)	ba (PC) ac (6) bac (T)	ba (PC) cb (10) cba (T)	ac (6) cb (10) acb (T)	b c a	ca (PC) ab (12) cab (T)
	c a b	ab (PC) bc (4) abc (5)	ba (2) ac (6) bac (T)	ba (2) cb (CP) cba (T)	ac (6) cb (PC) acb (T)	ca (PC) bc (4) bca (T)	c a b

Table 2.1: Romeo-Julia matrix for Arrow's theorem

The rows consist of the possible preferences of Romeo. The columns consist of the possible preferences of Julia. For these reasons we call it the Romeo-Julia matrix. We first explain some letters and numbers in this matrix:

- *T stands for 'Transitivity'. This property can be used because we are dealing with social welfare functions.*
- *(n) where n is a natural number, means that the corresponding result depends on the result that corresponds with the preceding number $n - 1$.*
- *We remember that IIA stands for Independence of Irrelevant Alternatives and PC for the Pareto Condition.*

The explanation of the succeeding numbers in the matrix is as follows:
(1): in the entry (abc, bac), we have ac because of PC and bc because of PC. Remains ab or ba. Suppose ba which is the preference of Julia. Then bac because of T.
(2): ba because of (1) and IIA.
(3): bac because of T.
(4): bc because of (3) and IIA.
(5): abc because of T.
(6): ac because of (5) and IIA.
(7): cba because of T.
(8): ca because of (7) and IIA.
(9): cab because of T.
(10): cb because of (9) and IIA.
(11): acb because of T.
(12): ab because of (11) and IIA.
Now looking in each entry of the matrix, we see that each social preference is the preference of Julia. Hence for this case, Julia is the dictator and, therefore, the theorem is true for $X = \{a, b, c\}$ and $N = \{Romeo, Julia\}$. Note that if we had decided in the entry (abc, bac) of the matrix to use ab instead of ba, then Romeo would have been the dictator.

Now we prove that Julia remains a dictator when expanding the set of alternatives. Let X" be a nonempty finite set of alternatives such that $X" \cap X = \emptyset$.

Step 1. *Pick an arbitrary alternative from X", say u, and substitute an alternative from $\{a, b, c\}$ by u, say c. Then by looking at the matrix with c substituted by u it is easy to verify that Julia is a dictator for $\{a, b, u\}$. The same is true for $\{a, u, c\}$ and $\{u, b, c\}$. But then Julia is able to determine the social preference between any pair from $\{a, b, c, u\}$ because of IIA. Hence Julia remains a dictator. Note that it is allowed in this step to use indifferent pairs.*

Step 2. *Take another alternative from X", say $v \neq u$. First substitute one of the alternatives in $\{a, b, c, u\}$ by v, say c. Then look at the matrices that correspond with the possible triple sets of alternatives from $\{a, b, v, u\}$. In the same way as in the preceding step, the matrices show that Julia can determine the social preference for each of these triple sets. But then Julia is also a dictator with respect to $\{a, b, v, u\}$ (see step 1). The same is true when substituting a, b, or u in the set $\{a, b, c, u\}$ by v. Because of IIA, Julia can then determine the social preference for each pair of alternatives form $\{a, b, c, u, v\}$.*

Step m. *Proceeding in this way we can exhaust the set X in m steps where $m = \mid X \mid$. Julia remains a dictator.*

Now let us make Julia happier by replicating Romeo a number of times, say n times. Since we have n Romeos, we now have n Romeo-Julia matrices, each of which shows that Julia is a dictator. That is, Julia is a dictator for each Romeo with respect to $X = \{a, b, c\}$. By using the arguments of the preceding steps, Julia remains a dictator for each Romeo when expanding $X = \{a, b, c\}$ with a nonempty finite set X" such that $X" \cap X = \emptyset$.
Does Julia therefore dictate all Romeos? Yes, she does. For let us suppose that she does not. Then there are x and y such that $xP_{\text{Julia}}y$ but not x is socially strictly preferred to y. Because of the Pareto condition, there must then be a Romeo with $yR_{\text{Romeo}}x$. But then, Julia does not dictate this Romeo which is in contradiction with the fact that she dictates each of them. Indeed, sweet Julia swallows them all. □.

What does Arrow's theorem exactly mean? Consider the hierarchical class of social choice systems as presented in section 6 of this chapter.

Arrow's theorem is a characterization of the class of social welfare functions, that is, of Φ_{SWF}. More precisely, the theorem is concerned with the existence of a subset of the set of SWFs in a negative way. It says that the subset of SWFs satisfying UD, PC, IIA and ND is empty. So, Arrow's theorem characterizes the most restrictive class of social choice rules. It says nothing about QSs and other systems which are not SWFs. In fact, there are QSs that belong to $\Phi_{QS} - \Phi_{SWF}$ and that satisfy the Arrow conditions. An example is the Pareto rule. See Sen (1970: Ch.5). The definition of this rule is as follows:

Definition 2.17 *A social choice rule F is called the* Pareto extension rule *:= for every x and y and for every $p \in \Pi$:*

1. *$xF(p)y$ when $xR_i^p y$ for all $i \in N$,*

2. *$xF(p)y$ and $yF(p)x$ otherwise.*

It is easy to check that the Pareto extension rule is a QS but not an SWF. It is also easy to verify that this rule satisfies UD, IIA, PC and ND.

Arrow's theorem has a particular methodological status. It expresses the *logical impossibility* of an SWF with the Arrow conditions. Since the set of empirically possible worlds (realities) is a proper subset of the set of logically possible worlds, that what is logically impossible, cannot exist in reality. Hence you need not look for an SWF with the required properties in the real world unless you do not believe in first order predicate logic. In this sense, Arrow's theorem is empirically irrefutable. It only can be tackled, perhaps, by using some deviant logic.

In the course of time several variations of Arrow's theorem have been presented (see Blair and Pollack 1982; Kelly 1978; Schwartz 1986; Sen 1970, 1977, 1986; Storcken 1989; Suzumura 1983). Most of these variations show an interplay between collective rationality and concentration of decision-making power (cf. Van Deemen 1988). In chapter 5 we shortly return to some of these variations.

2.9 Social Choice: The Functional Approach

The input set for an SCR is the set Π, the output set is $B(X)$. In social choice theory, another framework is in use to study social choice. Sen

(1977, 1986) calls this framework the *functional approach* to social choice. In this subsection we compare this approach with the relational approach as used in this work.

In the functional approach, social choices are studied directly in terms of choice functions without referring to social preferences. This approach is mainly the result of seeking escape routes to Arrow's impossibility theorem (cf. Sen 1977a: 166-7). The approach is used in the standard works of Kelly (1978, 1988) and Suzumura (1983). Fishburn uses a variation of it in his excellent *The Theory of Social Choice*. We shortly study this approach (including Fishburn's approach) and compare it with the relational approach.

Let Γ denote the set of all choice functions from Ω to $\mathcal{P}(X)$.

Definition 2.18 *G is a* Kelly-Suzumura social choice rule *(henceforward KS-SCR)* $:= G : \Pi \rightarrow \Gamma$.

In other words, a KS-SCR is a function that assigns a choice function to each preference profile. The final set of an agenda A produced by a choice function as a value of a preference profile is called the *social choice set of A*. A choice function associated to a preference profile $p \in \Pi$ by a KS-SCR G will be denoted by C_p^G. That is: $C_p^G := G(p)$.

The input set of a KS-SCR is, just as in the case of a relational SCR, the set Π. The output, however, is not a binary relation (social preference) but a choice function that has the set of agendas as argument. The final result is a nonempty social choice set for each agenda.

Using now the concept of rational (social) choice as developed in section 4 and 7 of this chapter, the conditions of collective rationality can be defined in the following adjusted way:

Definition 2.19 *Let G be a KS-SCR.*

- *G is* collective rational $:= C_p^G$ *is rational for each* $p \in \Pi$.
- *G is* collective Q-rational $:= C_p^G$ *is Q-rational for each* $p \in \Pi$.
- *G is* collective T-rational $:= C_p^G$ *is T-rational for each* $p \in \Pi$.

For the definition of the several rational choice functions consider section 4 above.

The most important difference between the relational approach and the functional approach as formulated by Kelly, Suzumura and others is the place and use of the notion of a social preference. In the relational approach a social preference is needed to yield a social choice. A best social choice is produced if a social preference satisfies a package of conditions that are sufficient and sometimes necessary to generate a rational choice function. The essence, however, is: first a preference, then a choice. A social preference is needed to reveal a choice. The functional approach is a reversed world. In this approach a social preference is not a starting point but a possible endpoint that may be revealed if some conditions of choice are satisfied. The essence in this approach is: first a choice, then a preference. A series of choices may reveal, under some conditions, a social preference. The conditions under which a social preference may be revealed by a series of choices has been a major theme in the functional approach. For summarizing studies of these conditions see Bordes (1975, 1979), Kelly (1978), Schwartz (1986), Sen (1977, 1986) or Suzumura (1983).

Fishburn (1973) uses a variation of the Kelly-Suzumura framework. He takes as input set of an SCR the cartesian product of Ω and Π and as output set $\mathcal{P}(X)$. An SCR as defined by Fishburn will be called a *Fishburn social choice rule*, in short, an F-SCR.

Definition 2.20 *H is a* Fishburn social choice rule *(henceforward F-SCR)* $:= H : \Omega \times \Pi \to \mathcal{P}(X)$ *such that for every* $(A, p) \in \Omega \times \Pi$,

1. $H(A,p) \subseteq A$
2. $H(A,p) \neq \emptyset$.

An element in the range of an F-SCR is called, of course, a social choice set. So an F-SCR maps an agenda together with a preference profile straight out into a social choice set. Hence, an F-SCR is a choice function with a different argument as the choice functions defined in section 4.1.

We formulate the conditions of collective rationality for F-SCRs as follows:

Definition 2.21 *Let H be an F-SCR.*

- *H is* collective rational := *there is a binary relation R such that* $H(A,p) = \beta(A,R)$ *for every* $(A,p) \in \Omega \times \Pi$.

- H *is* collective Q-rational := *there is a quasi-transitive binary relation R such that $H(A,p) = \beta(A,R)$ for every $(A,p) \in \Omega \times \Pi$.*

- H *is* collective T-rational := *there is a transitive binary relation R such that $H(A,p) = \beta(A,R)$ for every $(A,p) \in \Omega \times \Pi$.*

If an F-SCR is rational, then the underlying binary relation must be reflexive, complete and at least acyclic[4]. The similarity with the Kelly-Suzumura approach is obvious. Also for the Fishburn-framework a social preference must be revealed by choices. The differences are, firstly, that Fishburn uses the cartesian product of Ω and Π as the input set while Kelly and Suzumura use Ω as the input set, and, secondly, that Fishburn uses $\mathcal{P}(X)$ as the output set, while Kelly and Suzumura use Γ. However, the effect is the same. Either system produces the same choices for the same preference profiles and agendas. In this sense, they are equivalent (see Sen 1977: 166). The Fishburn-variation surely is flexible and elegant. Anyway, it is the most compact conceptual framework.

We choose for the relational approach for two reasons:

1. It is relatively the most simple to work with.

2. It ressembles the game-theoretical approach to collective decision-making processes.

According to Shubik (1982: 127):

> A theory of games can be regarded as composed of two parts, a descriptive theory and a solution theory. The descriptive part concerns the representation of the players and their preferences, the rules and strategic possibilities, and the outcomes and payoffs. The solution part concerns the end results of rationally motivated activities by the players.

A same distinction is made in Shapley (1962, 1967). This distinction between a descriptive part and a solution part runs parallel with the two stages of social choice processes as distinghuised in this chapter. The first stage can be considered as the descriptive part. This part entails a description

[4]This is only true if $H(A,p) \neq \emptyset$ for all A and p. Hence, this clause in the definition of an F-SCR.

of a social choice problem, of a preference profile and of a social choice rule. The second stage of producing a social choice (a solution of the social choice problem) on the basis of information of a social preference can be considered as the solution part. In this part, a social preference has the same function as a dominance relation in a game-theoretical solution theory.

In this chapter we presented some of the basic concepts of social choice theory. In the next chapter we illustrate the working of these concepts with the aid of some well known social choice rules.

Chapter 3

Examples of Social Choice Rules

In this chapter, we discuss some examples of social choice rules, namely, the system of majority decision, the Borda rule and the plurality rule. Further, we propose an extension of the system of majority decision and briefly discuss the Dutch electoral system.

In our discussion of the system of majority decision, the Borda rule and plurality rule we will use works of Marquis de Condorcet that are, we think, hitherto unknown in the field of social choice theory. An excellent presentation of the history of the theory of social choice and, especially, the contribution of Marquis de Condorcet to this theory, can be found in the second part of Black's *The theory of committees and elections*. However, Black mainly concentrates upon *Essai sur l'Application de l'Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix* (Paris 1785). In a recent work *Condorcet's Theory of Voting* by H.P. Young (1988), the *Essai sur l'Application* is also taken as a point of departure. Young summarizes, again, the proposals made by de Condorcet in this work, especially, just as Black, the proposal with respect to the solution of the problem of cyclic majorities. Instead of Condorcet's *Essai sur l'Application*, we will concentrate in this chapter on Condorcet's less known, if not unknown works *Essai sur la constitution et les fonctions des assemblées provinciales. Première partie* (1788), *Sur la forme des élections* and the journal paper *Sur les élections* that appeared in *Journal d'instruction sociale* (Saturday, July 1st 1789). With this we hope to give a valuable completion of the historical part of Black's exellent work.

3.1 The System of Majority Decision

The system of majority decision has much been studied. In order to be able to compare this system with other ones, we give an exact definition of it:

Definition 3.1 (System of Majority Decision) *An SCR F is called the* System of Majority Decision *(henceforward SMD) := for every $x, y \in X$ and for every $p \in \Pi$,*

$$xF(p)y \Leftrightarrow \mid \{i \in N \mid xR_i^p y\} \mid \geq \mid \{i \in N \mid yR_i^p x\} \mid .$$

If F is the SMD, then we write $xM(p)y$ instead of $xF(p)y$ and we will call $M(p)$ the majority relation *for p.*

For convenience, we write M instead of $M(p)$ when p is given.

The set of $M(p)$-best elements is named after the French social scientist, mathematician and philosopher Marquis de Condorcet (1743 - 1794) who has studied SMD extensively.

Definition 3.2 (Majority Choice) *1. Let $p \in \Pi$ and $A \in \Omega$. The set*

$$Con(A, p) := \{x \in A \mid \forall y \in A : xM(p)y\}$$

is called the Condorcet set *of A and p.*

2. An alternative $x \in Con(A, p)$ is called a majority choice.

Marquis de Condorcet already knows that SMD could lead to a cyclic majority relation. The following profile illustrates this. It can be found in De Condorcet (1789: 410):

Preference Profile 3.1

$$\begin{array}{rl} 23: & xyz \\ 2: & yxz \\ 17: & yzx \\ 10: & zxy \\ 8: & zyx \end{array}$$

Using the SMD we get the cycle xMy, yMz and zMx for this situation. Hence, there is no majority choice. Since a majority choice is an M-best element by definition, this shows that SMD is not an SDF and hence, by theorem 2.3, it violates collective rationality. The emptiness of the Condorcet set for a situation is often called, after its discoverer, the *Condorcet paradox*. We will take over this terminology.

Usually, the Condorcet paradox is demonstrated with preference profiles in which each individual has a linearly ordered preference. We give a situation with individual indifference:

Preference Profile 3.2

$$\begin{array}{ll} 1: & x(yz)w \\ 1: & w(xy)z \\ 1: & z(wx)y \\ 1: & y(zw)x \end{array}$$

Again, applying SMD we get a majority relation for which the Condorcet set is empty.

The emptiness of the Condorcet set has puzzled many social choice theorists including Condorcet himself[1]. It is seen as a failure that, in some way or another, has to be circumvented. Therefore, a number of efforts have been made to construct social choice systems that stay as close as possible to the SMD. These socalled *Condorcet systems* all produce a majority choice if it exists. Otherwise, they produce still a social choice, each system in its own way. A review of the most important Condorcet systems can be found in Fishburn (1977). Here we discuss the contribution of Marquis de Condorcet to this problem. Further we propose another extension of the system of majority decision-making that is based on a solution theory presented in chapter 4 of this work. But first we investigate the importance of SMD.

[1]Another paradox related with the system of majority decision is the socalled Ostrogorski paradox. This interesting paradox has received considerable less attention in social choice theory. For a study of this paradox, consider Bezembinder and Van Acker 1985.

3.1.1 Relevance of the System of Majority Decision

A very important reason to search for systems based on SMD is that SMD satisfies a number of properties that come close to the ideal of democracy. We give a brief account of some of these conditions.

Definition 3.3 (Anonymity) *An SCR $F \in \Phi$ satisfies* anonymity := *for all $p \in \Pi$ and for any permutation σ on $N = \{1, 2, \ldots, n\}$:*

$$F(R_1, R_2, \ldots, R_n) = F(R_{\sigma(1)}, R_{\sigma(2)}, \ldots, R_{\sigma(n)}).$$

This condition is really democratic in nature. It says that it does not matter who carries the preference. What counts is someone's preference, nothing else. The power positions of the individuals are perfectly symmetric. Other personal characteristics of the individuals do not count in determining a social choice. Anonymity prevents inequal treatment of individuals. It raises a barrier against any form of discrimination.

Let σ be a permutation on X and let $\sigma(X) = \{\sigma(x) : x \in X\}$. Letting R be a preference on X, we let R^σ denote the same preference on $\sigma(X)$. That is, if, for example, $(x, y) \in R$ and if $\sigma(x) = a$ and $\sigma(y) = b$, then $(a, b) \in R^\sigma$. The next condition, again, expresses a fundamental democratic requirement.

Definition 3.4 (Neutrality) *An SCR $F \in \Phi$ satisfies* neutrality := *for all $p \in \Pi$ and for any permutation σ on X:*

$$(F(R_1, R_2, \ldots, R_n))^\sigma = F(R_1^\sigma, R_2^\sigma, \ldots, R_n^\sigma).$$

This condition says that all alternatives are or should be treated in a symmetric or equal way. The labeling of the alternatives may not influence the making of a social preference and with that the production of a social choice. So this condition prohibits, for example, any favour of the status quo. Any opinion counts, whatever its content.

Anonymity and neutrality are conditions that every democratic social choice system must satisfy. Both are informational constraints that guarantee the equal treatment of preferences and opinions. Neither preferential information nor information about opinions may be lost. All information is equally worthy. These requirements are therefore, for example, incompatible with the condition that only the opinions and preferences of the experts (those who know) count, that is, with technocracy.

Both anonymity and neutrality are standard conditions in the theory of social choice since the works of Arrow (1963), May (1952) and Sen (1970). The conjunction of these conditions is called the symmetry condition.

Definition 3.5 (Symmetry) *An SCR $F \in \Phi$ satisfies* symmetry := *F satisfies anonymity and neutrality.*

The following condition has to do with the direction of the adjustment of a social preference after a change in someone's preference has occurred.

Definition 3.6 (Positive Reponsiveness) *An SCR F satisfies* positive responsiveness := *for all $p, q \in \Pi$ and for all $x, y \in X$: if*

1. $\forall i \in N[(xP_i^p y \Rightarrow xP_i^q y) \wedge (xI_i^p y \Rightarrow xR_i^q y)]$ *and*

2. $\exists k \in N[(xI_k^p y \wedge xP_k^q y) \vee (yP_k^p x \wedge xR_k^q y)]$,

then $xF(p)y \Rightarrow (xF(q)y \wedge \neg yF(q)x)$.

This difficult definition is a formalization of the one given in Sen (1970: 72). It says that if an alternative x is raised vis-a-vis an alternative y in someone's preference and x goes down in no one's preference vis-a-vis y, then x must also be raised vis-a-vis y in the social preference. Again, this is a fundamental requirement for democracy (for the requirement of responsiveness of democracies in general cf. Pennock 1979). To see the working of this condition, consider its negation. That is, suppose x raises in someone's preference and it goes down in no one's preference and suppose it remains on the same place or it goes down in the social preference. The system yielding this social preference could be accused of some form of inertia. It cannot register changes in preference profiles and adapt its output in accordance with these changes. Presumably, this will not be very conducive for the stability and durability of the concerned system. The individuals may lose their faith and even revolt when they discover that their preference changes are not reflected in the output of that system. Typical examples of social choice systems that violate this condition are systems that involve a religious code, magic, a chance mechanism, etc.

But, as we shall see later on, also the frequently used plurality rule violates this condition.

It is not difficult to see that SMD satisfies symmetry and positive responsiveness. May (1952) proves that these conditions are also sufficient for a social choice system to be the SMD.

Theorem 3.1 (May) *An SCR $F \in \Phi$ is the system of majority decision if and only if F satisfies symmetry and positive responsiveness.*

The proof of this theorem can be found in May (1952), Sen (1970), Suzumura (1983) or Kelly (1988). This theorem neatly characterizes SMD. It shows that SMD is uniquely determined by the three conditions that are both sufficient and necessary for a social choice rule to be a democracy. Of course, SMD satisfies more standard conditions. To mention a few, it satisfies IIA, PC, UD and ND. These conditions are discussed in section 8 of the previous chapter. Note that some of these conditions are derivable from the set of necessary and sufficient conditions as mentioned in May's theorem. Also note that SMD is not an SWF, since $M(p)$ need not be transitive.

3.1.2 Condorcet's Solution to the Condorcet Paradox

According to Black (1957, Part II: 175) there are at least three possible interpretations of De Condorcet's solution to the problem of the absence of a majority choice. Black chooses the following interpretation:

> It would be most in accordance with the spirit of Condorcet's previous analysis, I think, to discard all candidates except those with the minimum number of majorities against them and then to deem the largest size of minority to be a majority, and so on, until one candidate had only actual or deemed majorities against each of the others.

Now, the system that Black describes is mathematically a beautiful one. A definition of it is given in Fishburn (1977). To be in line with the forgoing definitions, we adapt Fishburn's in the following way. Define for every $x, y \in X$ and for all $p \in \Pi$,

$$n_p(x,y) := |\ \{i \in N \mid xR_i^p y\}\ | \,.$$

Thus $n_p(x,y)$ is the number of individuals that prefer x to y.

Definition 3.7 (Black-Condorcet System) *Let $x, y \in X$, $p \in \Pi$ and*

$$c(x,p) := min_{y \in X/\{x\}} n_p(x,y).$$

An SCR $F \in \Phi$ is the Black-Condorcet system, *henceforward BCS, := for all $p \in \Pi$ and for all $x, y \in X$:*

$$xF(p)y \Leftrightarrow c(x,p) \geq c(y,p).$$

If F is the BCS, then we write $xBC(p)y$ instead of $xF(p)y$. $BC(p)$ is called the Black-Condorcet relation *for p.*

Fishburn (1977), who works within the functional approach (see chapter 2, section 9 of this work), calls the defined system the Condorcet function. However, since we think that Condorcet had something else in mind, we prefer to call this nice system the Black-Condorcet system.

Definition 3.8 (Black-Condorcet Choice) *Let $F \in \Phi$ be the BCS and consider a preference profile p and an agenda A.*

$$BC(A,p) := \{x \in A \mid \forall y \in A : c(x,p) \geq c(y,p)\}.$$

The set $BC(A,p)$ is called the Black-Condorcet set *and an $x \in BC(A,p)$ is called a* Black-Condorcet choice, *shortly, a BC choice.*

Since the Black-Condorcet relation is reflexive, complete and transitive (and hence acyclic) for every $p \in \Pi$, $BC(A,p) \neq \emptyset$ for every $A \in \Omega$ and every $p \in \Pi$. That is, for every preference profile there exists a BC choice.

To illustrate the working of this system, we develop a special method. To every preference profile p we can associate a matrix $A_p = (a_{lm})$ with $l \neq m$ and where $a_{lm} :=\mid \{i \in N \mid x_l R_i^p x_m\} \mid$. Consider now the following situation:

Preference Profile 3.3

$$\begin{array}{rl} 4: & x_1 x_2 x_3 x_4 x_5 \\ 3: & x_5 x_4 x_1 x_3 x_2 \\ 2: & x_2 x_3 x_5 x_4 x_1 \end{array}$$

The matrix for this profile is:

	x_1	x_2	x_3	x_4	x_5
x_1	–	7	7	4	4
x_2	2	–	6	6	6
x_3	2	3	–	6	6
x_4	5	3	3	–	4
x_5	5	3	3	5	–

To determine the BC choice, select from each row the lowest value and, then, select the row(s) in which this lowest value is largest The finally selected row(s) are the BC choices. In the example, the lowest values for each row are, respectively, 4, 2, 2, 3, 3. From this, 4 is the largest and, hence x_1 is the BC choice. This working procedure clearly shows that the BCS is a maximin method (cf. Fishburn 1977). It selects the maximum of the minimum row value. This method also shows why the BCS always selects a majority choice if it exists. A majority choice will always have the maximin value.

The question is whether Condorcet had really this ingenious system in mind. In his journal paper of 1793 Condorcet writes:

> "Dans le cas d'une élection entre trois candidats, il est possible que les trois jugemens de la majorité sur ces concurrents comparés deux à deux, ne puissent subsister ensemble, quoique le résultat des jugements de chaque votant ne renferme aucune contradiction. .... Alors il faut abandonner la proposition qui a une moindre majorité, et s'en tenir aux deux autres.

With respect to the general case, Condorcet makes no difference: "Si ces jugemens ne peuvent subsister ensemble, on abandonneroit ceux qui ont obtenu la majorité la plus foible" (Condorcet 1789: 482). In his other main work, Condorcet gives an example of how to proceed in the case of a paradox.

Preference Profile 3.4 (Condorcet 1789)

23 :	*Pierre*	*Paul*	*Jacques*
2 :	*Paul*	*Pierre*	*Jacques*
17 :	*Paul*	*Jacques*	*Pierre*
10 :	*Jacques*	*Pierre*	*Paul*
8 :	*Jacques*	*Paul*	*Pierre*

According to this profile, Pierre has a 33 to 27 majority over Paul, Jacques has a 35 to 25 majority over Pierre and Paul has a 42 to 18 majority over Jacques. Hence there is no majority candidate. With respect to this result Condorcet (1788: 411) writes:

> " En effet, si nous en examinons le résultat de plus près, nous trouverons que, puisqu'il faut rejeter une proposition adoptée par la pluralité, il est plus naturel d'abandonner celle qui a la moindre pluralité; Nous rejeterons donc ici la première, et nous aurons un résultat en faveur de Paul."

This picture gives more room for interpretation. Since Pierre has the smallest majority vis-a-vis Paul, this majority pair must be rejected. What remains are that Jacques has a majority over Pierre and that Paul has a majority over Jacques. From these remaining pairs we may, apparently, deduce the fact that Paul is the ultimate majority candidate. In our opinion, what Condorcet does here is eliminating the pairs with the smallest majorities in the case of cycles and taking subsequently the transitive closure over the remaining pairs. If the result is xMy, yMz and zMx, and zMx has the smallest majority, then we may abandon zMx and take the transitive closure over xMy and yMz. This leads to the choice of x. If not, then it is difficult to explain why Paul should be chosen. But if this is true, then Condorcet had something else in mind as what Black asserts. To see this, consider the following profile:

Preference Profile 3.5

2 : *adbc*
2 : *abdc*
2 : *cbad*
1 : *bdca*
3 : *cbda*
1 : *dcba*

For this situation we have: a 6 to 5 majority of a over d; a 7 to 4 majority of b over a, a 8 to 3 majority of b over d; a 7 to 4 majority of c over a; a 6 to 5 majority of c over b and a 6 to 5 majority of d over c. According to the BC system, either b or c is the social choice. However, if we eliminate the pairs with the smallest majorities, namely aMd, cMb and dMc, and we take, subsequently, the transitive closure over the rest of the majority relation, then we get only b as the social choice.

However, we hasten to say that our opinion is as speculative as Black's one. Also in his other works, Condorcet gives too little information to be on sure grounds. The only thing we can say without too much risk is that Condorcet remains fragmentary and brief with respect to the problem of the absence of a majority choice (see Black 1957: 176).

3.1.3 Theory of Stable Majority Solutions

Most Condorcet systems are constructed on the basis of two design principles. Firstly, they must satisfy collective rationality and, secondly, they must produce a majority choice if it exists. However, if the first principle is used, then Arrow's theorem predicts that one of the conditions of UD, IIA, PC or ND will be violated[2]. For example, the BCS as discussed above, violates IIA. To see this consider profile 3 as treated in the previous subsection. According to the associated matrix, $c(x_3) = 2$ and $c(x_4) = 3$. Hence x_4BCx_3. Now consider the following situation:

Preference Profile 3.6

$$\begin{array}{rl} 4: & x_3x_2x_4x_1x_5 \\ 3: & x_2x_5x_1x_4x_3 \\ 2: & x_3x_4x_2x_5x_1 \end{array}$$

The position of x_3 vis-a-vis the position of x_4 has remained the same in each preference in comparison with profile 3. Hence, if IIA is valid, we may expect x_4BCx_3. However, look at the associated matrix of the

[2]Consider section 9 of the previous chapter for the definitions of the mentioned conditions.

changed situation:

	x_1	x_2	x_3	x_4	x_5
x_1	–	0	3	3	4
x_2	9	–	3	7	9
x_3	6	6	–	6	6
x_4	6	2	3	–	6
x_5	5	0	3	3	–

Now: $c(x_3) = 6$ and $c(x_4) = 2$. Hence $x_3 BC x_4$. This violates the condition of IIA. This is the price Arrow's theorem tells us to pay.

However, another design policy is possible. The basic principle in this policy is to take the SMD and the majority relations it produces, as the point of departure. We keep, so to say, the system untouched. With this we uphold the nice properties of this system. What we do instead is designing a method that shows how to deal with cyclic majority relations. This method must operate in such a way that it selects a majority choice if it exists. Of course, this implies that every condition of collective rationality as defined in section 7 of chapter 2 must be dropped. This is the price to be paid in this design policy. We present such a method. This method is an application of the theory of generalized stable sets as presented in the next chapter (also cf. Van Deemen 1988, 1990b).

Let A be an agenda and let $M(p)$ be the majority relation for the preference profile p. Let $\bar{a}M(p)$ denote[3] the asymmetric part of $M(p)$, that is, $x\bar{a}M(p)y$ if $xM(p)y$ but not $yM(p)x$. Suppose $x, x_1, \ldots, x_n, y \in A$. A *majority path* from x to y, is a sequence

$$x\bar{a}M(p)x_1, x_1\bar{a}M(p)x_2, \ldots, x_n\bar{a}M(p)y.$$

An alternative $x \in A$ is said to be *majority-dominant* vis-a-vis an alternative $y \in A$ if there is a majority path from x to y. This will be denoted by xD^My.

Definition 3.9 *Let A be an agenda and let $M(p)$ be the majority relation for $p \in \Pi$. A set $\omega(A, M(p)) \subseteq A$ is a* stable majority solution :=

1. for no $x, y \in \omega(A, M(p))$: xD^My,

[3]Storcken (1989: 36) uses the symbol $\bar{a}$ to denote the asymmetric part of a relation. Thus, $\bar{a}R$ denotes the asymmetric part of the relation R. We take over Storcken's notation here.

2. *for every* $y \in A - \omega(A, M(p))$ *there is an* $x \in \omega(A, M(p))$ *such that* xD^My.

The first condition expresses the property of internal stability of $\omega(A, M(p))$[4]. It says that no alternative in a stable majority solution dominates another alternative in the stable majority solution. For no element in a stable majority solution there is a path that goes to another element in that solution. The second condition expresses the property of external stability of a majority solution. It says that for every social state y not in $\omega(A, M(p))$ there is a social state x in $\omega(A, M(p))$ such that there is a majority path that goes from x to y. However, it is not precluded that there is a path going from an $y \in A - \omega(A, M(p))$ to an $x \in \omega(A, M(p))$. Inside a stable majority solution there is no majority dominance. The next theorem shows that there will be a stable majority solution when the asymmetric part of $M(p)$ is not empty. This theorem is a consequence of the existence theorem 4.16 (with $P = \bar{a}M(p)$) in the next chapter.

Theorem 3.2 *Let A be an agenda. For every* $p \in \Pi$, *if* $\bar{a}M(p)$ *restricted to A is nonempty, then there exists a nonempty stable majority solution of A.*

Hence, this method can deal with any possible cyclic majority relation. In general there will be several stable majority solutions for a profile. If there are majority choices, then it can be proven that a stable majority solution will contain these alternatives.

Theorem 3.3 *If* $Con(A, M(p)) \neq \emptyset$, *then* $Con(A, M(p)) \subseteq \omega(A, M(p))$.

This theorem is a consequence of theorem 4.17 as presented in the next chapter.

How to find the stable majority solutions of a preference profile? This can be done by using theorem 4.15 in the next chapter. The working of this theorem requires some knowledge of terms and techniques adopted from digraph theory. Consider chapter 4, section 2 for these terms and techniques. According to theorem 4.15, first the contraction of the majority relation of the profile under scrutiny must be formed. Then the theorem

[4]Cf. Von Neumann and Morgenstern 1953 for the notion of internal stable sets; also see chapter 4 of this study.

says to collect from each maximal vertex set of this contraction one and only one alternative. The set of alternatives formed in this way is a stable majority solution according to theorem 4.15.

To illustrate the working of the method, consider the following strict majority relation:
$x_1 \bar{a} M x_2, x_2 \bar{a} M x_3, x_3 \bar{a} M x_1, x_1 \bar{a} M x_4.$
$x_4 \bar{a} M x_5, x_5 \bar{a} M x_6, x_6 \bar{a} M x_4,$
$x_7 \bar{a} M x_4, x_7 \bar{a} M x_8, x_8 \bar{a} M x_9, x_9 \bar{a} M x_7, x_9 \bar{a} M x_5.$
The vertex sets of the strong components are:
$S_1 = \{x_1, x_2, x_3\},$
$S_2 = \{x_4, x_5, x_6\},$
$S_3 = \{x_7, x_8, x_9\}.$
The sets we need, are S_1 and S_3. Thus, any $\{x, y\}$ with $x \in S_1$ and $y \in S_3$ is a stable majority solution. To experiment, let us take $x_5 \bar{a} M x_9$ instead of $x_9 \bar{a} M x_5$. The vertex sets of the strong components now are:
$S_1 = \{x_1, x_2, x_3\},$
$S_2 = \{x_4, x_5, x_6, x_7, x_8, x_9\}.$
The set needed to determine a stable majority solution is now S_1. Thus, any $\{x\}$ with $x \in S_1$ is a stable majority solution.

To give another illustration, consider the famous voting paradox:

Preference Profile 3.7

$$\begin{array}{rl} 1: & xyz \\ 1: & zxy \\ 1: & yzx \end{array}$$

This leads to the majority relation xMy, yMz, zMx. It is easily verified that the algorithm leads to the stable majority sets $\{x\}$, $\{y\}$ and $\{z\}$.

Other solutions to the Condorcet paradox that are based on the same design principles as the theory of stable majority solutions are possible. The most important is the theory of generalized optimal choice as presented in Kalai and Schmeidler (1976), Schwartz (1986) and Shenoy (1979, 1980). This theory has several other names: theory of admissable solutions (Kalai and Schmeidler 1976), Generalized Optimal-CHoice Axiom or GOCHA (Schwartz 1986), the theory of dynamic solutions (Shenoy 1979, 1980). It will be studied extensively in the next chapter. There we also present the main differences of this theory with our theory of generalized stable solutions on which the theory of stable majority solutions is based.

Theories like the theory of stable majority solutions show that there may be reasonable outcomes even in the absence of a social optimum. Even if best social choices are absent, there might be social choices that are good enough to be acceptable as a possible outcome of a social choice process. Sets of such acceptable social states are called 'standards of behavior' by Von Neumann and Morgenstern (1953, Ch.I.4). In the next chapter we will discuss these standards of behavior more deeply.

3.2 The Plurality Rule

Sofar we have studied the system of majority decision. Another important rule is the plurality rule.

Definition 3.10 (Plurality Rule) *Let $x \in X$, $p \in \Pi$ and define*

$$t(x,p) := |\, \{i \in N \mid \neg\exists y \in X : yP_i^p x\} \,| \,.$$

An SCR $F \in \Phi$ is called the Plurality Rule *(henceforeward PR) := for every $x, y \in X$ and for every $p \in \Pi$:*

$$xF(p)y \Leftrightarrow t(x,p) \geq t(y,p).$$

If F is the PR, then we write $xPl(p)y$ instead of $xF(p)y$. $Pl(p)$ will be called the plurality relation for p.

According to this definition, an alternative x is socially preferred to y if the number of individuals that put x in the first place is at least as great as the number of individuals that put y in the first place.

Definition 3.11 (Plurality Choice) *Let $F \in \Phi$ be the PR and let $p \in \Pi$. Then, if $A \in \Omega$,*

$$Pl(A,p) := \{x \in A \mid \forall y \in A : t(x,p) \geq t(y,p)\}.$$

The set $Pl(A,p)$ is called the plurality choice set *for A and p. An $x \in Pl(A,p)$ is called a* plurality choice.

Hence, a plurality choice for an agenda is an alternative on that agenda for which there is no other alternative with more first places in the individual preferences. It is not difficult to see that $Pl(A, p)$ is nonempty for every $A \in \Omega$ and every $p \in \Pi$.

The PR is frequently used in practice. Variations of it can be found in many electoral systems. Despite its frequent use, this system has some serious shortcomings. First, it violates positive responsiveness. Therefore, this system cannot adjust a social preference and with that a social choice to comply with changes of individual preferences in society. This conflicts with any notion of democracy. Second, a plurality choice is not necessarily a majority choice. This was discovered, indeed, by Condorcet. To illustrate, Condorcet gives the following preference profile:

Preference Profile 3.8 (Condorcet 1789: 335)

$$\begin{array}{ll} 10: & xzy \\ 8: & yzx \\ 7: & zyx \end{array}$$

The plurality rule selects x as the social choice. However, since yMx, zMy and zMx, z is the majority choice. Condorcet concludes:

> On voit donc par cet example, comment, dans la méthode ordinaire d'élire, le jugement des électeurs n'est pas complet, et comment, par cette raison, le résultat d'une élection faite sous cette forme peut exprimer un vœréellement contraire à celui de la pluralité.

With "la pluralité" Condorcet means majority. Our terminology is more standard nowadays (see for example Fishburn 1973, Gärdenfors 1973, Kelly 1988). In those days of Condorcet, the PR was already so much in used that he calls it *la méthode ordinaire*.

Condorcet goes further. He gives an example in which the majority choice gets a zero plurality score.

Preference Profile 3.9 (Condorcet 1789: 403-4)

18 :	*Pierre*	*Jean*	*Jacques*	*Paul*
5 :	*Pierre*	*Jean*	*Paul*	*Jacques*
16 :	*Paul*	*Jean*	*Jacques*	*Pierre*
3 :	*Paul*	*Jean*	*Pierre*	*Jacques*
13 :	*Jacques*	*Jean*	*Paul*	*Pierre*
5 :	*Jacques*	*Jean*	*Pierre*	*Paul*

For this situation, Pierre is the plurality choice. However, Jean has a majority vis-a-vis every other candidate. But, clearly, Jean has plurality score 0. Because of this, Condorcet did not have much confidence in the plurality rule. Therefore he started to investigate an alternative system called the Borda rule.

3.3 The Borda Rule

The Borda Rule is proposed by the French academician Jean-Charles de Borda (1733-1799). Borda's original formulation of the rule was intended for individual preferences that are linear orders. In essence, his method works as follows (cf. Black 1957: 157-8):

1. If there are m alternatives, then assign the number m to the first place alternative in a linear order, the number $m-1$ to the second place alternative in the same linear order, the number $m-2$ to the next alternative in that order, etc. The last alternative (the least preferred) receives the number 1.

2. The total score of an alternative in a preference profile is the sum of the numbers as assigned to that alternative in the several individual preferences.

3. Take as a social choice the alternative with a total score that is at least as great as the total score of every other alternative.

In order to be able to deal with ties in individual preferences, this method must be adjusted. A logical adjustment is to assign the numbers to the alternatives such that the assigned numbers agree with the assignment for some linear order as described previously. The final number assigned to an alternative x in a weak order is then the average of the numbers of the alternatives that are indifferent to x (cf. Fishburn 1973: 164). Thus, in the weak order $x(yz)w$, x gets a score of 4, y and z get both $(3+2)/2 = 2.5$ and w gets 1. By summing the scores of x in each weak ordering in a preference profile p, we obtain the socalled *Borda-score* of x in profile p. Since the Borda-score of each alternative can be determined in this way, it is possible to compare mutually these Borda-scores. We declare x socially preferred to y if x has a larger Borda-score than y. Note that the Borda social preference obtained in this way is reflexive, complete and transitive. For example, if we have

Preference Profile 3.10

$$\begin{aligned} i &: \; x(yz)w \\ j &: \; w(zx)y \end{aligned}$$

then the Borda-scores of x, y, z, w equal, respectively, $4+2.5 = 6.5$, $2.5+1 = 3.5$, $2.5+2.5 = 5$ and $1+4 = 5$. So the Borda social preference for this profile is $x(zw)y$. Let us call the procedure that leads to this result the *classical Borda-rule.*

Let $r_i(x,p)$ denote the score of x in the weak order R_i^p obtained by means of the above sketched procedure. Let $b_i(x,p) := ar_i(x,p) - c$ be a linear transformation for each x with $a > 0$, $c \geq 0$ and a, c constant. Since
$b_i(x,p) \geq b_i(y,p) \Leftrightarrow (ar_i(x,p) - c) \geq (ar_i(y,p) - c)$,, we have

$$b_i(x,p) \geq b_i(y,p) \Leftrightarrow r_i(x,p) \geq r_i(y,p).$$

Taking $\sum_{i \in N} b_i(x,p)$ for every x, it is easily verified that the Borda social preference as obtained for a preference profile by the classical procedure is not changed if $b_i(x,p) = ar_i(x,p) - c$ for every x and every i. To illustrate this, consider the same profile as above. Then we obtain for $c = 5$ and $a = 2$, $b_i(x,p) = 8 - 5 = 3$, $b_i(y,p) = 5 - 5 = 0$, $b_i(z,p) = 5 - 5 = 0$ and $b_i(w,p) = 2 - 5 = -3$. For j, we obtain $b_j(x,p) = 5 - 5 = 0$, $b_j(y,p) = 2 - 5 = -3$, $b_j(z,p) = 5 - 5 = 0$ and $b_j(w,p) = 3$. So, the Borda score of x equals $3 + 0 = 3$, of y equals $0 + -3 = -3$, of z equals $0 + 0 = 0$, and of w equals $-3 + 3 = 0$. This yields the Borda social preference $x(zw)y$ which is the same as obtained by the classical Borda-rule. This shows that the following definition of the Borda-rule is equivalent to the classical Borda-procedure. A similar definition is presented in Fishburn (1973) and Gärdenfors (1973).

Definition 3.12 (Borda Rule) *Let $x \in X$ and $p \in \Pi$. We define:*

1. $b_i(x,p) := \mid \{y \in X \mid xR_i^p y\} \mid - \mid \{y \in X \mid yR_i^p x\} \mid$, *and*

2. $b(x,p) := \sum_{i \in N} b_i(x,p)$.

An SCR $F \in \Phi$ is called the Borda Rule *(in short BR) := for all $p \in \Pi$ and all $x, y \in X$:*

$$xF(p)y \Leftrightarrow b(x,p) \geq b(y,p).$$

If F is the BR, then we write $xB(p)y$ instead of $xF(p)y$. $B(p)$ will be called the Borda relation *of p.*

Thus, in determining the score of an alternative x, we not only count the number of alternatives that succeed x in a preference, but also the number of alternatives that precede x in that preference. The total score of an alternative is, just as in the original procedure, the sum of the scores of that alternative in the individual preferences.

Definition 3.13 *Let $F \in \Phi$ be the BR and let $p \in \Pi$. Then for every $A \in \Omega$,*

$$Bo(A,p) := \{x \in A \mid \forall y \in A : xB(p)y\}.$$

To illustrate this system, consider the following profile (Sen 1970: 39):

Preference Profile 3.11

$$\begin{array}{rl} 1: & xyz \\ 2: & zxy \end{array}$$

We have one individual, say i, with xyz. Hence $b_i(x) = 2 - 0 = 2$, $b_i(y) = 1 - 1 = 0$ and $b_i(z) = 0 - 2 = -2$. We have two individuals, say j and k, such that zxy. Then $b_j(z) = b_k(z) = 2 - 0 = 2$, $b_j(x) = b_k(x) = 1 - 1 = 0$ and $b_j(y) = b_k(y) = 0 - 2 = -2$. Adding the scores of the several alternatives gives:
$b(x) = b_i(x) + b_j(x) + b_k(x) = 2.$
$b(y) = b_i(y) + b_j(y) + b_k(y) = -4.$

$b(z) = b_i(z) + b_j(z) + b_k(z) = 2.$
Note that Borda's original method would give x a score of 7, y a score of 4 and z a score of 7. Thus, the ordering according to this method is the same as the ordering according to the method as proposed in the definition.

Condorcet, who called Borda "un géomètre célebre", has probably discovered the fact that SMD, PR and BR could yield different outcomes for the same preference profile. This is a remarkable discovery. It is the precursor of the general insight that different social choice rules may lead to different outcomes for the same preference profiles. We discuss some of the profiles given by De Condorcet that illustrate his discovery.

For the following situation the Borda choice and the majority choice agree but the plurality choice deviates. We use x, y and z instead of Condorcet's $Pierre$, $Jacques$ and $Paul$.

Preference Profile 3.12 (De Condorcet 1789: 402-3)

$$\begin{array}{ll} 18: & xyz \\ 5: & xzy \\ 16: & zyx \\ 3: & zxy \\ 13: & yzx \\ 5: & yxz \end{array}$$

The Borda scores of the alternatives are calculated with the aid of the original method. Of course, only this method was known to Condorcet. We get $b(x) = 23 \times 3 + 8 \times 2 + 29 \times 1 = 114$. In the same way we get $b(z) = 116$ and $b(y) = 130$ (cf. Condorcet 1788: 405). Hence, y is the Borda choice. SMD gives: zMx, yMz and yMx. Hence, y is also the majority choice. But x is the plurality choice. However, this profile did not increase Condorcet's confidence in the Borda system. He knew that the system could give a result that is not in agreement with the 'will of the majority'(le voeu de la pluralité). To illustrate this, he gives, among others, the following profile:

Preference Profile 3.13 (Condorcet (1789))

$$\begin{array}{ll} 9: & xyz \\ 3: & xzy \\ 4: & yxz \\ 6: & yzx \\ 4: & zxy \\ 4: & zyx \end{array}$$

The majority choice is x. The plurality system also gives x as the social choice. However, the Borda scores are:
$b(x) = 12 \times 3 + 8 \times 2 + 10 \times 1 = 62$.
$b(y) = 10 \times 3 + 13 \times 2 + 7 \times 1 = 63$.
$b(z) = 8 \times 3 + 9 \times 2 + 13 \times 1 = 55$.
Hence, y is the Borda choice.

Condorcet also presents a preference profile for which the Borda choice and the plurality choice agree but for which the majority choice deviates.

Preference Profile 3.14 (Condorcet 1789: 405)

$$\begin{array}{ll} 30: & xyz \\ 1: & xzy \\ 29: & yxz \\ 10: & yzx \\ 10: & zxy \\ 1: & zyx \end{array}$$

SMD gives: xMy, xMz and yMz. Hence x is the majority choice. The plurality choice is y. The Borda scores are:
$b(x) = 31 \times 3 + 39 \times 2 + 11 \times 1 = 182$.
$b(y) = 39 \times 3 + 31 \times 2 + 11 \times 1 = 190$.
$b(z) = 11 \times 3 + 11 \times 2 + 59 \times 1 = 114$.
Hence, y is the Borda choice.

It is very interesting to see why Condorcet rejects the Borda rule. It is not only because this rule can fail in yielding a majority choice. A very surprising argument is given in the following passage that is concerned with profile 14. In this passage, *Pierre* refers to x in the profile, *Paul* to y and *Jacques* to z (cf. Condorcet 1789: 408-409):

"Mais, dira-t-on, comment se peut-il que, tout paraisant d'ailleurs égal entre Pierre and Paul, si ce n'est que Pierre a eu trente-une fois la première place, et trente-neuf fois la seconde, tandis que Paul a eu trente-neuf fois la première, et trente-une la seconde, il n'en résulte pas évidemment un advantage en faveur de Paul? Le voici. Parmi les trente-neuf voix qui plaçaient Pierre à la seconde place, il y en avait dix qui le préféraient à Paul, et vingt-neuf qui le préféraient à Jacques. Mais, parmi les trente-une voix qui plaçaient Paul à cette seconde place il n'y en avait qu'une qui le préférât à Pierre. Or l'on a confondu, dans cette méthode d' évaluer les suffrages, les voix qui donaient la préférence à Pierre sur Paul , ou réciproquement avec celles qui donnaient la préférence à l'un ou à l'autre sur Jacques; on les a fait entrer de même dans les jugement qu'on voulait porter entre Pierre et Paul, et il a dû en résulter une erreur, puisque l'on faisait entrer dans ce jugement un élément qui ne devait pas y entrer, c'est- à-dire, la préférence donnée sur Jacques à Pierre ou à Paul."

What Condorcet shows here is that preferences between other alternatives as x and y are relevant in deciding the Borda scores of x and y. In other words, what he proves in this passage is the fact that the Borda system violates the IIA condition. In his view this is even the main cause of the failure of the Borda rule (Condorcet 1789: 409). He concludes:

"La méthode ordinaire (*the plurality rule, AvD*) trompe parce qu'on y fait abstraction de jugements qui devraient être comptés; la nouvelle méthode (*the Borda rule, AvD*), trompe parce qu'on a égard à des jugements qui ne devraient pas être comptés."

The Borda rule is a scoring or positional system. The position of the alternatives in each preference counts in determining the social preference. Young (1974) gives necessary and sufficient conditions for a social choice rule to be the Borda rule. The main drawbacks of this rule already has been given by Marquis de Condorcet. Firstly, it may fail in yielding a majority choice when such a choice exists. This is a serious drawback. It is very difficult to find reasons in a democracy that legitimate social choices which are not accepted by majorities. Secondly, the BR violates the condition of

IIA. This was, as we have tried to show, already discovered by Condorcet. This violation introduces an element of arbitrariness in yielding social preferences and social choices.

3.4 The Dutch Electoral System

In this section we show that the Dutch electoral system is not in agreement with the idea of majority as presented and studied in section 1 of this chapter. Especially, we shall prove that in the Netherlands it is possible that a political party that has a majority over every other political party may get the smallest number of seats in parliament. Conversely, it is shown that a party which has no majority over any of the other parties may get the largest number of seats.

The Dutch electoral system is based on the plurality rule. Each voter gives only one vote for the candidate of the party (s)he prefers most. The division of the seats in parliament is as much as possible in proportion with the number of votes a party gets. In this section we shall avoid mathematical formulations of this system. For a careful mathematical study of apportionment systems in general, consider Balinski and Young (1982, Appendix A).

Consider the following division of seats in the second chambre of the Dutch parliamentary system: GL (Green left): 6 seats; PvdA (Social Democrats): 49 seats; D66 (Left Liberals): 12 seats; CDA (Christian Democrats): 54 seats; VVD (Conservative Liberals): 22 seats; SR (Small Right parties): 7. This is the division of seats since the election of 6 September 1989. Suppose empirical research has revealed that this division of seats is the result of the following distribution of the voters' preference orders of the political parties:

Preference Profile 3.15

$$\begin{array}{lllllll}
36\%: & CDA & GL & SR & D66 & VVD & PvdA \\
33\%: & PvdA & GL & SR & D66 & VVD & CDA \\
15\%: & VVD & GL & SR & D66 & PvdA & CDA \\
8\%: & D66 & GL & SR & VVD & CDA & PvdA \\
5\%: & SR & GL & D66 & VVD & PvdA & CDA \\
4\%: & GL & SR & D66 & VVD & PvdA & CDA
\end{array}$$

We have rounded the percentages. As agreed[5], let $x\bar{a}My$ denote that party x has a strict majority over party y. Let us put the size of the majority between parentheses. The distribution matrix leads to the following result:

1. GL $\bar{a}M$ CDA (65-36); GL $\bar{a}M$ PvdA (68-33); GL $\bar{a}M$ VVD (86-15); GL $\bar{a}M$ D66 (93-8); GL $\bar{a}M$ SR (96-5).

2. SR $\bar{a}M$ CDA (65-36); SR $\bar{a}M$ PvdA (68-33); SR $\bar{a}M$ VVD (86-15); SR $\bar{a}M$ D66 (93-8).

3. D66 $\bar{a}M$ CDA (65-36); D66 $\bar{a}M$ PvdA (68-33); D66 $\bar{a}M$ VVD (86-15).

4. VVD $\bar{a}M$ CDA (65-36); VVD $\bar{a}M$ PvdA (68-33).

5. PvdA $\bar{a}M$ CDA (57-44).

Hence, the majority relation is
GL $\bar{a}M$ SR $\bar{a}M$ D66 $\bar{a}M$ VVD $\bar{a}M$ PvdA $\bar{a}M$ CDA. This is exactly the converse of the social preference relation as expressed by the original division of seats yielded by the Dutch system. That is,

1. GL has a strict majority over every other party. However, it has the smallest number of seats;

2. SR has a strict majority over every other party except GL. However, SR has less seats than every other party over which SR has a strict majority. SR has more seats than GL that has a strict majority over SR.

3. D66 has a strict majority over VVD, PvdA and CDA. However, D66 has less seats than each of these parties. However, it has more seats than the parties that have a strict majority over D66.

4. VVD has a strict majority over PvdA and CDA. However, VVD has less seats than each of these two parties. In contrast, VVD has more seats than the parties that have a strict majority over VVD.

[5] Cf. section 3.1 for the definition of the system of majority decision and the notation M; cf. section 3.1.4 for the notation $\bar{a}M$.

5. PvdA has a strict majority over CDA. However, CDA has more seats.

6. Every party has a strict majority over CDA. However, CDA has the largest number of seats.

This is a disturbing result. It shows that in Dutch politics a minority may rule over a majority. Anyway, it shows that the Dutch electoral system is not in agreement with the idea of majority decision-making as studied in section 1 of this chapter and as has been studied for centuries now in the theory of social choice[6].

Of course, preference profile 3.15 looks exceptional. However, preference profiles in which a party x that is preferred by a strict majority to party y still gets less seats than y are easy to construct. Apparently, in Dutch politics, the chance that a party x with a strict majority over party y gets a smaller number of seats than party y must not be underestimated. Probability calculations may show that such 'misrepresentations' are rather the rule than the exception.

Moreover, it may be argued that it is better to play on safety. A chance is a chance, whatever its size. What really counts is that the Dutch system does not preclude the possibility. It is therefore better to design a system that is based on an idea that precludes such possibilities. In order to meet these critical notes, the following design principle may be relevant. Since this design principle preserves the idea of majority decision making as studied in section 1 of this chapter, we call it the *majority principle for electoral systems.*

Design Principle 1 (Majority Principle for Electoral Systems) *Party x must obtain at least as much seats as party y if and only if a majority of the electorate prefers x to y.*

Clearly, this design principle might completely disturb the contemporary Dutch political scene. For example, this principle would not give the present-day government coalition $\{CDA, PvdA\}$ a majority when preference profile 15 would be actual.

[6]Cf. Black 1957, esp. Part II; also cf. section 1.3 for the contribution of the French social scientist and mathematician Marquis de Condorcet (1743 - 1794) who started the systematical study of majority decision-making.

The proposed design principle could lead to cyclic majorities. This is, as we have tried to show in this chapter, an ever returning problem since its discovery by Marquis De Condorcet in the eighteenth century. However, with the theory of stable majority solutions as presented in section 3.1.3, it is possible to deal with majority cycles. In the next chapter we turn to the solution part of social choice processes in general. We shall see in that chapter that the theory of stable majorities is but a version of a general theory of solutions for social choice problems. There, we also shall discover other possibilities to deal with cycles.

Chapter 4

General Solutions for Social Choice Problems

4.1 Introduction

In this chapter we study the solution part of social choice processes in a more extensive way. The aim is to solve problem 3 as formulated in the introductory chapter.

As we have seen in the previous chapter, the fact that cycles may occur among possible social choices is already discovered and studied by Marquis de Condorcet in the eightteenth century (Condorcet 1785, 1788, 1789). However, Condorcet, who speaks of 'contradiction' instead of 'cycle', is mainly preoccupied with the system of majority decision.

Arrow's impossibility theorem as presented and studied in chapter 2 not only deals with the system of majority decision, but with any social choice rule whatsoever. With respect to cycles Arrow is rigorous. He simply forbids cycles to occur by requiring that a social choice rule should always produce a transitive social preference. For Arrow this requirement is equivalent to the condition of collective rationality. However, the requirement of transitivity of social preferences has a clear price. Arrow's impossibility theorem (Arrow 1963, 1967, 1977; Kelly 1977, 1989; Sen 1970, 1977, 1986; Schwartz 1986, Suzumura 1983) shows that transitivity of social preference is inconsistent with the condition of nondictatorship given some other reasonable conditions of a social choice rule (cf. chapter

2).

A way to kill Arrow's dragon is to weaken the condition of transitive collective rationality. The first step coming then to mind is to require that only the asymmetric part of a social preference should be transitive (quasi-transitivity). Unfortunately, this leads to the existence of an oligarchy, that is, a coalition that has dictatorial decision-making power (cf. Gibbard 1969, Sen 1970, Suzumura 1983, Schwartz 1986). Also see chapter 5.4 for the socalled oligarchy theorem. A further-going step is to require that a social preference should be acyclic. However, also this step preserves the essence of Arrow's theorem. Acyclicity leads to the existence of a vetoer given some other reasonable conditions (Mas Collel 1972, Blau and Deb 1977, Blair and Pollak 1982, Schwartz 1986, Sen 1977, 1986). Also see chapter 5.4. If Arrow is right in concluding his famous work that "[c]ollective rationality in the social choice mechanism is ... an important attribute of a genuinely democratic system capable of full adaptation to varying environments" (Arrow 1963: 120), then, apparently, there must be a persistent and undemocratic concentration of decision making power in such a democratic system.

The following step is to give up even acyclicity of social preferences. However, this has an unpleasant implication. With this we must give up the requirement that social choices be best[1], since in the case of cycles a best social choice need not exist. The crucial question is then what to count as a social choice. What sets of alternatives may be considered as reasonable solutions of social choice problems when best choices are absent? To answer this question, alternative theories of collective rationality are needed that are able to deal systematically with cyclic social preferences. The main task of these solution theories is to specify what sets of social states may be seen as reasonable solutions of social choice problems when best social states do not exist. A desirable feature would be that these theories point to best social choices when these exist. They then are in line with the classical notion of collective rationality.

In the course of time a number of such solution theories have been proposed. We mention, among others, the concept of uncovered set (Miller 1980, 1983, McKelvey 1986, Cox 1987) and the equivalent notion of Fishburn set (Fishburn 1977); the notion of a Banks set which is a subset

[1]Cf. definition 2.4 in section 4 of chapter 2.

of the uncovered set (Banks 1985); the theory of GOCHA (Generalized Optimal CHoice Axiom) of Schwartz (1971, 1986) and the equivalent theories of dynamic solutions (Shenoy 1979, 1990) and admissible sets (Kalai and Schmeidler 1976, Kalai et.al. 1977); the related theory of GETCHA (Generalized Top-Choice Assumption) of Schwartz (1986); the theory of SOCON (SOlution CONdition) of Schwartz (1986); the notion of Copeland winner (Grofman (1987) etc. We do not claim this list to be complete. Most of these solution theories are studied for multi-dimensional choice spaces and with means of geometric methods. A rather informal review of most of these solution theories is given in Krehbiel (1988). Also see Schofield et.al. (1988). In both works the solution theories are discussed in relation with spatial models of social choice.

In this chapter we propose a theory that has a clear connection with the Von Neumann-Morgenstern theory of stable sets (Von Neumann and Morgenstern 1953). This latter theory has a long tradition in game theory. See Lucas (1977) for a review of the basic concepts of this theory and for an excellent discussion of the state of knowledge at that time concerning the existence of stable sets within game theory. The main shortcoming of Von Neumann-Morgenstern stable sets is that it can fail to yield a solution in the case of odd cycles. Our theory is designed to meet this shortcoming. It is able to yield a solution for any possible cyclic social preference. Another pleasant feature is that it is 'core-inclusive'. That is, if the set of best social choices is not empty, then the solution set as proposed will contain this set. Since our theory is a generalization of the Von Neumann-Morgenstern stable set theory, we call it the theory of *generalized stable sets*.

In this chapter, we will also investigate the relationship of the theory of generalized stable sets with the theory of generalized optimal choices of Schwartz (1972, 1986), Shenoy (1978, 1979, 1980), Kalai et.al. (1976), and Kalai and Schmeidler (1977). There are two reasons for this investigation. First, the theory of generalized optimal choices can be interpreted as a generalization of the game-theoretical core concept. Hence both the theory of generalized stable sets and the theory of generalized optimal choice are generalizations of solution theories that have a long and well established tradition in game theory. This also gives the second reason. In game theory, it is well known that the core of a game is contained in each Von Neumann-Morgenstern stable set of that game. It is worthwhile to discover whether this connection also holds for the generalized versions of the core and stable sets as presented in this chapter.

We use digraph theoretical methods instead of geometric ones. We think that the digraph theoretical orientation leads to results that might not be discovered by using geometric methods. In section 2 we discuss some elementary notions and techniques of digraph theory (Berge 1985, Behzad et.al. 1979, Harary et.al. 1965). The main technique introduced here is the contraction of a relation. This technique will be used to prove some of the results with respect to generalized stable sets. Further, it will play an important role in a theorem with respect to generalized optimal choice sets (the contraction theorem). In section 3 we present the theory of maximal choices. See for a definition of maximal choices chapter 2, section 4. We present this theory as a preparation for the theory of generalized optimal choices. As we shall see, also the theory of maximal choices can deal with cycles. However, these cycles are of a rather particular kind. In section 4 the theory of generalized optimal choices will be presented. In addition to the work of Schwartz (1986), Shenoy (1979, 1980), Kalai and Schmeidler (1976) and Kalai et.al. (1977), we will present some new results among which the contraction theorem is the most important. In section 5 we deal with the theory of Von Neumann-Morgenstern stable sets. This theory is presented as a preparation for the theory of generalized stable sets. In section 6, we finally present our theory of generalized stable sets. We prove that generalized stable sets are nonempty when the asymmetric part of R is nonempty. We also give a characterization of generalized stable sets by using the contraction technique. Finally, in the last section of this chapter we summarize the main commonalities and differences of the discussed theories, including the theory of maximal choices.

4.2 Mathematical Preliminaries

A *digraph* $D = (V, U)$ is a finite and nonempty set V together with an asymmetric binary relation U on V. The set V is called the vertex set and the set U the arc set. Each element in V is called a *vertex* or *point*. An element in U is called an *arc* or, also, a *directed edge*. Instead of $(x, y) \in U$ we write xUy.

If xUy, then y is called a *successor* of x while x is called a *predecessor* of y. A *subgraph* $S = (W, T)$ of D is a graph where $W \subseteq V$ and $T = U \cap (W \times W)$. What follows applies to a digraph $D = (V, U)$. Let $x, x_1, x_2, \ldots, x_m, y \in V$. A *path* from x to y in a digraph $D = (V, U)$ is a finite sequence of distinct points $(x, x_1, x_2, \ldots, x_m, y)$ such that $xUx_1, x_1Ux_2 \ldots x_mUy$. A *semipath* from x to y is a finite sequence of distinct points $(x, x_1, x_2, \ldots, x_m, y)$ such that $xUx_1 \vee x_1Ux, x_1Ux_2 \vee x_2Ux_1, \ldots, x_mUy \vee yUx_m$. A path from x to y becomes a *cycle* when the ordered pair yUx is added[2]. A *complete cycle* is a cycle that contains every element of V. The following definition is important. The finite sequence $(x_1, x_2, \ldots, x_m)$ is a *top cycle* in $D = (V, U)$:=

- $(x_1, x_2, \ldots, x_m)$ is a cycle in D,
- there is no $x \in V - \{x_1, x_2, \ldots, x_m\}$ and there is no $y \in \{x_1, x_2, \ldots, x_m$ such that xUy.

Obviously, a complete cycle is a top cycle. A digraph $D = (V, U)$ is

- *acyclic* := D does not contain any cycle.

The following lemma is useful. For a proof consider Bezhad et.al. (1979: 329).

Lemma 4.1 *Every acyclic digraph has at least one vertex that has no predecessors and at least one vertex that has no successors.*

Since this lemma will be extensively used, we give a proof.
Proof. *Let (V, U) be an acyclic digraph and suppose (V, U) has no maximal elements, i.e. $\neg\exists x \neg\exists y[yUx]$, i.e. $\forall x \exists y[yUx]$. But then (V, U) contains a cycle, which is in contradiction with the fact that (V, U) is acyclic.* $\square$.

The *transitive closure* U^τ of U is the set

$$U^\tau := \{(x, y) \in V \times V \mid \textit{there is a path from } x \textit{ to } y \textit{ or } x = y\}.$$

[2]See chapter 2.4.

U^τ is the intersection of all transitive relations on V that contain U (cf. Suzumura 1983: 12). U is transitive if and only if $U = U^\tau$.

Two distinct vertices x, y in V are

- *2-connected* := $xU^\tau y \wedge yU^\tau x$;
- *1-connected* := $xU^\tau y \vee yU^\tau x$;

i-Connectivity between two elements x and y will be denoted by $xCon^i y$ where $i = 1, 2$. A digraph $D = (V, U)$ is

- *strongly connected* := for all $x, y \in V : xCon^2 y$;
- *connected* := for all $x, y \in V$: $xCon^1 y$;

Clearly, a digraph $D = (V, U)$ is strongly connected if and only if U^τ is complete and symmetric. It is connected if and only if U^τ is complete.

Con^2 is reflexive, symmetric and transitive and thus an equivalence relation on V. It partitions D into subgraphs that are all strongly connected. The strongly connected subgraphs of D are called the *strong components* of D denoted by S_i^*. The vertex set of a strong component S_i^* is denoted by V_i^*. The symbol Ξ stands for the partition of V induced by the vertex sets of the strong components[3]. Obviously, a strong component of D is maximal in the sense that no proper subgraph of it is strongly connected. As a subgraph, it either consists of one vertex or it contains a complete cycle. To illustrate the notion of a strong component consider the following digraph (see Schwartz 1986: 142):

- $V = \{x, y, z, u, t, s\}$,
- $U = \{(x, y), (y, z), (z, x), (u, s), (s, t), (t, u), (x, t), (z, s), (y, u)\}$.

See figure 4.1. This digraph is connected since every pair of vertices is 1-connected. However, it is not strongly connected since for example x and s are not 2-connected. The strong components are $(\{x, y, z\}, \{(x, y), (y, z), (z, x)\})$ and $(\{u, s, t\}, \{(u, s), (s, t), (t, u)\})$. The vertex sets of these strong components constitute the partition $\{\{x, y, z\}, \{u, s, t\}\}$ of V.

[3]In mathematical terms: Ξ is the quotient set of V modulo Con^2.

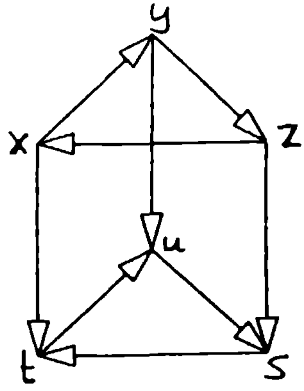

Figure 4.1: Digraph 1

The *contraction* D^{con} of D is the digraph (Ξ, U^{con}) where

$$U^{con} := \{(V_i^*, V_j^*) \in \Xi \times \Xi \mid \exists(x,y) \in U[x \in V_i^* \wedge y \in V_j^*]\}.$$

Intuitively, the contraction of a digraph is the collapsing of each strong component into one single point enriched with a relation among the resulting points. Two points are related with each other if and only if there is an arc in the original digraph with its first vertex in one point and the second vertex in the other one. The contraction of the digraph in figure 4.1 is the digraph with as vertex set the points $V_1^* = \{x, y, z\}$ and $V_2^* = \{u, s, t\}$ and with as arc set $\{(V_1^*, V_2^*)\}$. See figure 4.2. (V_1^*, V_2^*) is an arc in the contraction, because there is an arc in the original digraph going from a vertex in V_1^* to a vertex in V_2^*. Note that if one of the arcs (x,t), (y,u) or (z,s) is reversed, the resulting digraph gets strongly connected and collapses into one single point.

The following lemma will be used in the sequel:

Lemma 4.2 *Let D be a digraph. Then*

1. *the contraction of D is always acyclic;*

2. *the contraction of D is i-connected if D is i-connected, where* $i = 1, 2$.

$$V_1^* \longrightarrow V_2^*$$

Figure 4.2: Contraction of digraph 1

The proof of this elementary lemma is not difficult and can be found in Dehzad et.al. 1979: 327-8 or Harary et.al. 1965: 62-3.

Further, note that the contraction of a digraph $D = (V, U)$ is asymmetric. That is, for $i \neq j$, if $(V_i^*, V_j^*) \in U^{con}$, then not $(V_j^*, V_i^*) \in U^{con}$. Otherwise, $V_i^* \cup V_j^*$ would form a strong component and thus $(V_i^*, V_j^*) \notin U^{con}$ which is a contradiction. Also note that the contraction of a quasi-transitive digraph is transitive.

In the theory of social choice the set V in the digraph $D = (V, U)$ is commonly considered as a set of alternatives while the relation U serves as a strict preference relation on V. In the sequel, we pick up again the terminology as introduced in chapter 2 and 3. So, throughout this chapter we assume the existence of a social choice problem X and we will interpret V as an $A \in \Omega$ and U as a strict preference P restricted to A. In addition, we will assume in this chapter that $P \neq \emptyset$ In order to remember that we are dealing with strict preferences restricted to an agenda, we prefer to call a digraph $D = (A, P)$ a *preference structure*.

4.3 Maximal Social Choices

The notion of maximal choice set has already been defined in chapter 2, section 4. It has an equivalent in the game theoretical concept of a core (see Ordeshook 1986, Owen 1982 or Shubik 1982). Sometimes, it is also called the generalized Condorcet set (Blair et.al. 1976).

For completeness we give the definition again. Consider an agenda A and a preference R over X. An $x \in A$ is *maximal* for A given R if yPx for no $y \in A$, where P is the asymmetric part of R. The set of maximal elements of A is denoted by $\mu(A, R)$. Clearly, if R is complete, then $\mu(A, R) = \beta(A, R)$ where $\beta(A, R)$ is the set of R-best elements of A (see chapter 2, section 4). The main difference between R-best elements and R-maximal elements comes into the picture when giving up completeness. To see this, consider $\{x, y, z, w\}$ and the disconnected preference $yPwPz$. Since x is preferred by no alternative, it belongs to the maximal set. However, the set of R-best choices is empty.

A maximal element for a relation P is, in digraph-theoretical terms, an element that has no P-precedessors. Hence, according to lemma 4.1, a preference has a nonempty set of maximal choices if it is acyclic. However, $\mu(A, P)$ need not be empty if (A, P) contains a cycle. Acyclicity is not a necessary condition. To see this, consider the agenda $\{x, y, z, w\}$ with the connected relation $\{xPyPzPwPy\}$. Since there is no element that is preferred to x, it follows that x belongs to the set of maximal choices.

The next result gives a necessary and sufficient condition for the existence of a maximal choice. Consider a preference structure (A, P). Let (Ξ, P^{con}) denote the contraction of (A, P). We define:

Definition 4.1

$$\mu(\Xi, P^{con}) := \{V_i^* \in \Xi \mid \neg\exists V_j^* \in \Xi[i \neq j \wedge (V_j^*, V_i^*) \in P^{con}]\}.$$

That is, $\mu(\Xi, P^{con})$ is the set of the vertex sets of the strong components in P that are maximal in P^{con}.

Theorem 4.1 (Existence Theorem) *Let $D = (A, P)$ be a preference structure. x is a maximal element of D if and only if $\{x\}$ is a maximal element of the contraction of D.*

Proof. *Let x be a maximal element. Then x has no predecessors in P and hence the only strong component that contains x is $\{x\}$. Since there are no arcs going towards x in the original structure, there can be no arcs going towards $\{x\}$ in P^{con}. Hence $\{x\} \in \mu(\Xi, P^{con})$ and $\mid \{x\} \mid = 1$.*

Conversely, suppose $\{x\} \in \mu(\Xi, P^{con})$. Then there is no $V_i^* \in \Xi$ such that $(V_i^*, \{x\}) \in P^{con}$. Since every $y \in A$ must be an element of a strong *component of P, there is no $y \neq x \in A$ such that yPx.* □.

According to this simple result, a preference structure has a maximal element if and only if the contraction of this structure contains a maximal point that consists of exactly one vertex. Perhaps more important are the following implications of the existence theorem:

Corollary 4.1 *Let (A, P) be a preference structure. Then*

1. $\mu(A, P) = \emptyset$ *if and only if for each* $V_i^* \in \mu(\Xi, P^{con})$, $|V_i^*| \neq 1$.
2. $\mu(A, P) = \emptyset$ *if P is strongly connected.*

In the second part of this corollary the fact is used that an agenda $A \in \Omega$ contains at least three elements (cf. chapter 2.2). Since the proof is easy, we do not give it here. Note that strong connectedness of a (A, P) is not a necessary condition for the emptiness of the set of maximal social choices. Indeed, the relation $xPyPzPxPw$ is not strongly connected (since e.g. (x, w) is not 2-connected) but the set of maximal elements is empty.

In general, there may be $x \in A - \mu(A, P)$ for which there are no $y \in \mu(A, P)$ such that yPx. The core solution in n-person cooperative game theory also has this property (see a.o. Shubik 1982: 157). Clearly, if $x, y \in \mu(A, P)$, then neither xPy nor yPx. This property of maximal sets is called *internal stability*. A stable set also has this property (cf. section 5 of this chapter).

Theorem 4.1 and its corollary 4.1 say that maximal social choices exist when there are no top cycles in a social preference. The absence of top cycles in a social preference as a desirable condition for social choice processes is already discussed by In 't Veld (1975). He argues that the transitive collective rationality requirement of Arrow is too severe. In his view, the absence of top cycles in a social choice preference suffices for generating a social choice. In 't Veld gives the following example (1975, 64): $aPb, aPc, aPd, bPc, cPd, dPb$ where P is a strict social preference. Since a is strictly preferred to every other alternative, it is, according to In 't Veld, a satisfactory winner. The fact that b, c and d are involved in a cycle does not influence this result. Note that for this example, a is also best.

According to theorem 4.1 and corollary 4.1, maximal social choices do not exist when each $V_i^* \in \mu(\Xi, P^{con})$ contains at least two elements, thus, when there is a top cycle in a social preference. The question then is what to do in this case. In the next section we study a theory that provides an answer.

4.4 The Theory of Generalized Optimal Choices

The origin of this theory can be found in Schwartz (1971, 1986). The theory is equivalent to the theory of dynamic solutions of Shenoy (1977, 1979, 1980), and to the theory of admissible sets of Kalai and Schmeidler (1976) and Kalai et.al. (1977).

Crucial in the theory of generalized optimal choices is the notion of *minimal undominated set.* A set of possible social choices is undominated if there is no alternative outside this set that is socially strictly preferable to an alternative inside the set. An undominated set is minimal if none of its proper subsets has this property. The generalized optimal choice set of a set is the union of the minimal undominated subsets of that set.

Definition 4.2 *Let (A, P) be a preference structure and $B \subseteq A$.*

1. *B is an* undominated set *of A := for no $x \in B$ there is a $y \in A - B$ such that yPx,*

2. *B is a* minimal undominated set :=

   (a) B is an undominated set and

   (b) There is no $B^* \subset B$, $B^* \neq \emptyset$, such that B^* is undominated.

The first condition says that for no element in an undominated set B there is an alternative outside B that is more preferable. However, this does not mean that for every alternative $x \in A - B$ there is a $y \in B$ such that yPx. In this sense an undominated set satisfies, just like a maximal choice set, the property of *external incomplete stability* (cf. the preceding section). Also note that it is allowed that, eventually, an element x in B is preferred to another element y in B. The second requirement is a minimality property. To illustrate, consider figure 3 below (also see Schwartz 1986:

142). In this figure $X = \{x, y, w, z, x_1, y_1, z_1, w_1, y_2, x_2, z_2, w_2\}$. The subset $\{x_1, y_1, z_1\}$ of X is undominated since for no $u \in \{x_1, y_1, z_1\}$ there is a $v \in X - \{x_1, y_1, z_1\}$ such that vPu. However, $\{x_1, y_1, z_1\}$ is also minimal undominated, since it does not contain a nonempty proper subset which is undominated.

Schwartz (1986: 145) proves the following important result about minimal undominated sets:

Theorem 4.2 (Schwartz) *Let (A, P) be a preference structure. Let B be a minimal undominated set of A. Then $(B, P \cap (B \times B))$ is either a top cycle in P or a singleton set consisting of a maximal element.*

Definition 4.3 *Let (A, P) be a preference structure. A nonempty subset $\sigma(A, P)$ of A is the* generalized optimal choice set, *henceforward* GOCS, *of A :=*

$$\sigma(A, P) = \bigcup\{B \subseteq A \mid B \, is \, a \, minimal \, undominated \, subset \, of \, A\}.$$

Since, according to theorem 4.2, a minimal undominated set either is a top cycle or consists of one single maximal element, the GOCS is the union of top cycles and maximal elements in a structure (A, P) (cf. Schwartz 1986: 145).

The next theorem contains another equivalent formulation of a GOCS. This theorem is, again, formulated by Schwartz (1986: 146). In this theorem, $\bar{a}(P^{\tau})$ denotes the asymmetric part of the transitive closure of P. Hence, $x\bar{a}(P^{\tau})y$ means that there is a chain from x to y but not from y to x.

Theorem 4.3 (Schwartz) *Let (A, P) be a preference structure. Then*

$$\sigma(A, P) = \{x \in A \mid \neg\exists y \in A : y\bar{a}(P^{\tau})x\}.$$

This theorem shows the relation between the notion of a set of maximal choices and the notion of GOCHS. A set of maximal choices consists of elements which are maximal according to P while a GOCS consists of elements which are maximal according to $\bar{a}(P^{\tau})$. Clearly,

Theorem 4.4 *Let (A, P) be a preference structure.*

$$\mu(A, P) \subseteq \sigma(A, P).$$

Proof *Let x be maximal. Then there is no $y \in A$ such that yPx and hence $\{x\}$ is a minimal undominated set.* □.

If $x \in \sigma(A, P)$ then it is possible that there is a $y \in \sigma(A, P)$ such that yPx. Hence, $\sigma(A, P) \subseteq \mu(A, P)$ is not true in general. This result shows that the theory of GOCS is an extension of the classical notion of collective rationality as social maximality.

The next result gives another characterization of a GOCS. This theorem, again, is formulated by Thomas Schwartz (1986: 154).

Theorem 4.5 (Schwartz) *Let (A, P) be a preference structure. Then $\sigma(A, P)$ is a GOCS of A if and only if it satisfies the following conditions:*

1. *for no $x \in \sigma(A, P)$ there is a $y \in A - \sigma(A, P)$ such that yPx;*
2. *there is no $B \subset \sigma(A, P)$ that is minimally undominated in the substructure $(\sigma(A, P), R \cap (\sigma(A, P) \times \sigma(A, P)))$;*
3. *if B is an undominated set of A, then there is an $x \in B$ such that $x \in \sigma(A, P)$.*

Propery 1 is strong. It says that there is no y outside the GOCS that is better than some alternative x inside the GOCS. This also implies that there is no chain starting from a y outside GOCS and terminating at an x inside GOCS.

In general, there may be alternatives $y \in A$ for which there are no $x \in \sigma(A, P)$ such that xPy. To see this consider the preference $x_1Px_2, x_2Px_3, x_3Px_1, x_1Px_4, x_4Px_5$. For this case the GOCS is $\{x_1, x_2, x_3\}$. There is no $x \in \{x_1, x_2, x_3\}$ such that xPx_5. However, the following theorem shows that the GOCS can 'reach' every element outside the GOCS via a chain. Remember that $\bar{a}(P^\tau)$ is the asymmetric part of P^τ.

Theorem 4.6 *Let (A, P) be a preference structure and $\sigma(A, P)$ its GOCS. Then for every $y \in A - \sigma(A, P)$ there is an $x \in \sigma(A, P)$ such that $x\bar{a}(P^\tau)y$.*

In the proof of this theorem, we need the existence theorem (theorem 4.8). We therefore postpone it for later.

The next theorem is obtained by using the contraction technique (see the mathematical preliminaries for this chapter). It shows what relation is maximized by the theory of GOCS. Remember that $\mu(\Xi, P^{con})$ is the set of vertex sets that are maximal in P^{con} (see definition 4.1 in section 4.3).

Theorem 4.7 (Contraction Theorem) *Let (A, P) be a preference structure and (Ξ, P^{con}) its contraction. Then:*

$$\sigma(A,P) = \bigcup \mu(\Xi, P^{con}).$$

Proof. *Let B be a minimal undominated subset of A. Since, according to theorem 4.2, B either consists of a top cycle or is a singleton for which there is no alternative that is more preferable, $B \in \Xi$. Since there is no $y \in A - B$ such that yPx for some $x \in B$, $(V_i^*, B) \in P^{con}$ for no $V_i^* \in \Xi$. Hence B is maximal in P^{con}. Hence $\sigma(A,P) \subseteq \bigcup \mu(\Xi, P^{con})$.*

Let V_i^* be maximal in P^{con}. Then for no $V_j^* \in \Xi$ with $V_j^* \neq V_i^*$, $(V_j^*, V_i^*) \in P^{con}$ and therefore there is no $x \in A - V_i^*$ such that xPy for some $y \in V_i^*$. Hence, V_i^* is an undominated set. We now have to prove the minimality property of V_i^* as an undominated set. Two cases:
Case 1: V_i^* is a singleton. Then V_i^* is minimal undominated in A.
Case 2: V_i^* is the vertex set of a top cycle in P that is complete with respect to V_i^*. Let $W \subset V_i^*$ and $W \neq \emptyset$.. Then there is an $x \in V_i^* - W$ and a $y \in W$ such that xPy. So W is not undominated in A. □.

According to this theorem, the GOCS for a structure (A, P) is the union of the strong components in P that are maximal in P^{con}. Hence, what the theory of GOCS maximizes is the contraction of a social preference. To illustrate this point consider the following social preference given by Schwartz (1986: 142).

- xPy, yPz, zPw, wPx, wPy and zPx;
- x_1Py_1, y_1Pz_1 and z_1Px_1;
- x_2Py_2, y_2Pw_2, w_2Pz_2 and z_2Py_2;
- wPw_1, y_1Py_2, y_2Pw_1 and z_1Pw_1.

See figure 4.3. The contraction of the digraph in figure 4.3 is
$\{V_1^*, V_2^*, V_3^*, V_4^*, V_5^*\}$,
$\{(V_1^*, V_3^*), (V_2^*, V_3^*), (V_2^*, V_4^*), (V_4^*, V_3^*), (V_5^*, V_4^*)\}$, where
$V_1^* = \{x, y, z, w\}$, $V_2^* = \{x_1, y_1, z_1\}$, $V_3^* = \{w_1\}$,
$V_4^* = \{y_2, w_2, z_2\}$ and $V_5^* = \{x_2\}$. See figure 4.4. The set of maximal

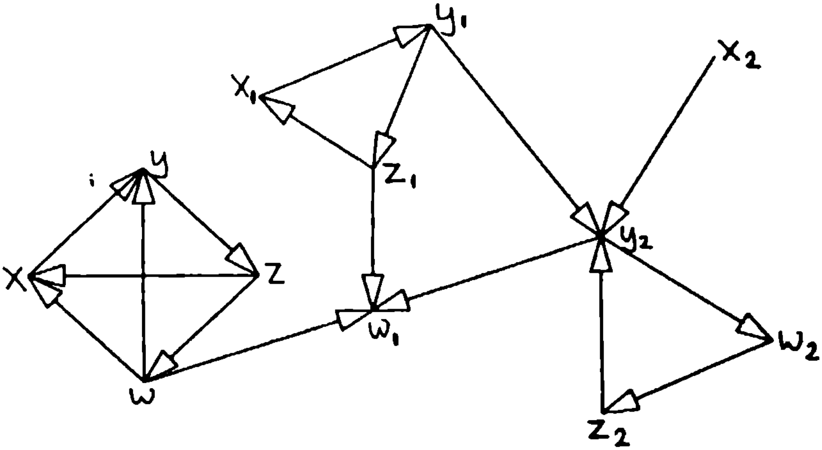

Figure 4.3: Digraph 2.

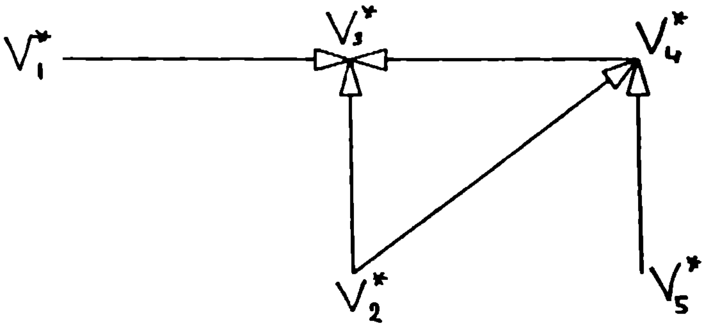

Figure 4.4: Contraction of digraph 2.

elements of this contraction is $\{V_1^*, V_2^*, V_5^*\}$. According to theorem 4.7, the GOCS is the union of V_1^*, V_2^* and V_5^*, that is, $\{x, y, z, w, x_1, y_1, z_1, x_2\}$. Note that the maximal elements of the contraction are exactly the minimal undominated subsets of the social preference.

Since a contracted relation is always acyclic (lemma 4.2) and since every acyclic relation has a nonempty set of maximal elements (lemma 4.1), we immediately have:

Theorem 4.8 (Existence Theorem) *For every preference structure (A, P) the GOCS is nonempty.*

Proof. *According to lemma 4.2, the contraction P^{con} is acyclic. Hence $\mu(\Xi, P^{con}) \neq \emptyset$. Hence $\bigcup \mu(\Xi, P^{con}) \neq \emptyset$. By theorem 4.7, $\sigma(S, P) \neq \emptyset$.* □.

An alternative proof of the existence theorem is given in Schwartz (1986). We are now ready to prove theorem 4.6:

Proof of theorem 4.6. *In this proof we use the existence theorem 4.8. Let $y \in A - \sigma(A, P)$. Then y belongs to no minimal undominated set and therefore there must be an $x_1 \in A$ such that $x_1 P y$. If x_1 belongs to no minimal undominated set, there must be an $x_2 \in A$ such that $x_2 P x_1$. If x_2 belongs to no minimal undominated set, there must be an $x_3 \in A$ such that $x_3 P x_2$. Proceeding in this way, since A is finite, there must be an $x_n \in A$ such that $x_n P x_{n-1}$ and $x_n \in \sigma(A, P)$. Because of theorem 4.8, $\sigma(A, P) \neq \emptyset$. If $y P x_n$, then $y \in \sigma(A, P)$. Contradiction. Hence $x \bar{a} P^\tau y$..*
□.

The next result is easily derived from the contraction theorem:

Corollary 4.2 *Let (A, P) be a preference structure. Then $\sigma(A, P) = A$ if and only if (A, P) is strongly connected.*

Hence, in the case of complete cycles the theory of GOCS loses its power to discriminate.

A GOCS satisfies external stability with respect to P^τ (cf. theorem 4.6). It even satisfies a stronger property: if $\sigma(A, P)$ is the GOCS for (A, P), then there are no $y \in A - \sigma(A, P)$ such that $y P^\tau x$ for some $x \in \sigma(A, P)$. This last property is called *strong external stability*. In this respect, a social choice process that produces elements from a GOCS as a social choice will never produce a social choice for which there is a better one outside the solution.

However, a GOCS does not satisfy the property of internal stability. That is, it is possible that there are $x, y \in \sigma(S, P)$ such that $yP^{\tau}x$. If this is the case, we also must have $yP^{\tau}x$. So, a social choice process that produces only elements from a GOCS may yield a solution for a social choice problem that is 'instable'. This instability is, so to say, neatly kept within that solution itself. The alternative that dominates the outcome, the produced social choice, is also in the solution. It may have been an outcome itself. However, it remains that a GOCS may contain alternatives that are involved in a cycle and, therefore, it may lead to a social choice process that does not terminate. The process then goes, so to say, 'round and round in cycles'.

We will formulate an alternative solution theory that yields solutions characterized by inner stability. This avoids going 'round and round in cycles'. The theory is a generalization of a well known solution theory from n-person cooperative game theory. We first study this solution theory before presenting its generalization.

4.5 Von Neumann-Morgenstern Theory of Stable Sets

This section is based on Van Deemen (1990b) The origin of the theory of Von Neumann-Morgenstern stable sets can be found in their celebrated work *Theory of Games and Economic Behavior* (1953). See also Luce and Raiffa 1957. Von Neumann and Morgenstern call this theory the theory of solutions. In order to avoid confusion with other solution concepts, we will use the term 'stable set' (cf. Shubik 1982). What we do here is to reformulate the game-theoretical theory of stable sets in a way that suits our aims.

In social choice theoretical terms, a subset B of a set A is a stable set if no alternative in B is socially strictly preferred to another alternative in B and if, in addition, for every alternative outside B there is an alternative inside B that is strictly preferred to this outside alternative.

Definition 4.4 *Let (A, P) be a preference structure. A nonempty $V \subseteq A$ is a* stable set of A :=

1. *for all $x, y \in V$: not xPy.*

2. *for all $y \notin V$ there is an $x \in V$ such that xPy.*

The first property is internal stability. As we have seen, also sets of maximal choices satisfy this property. It says that no element in a stable set is socially strictly preferred to another element in this stable set. The second property is external stability. It says that for every x outside a stable set there is a y in this stable set that is strictly preferred to x. However, the stable set solution concept does not preclude the possibility that there is a $y \notin V$ such that yPx for some $x \in V$. Hence, with respect to P, it does not satisfy the property of strong external stability.

Theorem 4.9 *Let (A, P) be a preference structure and V be a stable set of A. Then V is maximal with respect to internal stability and minimal with respect to external stability.*

Proof. *If V is not maximal with respect to internal stability, then there must be an $M \subseteq A$ such that $V \subset M$ and M satisfies internal stability. Take an $x \in M - V$. Then, since V satisfies external stability, there must be a $y \in V$ such that yPx. But then M does not satisfy internal stability. Contradiction.*

If V is not minimal with respect to external stability, then there must be a proper subset M of V that has the same property. Consider an $x \in V - M$. Since M satisfies external stability, there must be a $y \in M$ such that yPx. But then V does not satisfy internal stability. Contradiction. □.

Von Neumann and Morgenstern give a subtle interpretation of stable sets. In their view, a stable set is a characterization of what may be acceptable or established as a "standard of behavior" in society (Von Neumann and Morgenstern 1953: 41). Internal stability, then, expresses the fact that the standard of behavior has an inner consistency. It guarantees the absence of "inner contradictions". External stability has another function. It gives a reason to correct deviant behavior, that is, to correct behavior that is not conformable to the "standard of behavior" (1953: 41). Von Neumann and Morgenstern conclude (1953: 42):

> Thus our solutions S correspond to such "standards of behavior" as have an inner stability: once they are generally accepted they overrule everything else and no part of them can be overruled within the limits of the accepted standards.

This nice interpretation of Von Neumann and Morgenstern also shows why generalized optimal choice sets may not be acceptable. In their terms, such sets lack inner consistency. They may lead to collective behavior that is not free from contradicion.

In general, a stable set is not unique. A sufficient condition for its uniqueness is already given by Von Neumann and Morgernstern themselves (1953: section 65.8)

Theorem 4.10 (Von Neumann and Morgenstern) *Let (A, P) be a preference structure. If P is acyclic, then there is a unique stable set.*

The power of the theory of stable sets lies in the fact that it can handle cycles. To see this consider the agenda $\{x, y, z, w\}$ with the cycle $xPyPzPwPx$. A social choice rule satisfying one of the classical rationality conditions has nothing to tell about this structure. Further, the theory of maximal choices also fails. There are no maximal elements However, the stable sets are $\{x, z\}$ and $\{y, w\}$ since both sets are internally and externally stable.

The multiplicity of stable sets as solutions for a social choice problem[4] leads to a problem. Which stable set, which standard of behavior should be chosen by a society when a multiplicity of them are available? Von Neumann and Morgenstern do not claim to solve this problem. They are primarily interested in "where the equilibrium of forces lies" (Von Neumann and Morgenstern 1953: 42), not in what equilibrium actually will be arrived at. Neither do we claim to solve this problem. In general, we think that a multiplicity of solutions for social choice problems may reflect the complicated equilibrium forces which operate in human societies. It is also quite possible that in a dynamic context some stable sets may be excluded because, for example, they do not lie on suitable or historically acceptable equilibrium paths. Societies may 'grow' to particular stable sets as acceptable standards of behavior because of their history and the

[4]See chapter 2, section 2 for the notion of social choice problem.

dynamics of the operative systems of norms and values. So, societal dynamics may help to reduce the multiplicity of solutions of its social choice problems.

The theory of stable sets has a serious shortcoming. Consider the agenda $\{x, y, z\}$ with the cycle $xPyPzPx$. For this preference structure there are no R-stable sets. The same is true for a structure like $\{x_1, x_2, x_3, x_4, x_5\}$ with cycle $x_1Px_2Px_3Px_4Px_5Px_1$. Let us call a cycle $x_1Px_2P \dots Px_nPx_1$ *odd* if n is odd. If a cycle is not odd, then it is called *even.* It appears that the theory of stable sets can only deal with even cycles. As the following result shows, it cannot deal with complete odd cycles.

Theorem 4.11 *Let $A = \{x_1, x_2, \dots, x_n\}$ be an agenda with n odd. Then A has no stable set if there is a unique cycle*

$$x_1Px_2Px_3P \dots Px_nPx_1.$$

The proof of this theorem can be found in (Harary et. al. 1965: 177-8). It must be noticed that the conditions in this theorem are rather demanding. The cycle must be complete, i.e., every element in A must be in the cycle; the cycle must be unique and, finally, n must be odd. Note that the uniqueness and completeness of the cycle imply that (A, P) is strongly connected.

The following result is positive in nature. It shows that the theory of stable sets has no problems with even cycles. This theorem stems from Richardson (1953). It applies to any (A, P), irrespective of its connectedness degree.

Theorem 4.12 (Existence Theorem (Richardson)) *Let (A, P) be a preference structure. Then (A, P) has a stable set if P has no odd cycles.*

Somewhat stronger results can be found in Duchet (1987) and Galeana-Sanchez (1984). The necessary and sufficient conditions that guarantee the existence of a stable set for an arbritary preference structure (A, P) are still unknown.

Within the framework of n-person cooperative game theory, it can be shown that the core of a game is contained in each stable set of that game (Luce and Raiffa 1957, Shubik 1982). This is also true for the version presented here

Theorem 4.13 *Let (A, P) be a preference structure. If V is a stable set of A, then*

$$\mu(A, P) \subseteq V.$$

Proof. *Let x be maximal and suppose there is a V such that $x \notin V$. Then, by external stability, there must be a y in V such that yPx. But then x is not maximal. Contradiction.* □.

There may be preference structures with a nonempty set of maximal elements and no stable sets. To see this, consider the strict preference $x_1Px_2Px_3Px_4Px_5Px_3$. There is no stable set for this structure. The set of maximal elements is $\{x_1\}$. Note that $\{x_1\}$ also happens to be the GOCS.

A social choice process that always produces an element of a stable set as a social choice produces in fact a result that is conformable to some accepted standard of behavior. With the property of external stability it also is possible to correct for collective behavior that deviates from the accepted standard. However, the main flaw of stable sets and, hence, of social choice processes that produce stable sets as solutions, appears to be the handling of odd cycles. For this kind of structures, the Von Neumann-Morgenstern theory of stable sets may fail.

4.6 Generalized Stable Sets

This section is based on Van Deemen (1990a). The theory of stable sets has the asymmetric part of a social preference as its point of departure. In contrast, the theory of generalized stable sets will have the *transitive closure of the asymmetric part* of a social preference as its point of departure. In this section we shall see that this apparently little difference has rather great consequences.

Definition 4.5 *Let (A, P) be a preference structure. A nonempty subset V of A is a* Generalized Stable Set, *henceforward GESTS, of A :=*

1. *for all $x, y \in V$: not $xP^{\tau}y$,*

2. *for every $y \in A - V$ there is an $x \in V$ such that $xP^{\tau}y$.*

The first condition is *Generalized Internal Stability*. It says that at no element in a GESTS there starts a path toward another element in that GESTS. Note that this also implies that there are no x, y in a GESTS such that xPy. The second condition is *Generalized External Stability*. It says that for every x outside a GESTS there starts a path from some y inside that GESTS that terminates at x. In general, however, it is not precluded that there is a path starting from an element outside a GESTS and terminating at an element inside a GESTS. Inside a GESTS, however, there are no paths.

Theorem 4.14 *Let (A, P) be a preference structure. Then every GESTS of A is maximal with respect to generalized internal stability and minimal with respect to generalized external stability.*

Proof. *Let V be a GESTS of A. If V is not maximal with repect to generalized internal stability, there must be a nonempty $M \supset V$ such that M is generalized internally stable. Take an $x \in M - V$. Since V satisfies generalized external stability, there must be a $y \in V$ such that $yP^{\tau}x$. But then M does not satisfy generalized internal stability. Contradiction.*

If V is not minimal with repect to generalized external stability, then it must contain a nonempty M that has this property. Consider an $x \in V - M$. Since M satisfies generalized external stability, there must be a $y \in M$ such that $yP^{\tau}x$. But then V does not satisfy generalized internal stability. Contradiction. □.

Also generalized stable sets can be interpreted as accepted standards of behavior. The condition of generalized internal stability now applies to a generalized notion of domination, namely, to the transitive closure of the asymmetric part of a social preference. This includes the asymmetric part of a social preference. In this form it also works as a consistency condition in the sense that it leads to collective behavior that is free of inner contradictions (cf. Von Neumann and Morgenstern 1953: 41). In this respect the theory fundamentally differs from the theory of GOCS. As we have seen, the alternatives inside a GOCS may be part of a cycle. In terms of the behavioral interpretation of Von Neumann and Morgenstern, this means that this theory allows that an alternative that complies with an accepted standard of behavior can be dominated by an alternative that

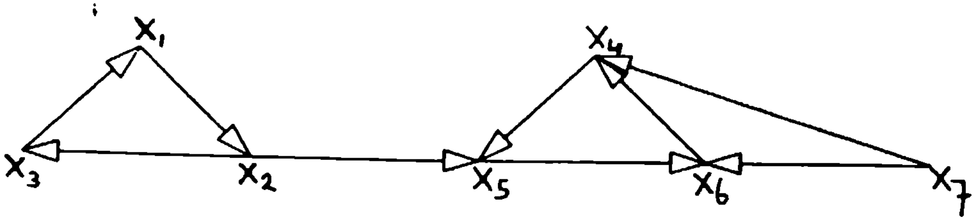

Figure 4.5: Digraph 3.

also complies with that standard of behavior. In this sense, a GOCS may lack internal consistency.

The condition of generalized external stability also operates as a kind of correction procedure. With this property it is possible to correct nonconformable behavior, i.e. behavior that deviates from the standard. A less convincing point of this correction procedure is that it does not preclude the possibility that an alternative that is not conformable to the accepted standard can dominate an alternative that is conformable to that standard. In this aspect, the theory of GOCS is stronger. This theory precludes the possibility that a disconformable alternative with respect to some accepted standard of behavior dominates a conformable alternative.

To give an illustration of the working of the theory of GESTS, consider the agenda $\{x, y, z\}$ and the cycle $xPyPzPx$. The GOCS for this set is $\{x, y, z\}$. There is no stable set. The GESTS are $\{x\}$, $\{y\}$ or $\{z\}$. A more intricate example is the following one:

- x_1Px_2, x_2Px_3, x_3Px_1,
- x_4Px_5, x_5Px_6, x_6Px_4,
- x_2Px_5, x_7Px_6, x_7Px_4.

See figure 4.5. For this case, the set of maximal elements is $\{x_7\}$. The GOCS is $\{x_1, x_2, x_3, x_7\}$. The GESTS-solutions are $\{x_1, x_7\}$, $\{x_2, x_7\}$ and

$\{x_3, x_7\}$.
As this example shows, a GESTS will generally not be unique. It is not difficult to prove that acyclicity, again, is sufficient for uniqueness.

The following result is fundamental. It gives a characterization of a GESTS:

Theorem 4.15 *Let (A, P) be a preference structure and (Ξ, P^{con}) its contraction. Let $\mu(\Xi, P^{con}) = \{V_1^*, V_2^*, \ldots, V_l^*\}$ and let $x_1, x_2, \ldots, x_l \in A$. Then $\{x_1, x_2, \ldots, x_l\}$ is a generalized stable set of A if and only if $x_1 \in V_1^*, x_2 \in V_2^*, \ldots, x_l \in V_l^*$.*

Proof. *Let $V_1^*, V_2^*, \ldots, V_l^*$ be the maximal elements in P^{con} and let $x_1 \in V_1^*, x_2 \in V_2^*, \ldots, x_l \in V_l^*$. We have to prove that $\{x_1, x_2, \ldots, x_l\}$ is a GESTS. Let $x_i, x_j \in \{x_1, x_2, \ldots, x_l\}$ and suppose $x_i P^\tau x_j$. Since $x_i \in V_i^*$ and $x_j \in V_j^*$ this means that there must be a chain from V_i^* to V_j^* in P^{con}. But then V_j^* cannot be maximal. Hence not $x_i P^\tau x_j$. In the same way we can prove that not $x_j P^\tau x_i$. Hence $\{x_1, x_2, \ldots, x_l\}$ satisfies generalized internal stability.*

Let $x \notin \{x_1, \ldots, x_l\}$. If x is in a maximal vertex set V_i^* of P^{con}, then V_i^* must contain the elements of a top cycle. Then $x_i P^\tau x$. Note that $x_i P x$ is not precluded. Suppose now that x is in a vertex set X of a strong component that is not maximal in P^{con}. Then there is a vertex set Y of a strong component such that $(Y, X) \in P^{con}$. But then there exists a $y \in Y$ and a $z \in X$ such that yPz by definition of P^{con}. Then $yP^\tau z$ and $zP^\tau x$ and hence $yP^\tau x$ by transitivity of P^τ. If Y is not maximal in P^{con}, then there must be a vertex set Y_1 of a strong component such that $(Y_1, Y) \in P^{con}$. Then there must be a $y_1 \in Y_1$ such that $y_1 P^\tau y$. If Y_1 is not maximal in P^{con}, there must be a vertex set Y_2 of a strong component such that $(Y_2, Y_1) \in P^{con}$. But then there must be a $y_2 \in Y_2$ such that $y_2 P^\tau y_1$. Proceeding in this way we must finally reach a maximal V_i^* since P^{con} is acyclic by lemma 4.2 and every acyclic digraph has a maximal element by lemma 4.1. Then we have
$x_i P^\tau y_n P^\tau y_{n-1} P^\tau \ldots P^\tau y P^\tau x$. Since P^τ is transitive, this implies $x_i P^\tau x$. Hence $\{x_1, x_2, \ldots, x_l\}$ satisfies generalized external stability.

Conversely, let W be a GESTS of A and let $x \in W$ be a member of the vertex set V_i^* of a strong component in P. We have to prove that V_i^* is maximal in P^{con}. Suppose therefore that V_i^* is not maximal. Then there must be a vertex set Y_1 of a strong component in P such that $(Y_1, V_i^*) \in P^{con}$

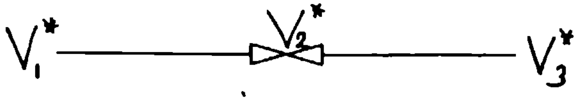

Figure 4.6: Contraction of digraph 3.

and hence there must be a $y_1 \in Y_1$ *such that* $y_1 P^\tau x$. *Proceeding in this way we must find a maximal vertex set* Y_n *of a strong component of P since* P^{con} *is acyclic by lemma 4.2. and by lemma 4.1 every acyclic digraph has a maximal element. But then* $y_n P^\tau x$ *for some* $y_n \in Y_n$. *Then* $y_n \notin W$. *Hence there is an* $u \in W$ *such that* $uP^\tau y_n$. *Since* $y_n P^\tau x$ *we must have* $uP^\tau x$ *by transitivity of* P^τ. *But this violates the condition of generalized internal stability, since* $x, u \in W$. *Hence* V_i^* *is maximal in* P^{con}.

Let $x, y \in W$ *with* $x \neq y$ *be members of respectively the maximal vertex sets* V_i^* *and* V_j^* *of* P^{con}. *We have to prove that* $V_i^* \neq V_j^*$. *Therefore let* $V_i^* = V_j^*$. *Since* $x \neq y$, x *and* y *are members of the top cycle that contains every element of* V_i^*. *But then W cannot be generalized internally stable. Contradiction. Hence* $V_i^* \neq V_j^*$. □.

If a preference P is strongly connected, its contraction consists of one point exactly. Hence, according to theorem 4.15, a GESTS exists of precisely one social state in this case. In this respect, the theory of GESTS differs from the theory of GOCS. According to corollary 4.2, the latter theory loses its discriminatory power in the case of strongly connected preferences.

To illustrate theorem 4.15, consider the contraction of the social preference in figure 4.5. See figure 4.6. The vertex sets of the strong components are $V_1^* = \{x_1, x_2, x_3\}$, $V_2^* = \{x_4, x_5, x_6\}$ and $V_3^* = \{x_7\}$. The arc set of the contraction is $P^{con} = \{(V_1^*, V_2^*), (V_3^*, V_2^*)\}$. The maximal vertex sets are V_1^* and V_3^*. According to theorem 4.15, a GESTS consists of one and only one element from V_1^* and one and only one element from V_3^*. Hence $\{x_1, x_7\}$ is a GESTS since $x_1 \in V_1^*$ and $x_7 \in V_3^*$.

The following implications of theorem 4.15 are stated without proof.

Corollary 4.3 *Let (A, P) be a preference structure.*

1. *Each alternative in a GESTS of A is contained in a top cycle;*

2. *The number of elements in a GESTS of A equals the number of strong components of (A, P) that are maximal in P^{con};*

3. *The possible GESTSs of (A, P) all have the same cardinal number.*

The next theorem deals with the existence of generalized stable sets.

Theorem 4.16 (Existence Theorem) *Let (A, P) be a preference structure. Then there exists a GESTS of A.*

Proof. *Let (Ξ, P^{con}) be the contraction of (A, P). By lemma 4.2, P^{con} is acyclic and hence by lemma 4.1 $\mu(\Xi, P^{con}) \neq \emptyset$. Hence, by theorem 4.15, there is a generalized stable set.* □.

This theorem also proves theorem 3.2 (see chapter 3). It also shows that the theory of generalized stable sets can handle, just like the theory of GOCS, every kind of cycle. In this respect, it improves the theory of stable sets.

Let $\Gamma(A, P)$ denote the set of generalized stable sets of (A, P). The following result connects the theory of generalized stable sets to the theory of maximal social choices and, furthermore, to the theory of generalized optimal choice sets.

Theorem 4.17 *Let (A, P) be a preference structure.*

$$\mu(A, P) = \bigcap \Gamma(A, P),$$

$$\sigma(A, P) = \bigcup \Gamma(A, P).$$

Proof. *1). Let* $x \in \mu(A, P)$ *and suppose there is a GESTS* V *such that* $x \notin V$. *Since* V *satisfies generalized external stability, there must be a* $y \in V$ *such that* $yP^{\tau}x$. *But then there must be a* $z \in A$ *such that* zPx. *Hence* x *is not maximal. Contradiction. Hence* $\mu(A, P) \subseteq \bigcap\Gamma(A, P)$.

Conversely, let $x \in \bigcap\Gamma(A, R)$. *Suppose* $\{x\}$ *is not maximal in* P^{con}. *Then there is a* Y *such that* $YP^{con}\{x\}$ *and hence there exists a* $y \in Y$ *such that* $yP^{\tau}x$. *For each* $S \in \Gamma$, $y \notin S$. *Otherwise there is an* S *with* $y \in S$ *which violates generalized internal stability. Then for each* $S \in \Gamma$ *there is a* $z \in S$ *such that* $z_S P^{\tau} y$ *by generalized external stability. By transitivity of* P^{τ}, $z_S P^{\tau} x$, *which violates generalized internal stability. Hence* $\{x\}$ *is maximal in* P^{con}. *By theorem 4.1,* $x \in \mu(A, P)$. *Hence* $\bigcap\Gamma(A, P) \subseteq \mu(A, R)$.
2). By theorem 4.15 $\bigcup\Gamma(A, P) = \bigcup\mu(\Xi, P^{con})$ *where* (Ξ, P^{con}) *is the contraction of* (A, P). *By theorem 4.7* $\bigcup\mu(\Xi, P^{con}) = \sigma(A, P)$. □.

This theorem proves theorem 3.3 (see chapter 3). Another consequence of it is:

Corollary 4.4 *Let* (A, P) *be a preference structure. Then for every GESTS* V *of* A:,

$$V \subseteq \sigma(A, P).$$

In n-person game theory the core of a game is contained in each stable set of that game. This corollary shows a reversed world. It is not the case that a generalized stable solution contains the generalized optimal choice set but just the opposite. Thus, the theory of generalized stable sets may yield more restrictive solutions than the theory of generalized optimal choice sets.

An alternative in a GESTS may be dominated by an alternative outside this GESTS. However, a social choice that is in a GESTS is characterized by internal stability. Within a GESTS there are no x and y that dominate each other. In this respect, a social choice process producing a GESTS as a solution differs from a social choice process producing a GOCHS.

The theory of GESTS is a generalization of the theory of stable sets. It is able to produce a nonempty social choice set for every possible preference structure, irrespective of its cycle structure. According to theorem 4.17, it is an extension of the classical notion of collective rationality as maximality.

4.7 Comparison

In this section we summarize the most important differences and commonalities of the presented solution theories. We first compare the properties of maximal choice sets, undominated sets and stable sets. After this we compare the generalized theories.

The differences and commonalities of maximal sets, minimal undominated sets and stable sets can be summarized as follows:

1. Sets of maximal choices and stable sets satisfy internal stability with respect to P. This means that no x in a stable set is strictly preferred to a y in that stable set. The same is true for sets of maximal choices. A minimal undominated set does not satisfy internal stability.

2. Minimal undominated sets and sets of maximal choices do not satisfy external stability with respect to P. There may be an x outside a set of maximal choices or outside a minimal undominated set for which there is no y in the maximal or minimal undominated set such that yPx. However, there is no y outside a maximal or minimal undominated set that is strictly preferred to an x inside a maximal or minimal undominated set, respectively. In contrast, stable sets satisfy external stability. For every x outside a stable set there is a y inside that stable set such that yPx. However, there may be x outside a stable set that are strictly preferred to a y inside a stable set.

3. Sets of maximal elements do not exist in the case of complete cycles. The theory of stable sets cannot handle odd cycles. In contrast, minimal undominated sets always exist, even in the case of odd cycles.

4. All three theories are extensions of the classical theory of collective rationality. That is, if a maximal social choice exists, then each set under scrutiny will contain it.

We summarize the main differences and commonalities of the theories of GESTS and GOCS:

1. A GESTS satisfies generalized internal stability. This means that for no elements x, y in a GESTS there is a path starting from x and terminating at y. This also implies that inside a GESTS no element is strictly preferred to another element in that solution. In contrast, the theory of GOCS does not satisfy generalized internal stability. This theory allows that an x in the GOCS is dominated by a y in the GOCS. This also implies that in the GOCS there might be x and y such that xPy.

2. GOCS satisfies strong external stability, i.e., no x outside the GOCS dominates an y inside the GOCS. This also implies that there is no x outside the GOCS that is strictly socially preferred to a y inside the GOCS. In contrast, a GESTS does not satisfy strong external stability. That is, it is allowed that an outside alternative dominates an inside alternative. As a consequence, it is possible that there is an x outside a GESTS that is strictly socially preferred to an element y inside a GESTS.

3. The GOCS for an agenda is unique. In contrast, there might be a multiplicity of GESTSs for an agenda.

Both theories have at least two common characteristics.

1. They both can handle all types of cycle structures. This is their main victory over i) the theory of best choice that is completely blind in the case of cyclic social preferences whatsoever, ii) the theory of maximal social choice that is blind in the case of complete cycles and iii) the theory of stable sets that may be blind in the case of odd cycles.

2. Both encompass the classical notion of collective rationality as maximality. See theorems 4.4 and 4.17.

The price of extending the classical notion of collective rationality appears from the first two items of this comparison between the theory of generalized stable sets and the theory of generalized optimal choice sets. Both theories encompass an aspect of instability in the sense of allowing an outcome to be dominated and hence to be strictly socially preferred. The theory of generalized optimal choice sets puts, so to say, this instability inside the solution itself. The alternative that may dominate a possible outcome is also in the solution. In contrast, the theory of generalized stable sets leaves the instability outside the solution. Outside a solution there may be a social state that is a dominant one. The choice of one of these theories is in fact the choice of one of these forms of instability.

In chapter 3, we presented the theory of stable majority solutions. This majority theory is based on the theory of generalized stable sets as presented in this section. In chapter 7 the theory of generalized stable sets will be used to produce predictions about coalition formation in social choice situations. In general we prefer the theory of generalized stable sets to the theory of generalized optimal choice sets because of its property of generalized internal stability. In fact, the violation of this condition by the theory of generalized optimal choice sets was an important reason for us to construct the theory of generalized stable sets.

Chapter 5

Coalition Formation in Simple Games

5.1 Introduction

So far we studied the procedural aspects of social choice processes. We discussed aggregation porcedures - social choice rules - in general and some packages of properties these aggregation procedures may satisfy (or not). We also studied some particular aggregation procedures, especially the system of majority decision. In this chapter we look at social choice processes in another way. As indicated in the introduction, the essence of politics is winning. Winning or losing - to enforce a social choice that is in accordance with one's preference in some degree or to be forced to accept a social choice that not to some extent at all accords with one's preference - will now be at the center of our attention.

Games in which winning and losing are basic notions are called simple games. Shapley (1967: 248), who has made a great contribution toward the development of the theory of simple games, describes a simple game as "an idealized power structure, a voting system, a legislature, or indeed any constituted procedure for arriving at group decisions." In his paper of (1981), Shapley even uses the terms 'political system' and 'simple game' interchangeably.

It is possible to associate a simple game with a social choice rule

Conversely, a simple game may be represented by a social choice rule[1]. Leaving this aside, however, in this chapter we are primarily interested in the formation of winning coalitions, that is, in coalitions that can enforce a social choice, and not so much in procedural aspects. So, we now abstract away from the nature of the rules for making social choices. The only assumption is that they exist and form an environment for winning and losing.

This abstract approch enables us to formulate coalition theories that are independent of the properties of aggregation. The coalition theories thus formulated apply to any simple game and not only to, e.g., the majority game. In chapter 8, when dealing with coalition formation in social choice games, we will explicitly use the notion of social choice rule again. Of course, the other concepts of social choice theory as presented in chapter 2 and 4 will be used again.

The aim of this chapter is to solve the first problem as stated in the introduction. We first study simple game theory and formulate some already existing political coalition theories as minimum size principle and conflict of interest theory (Axelrod 1970, De Swaan 1973) in terms of simple game theory. Further, we present Peleg's theory of dominated simple games and its associated coalition theory (Peleg 1981, Van Deemen 1989) as well as our theory of centralized policy games with the related theory of balanced coalitions (Van Deemen 1990, 1990a).

In section 2 of this chapter we present the basic concepts of simple game theory as provided in Shapley (1962, 1967, 1981) and Van Deemen (1989, 1990, 1991). In this section we also present and discuss Riker's minimum size theory (Riker 1962, Riker and Ordeshook 1973). In section 3 several types of simple games are studied, namely, veto-games, oligarchic games and dictatorial games. After this, we formulate some theorems that express Arrow's impossibility theorem and some of its variations in terms of simple game theory. In section 4, we pursue our study of extensions of simple game theory. In the then subsequent sections 5-7 the most important classical coalition theories and some new coalition theories will be presented. The first is the theory of dominated simple games. This theory is presented in Peleg (1981) and is further studied in Van Deemen

[1]Cf. Peleg (1984) for association and representation techniques of simple games and social choice rules. Peleg uses, by the way, the functional approach (see chapter 2).

(1989). After this we present our theory of centralized policy games. A first and tentative formulation of this theory can be found in Van Deemen (1987). A more elaborated version is in Van Deemen (1991). After this we present Axelrod's conflict of interest theory. We will formulate a connection with the theory of centralized policy games. After this, in section 8, the theory of balanced coalitions is presented. This theory deals with coalition formation in centralized policy games.

5.2 Basic concepts

The theory of simple games has its origin in the celebrated work on game theory by Von Neumann and Morgenstern (1947: Ch. X). They develop the theory within the general framework of n-person cooperative games. Within this framework, games are described in terms of a characteristic function (consider Luce and Raifa 1957, Riker and Ordeshook 1973, Shubik 1982, Ordeshook 1986) and are therefore called games in characteristic function form[2].

The theory is further developed and refined by Shapley (1962, 1967). However, in contrast with Von Neumann and Morgenstern, Shapley uses set-theoretical concepts without referring to the notion of characteristic function. Besides its relative simplicity and elegance, this has the advantage that it can be presented rather independently from the general cooperative framework (Shapley 1962: 59). For these reasons, we choose for the Shapley approach.

Like in chapter 2, we let N denote a nonempty and finite set of individuals. In this and following chapters, we will call members of N *players* or, sometimes, *actors*. As stated in chapter 2, any subset of N is

---

[2]A characteric function v is a function $v : \mathcal{P}(N) \rightarrow Re$ where Re is the set of reals. This function satisfies two axioms:

1. $v(\{i\}) = 0$ for every $i \in N$.
2. $v(S \cup T) \geq v(S) + v(T)$ if $S \cap T = \emptyset$.

The last axiom is the superadditivity condition. It expresses an incentive to cooperate. A cooperative game in charateristic function form is defined to be an ordered pair (N, v) where N is the set of players and where v is a characteristic function. If $v : \mathcal{P}(N) \rightarrow \{0, 1\}$, then the game (N, v) is called simple. A coalition S with $v(S) = 1$, then, is called winning and a coalition T with $v(T) = 0$ is called losing.

called a *coalition*. Coalitions will be denoted by S, T. If $i \in S$, then the subcoalition $S - \{i\}$ will be called the *internal opposition for* i in S. The complement of a coalition S, that is, the set of all players not in S, will be denoted with S^c. This set will be called *the external opposition for* S.

Game theory starts from the assumption of rational players. This means that a player always will choose the best option available. For the notion of 'rationality' and 'best option', consider again chapter 2.

Let $\mathcal{P}(N)$ denote the power set of N, that is, the set of all coalitions of N.

Definition 5.1 *G is a* simple game := $G = (N, W)$ *where* $W \subseteq \mathcal{P}(N)$ *such that*

1. Monotonicity: *if* $S \subset T$ *and* $S \in W$, *then* $T \in W$,
2. Non-triviality: $W \neq \emptyset$ *and* $\emptyset \notin W$.

In a simple game $G = (N, W)$, a coalition $S \in W$ is said to be *winning*. A coalition that is not winning is called *losing*. The first axiom expresses the fact that a winning coalition cannot change into a losing one by gathering up more members. This axiom is intuitively acceptable[3]. The second axiom precludes trivial games.

Let L denote the set of losing coalitions, that is, let

$$L = \mathcal{P}(N) - W.$$

Definition 5.2 *A simple game* $G = (N, W)$ *is*

1. proper := $S \in W$ *implies* $S^c \in L$,
2. strong := $S \in L$ *implies* $S^c \in W$,
3. decisive := G *is proper and strong*.

In words, a simple game is proper if the complement of any winning coalition is losing. A characterizing feature of proper simple games is that any pair of winning coalitions have some members in common:

[3]This axiom corresponds with the property of superadditivity of characteristic functions as is used in the original framework of Von Neumann and Morgenstern.

Theorem 5.1 *A simple game $G = (N, W)$ is proper if and only if $S \cap T \neq \emptyset$ for all $S, T \in W$.*

Proof.
1) Suppose G is proper and $S, T \in W$ but $S \cap T = \emptyset$. Then $T \subseteq S^c$ and hence - since $T \in W$ - also $S^c \in W$ by monotonicity. Since G is proper, $(S^c)^c = S$ is losing, which contradicts the hypothesis that $S \in W$..
2) Suppose $S \cap T \neq \emptyset$ for every $S, T \in W$ but G is not proper. Then there is an $S \in W$ such that $S^c \notin L$ and hence $S^c \in W$. But then $S \cap S^c = \emptyset$. Contradiction. □.

A game is strong if the complement of any losing coalition is winning. A losing coalition whose complement is also losing is called a *blocking coalition.* Such a coalition is not effective in forcing a decision (since it is losing). However, it can prevent the formation of a winning coalition and with that it can obstruct the decision-making process. A strong simple game is a game in which no blocking coalitions can occur. Therefore no obstruction with respect to the decision-making process can take place in such a game.

Decisive games are the counterpart of the socalled constant-sum games in the general cooperative framework. In fact, Von Neumann and Morgenstern (1947 chapter X) define and study only this class of simple games. Riker's minimum size principle is also formulated for this kind of games[4] (cf. De Swaan 1973). This aspect of Riker's theory is criticized in Grofman (1984). Also see De Swaan (1973).

A *minimal winning coalition* is a winning coalition of which every proper subcoalition is losing.

Definition 5.3 *Let $G = (N, W)$ be a simple game and let S be a coalition. S is minimal winning $:= S \in W$ and $(T \subset S \Rightarrow T \in L)$.*

That is, take out one or more players from a minimal winning coalition and the remaining coalition will lose. The set of minimal winning coalitions

[4]Remarkable enough Riker did not recognize that he was deriving his size principle in the context of simple game theory (cf. Riker 1962: Appendix I; Riker and Ordeshook 1973: chapter 7). With respect to these games he remarks that they are "probably rare in nature" and that "little of practical value is likely to result from studying them" (Riker 1962: 260).

is denoted by W^{min}. As a consequence of definition 5.3:

$$W^{min} = \{S \in W \mid \textit{ there is no } T \in W \textit{such that } T \subset S\}.$$

In words, minimal winning coalitions are the minimal elements of W with respect to proper set inclusion. Because of monotonicity, knowledge of W^{min} is sufficient to specify the whole game.

A *weighted majority game* is a simple game in which to each player a weight is assigned representing the voting strenght or decision-making power of this player. A coalition wins in a weighted majority game if the sum of the weights of the members in this coalition exceeds a certain prescribed number called the *threshold* or *quota* of the game. Formally:

Definition 5.4 *Let $G = (N, W)$ be a simple game where $N = \{1, 2, \ldots, n\}$. G is a* weighted majority game := *there exist a quota $q > 0$ and weights $w_i > 0$, $i \in N$, such that*

$$S \in W \Leftrightarrow \sum_{i \in S} \geq q.$$

The sum $w(S) := \sum_{i \in S} w_i$ is called the size *of coalition S. A weighted majority game $G = (N, W)$ with quota q and weights w_i, $i \in N$, will be denoted by the $n+1$-tuple*

$$[q; w_1, w_2, \ldots, w_n].$$

The first political coalition formation theory is formulated by Riker (1962) and deals, in fact[5], with coalitions in weighted majority games with a prescribed size.

Definition 5.5 *Let $[q; w_1, w_2, \ldots, w_n]$ be the representation of a weighted majority game. A coalition $S \subseteq N$ is of* minimum size :=

1. $S \in W$,

2. $w(S) \leq w(T)$ *for all* $T \in W$.

[5]Cf. De Swaan (1973).

In words, a coalition is of minimum size if it is winning and if its size does not exceed the size of any other winning coalition. Riker's principle of minimum size states that in decisive games only minimum size coalitions will be formed (Riker 1962, Riker and Ordeshook 1973). In Van Deemen (1989: 325), the basic ideas of Riker's principle are formulated as follows:

1. In decision-making situations the payoffs or gains of forming a coalition are divided proportionally to the weights of the members of that coalition.

2. Players are rational, that is, they strive for the highest gain possible.

With the aid of these ideas the formation of minimum size coalitions is easy to explain. If the bribe will be divided proportionally among the members of a coalition, then each member's share will be maximal if that coalition is of minimum size. In chapter 6, we will more subtily explain Riker's principle by power excess theory.

The set of minimum size coalitions will be denoted by W^{size}.

Theorem 5.2

$$W^{size} \subseteq W^{min}.$$

Proof. *If $S \in W^{size}$, then there is no $T \in W$ such that $T \subset S$. Otherwise, S cannot be of minimum size. Hence, $S \in W^{min}$.* □

Any decision-making situation or process in which power is unequally distributed and winning essential can be modelled as a weighted majority game. Examples are parliamentary voting situations, cabinet formation in a multi-party system, the formation of a policy program in political parties, stock holder meetings, authority and obedience relations in the family. For an excellent study of power in weighted majority games including a number of interesting applications, consider Lucas (1983).

A *symmetric game* is a simple game in which winning or losing only depends on the number of players in a coalition. Such a game can be represented as a weighted majority game $[q; 1, 1, \ldots, 1]$. Note that for this kind of simple games, we must have $W^{size} = W^{min}$.

5.3 Weak, Oligarchic and Dictatorial Games

In this section we introduce and study some particular classes of simple games.

Definition 5.6 *Let $G = (N, W)$ be a simple game. A player $i \in N$ is a* dummy *in G :=*

$$i \notin \bigcup W^{min}.$$

That is, i is a dummy if i does not belong to any minimal winning coalition. A dummy is a powerless player. Such a player can neither turn a losing coalition into a winning one, nor turn a winning coalition into a losing one. He is superfluous either for winning or losing. Dummies are readily found among the small parties in a multi-party system.

Definition 5.7 *Let $G = (N, W)$ be a simple game. A player $i \in N$ is a* veto player :=

$$i \in \bigcap W^{min}.$$

In other words, i is a veto player if i belongs to every minimal winning coalition. A veto player is a rather powerful player. Because of monotonicity (definition 5.1), a veto player must be a member of every winning coalition. A coalition cannot win without a veto player. Hence, such a player can obstruct the decision-making process by staying outside any possible winning combination. Clearly, a veto player on its own is a blocking solo-coalition and any losing coalition that contains a veto player is also blocking. Hence, the complement of a losing coalition with the veto player also must be losing. So, a simple game with a veto player can never be strong.

Definition 5.8 *A simple game $G = (N, W)$ is called* weak := *there exists an $i \in N$ such that i is a veto-player.*

The following result is evident. It explicates the structure of the set of winning coalitions in a weak game.

Theorem 5.3 *$G = (N, W)$ is a weak simple game if and only if*

1. *$\emptyset \notin W$ and $W \neq \emptyset$,*

2. *if* $S \in W$ *and* $S \subseteq T$, *then* $T \in W$,

3. $\bigcap W \neq \emptyset$.

A collection of sets with the properties as mentioned in this theorem is sometimes called a *prefilter* (cf. Brown 1976).

Definition 5.9 *Let* $G = (N, W)$ *be a simple game. The coalition* $\bigcap W^{min}$ *is an* oligarchy *in* $G :=$

$$\bigcap W^{min} \in W.$$

In other words, a coalition is an oligarchy if it is the intersection of the set of minimal winning coalitions and if this intersection itself is also winning. Clearly, an oligarchy is a minimal winning coalition. More important, each member of an oligarchy is a veto player by definition. A simple game with an oligarchy will be called an *oligarchic game*. The following theorem deals with the structure of oligarchic games.

Theorem 5.4 $G = (N, W)$ *is an oligarchic game if and only*

1. $\emptyset \notin W$ *and* $W \neq \emptyset$,

2. *if* $S \subset T$ *and* $S \in W$, *then* $T \in W$,

3. *if* $S, T \in W$, *then* $S \cap T \in W$.

A collection of sets with the properties as mentioned in this theorem is also called a *filter* (cf. Brown 1976, Hansson 1975, Kelly 1978). For a mathematical study of filters see van Dalen et.al. (1975: 263-266).

Proof of Theorem 5.4. *We only prove 3). Let* G *be an oligarchic game. Then* $\bigcap W \subset (S \cap T)$ *for every* $S, T \in W$. *Since* $\bigcap W \in W$ *and since* G *is monotonic,* $S \cap T \in W$ *for every* $S, T \in W$.
Conversely, let $S \cap T \in W$ *for every* $S, T \in W$. *Since* $\cap W$ *is a finite intersection of elements of* W, $\cap W \in W$. *By 1),* $\cap W \neq \emptyset$. □.

The study of oligarchies is very old. In fact, Aristotle (384-322 B.C.) in his *The Politics* already spent a lot of energy to this phenomenon. Theorem 5.4 shows the power structure behind oligarchies.

Definition 5.10 *Let $G = (N, W)$ be a simple game. A player $i \in N$ is a* dictator :=

$$\{\{i\}\} = W^{min}.$$

That is, i is a dictator if $\{i\}$ is the only minimal winning coalition. Of course, there can be only one dictator. Note the difference between a veto player and a dictator. A veto player cannot win on his own. In contrast, a dictator need no other players to form a winning coalition. Every other player is a dummy. There is no incentive to form coalitions. A simple game with a dictator is called a dictatorial game.

Theorem 5.5 *$G = (N, W)$ is a dictatorial game if and only*

1. *$\emptyset \notin W$ and $W \neq \emptyset$,*
2. *if $S \subset T$ and $S \in W$, then $T \in W$,*
3. *if $S, T \in W$, then $S \cap T \in W$*
4. *$S \in W$ or $S^c \in W$ for every $S \subseteq N$.*

A collection of sets that satisfies properties 1)-4) as mentioned in this theorem is called an *ultrafilter* (cf. Brown 1976, Hansson 1975, Kelly 1978). This theorem shows the power structure of a dictatorial game.

Proof of Theorem 5.5.
a). Let G be dictatorial and let i be the dictator. We only prove 3) and 4). Proof of 3). Since i is a dictator, $\{i\} \subseteq S$ for every $S \in W$. Hence, for every $S, T \in W : S \cap T \supseteq \{i\}$. Thus $S \cap T \in W$ for every $S, T \in W$. Proof of 4). Let S be a coalition. Then either $S \in W$ or $S \in L$. Let $S \in W$. Then, since i is a dictator, $i \in S$ and hence $i \notin S^c$. Therefore $S^c \in L$. Now, let $S \in L$. Then, since i is a dictator, $i \notin S$ and, hence, $i \in S^c$. But then $S^c \in W$.
b). Let W satisfy the properties 1)-4). We prove: there is an $i \in N$ such that $\bigcap W = \{i\}$ and $\{i\} \in W$. Repeated application of 3) gives $\bigcap W \in W$. Because of 1), $\bigcap W \neq \emptyset$. Let $i \in \bigcap W$ and let $j \in N$ such that $j \neq i$. If $j \in \bigcap W$, then $\{i, j\} \subseteq \bigcap W$ and hence, since it is impossible that $\{i, j\} \subseteq \{i\}$ when $i \neq j$, $\{i\} \notin W$. Then, because of 4), $N - \{i\} \in W$. But then $i \notin \bigcap W$. Contradiction. Hence $j \notin \bigcap W$ and therefore $\bigcap W = \{i\}$.

Clearly, $W^{min} = \{\{i\}\}$. □

5.4 Simple Games and Social Choice

We may formulate Arrow's theorem and some related results in terms of simple game theory. We present some results to that effect. The first theorem is an adjusted version of Arrow's theorem. First we need another definition.

Definition 5.11 *Let* $F \in \Phi$. *A coalition* $S \subseteq N$ *is* winning under F := *for all* $x, y \in X$ *and for every* $p \in \Pi$,

$$\forall i \in S[xP_i^p y] \Rightarrow (xF(p)y \wedge \neg yF(p)x).$$

The set of winning coalitions under F will be denoted by W_F.

Theorem 5.6 *Let* $F \in \Phi_{SWF}$ *and suppose* F *satisfies* Unrestricted Domain, Pareto condition *and* Independence of Irrelevant Alternatives. *Then the ordered pair* (N, W_F) *is a dictatorial game.*

A proof of this theorem can be found in Hansson (1975) and in Kelly (1978, chapter 8).

The following result is the socalled oligarchy theorem. The first version of this result is given by Gibbard.

Theorem 5.7 *Let* $F \in \Phi_{QS}$ *and suppose* F *satisfies* Unrestricted Domain, Pareto condition *and* Independence of Irrelevant Alternatives. *Then the ordered pair* (N, W_F) *is an oligarchic game.*

A proof of this theorem can be found in Suzumura (1983) and in Schwartz (1986).

The last result that expresses an interplay between concentration of decision-making power and collective rationality is the socalled veto-theorem. The origin of this theorem is Mas-Collel and Sonnenschein (1972). The theorem is extensively discussed in Blair and Pollak (1982), Schwartz (1986) and Suzumura (1983). In this context it can be formulated as follows:

Theorem 5.8 *Let* $|X| \geq 4$. *Let* $F \in \Phi_{SDF}$ *and suppose* F *satisfies* Unrestricted Domain, *the* Pareto condition *and* Positive Reponsiveness. *Then the ordered pair* (N, W_F) *is a weak game.*

For a proof, consider Blair and Pollack (1982), Schwartz (1986) or Suzumura (1983).

These results show that there is a relationship between social choice theory and simple game theory. In fact, a social choice rule determines the constraints for a coalition to win or loose. It determines when and how, under which conditions, a coalition may gain decisive power to enforce a social choice. Different social choice rules mean different constraints and therefore different winning and losing opportunities. Therefore, different social choice rules will, in general, lead to different sets of winning coalitions and hence to different simple games. As stated in the introduction to this chapter, we will not further explore the relationship between simple games and social choice rules in this monograph. Our aim is to formulate theories of coalition formation which are independent of the particular constraints of winning and loosing. So, we abstract away from the constraints (or rules) that determine winning and loosing. As already has been noted in the introduction to this chapter, this enables us to formulate general theories that work under any set of constraints of winning and loosing.

5.5 Theory of Dominated Simple Games

This section is based on Van Deemen (1987, 1989). Simple game theory as presented provides a first step to describe and explain the formation of coalitions in political systems. It is a kind of foundation on which other theories can be built. These other theories can be constructed by introducing additional concepts and specific assumptions. In this section we study a theoretical extension formulated by Peleg (1981).

A dominated simple game is a game in which a dominant player calls the tune. It is assumed that this player has the power to act in a decisive way on the coalition formation process which takes place in the game. Accordingly, only winning coalitions in which the dominant player bears the sceptre will be formed.

The terms 'dominant' and 'dominated' as used in this section and in the following chapters are introduced by Peleg himself. Since these terms are also used by other authors (Einy 1985) within this framework, we also use them. These terms, however, should not be confused with the dominance relation among payoff structures as defined and used in chapter 4 of this work and as used in standard n-person cooperative game theory (Shubik 1984).

The concept of a dominant player is derived from a binary relation between coalitions. Peleg calls this relation a desirability relation (Peleg 1980, 1981). This relation expresses the relative strenght of a coalition vis-a-vis another coalition.

Definition 5.12 *Let $G = (N, W)$ be a simple game and let S and T be coalitions.*

1. *S is* at least as desirable as *T, notation $S \succeq_D T$, := for every nonempty coalition $B \subseteq (N - (S \cup T))$:*

$$B \cup T \in W \Rightarrow B \cup S \in W.$$

2. *S is* more desirable than *T, notation $S \succ_D T$, := $S \succeq_D T$ but not $T \succeq_D S$.*

3. *S is* equally desirable as *T, notation $S \approx_D T$, := $S \succeq_D T$ and $T \succeq_D S$.*

The relation $\succ_D$ is the asymmetric part of the desirability relation. If it is nonempty, then it implies that there is a coalition $B \subseteq N - (S \cup T)$ such that S can win with B, while T cannot. The relation $\approx_D$ is the symmetric part of $\succeq_D$. $S \approx_D T$ means that for every coalition $B \subseteq N - (S \cup T)$, S can win with B if and only if T can win with B. Note that if both S and T are winning, then $S \approx_D T$.

Definition 5.13 *Let $G = (N, W)$ be a simple game and let S be a coalition.*

1. *$i \in N$* weakly dominates *S :=*

   (a) $i \in S$,

   (b) $\{i\} \succeq_D S - \{i\}$.

2. $i \in N$ dominates S :=

(a) $i \in S$,

(b) $\{i\} \succ_D S - \{i\}$.

The definition of the desirability relation leads to the following result of which the proof is evident.

Theorem 5.9 *Let $G = (N, W)$ be a simple game. A player $i \in S$ dominates a coalition S if and only if*

1. *for every nonempty coalition $B \subseteq S^c$: if $B \cup (S - \{i\}) \in W$, then $B \cup \{i\} \in W$,*

2. *there is a nonempty $B \subseteq S^c$ such that $B \cup \{i\} \in W$ and $B \cup (S - \{i\}) \in L$.*

The first condition says that if the internal opposition of S vis-a-vis i can make a winning deal with the external opposition, then so can i. Indeed, if i cannot win when joined with some external opposition, then neither can $S - \{i\}$. The second condition states that a player who dominates a coalition has strictly more opportunities to form a winning coalition with the external opposition than the internal opposition of the coalition he dominates.

Definition 5.14 *Let $G = (N, W)$ be a simple game. Player i is called* dominant *in G := there is a coalition $S \in W$ such that i dominates S. A simple game with a dominant player is called a* dominated simple game.

The set of dominant players of a simple game G will be denoted by $d(G)$. The following theorem is proved by Peleg. It states that if i is dominant, there is no other player $j \neq i$ who weakly dominates i whenever the game under consideration is proper.

Theorem 5.10 (Peleg 1981) *Let $G = (N, W)$ be a proper simple game and let $i \in N$. If there are $j \in N$, $j \neq i$, such that $\{j\} \succeq_D \{i\}$, then $i \notin d(G)$.*

The next propositions are easily derived from this theorem. They give an indication how big $d(G)$ might be.

Corollary 5.1 (Peleg 1981) *Let $G = [q; w_1, \ldots, w_n]$ be a representation of a proper weighted majority game. Then there is at most one dominant player.*

If a dominant player exists in a weighted majority game, then it must be the player with the highest weight.

If G is a weak game, every winning coalition S contains every veto player. Hence the complement of every winning coalition must be losing. Thus every weak game is proper. Therefore,

Corollary 5.2 (Peleg 1981) *Let $G = (N, W)$ be a weak game. Then there is at most one dominant player.*

Note that a weak game is not always a dominated game. To prove this we discuss the game representation of the United Nations Security Council (cf. Van Deemen 1989: 319). This council consists of five permanent members (France, Great Britain, China, the Soviet Union and the United States) and ten small countries whose membership is temporary. Each permanent member has a veto. This council can be represented by the game:

$$[39; 7, 7, 7, 7, 7, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1].$$

The first five players are the permanent members. Each of them is a veto player. However, none of them is a dominant player. Note that the permanent members need at least four nonpermanent members to form a winning vote. However, if a dominant player exists in a weak game, then it must be the veto player. This veto player must be unique because of corollary 5.2.

To sketch the control possibilities of a dominant player in a coalition he dominates we introduce the concept of subgame (see Peleg 1981).

Definition 5.15 *Let $G = (N, W)$ be a simple game and let $S \subset N$. A* subgame $G \mid S$ associated with S *is a simple game* (S, W_S), *where S is the players set and W_S is the set of winning coalitions.*

Intuitively, a subgame associated with a coalition is a game played within this coalition. The following theorem is proved in Van Deemen (1989).

Theorem 5.11 *Let $G = (N, W)$ be a dominated and proper simple game and let i be the dominant player. Suppose i dominates S. Then $G \mid S = (S, W_S)$ is dictatorial with i as the dictator.*

This theorem says that within a dominated coalition no decisive subcoalition can be formed against the dominant player. This result together with theorem 5.9 show that a dominant player is really powerful. Firstly, he has more opportunities to form new coalitions with the external opposition. Secondly, he fully controls the internal opposition of the coalition he dominates. No countervailing power is available for the rest of the coalition he dominates.

Definition 5.16 *Let $G = (N, W)$ be a simple game with exactly one dominant player. Let i be the dominant player.*

1. $D_w(G) := \{S \in W \mid i \text{ weakly dominates } S\}$,
2. $D(G) := \{S \in W \mid i \text{ dominates } S\}$.

The set $D_w(G)$ is presented for the first time in Peleg (1981) in which also the following hypothesis is formulated:

Hypothesis 5.1 *(Peleg 1981) Let G be a simple game with exactly one dominant player. Then only coalitions from $D_w(G)$ will be formed.*

The set $D(G)$ is presented in Van Deemen (1987, 1989). Since this set is more restrictive than Peleg's set, it is more interesting. The associated hypothesis is

Hypothesis 5.2 *(Van Deemen 1987, 1989) Let G be a simple game with exactly one dominant player. Then only coalitions from $D(G)$ will be formed.*

5.5.1 Computation Example

To illustrate the working of Peleg's theory we give a real life example. Consider the case of West Germany 1987. The game representation of the parliamentary system after the elections of 1987 is

$$[249; 174, 49, 186, 46, 42].$$

In this game, the players are political parties. The weights represent the number of seats of these parties in parliament. The quota of 249 is the number of seats needed to form a majority cabinet. The parties are, respectively, the Christian Democratic Union (CDU), the Christian Social

Union (CSU), the Social Democratic Party (SDP), the Free Democratic Party (FDP) and the Greens (GR).

We choose this example because it shows that political parties in a parliamentary system together can form one dominant alliance whenever no single party is dominant. In the form presented above this game has no dominant player. However, the parties CDU and CSU form in fact a Christian Democratic alliance. Together they can be conceived of as one actor that happens to be dominant. To see this consider the game representation

$$[249; 223, 186, 46, 42].$$

The first player of this game is the Christian Democratic alliance CDU+CSU. For the rest, this game is the same as the preceding one. Now consider coalition $\{CDU{+}CSU, SDP\}$. Clearly, this coalition is winning. Further, $\{CDU + CSU\} \succ_D \{SDP\}$, since CDU+CSU can win with FDP while SDP cannot. The set of all coalitions in which the Christian Democratic alliance is dominant is

$$\{\{CDU + CSU, SDP\}, \{CDU + CSU, FDP\}, \{CDU + CSU, GR\}\}.$$

According to hypothesis 5.2, one of these coalitions will be formed. Since both $\{CDU{+}CSU\}$ and $\{SDP, FDP\}$ can only win with $\{GR\}$, we have $\{CDU{+}CSU\} \approx_D \{SDP, FDP\}$. Hence, CDU+CSU weakly dominates $\{CDU + CSU, SDP, FDP\}$. Likewise, CDU+CSU weakly dominates $\{CDU + CSU, SDP, GR\}$ and $\{CDU + CSU, FDP, GR\}$. Collecting these coalitions together gives the set D_w :
$\{\{CDU + CSU, SDP\}, \{CDU + CSU, FDP\}, \{CDU + CSU, GR\},$
$\{CDU + CSU, SDP, FDP\}, \{CDU + CSU, SDP, GR\},$
$\{CDU + CSU, FDP, GR\}\}.$
According to Peleg's hypothesis 5.1, one of these coalitions will be formed.

A serious shortcoming of Peleg's theory is that the set of coalitions as predicted (the prediction set) is either empty or rather large. It is empty in the case that a dominant player does not exist. The theory then has nothing to say. In the other cases, the theory is rather poor in content. It will not generate very discriminating predictions. A new aspect of Peleg's coalition theory is that it takes a particular actor (namely the dominant player) as a point of departure. With this, Peleg introduces a new element in the game-theoretical study of coalition formation processes. His approach is actor-oriented. In the next section we will pursue this research line.

5.6 Theory of Centralized Policy Games

This section is based on Van Deemen (1991). In this section, we introduce three political concepts which are, in our view, important in describing political processes. These concepts are 'center' and the related concepts of 'right' and 'left'. These terms are abundantly used in political analysis. For a discussion of the relevance of the center of political systems, consider Daalder (1984).

The terms 'center', 'left' and 'right' will be defined here with the aid of the notion of 'policy position'. This last notion is a primitive term, that is, a term that will be left undefined[6].

Definition 5.17 *A* policy game *is a quadruple* $G_\Theta = (N, W, Pol, \Theta)$ *such that*

1. (N, W) *is a simple game,*
2. $Pol = \{p_i \mid i \in N\}$,
3. $Pol \neq \emptyset$, *and*
4. Θ *is a binary relation on* Pol *satisfying antisymmetry, completeness and transitivity.* Θ *is called the* policy order *for* G_Θ.

[6]Any axiomatic-deductive theory consists of a set *Prim* of primitive or undefined terms; a set *Def* of definitions comprising only these primitive terms; a set *Ass* of assumptions or axioms that use the primitive terms and definitions from *Prim* and *Def*; and a set *Theo* of propositions or theorems each of which is deducible from assumptions in *Ass* or from some previously deduced propositions by means of the rules of logic. In this work, the deduced propositions are called theorems. The terms 'player' and 'winning' as introduced above are examples of primitive terms. The terms 'veto player' and 'dominated simple game' are examples of definitions. The assumption of monotonicity is an example of an axiom. The proposition 'in a symmetric game there is no dominant player' is an example of a deducible proposition or theorem. An introduction to the axiomatic method is Suppes (1957, Ch. 12). A difficult account of this method is given in Braithwaite (1953).

Many theorems in this and the following sections and chapter apply to weighted majority games. Hence, we introduce an additional definition.

Definition 5.18 *A policy game G_Θ where $G = (N, W)$ is a weighted majority game is called a* weighted majority policy game.

A player $i \in N$ is said to be *to the left of* a player $j \in N$ if $p_i \Theta p_j$. Conversely, $i \in N$ is *to the right of* $j \in N$ if $p_j \Theta p_i$. Consider a coalition $S \subseteq N$. Given the properties of Θ, it is possible to assign to each player $i \in S$ a set of players of S who are to the left of i and a set of players of S who are to the right of i. The first set is denoted with $Le(i, S)$, the second set with $Ri(i, S)$. Of course, when there are no players to the left or to the right of i then $Le(i, S)$ or $Ri(i, S)$ are empty.

Definition 5.19 *Let G_Θ be a policy game and let S be a coalition. Then:*

$$Le(i, S) := \{j \in S \mid j \neq i \wedge p_j \Theta p_i\},$$

$$Ri(i, S) := \{j \in S \mid j \neq i \wedge p_i \Theta p_j\}.$$

Note that for no coalition S and for no $i \in N$ it is the case that $i \in Le(i, S)$ or $i \in Ri(i, S)$.

Definition 5.20 *Let G_Θ be a policy game. A player $i \in N$ is a* center player *in G_Θ :=*

1. *$Le(i, N) \in L$ and $Le(i, N) \cup \{i\} \in W$ and*

2. *$Ri(i, N) \in L$ and $Ri(i, N) \cup \{i\} \in W$.*

In words, a player i is a center player if the coalition of all players who are to the left of i are losing without i but winning with i *and* if all the players to the right of i are losing without him but winning with him. A similar definition of a center player can be found in Einy (1985). The set $Le(i, N)$ will be called *the left of G_Θ*, the set $Ri(i, N)$ *the right of G_Θ*. Clearly, i is a center player in G_Θ if no winning coalition from the left or from the right of G_Θ can be formed.

If a center player exists, then he is unique.

Theorem 5.12 *Let G_Θ be a policy game. Then G_Θ has at most one center player.*

Proof. *Suppose i and j, where $i \neq j$, are center players and suppose $p_i \Theta p_j$.*
Because j is a center player, $Le(j, N) \in L$. (1)
Because i is a center player, $Le(i, N) \cup \{i\} \in W$. (2)
But because of $p_i \Theta p_j$, $Le(i, N) \cup \{i\} \subset Le(j, N)$. (3)
From (2) and (3), $Le(j, N) \in W$. Contradiction with (1).
If $i \neq j$ and i, j are center players and $p_j \Theta p_i$, then a contradiction follows in a similar way. Therefore, if i and j are center players, then $i = j$. □.

The following result gives a sufficient condition for the existence of a center player:

Theorem 5.13 *Let $G_\Theta = (N, W, Pol, \Theta)$ be a policy game. Then G_Θ has a center player if $G = (N, W)$ is decisive (i.e. is proper and strong).*

Proof. *Let G be decisive. Let j be a player for which $Le(j, N) = \emptyset$. If $\{j\} \notin W$, take a player k for which $Le(k, N) = \{j\}$. If $\{j, k\} \notin W$, take a player l such that $Le(l, N) = \{j, k\}$. Proceed in this way until a winning coalition $S = \{j, k, l, \ldots, i\}$ has been obtained. Let i be the last added member. By construction, $Le(i, N) \in L$ and $Le(i, N) \cup \{i\} \in W$. Since G is proper, S^c is losing. For each $m \in S^c$, we have $p_i \theta p_m$. Hence $S^c = Ri(i, N)$. Since $S - \{i\} \in L$ and since G is strong, $S^c \cup \{i\} \in W$ and hence $Ri(i, N) \cup \{i\} \in W$..* □

The following definitions are important:

Definition 5.21 1. *A policy game G_Θ with a center player is called a* centralized policy game.

2. *In particular, a weighted majority policy game G_Θ with a center player is called a* centralized weighted majority policy game.

Since a decisive game need not be dominated, a center player is not necessarily a dominant player. Also, if a dominant player exists, it is not necessarily the center player. To see this consider the weighted majority policy game

$$[4; 3, 2, 1, 1].$$

The players are labeled from left to right with a, b, c and d. This is also the order of their policy position. In this game, a can form a winning coalition with c while b cannot. Therefore, a dominates, among others, the coalition $\{a, b\}$ and hence a is the dominant player. However, b is the center player since neither $\{a\}$ nor $\{c, d\}$ is winning, while $\{a, b\}$ and $\{c, d, b\}$ are winning. Further, note that a center player cannot be a dummy.

The following result shows that the unique position of the center player in a weighted majority game is also due to his relative weight. It further shows how closely the notion of a center player is related to De Swaan's notion of 'pivotal player' (De Swaan 1973: 89, 93-4). To distinguish the center player from the other players we will label this player in the sequel with c.

Theorem 5.14 *Let G_Θ be a centralized weighted majority policy game and $c \in N$ be the center player. Then:*

$$| w(Le(c, N)) - w(Ri(c, N)) | < w_c.$$

Proof. *Let c be the center player and suppose $w(Le(c, N)) - w(Ri(c, N)) > 0$. $Le(c, N) \in L$ and $Ri(c, N) \cup \{c\} \in W$. Hence, $w(Le(c, N)) < w(Ri(c, N)) + w_c$, i.e. $w(Le(c, N)) - w(Ri(c, N)) < w_c$. In the same way, it is proven that $w(Ri(c, N)) - w(Le(c, N)) < w_c$ in case that $w(Le(c, N)) - w(Ri(c, N)) < 0$.* □.

This theorem says that the weight of a center player is strictly greater than the absolute value of the difference between the size of the left and the size of the right. Accordingly, a center player is able to hold the balance in the game. He is in the position to form winning coalitions with the right, with the left or with both sides. Further, according to the previous two propositions, the position of a center player is unique. No other player can bent to the left, to the right or to both sides. For these reasons, a center player has more power to control the coalition formation process than any other player in the game. This justifies the following hypothesis.

Definition 5.22 *Let G_Θ be a centralized policy game and let c be the center player. Then*

$$C := \{S \in W \mid c \in S\}.$$

Hypothesis 5.3 *Let G_Θ be a centralized policy game. Then only coalitions from C will be formed.*

Since the set C will be relatively large in most cases, this hypothesis contains little information. To increase the information content, the theory of centralized policy games must be refined further. In the next section we present Axelrod's theory. We shall see there that this theory yields prediction sets contained by C. After this, we present another simple refinement of the theory of centralized policy games, namely, the theory of balanced coalitions. In the subsequent chapters we will present other, more complex theories.

5.7 Conflict of Interest Theory

The source of this theory is Axelrod (1970). The theory is further studied in De Swaan (1971) who calls it *Closed minimal range theory.* We prefer the original name of Axelrod. To formulate this theory, we need some additional concepts.

Definition 5.23 *Let G_Θ be a policy game.*

1. *A player k is* between *players i and j* $:= (p_i \Theta p_k \wedge p_k \Theta p_j) \vee (p_j \Theta p_k \wedge p_k \Theta p_i)$.

2. *Two players i and j are* neighbours := *there is no other player k between i and j.*

3. *A coalition $S \subseteq N$ is* closed := *for all $i \in S$ there is a $j \in S$ such that i and j are neighbours.*

4. *A coalition which is not closed is said to be* open.

The term *closed* is used in De Swaan. Axelrod uses the term *connected.* However, we already used this term in chapter 4. To avoid any confusion, we therefore use De Swaan's term. To illustrate the terms, consider the order $A\,B\,C\,D\,E$, where A, B, C, D and E are players. C is between B and D. The coalition $\{B, C, D\}$ is an example of a closed coalition. The coalition $\{B, D\}$ is open.

The basic assumption of Axelrod's theory is that each player strives to form a coalition in which the conflict of interest is minimal. However, according to Axelrod, this conflict of interest cannot be exactly measured in the case of an ordinal policy scale. What only seems possible is to say that "the less dispersion there is in the policy positions of the members of a coalition, the less conflict of interest there is" (Axelrod: 169). But this only leads to a problem shift. Neither the dispersion of the policy positions on an ordinal scale can be exactly measured, according to Axelrod. Nevertheless, the policy position of the most left player of a coalition and the policy position of the most right player of that coalition provide some information about the dispersion in that coalition. Under some condition this simple form of dispersion in a coalition can be compared with the dispersion in other coalitions. Axelrod illustrates this condition with the aid of the policy order $A\ B\ C\ D\ F\ G$. According to Axelrod, "the dispersion of the coalition consisting of the parties A, B and C cannot be compared to the dispersion of the coalition consisting of B, C and D. However, the closed coalition consisting of A, B and C is "certain to be less dispersed than the coalition consisting of A, B and D" (ibid). Axelrod then says that for that reason, a closed coalition "tends to have relatively low disperion and thus low conflict of interest" (ibid). However, this conclusion cannot be derived from the previous assumptions. As De Swaan (1973: 77) correctly observes, it is unclear why for example the open coalition $\{B, D\}$ should have a lower conflict of interest than the closed coalition $\{B, C, D\}$.

Following the line of Axelrod, a winning coalition that is closed and minimal in the sense " that it can lose no members without ceasing to be [closed] and winning" (Axelrod 1970: 170), would have a minimal conflict of interest. Therefore, according to Axelrod, only such coalitions will be formed. Let us denote the set of closed and winning coalitions of a policy game by W^{cl}.

Definition 5.24 *Let G_Θ be a policy game. A coalition $S \in W$ is* minimal closed := *$S \in W^{cl}$ and for every $i \in S$, $S - \{i\}$ is either losing or open. The set of minimal closed coalitions will be denoted by W^{MC}*

The hypothesis of Axelrod then is:

Hypothesis 5.4 (Axelrod) *Let G_Θ be a policy game. Then only coalitions from W^{MC} will be formed.*

How does this relate to the theory of centralized policy games? The next result shows that a minimal closed coalition always contains the center player.

Theorem 5.15 *Let G_Θ be a centralized policy game. If $S \in W$ and S is closed, then $c \in S$. In particular,*

$$W^{MC} \subseteq C.$$

Proof. *Let c be the center player. Since both $Le(c, N)$ and $Ri(c, N)$ are losing, every closed and winning coalition must contain c. If c is in every closed and winning coalition, then c must also be in every minimal closed coalition.* □

So, Axelrod's theory reduces the set C of a policy game and thereby refines the theory of centralized policy game. As has been indicated in this section, Axelrod's hypothesis is not correctly derived. However, it performs statistically very well when applied to cabinet formation processes in multi-party systems (cf. De Swaan 1973, Taylor and Laver 1973). However, also consider Browne et.al. (1984) who obtain a quite different and in comparison a rather disturbing research result[7].

5.8 Theory of Balanced Coalitions

This section is based on Van Deemen (1990). Fundamental in the theory of balanced coalitions is De Swaan's concept of pivotal player (De Swaan 1973: 89, 93-4).

Definition 5.25 *Let G_Θ be a weighted majority policy game and let S be a coalition. A player i is* pivotal *in $S := i \in S$ and*

$$| \, w(Le(i, S)) - w(Ri(i, S)) \, | \leq w_i.$$

Thus, a player i is pivotal in a coalition S if the absolute value of the difference between the size of the subcoalition of members of S to the left

[7]Different empirical tests give different results. The difficult problem in testing coalition formation theories as presented within the game-theoretical tradition is how to evaluate the often very different research strategies and test procedures that are in use and that may be used. In other words, the problem is how to evaluate the several research designs and methods that may be suitable to test empirically mathematical coalition theories.

of i and the size of the subcoalition of members of S to the right of i is equal to or less than the weight of i.

A pivotal player owes his power in a coalition to the fact that he is able to play off the left side of that coalition against the right side. If the left is in opposition to the right in a coalition and neither side can outvote the other, then the pivotal player can throw out this balance. He then has a decisive influence on the decision making process in that coalition.

5.8.1 Balanced coalitions

Definition 5.26 *Let G_Θ be a centralized weighted majority policy game and let c be the center player. A coalition S is* balanced *in G_Θ :=*

1. *S is winning and*
2. *c is pivotal in S.*

S is nonbalanced *:= S is not balanced. The set of balanced coalitions of G_Θ is denoted by $\mathbf{B}_\Theta$ or, if the context is clear, by $\mathbf{B}$.*

It is easy to verify that $\mathbf{B}$ is not empty.

If each member of such a coalition supports the policy proposals that best accord with his own policy position, then the policy proposal of the center player can never be outvoted in a balanced coalition. His policy position will, therefore, be decisive in such a coalition. For these reasons, it is plausible to assume that a center player will strive to form a balanced coalition.

Hypothesis 5.5 *Let G_Θ be a centralized weighted majority policy game Then only balanced coalitions will be formed.*

5.8.2 Maximally balanced coalitions

Intuitively, some coalitions will be more balanced than others. We formalize this intuition by using the concept of *balance excess*.

Definition 5.27 *Let G_Θ be a centralized weighted majority policy game, c be the center player and $S \in \mathbf{B}$. $bal(S)$ is the* balance excess *of S :=*

$$bal(S) = \mid w(Le(c,S)) - w(Ri(c,S)) \mid .$$

That is, if S is a coalition with center player c, then the balance excess of a coalition S is the absolute value of the difference between the size of the subcoalition of players in S who are to the left of c and the size of the subcoalition of players in S who are to the right of c[8]. The balance excess shows to what extent a coalition with the center player is in equilibrium. The greater the balance excess of a coalition, the easier it is to disturb the equilibrium of that coalition and, hence, the more instable this coalition will be. For this reason, it is plausible to assume that a center player will prefer a coalition with a lower balance excess to a coalition with a greater balance excess.

It is possible to determine the balance excess for each balanced coalition. Therefore, the set **B** can be ordered in a complete and transitive way. That is, for every balanced coalition S and T, it is possible to say whether $bal(S) \leq bal(T)$ or $bal(T) \leq bal(S)$. Further, it must be true that for all $S, T, U \in \mathbf{B}$, if $bal(S) \leq bal(T)$ and $bal(T) \leq bal(U)$, then $bal(S) \leq bal(U)$. A coalition S is said to be *maximally balanced* if S is balanced and there is no $T \in \mathbf{B}$ such that $bal(T) < bal(S)$. The set of maximally balanced coalitions for a policy game G_{Θ} will be denoted by $\mathbf{B}^{\max}$.

Definition 5.28

$$\mathbf{B}^{\max} := \{S \in \mathbf{B} \mid \neg\exists T \in \mathbf{B}[bal(T) < bal(S)]\}.$$

Of course, $\mathbf{B}^{\max}$ is a subset of **B**. Further, thanks to the properties of transitivity and completeness, the set $\mathbf{B}^{\max}$ is not empty[9].

If a center player is rational, he will strive to form a maximally balanced coalition. In such a coalition, he is in the best position to control the policy formation process. Thus,

Hypothesis 5.6 *Let G_{Θ} be a centralized weighted majority policy game. Then only maximally balanced coalitions will be formed.*

What about the preferences of the other players? Clearly, each player prefers a winning coalition in which he is pivotal to a winning coalition

[8]This concept is also used in De Swaan's policy distance theory (see De Swaan 1973: Assumption 5, p.96). There the name 'absolute excess' is used in stead of 'balance excess'.

[9]Cf. chapter 2, section 4 for the definition of maximal elements of a set. Cf. chapter 4 for a study of sets of maximal elements.

in which he is not. Therefore, the players who are not center will prefer coalitions in which the center player is not pivotal. However, the center player can, if the assumption of his control potential is plausible, block the formation of such coalitions. He is able to enforce the formation of maximally balanced coalitions. So the decision problem of the other players is reduced to the question whether they want to participate in a maximally balanced coalition or not. If they are rational, they will. Losing coalitions have nothing to offer.

A maximally balanced coalition is a balanced coalition and a balanced coalition is a winning coalition in which the center player is pivotal. Thus, if G_{θ} is a centralized weighted majority policy game, then $\mathbf{B}^{\max}$ is a subset of $\mathbf{B}$ and $\mathbf{B}$ is a subset of $\mathbf{C}$. The converse, however, is not true. Hence, the theory of maximally balanced coalitions, which yields $\mathbf{B}^{\max}$ as the prediction set, is more restrictive than the theory of balanced coalitions which is, in its turn, more restrictive than the theory of centralized policy games. More restrictive theories are more interesting, since such theories contain more empirical content and are, therefore, easier to falsify. In this sense, the theory of maximally balanced coalitions is the most interesting presented so far.

5.8.3 Closed (maximally) balanced coalitions

This is a variation of the theory of maximally balanced coalitions. The introduction of the notion of closed coalitions is in some sense a logical step within the center player perspective. So far, we assumed that a center player owes his potential of taking the initiative to form coalitions to his position in a relevant policy ranking. From this position, he is able to bend to the left, to the right or to both sides. However, it can be argued that all this has limited validity if the other players have little propensity to form closed coalitions. That is, if the left side or right side players of a center player are indifferent with respect to the open or closed character of coalitions or if they prefer, for some reason or another, open coalitions to closed ones, then they can do pretty well without the center player. They are not, then, inhibited from making policy jumps in order to form winning coalitions. The consequence of this will therefore be a decline of the center player's power potential.

As we have seen, another theory that uses the notion of closed coalitions is Axelrod's conflict of interest theory. De Swaan also formulates a closed version of his policy distance theory (De Swaan 1973: 117-119). In a later chapter on policy distance theory we also will discuss this closed version (see chapter 6). The notion of closed coalitions is especially relevant in the study of cabinet formation processes in multi-party systems. In the database presented in the classical work of De Swaan, 85 of the 108 cabinets are closed. Hence, the notion is not without relevance.

Let G_Θ be a centralized weighted majority policy game. A coalition which is simultaneously closed and (maximally) balanced will be called a *closed (maximally) balanced coalition.* Let $\mathbf{C}_{cl}$ denote the set of closed winning coalitions that contain the center player. It is easily verified that this set is not empty. Clearly, the propensity of the players from the left or the right to form closed coalitions only provides a power base for the center player. It does not imply that only closed coalitions will be formed. The center player might have other preferences. To present a real variation, we therefore have to assume that the propensity of forming closed coalitions is a general behavioral pattern that applies to each player, including the center player.

The relevant hypothesis then are:

Hypothesis 5.7 *Let G_Θ be a centralized weighted majority policy game. Then only closed balanced coalitions will be formed.*

and

Hypothesis 5.8 *Let G_Θ be a centralized weighted majority policy game Then only closed maximally balanced coalitions will be formed.*

Let $\mathbf{B}_{cl}$ and $\mathbf{B}_{cl}^{max}$ denote, respectively, the set of closed balanced coalitions and the set of simultaneously closed and maximally balanced coalitions. Clearly, there are balanced coalitions that are not closed. Therefore, hypothesis 5.7 is more restrictive than the corresponding hypothesis 5.5 in the open version. With respect to hypotheses 5.8 and 5.6 we note that these hypothesss may yield conflicting results. That is, a coalition from $\mathbf{B}_{cl}^{max}$ need not be a member of $\mathbf{B}^{max}$ or conversely.

5.8.4 Computation example: The Dutch election of 6 September 1989

This section is taken over from Van Deemen (1990). Consider the game representation of the Dutch parliament according to the election of 6 september 1989:

$$[76; 6, 49, 12, 54, 22].$$

The parties are, from left to right, GL (Green Left), PvdA (Social Democrats), D66 (Left Liberals), CDA (Christian Democrats) and VVD (Conservative Liberals). The policy positions of these parties are accordingly ordered from left to right. Parties with less than 2.5 % of the total number of votes have been left out. These parties, which are all to the right of the conservative liberals, are dummies that will have no influence on the cabinet formation process.

The CDA is the center party. To check this, take the sum of the weights of the parties which are to the left of CDA. This sum is less than 76. The sum of the parties to the right of the CDA is also less than 76. Hence any combination of parties to the left or to the right of the CDA needs the CDA to form a closed majority cabinet. Neither side can form a majority cabinet on its own. In contrast, the CDA can form a majority cabinet either with parties from the left or with parties from the right. Notice that

$$\mid w(Le) - w(Ri) \mid = 45 < w_{CDA} = 54,$$

where $w(Le)$ and $w(Ri)$ are, respectively, the sizes of left and right for the CDA.

Let us determine the preference of the CDA between two cabinets by using the theory of maximally balanced coalitions. Consider the coalition {CDA, VVD, D66}. This coalition is winning since its size is 54 + 22 + 12 = 88. It is also balanced since VVD is to the right of CDA and D66 is to the left of CDA and $\mid w_{D66} - w_{VVD} \mid < w_{CDA}$. The balance excess for the CDA in this cabinet is $\mid w_{D66} - w_{VVD} \mid = 10$. Compare this with coalition {CDA, VVD}. This also is a winning coalition and again the CDA is the pivotal player. The balance excess is 22. Therefore, according to the theory of maximally balanced coalitions, the CDA will prefer the cabinet {CDA, VVD, D66} to the cabinet {CDA, VVD}. Also compare the {CDA, D66, VVD} combination with the {CDA, PvdA} combination.

Cabinets with center party	Pivotal player	Balance excess
{CDA, VVD}	CDA	22
{CDA, VVD, D66}	CDA	10
{CDA, VVD, D66, PvdA}	CDA	39
{CDA, VVD, D66, PvdA, GL}	CDA	45
{CDA, VVD, PvdA}	CDA	27
{CDA, VVD, GL}	CDA	16
{CDA, VVD, D66, GL}	CDA	4
{CDA, D66, PvdA}	D66	n.d.
{CDA, D66, PvdA, GL}	D66	n.d.
{CDA, PvdA}	CDA	49
{CDA, PvdA, GL}	PvdA	n.d.

Table 5.1: Majority cabinets with the center party CDA. N.d. means 'not defined'.

In the last combination, the balance excess of the CDA is 49. Hence, for the CDA it is far more difficult to keep the balance in this cabinet than in a cabinet with D66 and VVD. However, the CDA prefers a cabinet {CDA, PvdA} to a cabinet {CDA, D66, PvdA}. In this last cabinet, D66 is the pivotal party and, hence, the CDA will prevent the formation of this combination.

The full set of cabinets with the center party CDA is given in table 1.

This table also indicates which of these cabinets is balanced and what the balance excess is of a balanced cabinet[10].

According to hypothesis 5.3[11], one of the cabinets in the first column of this table will be formed. In fact, the combination {CDA, PvdA} has been formed. Hence, this hypothesis is correct for this case.

From table 1, column 2 the set of balanced coalitions can be read off. This set is, in order of increasing balance of excess, { {CDA, VVD, D66, GL}, {CDA, VVD, D66}, {CDA, VVD, GL}, {CDA, VVD}, {CDA, VVD, PvdA}, {CDA, VVD, D66, PvdA}, {CDA, VVD, D66, PvdA,

[10]Since we defined the notion of balance excess only for combinations for which a center player is pivotal (see section 5.8.1), the combinations for which the CDA is not pivotal will not have a balance excess.

[11]Cf. section 5.6.

Closed cabinets with center party	Pivotal player	Balance excess
{CDA, VVD}	CDA	22
{CDA, VVD, D66}	CDA	10
{CDA, VVD, D66, PvdA}	CDA	39
{CDA, VVD, D66, PvdA, GL}	CDA	45
{CDA, D66, PvdA}	D66	n.d.
{CDA, D66, PvdA, GL}	D66	n.d.

Table 5.2: Closed majority cabinets with the center party CDA. N.d. means 'not defined'.

GL}, {CDA, PvdA} }. According to hypothesis 5.5, one of these cabinets will be formed. Since the cabinet {CDA, PvdA} is formed, hypothesis 5.5 is correct for this case.

Hypothesis 5.6 is more restrictive. Looking again at table 5.1, we see that {CDA, VVD, D66, GL} is the cabinet with the least balance excess and, hence, is maximally balanced. The set of maximally balanced cabinets consists only of this cabinet. So the theory of maximally balanced coalitions leads to the unique prediction of the cabinet {CDA, VVD, D66, GL}. Clearly, hypothesis 5.6 fails for this case. Unfortunately, it is more than just a failure. Notice that, according to column 2 of table 1, the formed cabinet {CDA, PvdA} has the greatest balance excess. Hence, it is the most difficult cabinet for the CDA to hold in balance. Therefore, according to the theory of maximally balanced coalitions, there is very little reason for the CDA to form this cabinet.

The picture changes when the notion of closed cabinets is introduced. The set of closed cabinets with the CDA is given in table 2 together with information about their balance excess.

According to hypothesis 5.7, only closed and balanced cabinets will be formed. Hence, the prediction is that one of the first four cabinets in the first column of table 2 will be formed. Since D66 is not a member, the formed cabinet of PvdA and CDA is open and, therefore, this hypothesis must fail for this case.

The closed maximally balanced cabinet is {CDA, VVD, D66}. According to hypothesis 5.8, this cabinet must have been formed. Also this hypothesis fails.

Chapter 6

Coalition Preferences

6.1 Introduction

In the coalition theories as presented in the previous chapter, the players' preferences with respect to winning coalitions are not explicated. In contrast, in this chapter we discuss two theories in which the formation of players' preferences with respect to winning coalitions is explicated. So, each of these theories lead to a preference profile concerning a set of winning coalitions. Both theories then show how to use these players' coalitional preferences in order to obtain a social preference with respect to the set of winning coalitions and how to obtain a solution, that is, a selection of one or more coalitions.

The first theory we discuss in this context is De Swaan's policy distance theory. The theory of De Swaan is, as far as we know, the first theory in which the formation of preferences has been made explicit. The second theory is new. In this theory the notion of power excess is basic. First, we re-examine policy distance theory.

The aim of this chapter is to solve the first part of our problem 2 (cf. chapter 1). The second part of this problem will be dealt with in the following chapter.

6.2 Policy Distance Theory

The origin of policy distance theory can be found in De Swaan (1973). The essence of this theory is "that an actor strives to be included in a winning coalition that he expects to adopt a policy which is as close as possible to his own most preferred policy position" (De Swaan 1973: 88).

The theory can be divided into two parts that correspond with what we have called in chapter 2 the descriptive part and the solution part. The descriptive part of the theory consists of axioms that describe and explain the choice behavior of the players and the expected policy positions of the possible coalitions. These axioms lead to a policy scale on which the policy positions of the possible coalitions are placed and to a preference profile which contains the players' preferences concerning the possible coalitions.

The second part of the theory is a solution part. This part predicts sets of coalitions on the basis of the preference profile constructed in the descriptive part. De Swaan uses a version of the game theoretical core-concept (cf. de Swaan 1973: 103-4) as a solution concept. For an explanation of the core-concept, consider Ordeshook 1986 or Shubik 1984. De Swaan does not prove that his theory always leads to a nonempty core.

De Swaan (1973: Ch. 5) formulates several versions of his policy distance theory. The most important are the open version and the closed one. A coalition is said to be closed if it consists only of players who are adjacent on a policy scale. Otherwise it is called open[1].The open version of policy distance theory predicts both open and closed coalitions; the closed version only closed ones. In de Swaan's work, both versions are confronted with data about nine European parliamentary systems. With respect to the open version of the theory the results of this empirical confrontation are deceptive. It cannot stand the test. With respect to the closed version there is more hope. The results for this version are "not entirely sufficient to accept, but too good to reject the theory" (De Swaan 1973: 153).

In this section, we re-examine De Swaan's policy distance theory. This re-examination will lead to rather negative results. First, it is proven that the open version of policy distance theory is inconsistent. The first who

[1]Cf. chapter 5.7, definition 5.20.

has noticed the inconsistency of the open version is Boute (1981). We also discuss the question how an inconsistent theory can pass a statistical test. At first sight this question seems absurd. From an inconsistent theory any proposition can be derived and thus such a theory is in principle untestable. However, *a posteriori*, this is a luxury point of view, since it only can be said when, indeed, the contradicions are discovered. The fact is that theories may be tested before its contradictions are discovered. The idea that an inconsistent theory always must confront us with its inconsistencies is a naive point of view. Usually it takes time to discover theoretical inconsistencies even in the mathematic field. To give a clear example, mathematicians working with Cantorian set theory in the end of the nineteenth and the beginning of the twentieth century did not know that this theory was inconsistent. They did not meet for example the Russell paradox until Russell discovered it. In practice, contradictions in a theory are not always easy to detect. Also, it is quite possible that the contradictions only are of minor importance and therefore hardly have any influence on the test procedure. Anyway, to find this out, we compute in a detailled way the set of coalitions predicted by the theory (the prediction set) for a "dubious" case in De Swaan's dataset. We then investigate the closed version of the theory. It appears that also this version is inconsistent. As we shall see, this inconsistency is due to the two in essence incompatible ideas on which this version is based.

De Swaan's theory has been a major inspiration source of this work. The in essence negative results do not haggle anything of this. On the contrary. In all honesty, we think that neither inconsistency nor refutation by an empirical test are sufficient grounds to throw a theory away. We agree with the philosopher of science I. Lakatos who writes that " [a] theory can only be eliminated by a *better* theory, that is, by one which has excess empirical content over its predecessors ... (Lakatos 1978: 150, italics in original). Hence, if we want to eliminate policy distance theory we first have to construct better alternatives. This work is both a reflection and a result of this research stratagem. In the next section we formulate a new theory that continues the theoretical research path as paved by De Swaan. This theoretical path will be pursued in the next chapter.

6.2.1 Policy Distance Theory: Open Version

The theory of policy distance theory is formulated for weighted majority games with a relevant policy order (see chapter 5 for these concepts). However, in chapter 5 we defined a policy order as a linear order. Instead, De Swaan defines a policy order as a weak order, that is, as a binary relation over the set of policy positions of the players satisfying reflexivity, completeness and transitivity (Cf. De Swaan 1973: 68, 91). In order to stay as close as possible to De Swaan's formulation of the theory we will assume throughout this section but only in this section that Θ is a weak order. This assumption will be formulated explicitly (see assumption 1 below).

In the rest of this section, all definitions, assumptions and discussion apply to a weighted majority game $G_\Theta = [q; w_1, w_2, \ldots, w_n]$ with policy order Θ. $N = \{1, 2, \ldots, n\}$ is the set of players. As ususal, the set of winning coalitions is denoted by W. That is, $W := \{S \subseteq N \mid \sum_{i \in S} w_i \geq q\}$.

The key idea of the theory is the minimization of the distance between the policy position of a player and the expected policy position of a coalition. The expected policy position of a coalition is supposed to be determined by the weights and the policy positions of the players that are member of that coalition. In words of De Swaan: "it is assumed that all actors have some kind of expectation about the future policy of the coalition and that they base that expectation on the policies and weights of the actors that are members of that coalition" (De Swaan 1973: 91). The expected policy position of a coalition S will be denoted by p_S. In fact, the theory is meant to extend Θ, the weak order of policy orders of the players (or, equivalently, of one-member coalitions), in such a way that the distances between the players and the coalitions can be used in order to construct coalitional preferences.

The following assumption formulated by De Swaan says that the policy positions of the players can be weakly ordered.

Assumption 1 *$G = (W, N, Pol, \Theta)$ is a policy game*[2] *where Θ is a weak order, that is:*

1. for all $p_i, p_j \in Pol$, $p_i \Theta p_j$ or $p_j \Theta p_i$.

[2]See chapter 5 for the definition of policy game.

2. *for all* $p_i, p_j, p_k \in Pol$, *if* $p_i \Theta p_j$ *and* $p_j \Theta p_k$, *then* $p_i \Theta p_k$.

See De Swaan (1973: 91) for this assumption.

Crucial in policy distance theory is the concept of distance[3]. The distance concept used by De Swaan does not satisfy any of the usual properties of a metric. It is *not* a distance function on the set of the policy positions of the players and coalitions in the usual sense of distance function. What then does De Swaan mean with distance? In an earlier chapter, he formulates the following set of statements on the distance of players' policy positions (De Swaan 1973, 69):

- The "distance" between the policy positions of actors i and j is written as "$d(p_i, p_j)$".
- If $p_i \Theta p_j \Theta p_k$, and only then
  $d(p_i, p_j) \leq d(p_i, p_k)$
  and $d(p_i, p_k) \geq d(p_j, p_k)$
  and $d(p_i, p_j)$ and $d(p_j, p_k)$ cannot be compared.

There is a problem with this definition. What is the precise meaning of the expression "if ... and only then"? This is important. If the expression is intended to be an implication, then it is not allowed to say that $p_i \Theta p_j \Theta p_k$ when $d(p_i, p_j) \leq d(p_i, p_k)$ and $d(p_i, p_k) \geq d(p_j, p_k)$. In contrast, if it is an equivalence, then it is allowed. Kleene (1967: 63-4) does not mention the expression in his list of expressions that may be translated by one of the logical connectives. We are inclined to read it as "exactly if", thus as an equivalence. This is supported by the fact that De Swaan sees the statements as a definition of distance (see De Swaan 1973: 90) and by the fact that he uses equivalence on several places in his work (see De Swaan 1973: 95, 105). Later, we give another argument for our interpretation.

De Swaan does not make explicit whether his notion of distance between the policy positions of players may be extended to include the distances between the expected policy positions of coalitions. However,

[3] In his discussion of the concept of distance, De Swaan remarks in a footnote: "It may be seen from the rest of the argument that the notion of "distance" can be dispensed with entirely; its use may help understanding, however." (De Swaan 1973: 69 (footnote)). However, in policy distance theory, the notion is, as the name of the theory already suggests, indispensable. See e.g. assumptions 2 and 4 of the theory as presented here (assumptions 3 and 5 in De Swaan's work).

his intention is clear enough. So we use his definition of distance also for the expected policy positions of the 'more-than-one-member' coalitions.

Essential in the determination of an expected policy position p_S is the notion of pivotal player. This notion is defined in definition 5.22, chapter 5, section 8. Since Θ is supposed to be a weak order according to assumption 1, we give a slightly different definition. Further, we define the basic notion of 'excess[4].

Definition 6.1 *1. Let G_Θ be a weighted majority policy game where Θ is a weak order. Let S be a coalition. A player $i \in S$ is* pivotal *in S* :=

$$| \, w(Le(i,S)) - w(Ri(i,S)) \, | \leq w_i.$$

2. *The difference $e(i,S) := w(Le(i,S)) - w(Ri(i,S))$ is called the* excess *of i in S.*

For the definitions of the sets $Le(i,S)$ and $Ri(i,S)$, consider chapter 5.6, definition 5.17. Thus, a player is pivotal in a coalition S if the absolute value of the difference between the size of the subcoalition of the members from S to the left of i and the size of the subcoalition of the members from S to the right of i is equal to or less than the weight of i. A pivotal player is able to play off the left side of a coalition against the right side. This ability is the root of his power. If the left side of that coalition is in opposition to the right side of that coalition and if neither side can outvote the other, then the pivotal player can throw out this balance. He then has a decisive influence on the decision-making process in that coalition. Hence, coalitions in which a player i is pivotal are important for i. The set of all coalitions in which player i is pivotal is denoted by $\Sigma(i)$.

Definition 6.2

$$\Sigma(i) := \{S \subseteq N \mid i \textit{ is pivotal in } S\}.$$

Let $d(p_S, p_T)$ denote the distance between the policy positions of S and T. If we are dealing with one person, say i, then we write p_i instead of $p_{\{i\}}$. Thus, the distance between the policy position of player i and coalition S is denoted by $d(p_i, p_S)$.

[4]This term is taken over from De Swaan. It should not be confused with 'balance excess' (chapter 5, section 8) or 'power excess' (this chapter).

Assumption 2 *Let G_Θ be a weighted majority policy game. For all $i \in N$, for all $S, T \subseteq N$, if $S \in \Sigma(i)$ and $T \notin \Sigma(i)$, then $d(p_i, p_S) < d(p_i, p_T)$.*

This assumption can be found in De Swaan (1973: 94). According to De Swaan (1973: 95), it

> ""'groups" the expected policies of coalitions with the same pivotal actor, the sets [$\Sigma(k)$], around the policy position of these pivotal actors and in this manner among the sets of coalitions, [Σ_k], the same order prevails as among the actors, k, that are pivotal for the coalitions in those sets."

Note that in the assumption itself distances between players' policy positions and coalitional policy positions are used while in the elucidation only the policy positions are used. Therefore, we again may conclude that 'if ... and only then' in De Swaan's formulation of the distance concept must be read as an equivalence.

The next assumption of policy distance theory can be found in De Swaan (1973: 95).

Assumption 3 *Let G_Θ be a weighted majority policy game. For all $i \in N$, for all $S, T \in \Sigma(i)$,*

$$p_S \Theta^{ex} p_T \Leftrightarrow e(i, S) \geq e(i, T).$$

We must formulate two notes here. First, De Swaan makes no distinction between Θ and the extension of Θ. That is, he uses Θ[5] both to denote a binary relation on $Pol = \{p_i \mid i \in N\}$ and on $Pol^{ex} = \{p_S \mid S \subseteq N\}$. In order to avoid confusion we have introduced Θ^{ex} as the extension of Θ, that is, as a binary relation on $Pol^{ex} = \{p_S \mid S \subseteq N\}$. Of course, Θ^{ex} must include Θ.

Secondly, De Swaan uses strict inequality in the original formulation of this assumption. However, this must be an error. Firstly, Θ is defined as a weak order (see assumption 1). Seeking an extension of Θ in order to place the policy positions of the coalitions, cannot make a linear order out of it since individuals are solo coalitions, i.e. if $p_i \Theta p_j$ then $p_{\{i\}} \Theta p_{\{j\}}$. Secondly, there may be coalitions such that $e(i, S) = e(i, T)$ with $S, T \in \Sigma(i)$. So,

[5] In fact, De Swaan uses the symbol $\leq$. Cf. De Swaan 1973).

ties of the expected policy positions of coalitions cannot be precluded[6]. Assumption 3 orders the policy positions of the coalitions in $\Sigma(i)$ from left to right for each $i \in N$.

The next assumption can be found in (De Swaan 1973: 96):

Assumption 4 *Let G_Θ be a weighted majority policy game where Θ is a weak order. For all $i \in N$, if $S, T \in \Sigma(i)$, then*

$$d(p_i, p_S) \leq d(p_i, p_T) \Leftrightarrow | e(i, S) | \leq | e(i, T) | .$$

Again, De Swaan uses strict inequality, both for distances and for the absolute values in this assumption. Again, this precludes possible ties and hence possible indifference of the players with respect to coalitions. Since this is not the aim of De Swaan, we have adjusted the assumption by using $\leq$. The idea expressed by assumption 4 is plausible. The smaller the absolute excess of a player in a coalition for which he is pivotal, the better he can outvote both sides and hence the better he can throw out the balance. The better he can throw out the balance, the better he can influence the policy making process in that coalition and, hence, the closer the policy position of that coalition will be to his policy position.

With respect to the coalitional preferences of the players De Swaan formulates two assumption. The first assumption is that if $S, T \in W_i$, then i will prefer S to T if the policy position of S is closer to his policy position than that of T.

Assumption 5 *Let G_Θ be a weighted majority policy game. Let $i \in N$ and let S, T be coalitions.* i strictly prefers S to T, *notation $S \mathbf{P}_i T$ if and only if $S, T \in W_i$ and $d(p_i, p_S) < d(p_i, p_T)$.*

This is assumption 1 in De Swaan 1973,p. 91. The second assumption De Swaan formulates with respect to preferences concerning coalitions is that each player prefers a winning coalition to a losing one or to a coalition he is not a member of.

[6]This is in line with the intention of De Swaan. On page 100 of his work he states: " In this manner, assumptions (2) - (4) establish a *complete* and transitive order of all coalitions, including losing coalitions, and of all actors on a single scale." (My emphasis). However, if strict inequality is used in assumption 3, then the expected policy positions of two coalitions S and T such that $e(i, S) = e(i, T)$ and $S, T \in \Sigma(i)$ cannot be scaled, and thus the "order of all coalitions" cannot be complete.

Assumption 6 *Let G_{Θ} be a weighted majority policy game. For all $i \in N$, if $S \in W_i$ and $T \notin W_i$ then $S\mathbf{P}_iT$.*

This is assumption 6 in De Swaan 1973, p. 99. De Swaan does neither define indifference nor weak preference. However, since ties between policy positions of coalitions are not ruled out, it is possible that players are indifferent with respect to some coalitions. In fact, De Swaan uses implicitly an assumption about indifference that we will formulate explicitly in the following way:

Assumption 7 *Let G_{Θ} be a weighted majority policy game. Let $i \in N$ and S and T be coalitions.*

1. *i is* indifferent between S and T, *notation $S\mathbf{I}_iT$, if and only if*
   (a) *$S, T \notin W_i$ or*
   (b) *$S, T \in W_i$ and $d(p_i, p_S) = d(p_i, p_T)$.*
2. *i* weakly prefers S to T, *notation $S\mathbf{R}_iT$, if and only if $S\mathbf{P}_iT$ or $S\mathbf{I}_iT$.*

Note that $d(p_i, p_S) = d(p_i, p_T)$ if the positions of S and T are tied and i is pivotal in S and T.

The assumptions presented sofar lead to a preference profile concerning the set of possible coalitions. That is, the assumptions of policy distance theory imply a preference for each player concerning the possible coalitions.

In the solution part, De Swaan defines a dominance relation on the basis of the preferences as constructed with the assumptions of the descriptive part. In order to find a prediction set, De Swaan uses the core concept for this dominance relation.

Definition 6.3 *Let G_{Θ} be a weighted majority policy game. Coalition S* dominates *coalition T, notation $S\Delta T$,* $:= \forall i \in S[S\mathbf{P}_iT]$. *$S$ is* undominated *:= there is no coalition T such that $T\Delta S$. The* core *is the set of undominated coalitions.*

The core is the set of coalitions which are Δ-maximal. Clearly, if a coalition S belongs to the core, then for all coalitions T there must be $i \in T$ such that $S\mathbf{R}_iT$.

The set $\mu(W, \mathbf{R}_i)$ of maximal choices for an $i \in N$ is the set of coalitions for which there are no better ones according to the preference $\mathbf{R}_i$.

Definition 6.4

$$\mu(W, \mathbf{R}_i) := \{S \subseteq W \mid \neg\exists T \subseteq N[T\mathbf{P}_i S]\}.$$

For a study of sets of maximal choices see chapter 4, section 3. Of course, we would like to prove that $\mu(W, \mathbf{R}_i) \neq \emptyset$ for every $i \in N$ and that this set is the set of minimum distance coalitions for every i. Unfortunately, this is not the case. The theory as presented sofar happens to be inconsistent. The first who has noticed this is, as far as we know, Boute (1981). To prove the inconsistency, we provide a calculation for a game that leads to different policy positions for the same coalitions. This allows the contradictory propositions that for a player i, $d(p_i, p_S) < d(p_i, p_T)$ and $d(p_i, p_T) < d(p_i, p_S)$.

6.2.2 The Inconsistency of the Open Version of Policy Distance Theory

Consider the decisive game $[2; 1, 1, 1]$. The players are, in the given order, a, b and c. Without loss of information, we will restrict ourselves to the set of winning coalitions for this game. Table 6.1 gives the size of each winning coalition, the pivotal players in each winning coalition and the excess of the pivotal player in each winning coalitions. To see how the excesses are calculated, consider coalition $\{a, b\}$. Player a has 0 players to the left of him and 1 player, namely b, to the right of him. Hence the excess of a in $\{a, b\}$ is $0 - 1 = -1$. The remaining excesses can be calculated in the same way. According to this table:
$\Sigma(a) = \{\{a, b\}, \{a, c\}\}$,
$\Sigma(b) = \{\{b, c\}, \{a, b\}, \{a, b, c\}\}$,
$\Sigma(c) = \{\{b, c\}, \{a, c\}\}$.
We prove the contradiction:
Since a is pivotal in $\{a, c\}$ but not in $\{a, b, c\}$, we conclude from assumption 2 that $d(p_a, p_{\{a,c\}}) < d(p_a, p_{\{a,b,c\}})$ and, hence, that the policy position of $\{a, c\}$ lies between those of a and $\{a, b, c\}$. Since $e(b, \{a, b, c\}) = e(b, \{b\}) = 0$, we conclude from assumption 3, $p_b \Theta^{ex} p_{\{a,b,c\}}$ and $p_{\{a,b,c\}} \Theta^{ex} p_b$, i.e., that the policy position of b ties with the policy position of $\{a, b, c\}$.

Winning coalitions	Size	Pivotal player	Excess
$\{a, b\}$	2	a	-1
		b	+1
$\{b, c\}$	2	b	-1
		c	+1
$\{a, c\}$	2	a	-1
		c	+1
$\{a, b, c\}$	3	b	0

Table 6.1: Table for the game $[2; 1, 1, 1]$.

Since $p_{\{a,c\}}$ is to the left of $p_{\{a,b,c\}}$ we conclude that $p_{\{a,c\}}$ is to the left of p_b, i.e. $p_{\{a,c\}}\Theta^{ex}p_b$ and not $p_b\Theta^{ex}p_{\{a,c\}}$ $(*)$.
Since c is pivotal in $\{a, c\}$ but not in $\{a, b, c\}$, we conclude from assumption 2 that $d(p_c, p_{\{a,c\}}) < d(p_c, p_{\{a,b,c\}})$. Hence, since $p_{\{a,b,c\}}$ and p_b are tied, we conclude that $p_{\{a,c\}}$ is to the right of p_b, that is, $p_b\Theta^{ex}p_{\{a,c\}}$ and not $p_{\{a,c\}}\Theta^{ex}p_b$, which is in contradiction with $(*)$.

Note that assumption 5 leads to the anomaly $\{a, b, c\}\mathbf{P_c}\{a, c\}$ and $\{a, c\}\mathbf{P_c}\{a, b, c\}$ for this case.

The source of the inconsistency of De Swaan's theory is not difficult to detect. It is due to the fact that the sets $\Sigma(i)$ with $i \in N$ may not be mutually disjoint. This is, in its turn, caused by the fact that some coalitions may contain more than one pivotal player[7]. And this is allowed by the 'less than that or equal to' relation in the definition of pivotal player[8]. Therefore, a remedy for this inconsistency seems easy to formulate: forbid the possibility of more than one pivotal player for a coalition[9]. This result

[7] Also cf. Boute (1981).

[8] See De Swaan 1973: 94, footnote * for the simple proof.

[9] Boute (1984) does not want to preclude the existence of coalitions with two pivotal actors. He assumes that a nonpivotal actor "follows the point of view of the pivot which is closer to his own position" (Boute 1984: 126). Also see his assumption 10. However, looking at the game as given above, it is easily seen that this does not lead to a solution. The inconsistency with respect to the position of coalition $\{a, c\}$ and therefore the inconsistency of c's preference with respect to this coalition does not disappear when applying Boute's assumption. The only remedy is, in our view, to forbid the existence of coalitions with two pivotal players. Another argument is that any rational player will prefer a coalition in which he is the only pivotal player to any coalition in which he and

can be obtained by using the strict inequality instead of the 'less than or equal to' relation in the definition of pivotal player.

But, as may be guessed, for this interdiction a price must be paid. It is possible that, when using strict inequality, the expected policy positions of some winning coalitions cannot be determined, simply because these coalitions do not have a pivotal player. As a consequence, these coalitions cannot be ordered or scaled and therefore the distances of the players' policy positions and the policy positions of such coalitions cannot be determined. To illustrate this, consider, again, the game $[2; 1, 1, 1]$ (see table 6.1). If we determine the pivotal player of each coalition with the aid of $<$ instead of $\leq$, then the coalitions $\{a, b\}, \{a, c\}$ and $\{b, c\}$ do not have a pivotal player. Since a is a member of $\{a, b, c\}$ but not of $\{b, c\}$, we may infer that a strictly prefers $\{a, b, c\}$ to $\{b, c\}$. Hence, we may conclude that $p_{\{a,b,c\}}$ lies somewhere between p_a and $p_{\{b,c\}}$. But what about $p_{\{a,c\}}$? Is it to the left of $p_{\{a,b,c\}}$ or to the right of it? We do not have sufficient information to answer this question.

Boute (1984) proposes another way to avoid the inconsistency. He wants to make a distinction between an "objective scale" and a "subjective scale" for the policy positions of the players and coalitions. The scale is objective if "every actor places all parties in exactly the same way on the scale". Otherwise, it is subjective (Boute 1984, 114). Then, he formulates in assumption 2a the existence of "objective positions for actors" and in assumption 2b the existence of "objective positions for coalitions" (Boute, 1984: 115). Boute asserts that De Swaan implicitly uses the assumption of objective positions for coalitions, hence assumption 2b. He proposes to maintain the assumption of an objective actors' policy scale, hence assumption 2a, but to "define the expected policy positions of coalitions to be subjective" (ibid, 115). However, Boute neither gives a definition nor an assumption that show how different actors may place the coalitions in a different way on a scale. Consequently it is impossible to construct the subjective policy scales and thus to calculate the actors' coalition preferences on the basis of these scales. Moreover, the policy order for coalitions is a *consequence* of policy distance theory, not an implicit assumption[10].

another player are pivotal, since in a coalition with two pivotal players each of these players can undermine the power position of the other.

[10]Also see page 125 of Boute's work. There he writes: "De Swaan's assumption of a complete ordering of coalitions on the policy scale ...." Again, the completeness (and

Apparently, Boute's proposal neither is a solution to the inconsistency problem of the open version of policy distance theory.

De Swaan statistically tests the open version of his theory with data about cabinet formation processes in nine European multi-party systems (cf. De Swaan 1973: Part II). The question arises how an inconsistent theory can be tested empirically. It is well known that from an inconsistent theory any proposition can be derived. Thus, from a logical point of view, such a theory is untestable (cf. Popper 1959: 92). However, this is an a posteriori point of view, that is, a view that works when the inconsistencies are known. In practice, the inconsistencies first have to be discovered. And it is false to believe that empirical tests can do this job.

We have two arguments for the assertion that empirical tests need not reveal the logical inconsistency of a theory.

1. There may be no data that reveal the contradictions. For example, in the case of De Swaan's policy distance theory the data may be such that there are no coalitions with more than one pivotal player.

2. Even if the data set contains information that leads to contradictions, then still most empirical tests will be blind for this information. Empirical tests are designed to confront the theory with reality. Inconsistency has to do with logic. To discover it requires *logical tests*, not empirical ones.

With respect to the inconsistency of policy distance theory, we may argue in the same way. It is not necessary that the data in De Swaan's dataset lead to contradictions. The data need not lead to coalitions with two pivotal players. Moreover, even when there are cases in which coalitions have two pivotal players, the contradictions need not be revealed by the empirical test employed by De Swaan. Contradictions only can be discovered by logic. Of course, we cannot prove this. However, we can give this view more support by looking at a case in De Swaan's dataset that *a posteriori* surely must lead to contradictions. Let us keep in mind that we already know that the theory is inconsistent.

---

transitivity) of the "scale" of coalitional policy positions is intended to be a *consequence* of policy distance theory, not an assumption. See De Swaan, 1973: 100. Also cf. footnote 6 in this chapter.

The case concerns the data as given for The Netherlands 1952 (De Swaan 1973: 220). The game representation is:

$$[51; 6, 30, 30, 12, 9, 9]$$

The parties are, from left to right,

1. the Communists, henceforward a,
2. the Social Democrats, henceforward b,
3. the Catholics, henceforward c,
4. the Anti-Revolutionaries, henceforward d,
5. the Christian Historicals, henceforward e and
6. the Liberals henceforward f.

The policy positions of the parties are ordered in the indicated way from left to right. The positions of the c and d are tied. Clearly, this game is decisive. Therefore, there are $2^5 = 32$ winning coalitions. However, since the parties with less than 2.5 percent of votes have been left out (cf. De Swaan 1973: 131), there are only 27. The next table contains all winning coalitions including the pivotal players and including the excesses of these pivotal players in the several coalitions. The pivotal player is determined with the aid of the original definition as is also used by De Swaan. In order to detect what exactly happens with the anomalies when computing the prediction set, we present a very detailled computation. Without loss of information we only will work with winning coalitions. The sets of coalitions with pivotal players are:

1. $\Sigma(a) = \emptyset$,
2. $\Sigma(b) = \{\{b,c\}, \{b,c,a\}, \{b,d,e\},$
$\{b,d,e,f\}, \{b,d,f\}, \{b,d,f,a\}, \{b,d,e,a\}, \{b,e,f,a\},$
$\{b,d,e,f,a\}\},$
3. $\Sigma(c) = \{\{b,c\}, \{b,c,d\}, \{b,c,e\}, \{b,c,f\},$
$\{b,c,d,f\}, \{b,c,d,e\}, \{b,c,e,f\}, \{b,c,d,a\}, \{b,c,e,a\},$
$\{b,c,a,f\}, \{b,c,d,e,f\}, \{b,c,d,e,f,a\}, \{c,d,e\}, \{c,d,f\},$
$\{c,d,e,f\}, \{c,d,e,a\}, \{c,d,f,a\}, \{c,d,e,f,a\}, \{c,e,f,a\}\},$

Winning coalition	Pivotal player	Excess
$\{b,c\}$	b	-30
	c	+30
$\{b,c,d\}$	c	+18
$\{b,c,e\}$	c	+21
$\{b,c,f\}$	c	+21
$\{b,c,a\}$	b	-24
$\{b,c,d,f\}$	c	+9
$\{b,c,d,e\}$	c	+9
$\{b,c,e,f\}$	c	+12
$\{b,c,d,a\}$	c	+24
$\{b,c,e,a\}$	c	+27
$\{b,c,a,f\}$	c	+27
$\{b,c,d,e,f\}$	c	0
$\{b,c,d,e,f,a\}$	c	+6
$\{b,d,e\}$	b	-21
$\{b,d,f\}$	b	-21
$\{b,d,e,f\}$	b	-30
	d	+12
$\{b,d,e,a\}$	b	-15
$\{b,d,f,a\}$	b	-15
$\{b,e,f,a\}$	b	-12
$\{b,d,e,f,a\}$	b	-24
$\{c,d,e\}$	c	-21
$\{c,d,f\}$	c	-21
$\{c,d,e,f\}$	c	-30
	d	+12
$\{c,d,e,a\}$	c	-15
$\{c,d,f,a\}$	c	-15
$\{c,d,e,f,a\}$	c	-24
$\{c,e,f,a\}$	c	-12

Table 6.2: Table for the Netherlands 1952

4. $\Sigma(d) = \{\{b, d, e, f\}, \{c, d, e, f\}\}$,

5. $\Sigma(e) = \emptyset$,

6. $\Sigma(f) = \emptyset$,

Assumption 3 orders the coalitions in the sets $\Sigma(i)$ for party i by means of the notion of excess. The greater the excess, the more left the policy position of the coalition is. We first order each set of coalitions for which a party is pivotal. To avoid itching notation, we write S instead of p_S. We put the excess of the concerned player in the relevant coalition in subscript between parentheses. Strict ordering is denoted by $\prec$, ties are denoted by $\approx$.

- $\Sigma(a) : a$
- $\Sigma(b) : b \prec \{b, e, f, a\}_{(-12)} \prec \{b, d, e, a\}_{(-15)} \approx \{b, d, f, a\}_{(-15)}$
  $\prec \{b, d, e\}_{-(21)} \approx \{b, d, f\}_{(-21)} \prec \{b, c, a\}_{(-24)} \approx \{b, d, e, f, a\}_{(-24)}$
  $\prec \{b, c\}_{(-30)}$
  $\approx \{b, d, e, f\}_{(-30)}$
- $\Sigma(c) : \{b, c\}_{(30)} \prec \{b, c, e, a\}_{(27)} \approx \{b, c, f, a\}_{(27)} \prec \{b, c, d, a\}_{(24)}$
  $\prec \{b, c, e\}_{(21)} \approx \{b, c, f\}_{(21)} \prec \{b, c, d\}_{(18)} \prec \{b, c, e, f\}_{(12)}$
  $\prec \{b, c, d, f\}_{(9)}$
  $\approx \{b, c, d, e\}_{(9)} \prec \{b, c, d, e, f, a\}_{(6)} \prec \{b, c, d, e, f\}_{(0)}$
  $\approx c \prec \{c, e, f, a\}_{(-12)} \prec \{c, d, e, a\}_{(-15)} \approx \{c, d, f, a\}_{(-15)}$
  $\prec \{c, d, e\}_{(-21)} \approx \{c, d, f\}_{(-21)} \prec \{c, d, e, f, a\}_{(-24)} \prec \{c, d, e, f\}_{(-30)}$
- $\Sigma(d) : \{b, d, e, f\}_{(12)} \approx \{c, d, e, f\}_{(12)} \prec d$
- $\Sigma(e) : e$
- $\Sigma(f) : f$

According to assumption 2, $d(p_{x_i}, p_S) < d(p_{x_i}, p_T)$ with $x_i = a, b, c, d, e, f$, if $S \in \Sigma(x_i)$ and $T \notin \Sigma(x_i)$. In plain language, the distance between a party and a coalition for which this party is pivotal is smaller than the distance between this party and a coalition for which it is not pivotal. This allows to put side by side the sets $\Sigma(i)$ and subsequently to stick together the ordering of coalitions in those sets. This leads to the following ordering

of coalitional policy positions:
$a \prec b \prec \{b,e,f,a\} \prec \{b,d,e,a\} \approx \{b,d,f,a\}$
$\prec \{b,d,e\} \approx \{b,d,f\} \prec \{b,c,a\} \approx \{b,d,e,f,a\}$
$\prec \{b,c\} \approx \{b,d,e,f\} \prec \{b,c\} \prec \{b,c,e,a\}$
$\approx \{b,c,f,a\} \prec \{b,c,d,a\} \prec \{b,c,e\} \approx \{b,c,f\}$
$\prec \{b,c,d\} \prec \{b,c,e,f\} \prec \{b,c,d,f\} \approx \{b,c,d,e\}$
$\prec \{b,c,d,e,f,a\} \prec \{b,c,d,e,f\} \approx c \prec \{c,e,f,a\}$
$\prec \{c,d,e,a\} \approx \{c,d,f,a\} \prec \{c,d,e\} \approx \{c,d,f\}$
$\prec \{c,d,e,f,a\} \prec \{c,d,e,f\} \prec \{b,d,e,f\}$
$\approx \{c,d,e,f\} \prec d \prec e \prec f.$
Clearly, a number of coalitions are "double-scaled", namely, $\{b,c\}$, $\{b,d,e,f\}$ and $\{c,d,e,f\}$. This is unavoidable, since these coalitions contain two different pivotal players. This leads to a number of anomalies, for example,
$d(b,\{b,d,e,f) < d(b,\{b,c\})$ and $d(b,\{b,c\}) < d(b,\{b,d,e,f\})$. However, we do as if we are blind for these anomalies and proceed in a consequent way.

Assumption 4 allows the comparison of the distance between the policy position of a party and the policy position of a coalition for which that party is pivotal and the distance of that party between a coalition for which it is pivotal and that is on its other side. The measure for this comparison is the absolute value of the excess of the concerned pivotal party in the relevant coalitions. If the absolute value of the excess of a player in a coalition is smaller, then the distance to this coalition will be smaller. By definition of preference this coalition will, then, be more preferred. For example, consider the position of the Catholics (c) in the constructed ordering. Coalition $\{b,c,d\}$ is to the left of c while coalition $\{c,d,e,a\}$ is to the right of it. The Catholic party is pivotal in both coalitions. To determine the preference between these coalitions look at the absolute value of the excess of this party in the coalitions. We find that $\mid e(c,\{c,d,e,a\}) \mid= 15$ and that $\mid e(c,\{b,c,d\} \mid= 18$. Hence, according to assumption 4, $d(c,\{c,d,e,a\}) < d(c,\{b,c,d\})$. Applying the definition of preference we obtain $\{c,d,e,a\}\mathbf{P_c}\{b,c,d\}$. The distances between a party and coalitions which are on either side of that party but for which this party is not pivotal cannot be compared. Hence, the preference between such a pair of coalitions cannot be determined for this party.

By using assumption 4 together with the position of the coalitions in the constructed ordering and the definition of preference the coalitional preference of each party can be determined. We only construct the preferences for W_i, where i is one of the parties. These coalitional preferences are:

1. Communists (a):
$\{b,e,f,a\}\mathbf{P}\{b,d,e,a\}\mathbf{I}\{b,d,f,a\}$
$\mathbf{P}\{b,c,a\}\mathbf{I}\{b,d,e,f,a\}\mathbf{P}\{b,c,e,a\}$
$\mathbf{I}\{b,c,f,a\}\mathbf{P}\{b,c,d,a\}\mathbf{P}\{b,c,d,e,f,a\}$
$\mathbf{P}\{c,e,f,a\}\mathbf{P}\{c,d,e,a\}\mathbf{I}\{c,d,f,a\}$
$\mathbf{P}\{c,d,e,f,a\}$.

2. Social Democrats (b):
$\{b,e,f,a\}\mathbf{P}\{b,d,e,a\}\mathbf{I}\{b,d,f,a\}$
$\mathbf{P}\{b,d,e\}\mathbf{I}\{b,d,f\}\mathbf{P}\{b,c,a\}$
$\mathbf{I}\{b,d,e,f,a\}\mathbf{P}\{b,c\}\mathbf{I}\{b,d,e,f\}$
$\mathbf{P}\{b,c\}\mathbf{P}\{b,c,e,a\}\mathbf{I}\{b,c,f,a\}$
$\mathbf{P}\{b,c,d,a\}\mathbf{P}\{b,c,e\}\mathbf{I}\{b,c,f\}$
$\mathbf{P}\{b,c,d\}\mathbf{P}\{b,c,e,f\}\mathbf{P}\{b,c,d,f\}$
$\mathbf{I}\{b,c,d,e\}\mathbf{P}\{b,c,d,e,f,a\}\mathbf{P}\{b,c,d,e,f\}$
$\mathbf{P}\{b,d,e,f\}$.

3. Catholics (c):
$\{b,c,d,e,f\}\mathbf{P}\{b,c,d,e,f,a\}\mathbf{P}\{b,c,d,e\}$
$\mathbf{I}\{b,c,d,f\}\mathbf{I}\{b,c,e,f\}\mathbf{I}\{c,e,f,a\}$
$\mathbf{P}\{c,d,e,a\}\mathbf{I}\{c,d,f,a\}\mathbf{P}\{b,c,d\}$
$\mathbf{P}\{c,d,e\}\mathbf{I}\{c,d,f\}\mathbf{I}\{b,c,f\}$
$\mathbf{I}\{b,c,e\}\mathbf{P}\{c,d,e,f,a\}\mathbf{I}\{b,c,d,a\}$
$\mathbf{P}\{b,c,f,a\}\mathbf{I}\{b,c,e,a\}\mathbf{P}\{b,c\}$
$\mathbf{I}\{c,d,e,f\}$
1)$\mathbf{P}\{b,c\}\mathbf{P}\{b,c,a\}$.
2)$\mathbf{P}\{c,d,e,f\}$.

4. Anti-Revolutionaries (d):
$\{c,d,e,f\}\mathbf{I}\{b,d,e,f\}\mathbf{P}\{c,d,e,f\}$
$\mathbf{P}\{c,d,e,f,a\}\mathbf{P}\{c,d,f\}\mathbf{I}\{c,d,e\}$
$\mathbf{P}\{c,d,f,a\}\mathbf{I}\{c,d,e,a\}\mathbf{P}\{b,c,d,e,f\}$
$\mathbf{P}\{b,c,d,e,f,a\}\mathbf{P}\{b,c,d,e\}\mathbf{I}\{b,c,d,f\}$

$\mathbf{P}\{b,c,d\}\mathbf{P}\{b,c,d,a\}\mathbf{PP}\{b,d,e,f\}$
$\mathbf{P}\{b,d,e,f,a\}\mathbf{P}\{b,d,f\}\mathbf{I}\{b,d,e\}$
$\mathbf{P}\{b,d,f,a\}\mathbf{I}\{b,d,e,a\}$.

5. Christian Historicals (e):
$\{c,d,e,f\}\mathbf{I}\{b,d,e,f\}\mathbf{P}\{c,d,e,f\}$
$\mathbf{P}\{c,d,e,f,a\}\mathbf{P}\{c,d,e\}\mathbf{P}\{c,d,e,a\}$
$\mathbf{P}\{c,e,f,a\}\mathbf{P}\{b,c,d,e,f\}\mathbf{P}\{b,c,d,e,f,a\}$
$\mathbf{P}\{b,c,d,e\}\mathbf{P}\{b,c,e,f\}\mathbf{P}\{b,c,e\}$
$\mathbf{P}\{b,c,e,a\}\mathbf{P}\{b,d,e,f\}\mathbf{P}\{b,d,e,f,a\}$
$\mathbf{P}\{b,d,e\}\mathbf{P}\{b,d,e,a\}$
$\mathbf{P}\{b,e,f,a\}$.

6. Liberals (f):
$\{c,d,e,f\}\mathbf{I}\{b,d,e,f\}\mathbf{P}\{c,d,e,f\}$
$\mathbf{P}\{c,d,e,f,a\}\mathbf{P}\{c,d,f\}\mathbf{P}\{c,d,f,a\}$
$\mathbf{P}\{c,e,f,a\}\mathbf{P}\{b,c,d,e,f\}\}\mathbf{P}\{b,c,d,e,f,a\}$
$\mathbf{P}\{b,c,d,f\}\mathbf{P}\{b,c,e,f\}\mathbf{P}\{b,c,f\}$
$\mathbf{P}\{b,c,f,a\}\mathbf{P}\{b,d,e,f\}\mathbf{P}\{b,d,e,f,a\}$
$\mathbf{P}\{b,d,f\}\mathbf{P}\{b,d,f,a\}\mathbf{P}\{b,e,f,a\}$.

Note the several anomalies in these preferences. For example, the Catholics strictly prefer $\{b,c\}$ to, indeed, $\{b,c\}$ and $\{c,d,e,f\}$ to $\{c,d,e,f\}$. However, we stick to our task and continue the computation. The last task is to determine the prediction set. Therefore the core-concept as defined in definition 6.3 must be applied. De Swaan (1973: 107) argues:

> "In order to determine what coalitions are undominated, it is simplest to eliminate those that are dominated. A coalition - or, more precisely, the preference vector that goes with it - is dominated when there exists some other coalition all the members of which are better off than they would be if the former were to form."

However, some care must be taken with eliminating dominated coalitions since dominated coalitions may dominate other coalitions. Thus the procedure is as follows: take a coalition S and compare it with other coalitions until one is found that according to the preferences of its members is better than S. Then S is dominated. If no other coalition is found that is better, then S is undominated and it then belongs to the prediction set. We assume that this procedure only takes the first occurrence of a coalition in a preference. We are not sure whether De Swaan uses this assumption, but it seems plausible[11]. Why looking for more when only one is expected? To illustrate, take a look at the preference of the Social Democrats (b). In this preference, the coalition $\{b, c\}$ occurs two times. The procedure only takes the first time it occurs in consideration and neglects the other one. In this way the anomalies are not discovered.

To see how the procedure works, consider coalition $\{b, c\}$ and compare it with the coalitions as presented in table 6.2 until one is founded that dominates it. Going from top to bottom in table 6.2, we take coalition $\{b, c, d\}$. Clearly, d strictly prefers this coalition to $\{b, c\}$. Party b strictly prefers $\{b, c\}$ to $\{b, c, d\}$, while c strictly prefers $\{b, c, d\}$ to $\{b, c\}$. Hence, neither $\{b, c\}$ dominates $\{b, c, d\}$, nor $\{b, c, d\}$ dominates $\{b, c\}$. Now take coalition $\{b, c, e\}$ and compare it with $\{b, c\}$. Again we find that neither $\{b, c, e\}$ dominates $\{b, c\}$, nor $\{b, c\}$ dominates $\{b, c, e\}$. Proceeding in this way we arrive at coalition $\{b, d, e\}$. Now b strictly prefers $\{b, d, e\}$ to $\{b, c\}$. Clearly, also d and e strictly prefer $\{b, d, e\}$ to $\{b, c\}$. Hence, since every player in $\{b, d, e\}$ strictly prefers $\{b, d, e\}$ to $\{b, c\}$, $\{b, d, e\}$ dominates $\{b, c\}$ and therefore $\{b, c\}$ is dominated. Proceeding in this way we find the following undominated coalitions: $\{b, c, d, f\}$, $\{b, c, d, e\}$, $\{b, c, e, f\}$, $\{b, c, d, e, f\}$, $\{b, c, d, e, f, a\}$, $\{c, d, e, f\}$ and $\{c, e, f, a\}$. Thus the prediction set contains these seven coalitions.

Sticking to the job, we have computed a nonempty prediction set for a case for which the theory clearly yields contradictory propositions. It is quite possible to compute a nonempty prediction set without noticing that there are anomalies. When, a priori, inconsistencies are not expected, we may not hope that an empirical test may reveal them.

There are other cases in De Swaan's dataset that must lead to contradictions (e.g. Finland 1924).

[11]It is impossible to conclude from the computing example given by De Swaan (1973, especially 107-8) or from his discussion of the core concept whether he uses this assumption.

6.2.3 Policy Distance Theory: Closed Version

The notion of closed coalition has been extensively discussed in chapter 5[12]. We do not discuss it here again. In the open version of policy distance theory, assumption 6 states that for every $i \in N$ and every $S, T \subseteq N$:

$$S\mathbf{P}_i T \Leftrightarrow S \in W_i, T \notin W_i.$$

In the closed version of his theory, De Swaan substitutes this assumption for a version in which closed coalitions are basic (see De Swaan 1973: 117-119).

Assumption 8 *Let G_Θ be a weighted majority game with relevant policy order Θ. Let W^{cl} denote the set of all winning coalitions in G_Θ which are closed according to Θ and let*

$$W_i^{cl} = \{S \in W \mid S \in W^{cl}, i \in S\}.$$

Then $\forall i \in N$ and $S, T \subseteq N$:

$$S \in W_i^{cl}, T \notin W_i^{cl} \Rightarrow S\mathbf{P}_i T.$$

We shall prove that this assumption is inconsistent with respect to the other assumptions of policy distance theory. First we shall proceed in a logical way, that is, without performing any computation. After this, we illustrate the contradiction on the basis of a concrete case.

1. *Contradiction 1.*
   Let $S \in W_i^{cl}$ and $T \in (W_i - W_i^{cl})$. Let $S, T \in \Sigma(i)$ and suppose $\mid e(i,T) \mid < \mid e(i,S) \mid$. Then by assumption 8, $S\mathbf{P}_i T$ so that, by assumption 5, $d(p_i, p_S) < d(p_i, p_T)$. However, since $\mid e(i,T) \mid < \mid e(i,S) \mid$, $d(p_i, p_T) < d(p_i, p_S)$ by assumption 4. Contradiction.
   In words: if

   (a) player i is in S and S is a closed winning coalition and

   (b) player i is in T and T is a winning but not closed,

   then i strictly prefers S to T by assumption 8, and hence $d(p_i, p_S) < d(p_i, p_T)$ by assumption 5. However, if

[12]Cf. 5.7 and 5.8.3.

(a) i is pivotal in S and in T and the absolute value of the excess of i in T is strictly less than the absolute value of the excess of i in S,

then $d(p_i, p_T) < d(p_i, p_S)$ by assumption 4.

2. *Contradiction 2.*
Let $S \in W_i^{cl}$, $T \in W_i - W_i^{cl}$. Let $T \in \Sigma(i)$ and $S \notin \Sigma(i)$. Then $S\mathbf{P}_i T$ by assumption 8 and hence, by assumption 5, $d(p_i, p_S) < d(p_i, p_T)$. By assumption 2, since i is pivotal in T but not in S, $d(p_i, p_T) < d(p_i, p_S)$. Contradiction.
In words: if

(a) player i is in S and S is a closed winning coalition,

(b) player i is in T and T is winning but not closed,

then $S\mathbf{P}_i T$ by assumption 8, and hence $d(p_i, p_S) < d(p_i, p_T)$. However, if

(a) i is pivotal in T but not in S, then $d(p_i, p_T) < d(p_i, p_S)$ by assumption 2. Contradiction.

We will illustrate these contradictions on the basis of a real-life example. Consider the results of the elections of 21 may 1986 for the Netherlands

$$[76; 3, 52, 9, 54, 27, 5].$$

The parties are, from left to right, SL (Small Left), PvdA (Social Democrats), D66 (Democrats), CDA (Christian Democrats), VVD (Liberals) SR (Small Right). Clearly, this game is decisive. Therefore there are $2^5 = 32$ winning coalitions. However, to illustrate the contradictions we do not need to compute every winning coalition.

1. *Illustration of contradiction 1.*
Consider coalitions $\{CDA, VVD\}$ and $\{SL, D66, CDA, VVD\}$. Since $w(\{CDA, VVD\})$ and $w(\{SL, D66, CDA, VVD\})$ are , respectively, 81 and 93, both coalitions are winning. The CDA is pivotal player in both coalitions. We have: $\mid e(CDA, \{CDA, VVD\}) \mid = 27$ and $\mid e(CDA, \{SL, D66, CDA, VVD\}) \mid = 15$.

Hence, (ommitting the p in order to avoid itching notation), $d(CDA, \{SL, D66, CDA, VVD\}) < d(CDA, \{CDA, VVD\})$ according to assumption 4. Therefore the CDA strictly prefers $\{SL, D66, CDA, VVD\}$ to $\{CDA, VVD\}$. Now, the coalition $\{CDA, VVD\}$ is closed but the coalition $\{SL, D66, CDA, VVD\}$ is open since the party $PvdA$ that is between party $D66$ and SL is not a member. Therefore, by assumption 8, the CDA strictly prefers coalition $\{CDA, VVD\}$ to coalition $\{SL, D66, CDA, VVD\}$ which is a contradiction.

2. *Illustration of contradiction 2.*
   Consider coalition $\{PvdA, D66, CDA\}$ and coalition $\{PvdA, D66, VVD\}$. Both coalitions are winning. $D66$ is the pivotal player for the first coalition and $PvdA$ for the second coalition. Hence, by assumption 2,
   $d(PvdA, \{PvdA, D66, VVD\}) < d(PvdA, \{PvdA, D66, CDA\})$
   and therefore $\{PvdA, D66, VVD\}\mathbf{P}_{\mathbf{PvdA}}\{PvdA, D66, CDA\}$.
   However, $\{PvdA, D66, CDA\}$ is closed but
   $\{PvdA, D66, VVD\}$ is open. Therefore,
   $\{PvdA, D66, CDA\}\mathbf{P}_{\mathbf{PvdA}}\{PvdA, D66, VVD\}$ by assumption 8 and, hence,
   $d(PvdA, \{PvdA, D66, CDA\}) < d(PvdA, \{PvdA, D66, VVD\})$,
   which is a contradiction.

The inconsistency of the closed version of De Swaan's theory is due to the fact that assumption 7 with respect to closed coalitions expresses an idea that is irreconcilable with the basic idea of policy distance theory. The basic idea of policy distance theory is that a player strives to form a coalition with a policy position which is as close as posible to his own position. The new introduced assumption 7, however, says that a player prefers a closed coalition to an open one. The point is that it is not precluded by the theory that an open coalition has an expected policy position that is closer to the position of a player than the expected policy position of any closed coalition with that player. In this case, assumption 7 and the other assumptions of the theory clash. The original assumptions, with the exception of the subsituted one, say to choose for the coalition with the minimum distance from the set of possible coalitions. However, the added assumption 7 says to choose then for the closed coalition. The idea of minimal distance coalitions and the idea of closed coalitions as formulated by De Swaan are in this sense incompatible.

Because of its fundamental character it is very difficult to repair the inconsistency of the closed version. A first thought is to partition W_i in the sets W_i^{cl} and W_i and to put a priority on the set W_i^{cl}. That is, define the assumption that every player prefers any coalition in W_i^{cl} to any coalition not in that set, irrespective of whether it is pivotal in such a closed coalition or not. Then make a partition of W_i^{cl} into a set of closed coalitions in which i is pivotal and a set of closed coalitions in which i is not pivotal. Then apply the assumptions of policy distance theory first to this partitioning and subsequently to the set of remaining open coalitions. That is, order the set of closed coalitions in which i is pivotal on the basis of excess and use the assumption that the distance between i and a closed coalition for which i is pivotal is smaller than the distance between i and a closed coalition for which i is not pivotal. Then do the same with respect to the open coalitions. Note that this implies that a player may prefer a closed coalition in which he is not pivotal to an open coalition in which he is pivotal.

However, this reparation procedure leads to another inconsistency. Since a closed coalition $S \in W$ will, in general, occur in every W_i^{cl} when $i \in S$, and since not every $i \in S$ will be pivotal in S, the conclusion is that p_S must be scaled several times. And this leads to anomalies with respect to distance.

With respect to the open version of policy distance theory we have tried to show how small the basis is to free the theory from the source of inconsistency. With a plausible adjustment, i.e., with leaving coalitions with more than one pivotal player outside the calculations, the theory becomes incomplete. It is no longer possible to place all coalitions on the extended policy scale. With respect to the closed version, we must conclude that there is a clash of ideas. We did not found any room for reparation in this case. One of the two ideas must be dropped. However, dropping the minimum distance idea leads to a rest assumption that may work on its own[13] that, altogether, is a real impoverishment of the theory.

[13]In fact, De Swaan also tests this 'rest assumption' in his work. There, he calls it the "the closed coalition proposition". His conclusion is that "though the results are not clearly insignificant, they are insufficient to accept the theory as an independent explanatory construct." (De Swaan 1973: 155).

Dropping the idea of closed coalitions leads to the open version including the indicated problems. Apparently, the policy that remains is trying to find better theories.

6.3 Power Excess Theory

The following part is based on Van Deemen (1991). The basic idea of this theory is that each player seeks to form coalitions in which the size of the internal opposition is minimal. He can then maximally exert his control potential and hence maximally influence the decision-making processes in such a coalition.

The theory will first be presented for coalition formation processes in a non-policy context and without employing the actor-oriented approach. We then restore the actor-oriented approach. It will then be assumed that the formation of a coalition takes place in a political system that can be modelled as a centralized policy game.

6.3.1 Power Excess

The concept of power excess of a player in a coalition is essential:

Definition 6.5 *Let G be a weighted majority game, let $i \in N$ and let $S \in W$.*

$$pow(i, S) := w_i - w(S - \{i\})$$

is called the power excess *of i in S.*

In words, the power excess of a player in a coalition is the difference between the weight of that player and the size of the internal opposition for that player in that coalition. It is a simple measure of the control potential of a player in a coalition. The greater his power excess, the greater his control potential.

Instead of w_i a power index $p(i)$ can be used. Consider a simple game $G = (N, W)$. A power index p is a function from N into the set of positive real numbers Re^+. The best known power indices are the Shapley-Shubik power index (Shapley and Shubik 1954, Shapley 1981) and the Banzhaf index (Banzhaf 1965)[14]. Other indices are the Deegan-Packel index (Deegan and Packel 1983, Packel and Deegan 1980), the Holler index (Holler 1982, Holler and Packel 1983) and the Curiel index (Curiel 1987). Some care must be taken when working with one of these measures. They express the a priori power of a player in the complete game. Within the power excess theory a measure is needed that indicates the (expected) power of a player in a coalition once this coalition has been formed. A possible approach to this problem is to calculate the power index for a player i in each of the subgames $G_S = (S, W_S)$ associated with coalitions $S \in W$ when $i \in S$. This means that W_S is the set of winning subcoalitions of S. Working with a power index in the way indicated has the advantage that the theory of power excess coalitions can be generalized to the whole class of simple games. However, because of its simplicity we prefer to use the notion of 'weight' instead of the notion of power index[15]. Note that if we had worked with $p(i)$ instead of w_i, we would have had a bundle of theories, namely one for each particular index.

A dummy is a player who cannot possess any control potential. Such a player is not able to exert any power in any coalition.

Theorem 6.1 *Let G be a simple game and let $i \in N$ be a dummy. Then for every $S \in W$ with $i \in S$, $pow(i, S) < 0$.*

Proof. *Let i be a dummy and suppose there is an $S \in W$ such that $pow(i, S) > 0$. Then $w_i > w(S - \{i\})$. Since i is a dummy, $S - \{i\} \in W$ and thus $w(S - \{i\}) > q$. Hence, $w_i > q$. Contradiction.* □.

With respect to dominated simple games, the following result is relevant.

Theorem 6.2 *Let $G = [q; w_1, w_2, \ldots, w_n]$ be a dominated weighted majority game. If i dominates S, then $pow(i, S) > 0$.*

Proof. *Suppose i dominates S and $pow(i, S) \leq 0$. Then $w_i \leq w(S - \{i\})$. But then i cannot dominate S. Hence $pow(i, S) > 0$.* □.

In words, a dominant player has positive power excess in each coalition he dominates.

[14] An excellent study of these indices is Straffin (1983).

[15] Most of the results presented below also will hold for the power indices when used in the way as indicated.

6.3.2 Power Excess Coalitions in Weighted Majority Games

Let W_i be the set of all winning coalitions with player i, i.e. $W_i := \{S \in W \mid i \in S\}$. In general, i will prefer a coalition in W_i to any coalition not in W_i. Losing coalitions or winning coalitions he is not a member of, have nothing to offer. Further, the greater the power excess of i in a winning coalition, the better he can control the internal opposition in that coalition. The better he can control the internal opposition, the greater his influence on the decision-making process in that coalition and therefore the better he can enforce his own wants. Therefore, a player $i \in N$ will prefer a coalition with a greater power excess for him.

Definition 6.6 *Let $G = (N, W)$ be a weighted majority game, $S, T \in W$ and $i \in N$.*

1. i strictly prefers S to T, *notation* $S\pi_i T :=$
   (a) $S \in W_i$, $T \notin W_i$ *or*
   (b) $S, T \in W_i$ *and* $pow(i, S) > pow(i, T)$.
2. i is indifferent between S and T, *notation* $S\iota_i T :=$
   (a) $S, T \notin W_i$ *or*
   (b) $S, T \in W_i$ *and* $pow(i, S) = pow(i, T)$.
3. i weakly prefers S to T, *notation* $S\rho_i T, := S\pi_i T$ *or* $S\iota_i T$.

Given ρ_i, π_i and ι_i are also called, respectively, the asymmetric part and the symmetric part of the weak preference ρ_i. The strict preference may be read as 'is strictly better than', the indifference as 'is as good as' and the weak preference as 'is at least as good as'. It is easy to verify that ρ_i is complete and transitive with respect to W for every $i \in N$. The set $\mu(W, \rho_i)$ of ρ_i-maximal choices is:

Definition 6.7

$$\mu(W, \rho_i) := \{S \in W \mid \neg\exists T \in W[T\pi_i S]\}.$$

In words, $\mu(W, \rho_i)$ is the set of all coalitions for which there are no better ones for i. Because of the completeness and transitivity of ρ_i, this set is not empty[16]. Clearly, coalitions in $\mu(W, \rho_i)$ have maximal power excess

[16]See chapter 2.4 and 4.3 for the definition and relevant results of sets of maximal choices.

for i. That is, if $S \in \mu(W, \rho_i)$, then there is no $T \in W_i$ such that $pow(i, T) > pow(i, S)$. The converse also happens to be true. Therefore, each rational player will strive to form a coalition from $\mu(W, \rho_i)$.

So far the descriptive part of the theory of power excess coalitions for weighted majority games has been presented. The next step is to formulate a predictive part that is logically related to the descriptive part. In game theory, the predictive part is the subject of the socalled solution theories[17]. In the theory of power excess coalitions, we exclusively use the solution theory of the core. For this we need to explicate a domination relation over W.

Definition 6.8 *Let $G = (N, W)$ be a weighted majority game and $S, T \in W$.*

1. S dominates T :=
   (a) *for every $i \in S$: $S \rho_i T$,*
   (b) *there is at least one $i \in S$ such that $S \pi_i T$.*
2. $S \in W$ *is* undominated := *there is no $T \in W$ such that T dominates S. The set of undominated coalitions is called the* coalitional core of *G, in short, $Co(G)$. An $S \in Co(G)$ will be called a* core-coalition.

The core is a frequently applied solution concept in game theory. In this case it is equivalent to our concept of a set of maximal choices of W. Then, the relation that establishes the maximal elements is the dominance relation. Unfortunately, for many games the core is empty. This will, however, not be the case for the coalitional core in the case of proper weighted majority games:

Theorem 6.3 *Let G be a proper weighted majority game. Then*

$$Co(G) \neq \emptyset.$$

Proof. *Let G be a weighted majority game. Then $W^{size} \neq \emptyset$. Hence, by theorem 6.4, $Co(G) \neq \emptyset$.* □.

[17]The distinction between a descriptive part and a solution part is explicitly made in Shubik (1982: 127).

If S is a core-coalition, then for every $i \in S$, $S \in \mu(W, \rho_i)$, and, hence, every $i \in S$ must have maximal power excess. Otherwise, S would be dominated. So, there is no better coalition for the members of that coalition.

Hypothesis 6.1 *Let G be a proper weighted majority game. Then only core-coalitions of G will be formed.*

The next result expresses a clear connection between Riker's minimum size principle and the theory of power excess coalitions for weighted majority games.

Theorem 6.4 *Let G be a proper weighted majority game. Then*

$$Co(G) = W^{size}.$$

Proof.
1) Let $S \in Co(G)$, but suppose S is not of minimum size. Then there is a coalition $T \in W$ such that $w(T) < w(S)$. Since G is proper, $S \cap T \neq \emptyset$. So, there is an $i \in S \cap T$ such that $T\pi_i S$. Then $w(T - \{i\}) < w(S - \{i\})$. Consequently, for all $i \in T$, $T\rho_i S$. Since $i \in S \cap T$, $T\pi_i S$ by definition and therefore $T dominates S$. Contradiction with $S \in Co(G)$. Therefore, S is of minimum size. Thus $Co(G) \subseteq W^{size}$.
2) Let $S \in W^{size}$, i.e. $w(S) \leq w(T)$ for every $T \in W$ and hence $pow(i, S) \geq pow(i, T)$ for every $i \in S$ and every $T \in W$. Therefore, $S\rho_i T$ for every $i \in S$ and $T \in W$ and hence, since ρ_i is complete, $S \in \mu(W, \rho_i$ for every $i \in S$. Since G is proper, $S \cap T \neq \emptyset$ for every $T \in W$. Hence, if $S \in \mu(W, \rho_i)$, then there is no coalition $T \in W$ that dominates S. Therefore, $S \in Co(G)$ and thus $W^{size} \subseteq Co(G)$. □.

According to this theorem, a core-coalition is of minimum size and a minimum size coalition is a core-coalition. This provides an explanation of the entry of minimum size coalitions in systems that can be modelled as weighted majority games. Players prefer such a coalition because there is no other coalition in which they can better control the internal opposition.

6.3.3 Power Excess Coalitions in Centralized Weighted Majority Policy Games

In this section we use again the actor-oriented approach as introduced in Peleg (1981) and apply it to centralized weighted majority games[18]. Again, the fundamental behavioral assumption will be that each player strives to form a coalition in which his power excess is maximal. However, within the actor-oriented context, each player must reckon with the constraint that only coalitions with the center player will be formed.

Let C_i be the set of all winning coalitions with the center player that contain i. That is, $C_i := \{S \in C \mid i \in S\}$. A player i will prefer a coalition $S \in C_i$ to a coalition not in C_i. Further, he will prefer $S \in C_i$ to $T \in C_i$ if he has a greater power excess in S.

Definition 6.9 *Let G_Θ be a centralized weighted majority policy game and $i \in N$, and suppose $S, T \in C$.*

1. i strictly prefers S to T, *notation* $S\pi_i^\theta T$, :=
   (a) $S \in C_i$, $T \notin C_i$,
   (b) $S, T \in C_i$ *and* $pow(i, S) > pow(i, T)$.
2. i is indifferent between S and T, *notation* $S\iota_i^\theta T$, :=
   (a) $S \notin C_i$, $T \notin C_i$,
   (b) $S, T \in C_i$ *and* $pow(i, S) = pow(i, T)$,
3. i weakly prefers S *to* T, *notation* $S\rho_i^\theta T$, := $S\pi_i^\theta T$ *or* $S\iota_i^\theta T$.

The superscript θ in these notations is used to remind us that we are dealing with coalition preferences of players in a policy game G_Θ. The binary relation π_i^θ is the strict preference of i for the coalitions in C; the relation ι_i^θ is the indifference of i over C and ρ_i^θ is his weak preference. This latter relation is, in fact, the union of π_i^θ and ι_i^θ.

It is not difficult to prove that each ρ_i^θ is complete and transitive with respect to C. Therefore, the set $\mu(C, \rho_i^\theta)$ defined by

[18]See chapter 5 for the definition of centralized weighted majority policy games.

Definition 6.10

$$\mu(C, \rho_i^\theta) := \{S \in C \mid \neg \exists T \in C[T\pi_i^\theta S]\},$$

is not empty. If $S \in \mu(C, \rho_i^\theta)$, then there is no $T \in C$ such that $pow(i, T) > pow(i, S)$. The converse is also true: if there is no coalition $T \in C$ such that $pow(i,T) > pow(i, S)$, then $S \in \mu(C, \rho_i^\theta)$. Clearly, a player i will, if rational, strive to form a coalition from $\mu(C, \rho_i^\theta)$ within this context.

Definition 6.11 *Let G_Θ be a centralized weighted majority policy game and c be the center player.*

1. *A coalition $S \in C$* dominates *a coalition $T \in C$ in G_Θ :=*
   (a) *for each $i \in S$: $S\rho_i^\theta T$,*
   (b) $S\pi_c^\theta T$.
2. *A coalition $S \in C$ is* undominated := *there is no other coalition $T \in C$ that dominates S. The set of undominated coalitions in a centralized policy game G_Θ is called the* coalitional core *of that game. This set will be notated with $Co(G_\Theta)$. An $S \in G_\Theta$ is called a* core-coalition of G_Θ. *If the context is clear, we only speak of core-coalitions.*

In words, a coalition S dominates a coalition T if each player in S finds S at least as good as T and if the center player finds S strictly better than T.

Theorem 6.5 *Let G_Θ be a centralized policy game. Then*

$$Co(G_\Theta) \neq \emptyset.$$

Proof. *This is a consequence of theorem 6.6 and the fact that $\mu(C, \rho_c^\theta) \neq \emptyset$.* □.

If S is a core-coalition of G_Θ, then for each $i \in S$, $S \in \mu(C, \rho_i^\theta)$. Otherwise, S would be dominated. Hence, there is no better coalition for the members of a core-coalition. Therefore:

Hypothesis 6.2 *Let G_Θ be a centralized policy game. Then only core-coalitions of G_Θ will be formed.*

The computation of the core for centralized policy games will, in general, be rather involved. First, the coalition preference of each player must be constructed. Next the dominance relation between each pair of winning coalitions with the center player must be computed on the base of these preferences. Then the coalitional core can be determined. The following result, which is as such perhaps somewhat surprising, simplifies the computation process of the coalitional core.

Theorem 6.6 *Let G_Θ be a centralized weighted majority policy game and let c be the center player. Then*

$$Co(G_\Theta) = \mu(C, \rho_c^\theta).$$

Proof.
1) Let $S \in \mu(C, \rho_c^\theta)$. Then $pow(c, S) \geq pow(c, T)$ for all $T \in C$. If there is a $T \in C$ such that T dominates S, then $pow(c, T) > pow(c, S)$. But then not $pow(c, S) \geq pow(c, T)$ for all $T \in C$.. Contradiction. Hence, $S \in Co(G_\Theta)$.
2) If $S \in Co(G_\Theta)$, then $pow(c, S) \geq pow(c, T)$ for all $T \in C$. Hence, $S \rho_c^\theta T$ for all $T \in C$ and therefore $S \in \mu(C, \rho_c^\theta)$. □.

That is, the coalitional core equals the set of coalitions with maximal power excess for the center player c. Hence, to compute the coalitional core of a centralized policy game, it suffices to compute the preference of the center player and to determine the set of maximal elements for this preference. Thus, if the players maximize their power excess under the constraint that the center player must be a member of each winning coalition, then the center player will firstly be in each core-coalition and secondly have maximum power excess in each core-coalition.

If a core-coalition in G_Θ is not minimal winning, then the power excess is maximal for no member. Therefore a core-coalition is minimal winning. The connection with Riker's minimum size theory, however, is ompletely cut off in this context. In general, a core-coalition is not necessarily a minimum size coalition and a minimum size coalition is not necessarily a core-coalition in a centralized policy game. A reason for this is that a minimum size coalition need not contain the center player. However, it is possible to adjust the size parameter to the policy context.

Definition 6.12 *Let G_Θ be a centralized weighted majority policy game. A coalition $S \in C$ is* of minimum size in G_Θ :=

$$w(S) \leq w(T) \; for \; all \; T \in C.$$

In words, a coalition in a centralized weighted majority policy game is of minimum size if it contains the center player and if its size is at least as small as every other coalition that contains the center player. Let us denote the set of minimum size coalitions for G_Θ by C^{size}.

Theorem 6.7 *Let G_Θ be a centralized weighted majority policy game. Then*

$$Co(G_\Theta) = C^{size}.$$

Proof. *This proof goes in the same way as the proof of theorem 6.4. In stead of W, the set C is used.* □.

Hence, a core coalition for G_Θ is a coalition with the center player that has a size less than or equal to any other winning coalition with the center player.

6.3.4 Theory of Power Excess Coalitions: Closed Version

This is a variation of the theory of power excess coalitions. To formulate this variation, we need the concepts as defined in definition 5.23. We will repeat them informally. Let G_Θ be a policy game. A player k is said to be *between* players i and j if $p_i \theta p_k$ and $p_k \theta p_j$. Two players i and j are *neighbours* if there is no other player k between them. A coalition S is said to be *closed* if it consists only of neighbours. A coalition which is not closed is said to be *open*.

The fundamental behavioral assumption now will be that each player maximizes his power excess under the constraint that the coalition to be formed must contain the center player and, simultaneously, must be closed. For a discussion of the relevance of closed coalitions, consider 5.8.3.

In this version of power excess theory, the coalition preferences of the players will have a more complicated structure.

Definition 6.13 *Let G_Θ be a centralized weighted majority policy game, c be the center player and C^{cl} be the set of closed winning coalitions with the center player. Let $C_i^{cl} := \{S \in C^{cl} \mid i \in S\}$. For each $S, T \in C$,*

1. $S\pi_i^{cl}T :=$

   (a) $S \in C_i, T \notin C_i$,

   (b) $S \in C_i^{cl}, T \in (C_i - C_i^{cl})$,

   (c) $S, T \in (C_i - C_i^{cl})$ *and* $pow(i, S) > pow(i, T)$,

   (d) $S, T \in C_i^{cl}$ *and* $pow(i, S) > pow(i, T)$.

2. $S\iota_i^{cl}T :=$

   (a) $S, T \notin C_i$,

   (b) $S, T \in (C_i - C_i^{cl})$ *and* $pow(i, S) = pow(i, T)$,

   (c) $S, T \in C_i^{cl}$ *and* $pow(i, S) = pow(i, T)$.

3. $S\rho_i^{cl}T := S\pi_i^{cl}T$ *or* $S\iota_i^{cl}T$.

These relations may be interpreted in the same way as in the open version. That is, $S\pi_i^{cl}T$ means that i finds S better than T, $S\iota_i^{cl}T$ means that i is indifferent between S and T, and $S\rho_i^{cl}T$ means that i finds S at least as good as T. The superscript *cl* in these notations reminds us that we are dealing with coalition preferences in the closed version of the theory of power excess coalitions.

It is easy to prove that for each $i \in N$, ρ_i^{cl} is complete and transitive with respect to C^{cl}. Therefore, for each i the set

$$\mu(C^{cl}, \rho_i^{cl}) = \{S \in C^{cl} \mid \neg\exists T \in C[T\pi_i^{cl}S\}],$$

is not empty. This set contains the maximal elements for i according to his closed coalition preference. If a player is rational, then he will strive to form a coalition from $\mu(C^{cl}, \rho_i^{cl})$.

A coalition $S \in \mu(C^{cl}, \rho_i^{cl})$ is closed and has maximal power excess for i. That is, there is no other closed coalition with a greater power excess for i. Note, however, that $S \in \mu(C^{cl}, \rho_i^{cl})$ does not imply that $S \in \mu(C, \rho_i^{\theta})$. Conversely, a coalition $S \in \mu(C, \rho_i^{\theta})$ need not be a member of $\mu(C^{cl}, \rho_i^{cl})$.

In defining a dominance relation for C, we proceed in the same way as in the open version.

Definition 6.14 *Let G_Θ be a centralized weighted majority policy game and S,T be coalitions.*

1. $S \in C$ dominates $T' \in C$:=

   (a) for every $i \in S$, $S\rho_i^{cl}T$ and

   (b) there is at least one player $i \in S$ such that $S\pi_i^{cl}T$.

2. $S \in C$ *is* undominated := *there is no other coalition that dominates S. The set of undominated coalitions will be called the* core of closed coalitions. *Notation: $Co^{cl}(G_\Theta)$. An $S \in Co^{cl}(G_\Theta)$ is called a* closed core-coalition.

The following result runs parallel to that in the open version:

Theorem 6.8 *Let G_Θ be a centralized weighted majority policy game. Then*

$$Co^{cl}(G_\Theta) \neq \emptyset.$$

Proof. *This is a consequence of theorem 6.9 and the fact that $\mu(C^{cl}, \rho_c^{cl}) \neq \emptyset$.* □.

If S is a closed core-coalition, then for each $i \in S$, $S \in \mu(C^{cl}, \rho_i^{cl})$. Otherwise, S would be dominated. Hence, there is no better coalition for the members of S.

Hypothesis 6.3 *Let G_Θ be a centralized policy game. Then only coalitions from $Co^{cl}(G_\Theta)$ will be formed.*

The computation process for the core of closed coalitions can be simplified in a similar way as in the open version.

Theorem 6.9 *Let G_Θ be a centralized policy game and let c be the center player. Then*

$$Co^{cl}(G_\Theta) = \mu(C^{cl}, \rho_c^{cl}).$$

Proof. *This proof goes in an analogues way as the proof of 6.6.* □.

However, the coalitional core must not be confused with the core of closed coalitions. Since $\mu(C, \rho_i^\theta$ will, in general, deviate from $\mu(C^{cl}, \rho_i^{cl}$, the coalitional core is definitely not the same as the core of closed coalitions. Also note that a closed core-coalition is not necessarily a minimal winning coalition.

A closed winning coalition within this context is called of minimum size if it contains the center player and if its size is not greater than any other closed winning coalition with the center player.

Definition 6.15 *Let G_Θ be a centralized weighted majority policy game. Coalition S is of* closed minimum size *in G_Θ :=*

1. *$S \in C^{cl}$ and*

2. *$w(S) \leq w(T)$ for every $T \in C^{cl}$.*

The set of closed minimum size coalitions will be denoted by C_{cl}^{size}.

Theorem 6.10 *Let G_Θ be a centralized weighted majority policy game. Then*

$$Co^{cl}(G_\Theta) = C_{cl}^{size}.$$

Proof. *This proof goes in the same way as the proof of theorem 6.4. In stead of W, the set C is used.* □.

That is, a closed core coalition for a policy game G_Θ is a closed coalition that contains the center player and that has a size that is less than or equal to the size of any other closed winning coalition with the center player.

6.4 Center Parties and Cabinet Formations in Parliamentary Systems: Some Hypotheses

In this section we translate the hypotheses as deduced in the theory of power excess coalitions in empirically testable hypotheses about cabinet formation in parliamentary systems. To illustrate the working of the theories, we compute a prediction for the Dutch election of 6 September 1989.

A *parliamentary system* can be seen as a policy game
$G_\Theta = [q; w_1, w_2, \ldots, w_n]$, where the players are political parties and where

the weights are the number of seats of a political party in parliament. The policy order corresponds with the order in which the parties are given. Thus, party i is to the left of party $i+1$ etc. A coalition in a parliamentary system is called a *cabinet*. The quota is the number of seats necessary to form a majority cabinet.

The theory of power excess coalitions for weighted majority games without policy order says, in fact, that only minimum size cabinets will be formed in parliamentary systems. This hypothesis has been thoroughly investigated and discussed (for example: De Swaan 1973, Taylor and Laver 1973, Browne and Dreijmanis 1982), and has been coupled to hypotheses about the durability of cabinets (especially Dodd 1976, see also Grofman 1989). We do not discuss it here.

From the theory of power excess coalitions the following hypotheses can be derived:

Hypothesis 6.4 *In centralized parliamentary systems only core-cabinets will be formed.*

In such cabinets, each party, including the center party, has maximal power excess. According to the closed version of the theory, the hypothesis is

Hypothesis 6.5 *In centralized parliamentary systems only closed core-cabinets will be formed.*

As the theory indicates, a closed core-cabinet need not coincide with a core-cabinet.

6.4.1 Computation example

Consider the game representation of the Dutch parliament according to the election of 6 september 1989:

$$[76; 6, 49, 12, 54, 22].$$

The parties are, from left to right, GL (Green Left), PvdA (Social Democrats), D66 (Left Liberals), CDA (Christian Democrats) and VVD (Conservative Liberals). The policy positions of these parties are accordingly ordered from left to right. Parties with less than 2.5 % of the total number of votes have been omitted. These parties, which all are to the right of the

Cabinets with CDA	CDA	VVD	D66	PvdA	GL
{CDA, VVD}	32	-32			
{CDA, VVD, D66}	20	-44	-64		
{CDA, VVD, D66, PvdA}	-29	-93	-113	-39	
{CDA, VVD, D66, PvdA, GL}	-35	-99	-119	-45	-131
{CDA, VVD, PvdA}	-17	-81		-27	
{CDA, VVD, GL}	26	-38			-70
{CDA, VVD, D66, GL}	14	-50	-70		-82
{CDA, D66, PvdA}	-7		-91	-17	
{CDA, D66, PvdA, GL}	-13		-97	-23	-109
{CDA, PvdA}	5			-5	
{CDA, PvdA, GL}	0			-12	-97

Table 6.3:

conservative liberals, are dummies which have no influence on the cabinet formation process[19].

Clearly this game is decisive. Hence there must be a center party (cf. theorem 5.13). This is the CDA. To see this, take the sum of the weights of the parties which are to the left of the CDA. This sum is less than 76. Also, the sum of the parties to the right of the CDA is less than 76. Hence any combination of parties to the left or to the right of the CDA needs the CDA in order to form a closed majority cabinet. Neither side can form a majority cabinet on its own. In contrast, the CDA can form a majority cabinet either with parties from the left or with parties from the right. To check, note that

$$\mid w(L) - w(R) \mid = 45 < w_{CDA} = 54,$$

where w(L) and w(R) are, respectively, the sizes of Left and Right.

Thanks to theorem 6.6, the preference of the CDA suffices to compute the coalitional core. The full set of cabinets with the center party CDA is given in table 1. This table also indicates the power excess of the several parties in the several possible winning combinations. From this table, the power excess of the CDA in the several cabinets can be read off (column two). It has greatest power excess in the combination {CDA, VVD}.

[19]A similar convention is adopted in De Swaan (1973).

Hence, this coalition is the most preferred by this party. By theorem 6.6, it is the core coalition and, according to hypothesis 6.4, a {CDA, VVD} cabinet will be formed. To check, note that this is also the cabinet that is mostly preferred by the VVD. In this cabinet, this party too has maximal power excess. With respect to the first column of table 6.3, the {CDA, VVD} cabinet is, as expected, of minimum size (cf. theorem 6.7).

Consider the closed cabinets in the first column of table 6.3. Since {CDA, VVD} is a closed cabinet, the picture does not change. According to hypothesis 6.5, this cabinet will be formed. Note that it is, again, the best possibility for the VVD. In this combination, the VVD has maximal power excess. Adding more parties (f.e. D66) only would increase the internal opposition for the VVD and hence decrease its power excess. With respect to the set of closed cabinets, the {CDA, VVD} combination is, as expected, of minimum size. This cabinet even happens to be of minimum size with respect to the set of all winning cabinets. This, however, is an accidental fact.

Chapter 7

Coalition Formation in Social Choice Games

7.1 Introduction

In this chapter we will solve the second part of problem 2 formulated in chapter 1. Consider a nonempty and finite set $N = \{1,2,\ldots,n\}$ of players and a nonempty set X of social states. Let $p = (R_1, R_2, \ldots, R_n)$ be a preference profile concerning X. Now suppose that a social choice only can be made by cooperation, that is, by forming a coalition. What coalitions will be formed? What coalition preferences will each player form in order to realize a social choice that is as close as possible to his most preferred alternative?

A social choice game is a 4-tuple (N, W, X, p) where

1. (N, W) is a simple game,
2. X is a social choice problem (a nonempty set of social states), and
3. p is a preference profile $(R_1, R_2, \ldots, R_n)$, where $R_i \in O(X)$ for every $i \in N$.

More precisely, the above mentioned problem now is how players in a social choice game form preferences concerning W on the basis of their preferences in p and how they use their thus formed coalitional preferences in order to form a winning coalition.

In this chapter two theories of coalition formation in social choice games are presented. In both theories the concept of distance is crucial. In the first theory distance is used as a measure of conflict between two preferences. The basic idea is that the greater the distance between two preferences R_i and R_j, the greater the conflict of interest between individuals i and j. By defining a measure of dispersion for the set of distances for each pair of preferences in an arbitrary preference profile, a conflict index can be obtained. By using the preferences of the individuals in a coalition, also the conflict index of each coalition can be determined. The basic behavioral assumption of this theory is that each player strives to form a coalition with a minimal conflict index. Because of this assumption this theory will be called *conflict minimization theory*.

The other theory is related to De Swaan's policy distance theory. De Swaan's theory is treated in chapter 6. Point of departure of our theory is the distance between a player's preference with respect to X and the social preference that he expects to be formulated when that coalition will be formed. As we have tried to make clear in chapter 2, generating a social preference requires a social choice rule. Thus, it is assumed that each player knows the social choice rule to be used in each relevant coalition. Further it is assumed that each player exactly knows the other players'preferences with respect to X. A player then can calculate the social preference for each coalition and therefore he can compute the distance between his preference and the social preference of each winning coalition. The basic behavioral assumption is that each player strives to form a winning coalition such that the distance between his preference and the social preference of that coalition is minimal. Because distances between preferences are fundamental in this theory, it will be called *preference distance theory*.

Distance in the first place is a geometrical concept and therefore it is natural to employ the usual geometrical methods to develop a theory of preference distances. In fact, these geometrical methods are employed in the theory of spatial voting (cf. Elenow and Hinich 1984; Riker and Ordeshook 1973: Ch. 11 and 12; for an intuitive introduction to these methods see Grofmann et. al. 1988). However, crucial for the methods in that field of application is the presupposition of the continuity of space. This makes these methods less appropriate for our purpose. The sets we are dealing with all have a discrete character. In chapter 2, we have assumed that X is finite. Consequently, individual preferences, which are relations on X are also finite. Further, the set of players is assumed to be finite, so that the set W is finite (cf. Chapter 5). Clearly, each set derived in some way or another from these basic sets must be finite (cf. for example the several prediction sets in chapters 5 and 6). Methods for analyzing continuous space will ignore the essentially discrete character of these sets and are therefore less appropriate. The use of discrete mathematical methods seems to be more natural.

Another argument for the use of discrete mathematics is more filosophical in nature. We think that the social and political realm in essence is inhabited by discrete combinatorial structures. The social and political world itself is combinatorial in nature. This 'Weltanschauung' has some consequences. If social and political life is in essence discrete and combinatorial in nature, then the empirical domain for the social and political science must consist mainly of discrete variables. In order to be meaningful, social and political theories must reflect the discrete and combinatorial character of the empirical domain. This only can be done by using discrete mathematical methods.

For these reasons we explicitly choose for the use of discrete mathematics in this chapter. This decision is in line with the research policy as used in chapter 4 of this monograph. Also in that chapter we decided to use discrete mathematics, especially digraph theory, for developing solution theories for preference profiles.

In section 2 of this chapter we define the concept of social choice game. This concept is very important since it is the starting point of the coalition theories presented in this chapter. In section 3 a theory of preference distances is presented that is entirely formulated in terms of discrete mathematics. With the aid of the set-theoretical operation of symmetric difference we define a discrete metric space in which the distance between each pair of preferences can be measured. This metric space is called Hamming space. In the thereupon following section - section 4 - we present conflict minimization theory. First the descriptive part of this theory is presented. After this the solution part of the theory is presented. Finally, in order to illustrate the working of conflict minimization theory, we provide a computation example. In section 5 preference distance theory is presented. Again the presentation has a similar structure: first we present the descriptive part and then the solution part. Since the concept of a set of maximal elements will not work for this theory, we employ the theory of generalized stable sets as developed in chapter 4. Finally, in order to illustrate the working of preference distance theory, we provide a computation example for this theory too. In the final section we discuss the possibility of empirical applications of the presented theories. Further, we discuss other application fields of the concepts of Hamming space and Hamming conflict indices.

7.2 Social Choice Games

In this section simple game theory will be complicated by introducing explicitly the preferences of the individual players concerning a social choice problem. The games will be called *social choice games*. An important impuls to the study of this kind of games is given in Nagamura (1975, 1979). Nagamura calls these games 'simple games with ordinal preferences'.

Definition 7.1 *A 4-tuple* (N, W, X, p) *is a* social choice game :=

1. (N, W) *is a simple game,*
2. X *is a social choice problem,*
3. $p \in \Pi$ *is a preference profile concerning* X.

We will abstract away, just as in the other types of simple games, from the rules that specify winning and losing coalitions[1]. We are thus able to formulate propositions that are independent of the rules of winning and losing.

7.3 Hamming Space

The origin of the distance theory presented in this section is Kemeny and Snell (1963) and Bogart (1973). Also see Barthelemy and Monjardet (1981). Kemeny and Snell and also Bogarts use matrix-theoretical

[1]Cf. chapter 5.1.

terms to define distances between preferences. In contrast we will use set-theoretical terms without referring to matrices.

Intuitively, the notion of distance between two preferences has to do with the dissimilarity between those preferences. The greater their distance, the greater their dissimilarity. However, before formalizing this intuition, we first have to specify exactly what distance on a set of preferences is. This specification will be done by using the notion of preference distance function.

Definition 7.2 *Let Re be the set of real numbers. The function $d : B(X) \times B(X) \rightarrow Re$ is a* preference distance function *or, equivalently, a* preference metric :=

1. *$d(R_i, R_j) \geq 0$ for every $R_i, R_j \in B(X)$ and $d(R, R) = 0$ for every $R \in B(X)$;*

2. *$d(R_i, R_j) = d(R_j, R_i)$ for every $R_i, R_j \in B(X)$;*

3. *$d(R_i, R_j) + d(R_j, R_k) \geq d(R_i, R_k)$ for every $R_i, R_j, R_k \in B(X)$.*

Remember that $B(X)$ is the set of complete and reflexive preferences on X (see chapter 2). An ordered pair $(B(X), d)$, where d is a preference distance function, is called a *metric preference space*. Since X is finite, $(B(X), d)$ where d is a preference distance function is a discrete metric space[2].

The first property of a preference distance function states that the distance between two preferences is always nonnegative and that the distance between two preferences is zero if the two preferences are identical. Both parts of this property are intuitively acceptable. The second property is a *symmetry* condition. It states that the distance between R_i and R_j equals the distance between R_j and R_i. Again this corresponds with our intuition of distance and dissimilarity. The third property is called *triangle inequality*. This property is somewhat more difficult to understand. It asserts, in fact, a kind of transitivity of distance.

[2]In general, a point $v \in V$ in a metric space (V, d) is isolated if the set $\{v\}$ consisting of v alone is open. A metric space (V, d) is discrete if every point in V is isolated. An extensive study of discrete metric spaces can be found in Storcken (1989, esp. chapter 3).

In the following sections we want every preference profile to be a metric space. In order to make this possible, we need the concept of a subspace. Let $(B(X), d)$ be a metric space. A subspace is obtained by taking a subset $B_1(X)$ from $B(X)$ and subsequently by restricting d to $B_1(X) \times B_1(X)$ (cf. Mendelson 1973: 56).

Definition 7.3 *Let $(B(X), d)$ be a metric preference space. Let $B_1(X) \subseteq B(X)$ and let*

$$d' = d \mid B_1(X) \times B_1(X).$$

Then $(B_1(X), d')$ is said to be a subspace of $(B(X), d)$.

Though we now know what properties a preference distance function must satisfy, we do not yet know how to determine the distances between preferences. For this it is necessary to specify a particular preference distance function. The general idea we will use for this purpose is that the smaller the number of ordered pairs of social states two preferences have in common, the more dissimilar the two preferences are. A distance function that captures this idea in an exact way is the *Hamming distance* function. In general, the Hamming distance between two sets is the number of elements in the symmetric difference of these sets (cf. Barthelemy and Monjardet 1981, Bollobas 1986); that is, the number of elements that are in either one of the two sets but not in both. Since preferences are binary relations and since binary relations are sets, this notion can also be applied to preferences.

Definition 7.4 *Let $R_i, R_j \in B(X)$. The* symmetric difference *of R_i and R_j, denoted by $R_i \oplus R_j$, is the set of all ordered pairs of social states that are in R_i or in R_j but not both, i.e.*

$$R_i \oplus R_j := (R_i \cup R_j) - (R_i \cap R_j).$$

The precise definition of the Hamming preference distance function is:

Definition 7.5 *The mapping $h : B(X) \times B(X) \to Re$ such that*

$$h(R_i, R_j) = \mid (R_i \oplus R_j) \mid,$$

for every $(R_i, R_j) \in B(X) \times B(X)$ is called the Hamming preference distance function.

It must be proven that h really is a preference distance function.

Theorem 7.1 *The ordered pair $(B(X), h)$ is a metric preference space, that is, h is a preference distance function.*

Proof. *We only prove triangle inequality, that is,*

$$(R_i \oplus R_k) \subseteq ((R_i \oplus R_j) \cup (R_j \oplus R_k)).$$

Let $(x, y) \in R_i \oplus R_k$ but suppose $(x, y) \notin ((R_i \oplus R_j) \cup (R_j \oplus R_k))$. That is, $(x, y) \notin R_i \oplus R_j$ and $(x, y) \notin R_j \oplus R_k$. But then $(x, y) \in R_i \cap R_j$ and $(x, y) \in R_j \cap R_k$, that is, $(x, y) \in \bigcap\{R_i, R_j, R_k\}$. This implies $(x, y) \notin R_i \oplus R_k$. Contradiction. Therefore, $(x, y) \in ((R_i \oplus R_j) \cup (R_j \oplus R_k))$. $\square$.

Thus, h is a metric on $B(X)$. As we wanted, h shows how many different ordered pairs of social states a couple of preferences have and therefore how dissimilar they are. The less they have in common, the greater the h-distance between them[3].

Since the space $(B(X), h)$ is so fundamental in our story, it is useful to give it a particular name:

Definition 7.6 *The metric preference space $(B(X), h)$ is called* Hamming space.

In order to avoid cumbersome notation we also will use the notation h for every subspace of a Hamming space unless confusion may be raised[4].

To illustrate the working of the concept of Hamming distance, consider the preference profile (R_1, R_2) where $R_1 = \{(x_1, x_2), (x_2, x_3), (x_1, x_3)\}$ and

[3]Barthelemy and Monjardet (1981) use the Hamming distance to study median relations of preference profiles. In their study, the remoteness $\Delta(R, p)$ between a relation R and a profile p is defined as $\Delta(R, p) = \sum_{i=1}^{n} h(R, R_i)$. A relation M is defined to be a median for a preference profile if and only if $\Delta(M, p) = min_{R \in B(X)} \Delta(R, p)$. That is, medians are relations that minimize the remoteness from a profile. A social choice system that selects a median for each profile is the system of majority decision (see chapter 3 for a definition of this system). The socalled Kemeny procedure (Kemeny 1959, Fishburn 1977) selects a linear ordering as a median for every preference profile, hence also for profiles for which the majority relation is cyclic.

[4]For the concept of subspace, cf. definition 2.2 of this section.

$R_2 = \{(x_3, x_1), (x_1, x_2), (x_3, x_2)\}$. In compact notation (cf. chapter 2, section 5), this profile can be written as

$$\begin{array}{ll} R_1: & x_1x_2x_3 \\ R_2: & x_3x_1x_2 \end{array}$$

$R_1 \cup R_2 = \{(x_1, x_2), (x_2, x_3), (x_1, x_3), (x_3, x_1), (x_3, x_2)\}$. $R_1 \cap R_2 = \{(x_1, x_2)\}$. $(R_1 \cup R_2) - (R_1 \cap R_2) = \{(x_2, x_3), (x_1, x_3), x_3, x_1), (x_3, x_2)\}$. Therefore, $h(R_1, R_2) = 4$.

Since h is a preference distance function, it satisfies symmetry. Therefore, for each preference profile, we only have to compute $(n(n-1))/2$ Hamming distances, where n is the number of individuals.

With the aid of h other useful concepts can be defined[5].

Definition 7.7 *Let $B_1(X) \subseteq B(X)$.*

1. *The* diameter *of $B_1(X)$ is*

$$max\{h(R_i, R_j) \mid R_i, R_j \in B_1(X)\}$$

and is denoted by $diam(B_1(X))$.

2. *The* meshwidth *of $B_1(X)$ is*

$$min\{h(R_i, R_j) \mid R_i, R_j \in B_1(X)\}.$$

and is denoted by $mesh(B_1(X))$.

The diameter and the meshwidth of a set of preferences are, respectively, the largest distance between two preferences in that set and the smallest

[5]The following definitions of course can be given for any distance function on a set. However, since we will not work with other distance functions, we only define these concepts for Hamming space (and subspaces of Hamming space). A similar research procedure is employed in the field of spatial voting. In that field, the relevant space is Euclidian space. The distance measures mostly used are Euclidian Distance and its generalization Weighted Euclidian Distance (WED). Preferences of voters and other concepts are defined with the aid of these measures. Cf. Riker and Ordeshook (1973: Chapters 11 and 12), Enelow and Hinich (1984). In contrast, Storcken (1989: chapter 3) studies discrete metric spaces in general without reference to any particular distance function.

distance between two different preferences in that set. Both terms are standard in topology and digraph theory. In Storcken (1989: 149) the definitions of diameter and meshwidth are given with respect to an arbitrary set V and a distance function d on V. Instead of the operations of max and min, Storcken uses, respectively, the operations of sup (least upper bound) and inf (greatest lower bound). The operation of max is also used in Bollobas (1986: 102). Since we are exclusively dealing with finite sets, there is no difference in this respect.

7.3.1 Normalization of Dissimilarity

The Hamming distance is a measure of the dissimilarity between two arbitrary preferences. In order to compare the dissimilarities between pairs of arbitrary preferences, it may be necessary to normalize the Hamming distances of these pairs of preferences. For this, we propose the following normalization procedure.

Let R_i and R_j be linear orders and suppose R_j is the converse of R_i, that is, xP_jy if yP_ix for every $x, y \in X$. Then $R_i \cap R_j = \emptyset$. Therefore, since $\mid R_i \mid = \mid R_j \mid = (m(m-1))/2$ where m is the number of alternatives, $h(R_i, R_j) = m(m-1)$. If one of the preferences is a weak order or if their intersection is not empty, then, clearly, the Hamming distance must be less than $m(m-1)$. Therefore, $m(m-1)$ is the maximal possible Hamming distance between preferences. This justificates the following definition:

Definition 7.8 *Let m be the number of alternatives in X and let R_i and R_j be preferences.*

$$h_{norm}(R_i, R_j) := \frac{h(R_i, R_j)}{m(m-1)}.$$

$h_{norm}(R_i, R_j)$ is called the normed Hamming distance *between R_i and R_j.*

According to this definition, the greater h_{norm} between two preferences, the greater their dissimilarity or the less their similarity defined as $1 - h_{norm}$. The dissimilarity is maximal, if $h_{norm} = 1$. Then, the similarity is minimal. It is now also possible to compare the dissimilarities between arbitrary preferences from different preference profiles. To be more specific, let p and q be two different preference profiles. If $h_{norm}(R_i^p, R_j^p) \leq h_{norm}(R_k^q, R_l^q)$, then it is allowed to say that R_k^p and R_l^p are at least as dissimilar as R_i^q and R_j^q.

7.3.2 Betweenness of Preferences

In general, three elements in a set will satisfy the triangle equality if one element is, in some way or another, *between* the two other elements. The same is true for a set of preferences if we use the following notion of betweenness[6]:

Definition 7.9 *Let* $R_i, R_j, R_k \in B(X)$. R_j *is* between R_i and R_k :=

- $R_j \subseteq R_i \cup R_k$;
- $R_j \supseteq R_i \cap R_k$.

According to this definition, a preference R_j is between two other preferences R_i and R_k if, firstly, it is contained in the union of R_i and R_k and if, secondly, it contains all ordered pairs of social states common to R_i and R_k. This notion has some resemblance with our intuition of betweenness of a point between two other points on a straight line. However, some care is necessary. For the straight line, betweenness is a transitive notion. If a point b is between points a and c and if c is between b and d, then both b and c are between a and d. In order to obtain transitivity of betweenness of preferences we need a similar concept as 'straight line'. Let us first study an example.

Preference Profile 7.1

$$\begin{aligned} R_1 &: \; x_1x_2x_3x_4 \\ R_2 &: \; x_3x_1x_4x_2 \\ R_3 &: \; x_3x_4x_1x_2 \\ R_4 &: \; x_4x_2x_3x_1 \end{aligned}$$

Since $R_2 \subset R_1 \cup R_3$ and $R_2 \supset R_1 \cap R_3 = \{(x_1, x_2), (x_3, x_4)\}$, R_2 is between R_1 and R_3. Since $R_3 \subset R_2 \cup R_4$ and $R_3 \supset R_2 \cap R_4 = \{(x_3, x_1), (x_4, x_2)\}$,

[6] This notion is used in another formulation in the work of Restle (1959). There a set S_j is defined to be between two sets S_j and S_k if and only if

1. $S_i \cap S_j^c \cap S_k = \emptyset$,
2. $S_i^c \cap S_j \cap S_k^c = \emptyset$,

where S^c denotes the complement of S. As the reader can easily verify, this definition is equivalent to ours.

R_3 is between R_2 and R_4. However, neither R_2 nor R_3 is between R_1 and R_4 since, for example, (x_2, x_3), which is in the intersection of R_1 and R_4, is neither in R_2 nor in R_3.

The Hamming distances of the preferences in profile 7.1 are as follows. $R_1 \cap R_2 = \{(x_1, x_2), (x_1, x_4), (x_3, x_4)\}$. This makes: $(R_1 \cup R_2) - (R_1 \cap R_2) = \{(x_2, x_3), (x_2, x_4), (x_3, x_1), (x_3, x_2), (x_1, x_3), (x_4, x_2)\}$. Therefore, $h(R_1, R_2) = 6$. In the same way, we compute $h(R_1, R_3) = 8$, $h(R_1, R_4) = 10$, $h(R_2, R_3) = 2$, $h(R_2, R_4) = 8$ and $h(R_3, R_4) = 6$. This leads to the observation that $h(R_1, R_2)+h(R_2, R_3) = h(R_1, R_3)$ and $h(R_2, R_3)+h(R_3, R_4) = h(R_2, R_4)$, but $h(R_1, R_2)+h(R_2, R_4) > h(R_1, R_4)$ and $h(R_1, R_3)+h(R_3, R_4) > h(R_1, R_4)$. The next proposition shows that this observation holds in general.

Theorem 7.2 *For every $R_i, R_j, R_k \in B(X)$: R_j is between R_i and R_k if and only if $h(R_i, R_j) + h(R_j, R_k) = h(R_i, R_k)$.*

Proof.
1) Let R_j be between R_i and R_k. It suffices to prove that $\mid R_i \oplus R_j \mid + \mid R_j \oplus R_k \mid = \mid R_i \oplus R_k \mid$, that is, $(R_i \oplus R_j) \cup (R_j \oplus R_k) = R_i \oplus R_k$. Now:

$$(R_i \oplus R_j) \cup (R_j \oplus R_k) = (R_i \cup R_j \cup R_k) - (R_i \cap R_j \cap R_k)).$$

Since $R_j \subseteq (R_i \cup R_k)$, we have:

$$((R_i \cup R_j \cup R_k) - (R_i \cap R_j \cap R_k)) = ((R_i \cup R_k) - (R_i \cap R_j \cap R_k)).$$

Since $R_j \supseteq (R_i \cap R_k)$, we must have:

$$(R_i \cap R_j \cap R_k) = (R_i \cap R_k).$$

Therefore,

$$(R_i \oplus R_j) \cup (R_j \oplus R_k) = ((R_i \cup R_k) - (R_i \cap R_k) = R_i \oplus R_k.$$

2) Let $h(R_i, R_j) + h(R_j, R_k) = h(R_i, R_k)$,
i.e. $R_i \oplus R_j \cup R_j \oplus R_k = R_i \oplus R_k$,
i.e.

$$((R_i \cup R_j \cup R_k) - (R_i \cap R_j \cap R_k)) = ((R_i \cup R_k) - (R_i \cap R_k)). \quad (*)$$

We have to prove:
a) $R_j \subseteq R_i \cup R_k$, *and*
b) $R_j \supseteq R_i \cap R_k$.
First we prove a). Let $(x,y) \in R_j$ *and suppose* $(x,y) \notin R_i \cup R_k$. $(x,y) \notin R_i \cup R_k$, *implies* $(x,y) \notin R_i \cap R_k$. *Hence,* $(x,y) \notin ((R_i \cup R_k) - (R_i \cap R_k))$. *Since* $(x,y) \in R_j$, $(x,y) \in (R_i \cup R_j \cup R_k)$. *Clearly,* $(x,y) \notin (R_i \cap R_j \cap R_k)$. *Hence,* $(x,y) \in ((R_i \cup R_j \cup R_k) - (R_i \cap R_j \cap R_k))$. *Since* $(x,y) \notin ((R_i \cup R_k) - (R_i \cap R_k))$,

$$((R_i \cup R_j \cup R_k) - (R_i \cap R_j \cap R_k)) \neq ((R_i \cup R_k) - (R_i \cap R_k)),$$

which is in contradiction with (*). Hence $(x,y) \in R_i \cup R_k$.
Now we prove b). Let $(x,y) \in R_i \cap R_k$ *and suppose* $(x,y) \notin R_j$. *Then,* $(x,y) \in R_i \cap R_k)$ *implies* $(x,y) \notin (R_i \cup R_k) - (R_i \cap R_k)$.
Further, $(x,y) \in R_i \cap R_k)$ *also implies* $(x,y) \in (R_i \cup R_k)$. *Hence,* $(x,y) \in (R_i \cup R_j \cup R_k)$. *Since* $(x,y) \in (R_i \cap R_k)$ *but* $(x,y) \notin R_j$, $(x,y) \notin (R_i \cap R_j \cap R_k)$. *Hence,*

$$((R_i \cup R_j \cup R_k) - (R_i \cap R_j \cap R_k)) \neq ((R_i \cup R_k) - (R_i \cap R_k)),$$

which is in contradiction with (*). Hence, $(x,y) \in R_j$. □.

In order to obtain transitivity of betweenness of preferences as defined in definition 7.9, we use the concept of 'linear profile'. This concept is obviously inspired by the term 'interval' on the real line.

Definition 7.10 *Let* $p = (R_1, R_2, \ldots, R_n)$ *be a preference profile.* p *is called a* linear profile := *for all* $i, j, k = 1, 2, \ldots, n$ *and* $i \leq j \leq k$, R_j *is between* R_i *and* R_k.

In terms of Restle (1959), a linear profile is a linear array of preferences[7]. Note that the preferences in preference profile 7.1 cannot be re-arranged such that they form a linear profile.

Linear profiles will play a curious role in the solution part of preference distance theory of coalition formation (cf. section 7.5 below). Before turning to the coalition formation theories we first formulate a connection between linear profiles and majority choice.

[7]Consider Restle's definition of linear array (Restle 1959). To see that his definition of linear array applied to preference profiles is equivalent to our definition of linear profile, use theorems 5 and 6 in his work.

7.3.3 A Sufficient Condition for the Existence of Majority Choice

The system of majority decision has been discussed in chapter 3 of this monograph. In this subsection we show that there is a majority choice for every linear profile.

Lemma 7.1 *For every $i, j \in N$: $M(R_i, R_j)$ is between R_i and R_j.*

Proof. *If $(x, y) \in M(R_i, R_j)$, then $(x, y) \in R_i \cup R_j$. If $R_i \cap R_j \neq \emptyset$, then there is unanimity between i and j vis-a-vis x and y and hence $(x, y) \in M(R_i, R_j)$.* □

Theorem 7.3 *Let $p = (R_1, R_2, \ldots, R_n)$ be a linear profile. Then*

1. *$M(p) = R_{(n+1)/2}$ when n is odd,*
2. *$M(p) = M((R_{n/2}, R_{(n/2)+1}))$ when n is even.*

Proof.
1a). *Let p be a linear profile, let n be odd and let $(x, y) \in R_{(n+1)/2}$. Since p is a linear profile, $R_{(n+1)/2}$ is between R_j and R_k for every $j < (n+1)/2$ and $k > (n+1)/2$. Therefore, $(x, y) \in R_j \cup R_k$ for $j < (n+1)/2$ and $k > (n+1)/2$ by definition of betweenness. Suppose $(x, y) \notin R_j$ for $j < (n+1)/2$. Then $(x, y) \in R_k$ for every k such that $((n+1)/2) \leq k \leq n$ and hence, since $(x, y) \in R_{(n+1)/2}$ by hypothesis, $(x, y) \in M(p)$. In the same way, if $(x, y) \notin R_k$, then $(x, y) \in R_j$ for every j such that $1 \leq j \leq ((n+1)/2) - 1\}$ and hence, since $(x, y) \in R_{(n+1)/2}$ by hypothesis, $(x, y) \in M(p)$. If $(x, y) \in R_j$ and $(x, y) \in R_k$, then there is unanimity and, hence, $(x, y) \in M(p)$.*
1b). *Let p be a linear profile and suppose $(x, y) \in M(p)$. Suppose $(x, y) \notin R_{(n+1)/2}$. But then, since $| \{i \in N \mid xR_iy\} | \geq (n+1)/2$, there must be R_j and R_k with $j < (n+1)/2$ and $k > (n+1)/2$ such that $(x, y) \in R_j \cap R_k$. This implies that $R_{(n+1)/2}$ is not between R_j and R_k. Contradiction. Hence, $(x, y) \in R_{(n+1)/2}$.*
2a). *Let p be a linear profile, let n be even and let $(x, y) \in M(R_{n/2}, R_{(n/2)+1})$. Since $M(R_{n/2}, R_{(n/2)+1})$ is between $R_{n/2}$ and $R_{(n/2)+1}$ by lemma 7.1, and since p is a linear profile, $M(R_{n/2}, R_{(n/2)+1})$ is between every R_j and R_k*

with $j \leq n/2$ and $k \geq (n/2)+1$. Therefore, $(x,y) \in R_j \cup R_k$ for every $j \leq n/2$ and $k \geq (n/2)+1$. Suppose $(x,y) \notin R_j$ for $j \leq n/2$. Then $(x,y) \in R_k$ for every $(n/2)+1 \leq k \leq n$. Hence, $(x,y) \in M(p)$. In the same way we prove that $(x,y) \in M(p)$ when $(x,y) \notin R_j$ with $1 \leq j \leq n/2$.
2b). *Let p be a linear profile, let n be even and suppose $(x,y) \in M(p)$. If $(x,y) \notin M(R_{n/2}, R_{(n/2)+1})$, then $(x,y) \notin R_{n/2}$ and $(x,y) \notin R_{(n/2)+1}$. Hence, since $(x,y) \in M(p)$, there must be R_j and R_k with $j < n/2$ and $k > (n/2)+1$ such that $(x,y) \in R_j \cup R_k$. However, then neither $R_{n/2}$ nor $R_{(n/2)+1}$ between R_j and R_k and, hence, p is not a linear profile. Contradiction and thus $(x,y) \in M(R_{n/2}, R_{(n/2)+1})$.* □

This theorem provides us with a sufficient condition for the existence of majority choice.

Corollary 7.1 *Let $p = (R_1, R_2, \ldots, R_n)$ be a linear profile. Then there exists a majority choice.*

Proof. *If n is odd, then by theorem 7.3, $M(p) = R_{(n+1)/2}$. Since this individual preference is complete and transitive, $M(p)$ is complete and transitive and hence $Con(X, M(p)) \neq \emptyset$. If n is even, then by theorem 7.3, $M(p) = M(R_{n/2}, R_{(n/2)+1})$. Since $M(R_{n/2}, R_{(n/2)+1})$ is complete and acyclical, $M(p)$ is complete and acyclical and, hence, $Con(X, M(p)) \neq \emptyset$.* □.

Note that this result differs from the well known result of Black that a majority choice exists when a preference profile is single peaked (Black 1957: 16; Fishburn: 105). In the case of single peakedness, there is a linear order on the set of alternatives X such that "each individual's preference increases up to a peak or to an indifference plateau, and then decreases thereafter" (Fishburn 1973: 101). In the case of linear profiles, there is a weak ordering on the set of individuals N such that the preferences are between each other. Either condition, however, is sufficent for the existence of majority choice.

7.4 Conflict Minimization and Coalition Formation

In chapter 5, we discussed Axelrod's conflict of interest theory. The basic behavioral assumption of that theory is that each player strives to form a coalition with minimal conflict of interest. But, as we have discussed in chapter 5, there are some problems with the measurement of conflict of interest. In Axelrod's view, the conflict of interest notion has to do with the dispersion of the policy positions of the players. Since these policy positions are on an ordinal scale, the dispersion hardly can be measured. What only seems possible to say is that the range of the dispersion of the policy positions of the members in a coalition positively correlates with the degree of conflict in that coalition (Axelrod 1970: 169; also cf. chapter 5, section 7).

In this section we present a theory that is based on Axelrod's conflict of interest theory of coalition formation. The theory starts from the same behavioral assumption as Axelrod's one, namely, that each player strives to form a coalition with minimal conflict. However, our theory will go further. To solve the problem of the measurement of conflict in a coalition, we will use the theory of Hamming spaces as developed in the previous section. With the aid of this theory the notion of *conflict index* of a coalition will be developed. With this notion, the degree of conflict in each coalition can be measured. Another important difference with Axelrod's theory is that we assume that coalition formation processes take place in political systems that can be modelled as social choice games. Thus, the theory presented in this section uses more parameters than Axelrod's one (and is therefore comparatively more complex). In this sense, our theory may be thought of as a refinement of Axelrod's theory. Since conflict minimization is so important for our theory, it will be called the *conflict minimization theory of coalition formation.* Before presenting this theory, we first define and discuss the concept of conflict index.

7.4.1 Conflict Indices for Sets of Preferences

A *conflict index* is a mapping which assigns a real number to each subset of $B(X)$. The number assigned to a set of preferences $B_1 \subseteq B(X)$ is called the *conflict index of* B_1. Clearly, not every mapping is suitable. Firstly, a conflict index of a set of preferences must at least be zero. Either there is some conflict or else there is none. Secondly, if all preferences in a set of preferences are identical then the conflict index of this set must be equal to zero. Then, there is absolute consensus and hence no conflict. Formally:

Definition 7.11 *A* conflict index C *is a mapping from the power set of* $B(X)$ *into the set of real numbers* Re *such that for all* $B_1 \subseteq B(X)$,

1. $C(B_1) \geq 0$,
2. $C(B_1) = 0$ *if and only if for every* $R, Q \in B_1 : R = Q$.

The real number $C(B_1)$ *assigned to a* $B_1 \subseteq B(X)$ *by a conflict index* C *is called the* conflict index of B_1.

A conflict index using information about the Hamming distances between preferences in a set of preferences is called a *Hamming conflict index.*

Definition 7.12 *Let* $(B(X), h)$ *be a Hamming space. A function* $H : \mathcal{P}(B(X)) \rightarrow Re$ *is a* Hamming conflict index :=

1. *H is a conflict index and*
2. *there is a function f that assigns a real number to each set of Hamming distances such that for every* $B_1 \subseteq B(X)$, $H(B_1) = f(\{h(R, Q) \mid R, Q \in B_1\})$.

Clearly, there are many possible Hamming conflict indices. For example, the usual measures of dispersion like the standard deviation, the variance or the coefficient of variation of the Hamming distances for a set of preferences may be used. Another possibility is the ratio between the sum of Hamming distances $\sum h(R, Q)$ for $R, Q \in B_1 \subset B(X)$ and the maximum possible value this sum can take for B_1.

A Hamming measure that is most in accordance with the work of Axelrod (1970) and De Swaan (1973) is what we call the *conflict range.* Since our theory owes a great deal to the conflict of interest theory we shall explicate this concept. In section 7.4.4 we will use it in a computation example. The notion of dispersion as used in Axelrod's conflict of interest theory has been made explicit in the work of De Swaan (1973). There it is explicitly associated with the notion of range. De Swaan defines the range of a coalition S as the segment of the policy order that is between the most left player in S and the most right player of S. With the next definition, we keep in line with this approach. Remember that the diameter and the meshwidth of a set of preferences are defined in section 2 of this chapter.

Definition 7.13 *Let* $B_1(X) \subseteq B(X)$.

The conflict range of $B_1(X)$ *is the difference*

$$diam(B_1(X)) - mesh(B_1(X)).$$

Notation: $r(B_1(X))$.

2. *The* normalized conflict range of $B_1(X)$ *is equal to*

$$\frac{diam(B_1(X)) - mesh(B_1(X))}{diam(B_1(X))}.$$

Notation: $r_{norm}(B_1(X))$

It is easy to verify that for every $B_1 \subseteq B(X)$ the conflict range is nonnegative and that the conflict range is zero if for every $R, Q \in B_1$, $R = Q$.. The same is true for the normalized conflict range. Further, $0 \leq r_{norm}(B_1) \leq 1$ for every $B_1 \subseteq B(X)$. The normalized conflict range can be used for comparison purposes.

The concept of Hamming conflict index is basic in conflict minimization theory. However, the theory will be formulated such that it works for any Hamming conflict index. No reference to a particular index will be made.

7.4.2 Conflict Minimization Theory: Descriptive Part

Starting point is that each player wants to be a member of a winning coalition with a minimal Hamming conflict index. Consider a social choice game $G = (N, W, X, p)$. Remember that

$$W_i := \{S \in W \mid i \in S\}.$$

Define

$$R^p(S) := \{R_i^p \mid i \in S\}.$$

Now consider a Hamming conflict index H and let $H(R^p(S))$ denote the Hamming conflict index for S. A player $i \in N$ will strictly prefer a coalition S to T if and only if $S, T \in W_i$ and $H(R^p(S)) < H(R^p(T))$. Formally:

Definition 7.14 *Let $G = (N, W, X, p)$ be a social choice game, let $S, T \in W$, let $i \in N$ and let H be a Hamming conflict index.*

1. i strictly prefers S to T, *notation:* $S\pi_i^c T, :=$
   (a) $S \in W_i, T \notin W_i$ *or*
   (b) $S, T \in W_i$ *and* $H(R^p(S)) < H(R^p(T))$.
2. i *is* indifferent between S and T, *notation:* $S\iota_i^c T :=$
   (a) $S, T \notin W_i$ *or*
   (b) $S, T \in W_i$ *and* $H(R^p(S)) = H(R^p(T))$.
3. i weakly prefers S to T, *notation:* $S\rho_i^c T, := S\pi_i^c T$ *or* $S\iota_i^c T$.

The superscripts in the notation of these preferences are used to remind us that we are dealing with preferences as defined in conflict minimization theory. Since ρ_i^c is complete and transitive with respect to W, the set

$$\mu(W, \rho_i^c) = \{S \in W \mid \neg\exists T \in W[T\pi_i^c S]\}$$

is not empty[8].

7.4.3 Conflict Minimization Theory: Solution Part

The following step is to specify the solution part of conflict minimization theory. In accordance with De Swaan (1973, also see chapter 6), we may say that coalition S dominates coalition T if and only if every i in S strictly prefers S to T. Since every $i \in S - T$ strictly prefers S to T and

[8] For a general definition of maximal choice sets, consider chapter 2 and 4. For a study of the properties of such sets, consider chapter 4.

since every $i \in T - S$ strictly prefers T to S by assumption 6 of policy distance theory[9], only the strict preferences of the members in $S \cap T$ must be investigated in order to determine the dominance relation between S and T.

In this chapter we explicitly use the idea that only the preferences of the players in the intersection of two coalitions are decisive. This idea is worked out in In McKelvey, Ordeshook and Winer (1978) and in McKelvey and Ordeshook (1978). In their view, if $\exists i \in S \cap T : S\rho_i T$, then S is said to be viable against T. Clearly, if $S \cap T = \emptyset$, then there is no $i \in S \cap T$ such that $T\pi_i S$ and, hence, S is viable against T. Thus, the notion of viability is explicitly based on the idea that players in the *intersection* of two coalitions are 'critical' in the determination of the survival chances of the respective coalitions. This idea is attractive. To feel this intuitively, suppose that everyone in $S \cap T$ weakly prefers S to T and that at least one individual in $S \cap T$ strictly prefers S to T. Then T has little chance to be formed since the members in $T \cap S$ all prefer S to T. The players in $T - S$ are not able to realize T since all the players who are in $S \cap T$ will immediately leave T in order to participate in S (since they prefer S). Therefore, T is not viable against S. Now suppose there are $i \in S \cap T$ who prefer S to T and there are $i \in S \cap T$ who prefer T to S. Then, clearly, S and T both have a chance to be formed and thus S is viable against T and T is viable against S.

Definition 7.15 *Let $G = (N, W, X, p)$ be a social choice game and let $S, T \in W$.*

1. S is viable against T, *notation:* $S \succeq_c T$, := *there are* $i \in S \cap T$: *such that* $S\rho_i^c T$.

2. S *and* T *are* viable with respect to each other, *notation* $S \approx_c T$, := $S \succeq_c T$ *and* $T \succeq_c S$.

3. S is strictly viable *against* T, *notation:* $S \succ_c T$, := $S \succeq_c T$ *but not* $T \succeq_c S$.

4. *The set*

$$Co^c(G) = \mu(W, \succeq_c)$$

[9] See chapter 6.2.1.

is called the c-core of G.

A minimum conflict coalition in a social choice game is a coalition that has the smallest Hamming conflict index of all winning coalitions. Formally:

Definition 7.16 *Let* $G = (N, W, X, p)$ *be a social choice game and let* H *be a Hamming conflict index. A coalition* $S \in W$ *is a* minimum conflict coalition in G :=

$$H(R^p(S)) \leq H(R^p(T))$$

for every $T \in W$. *The set of all minimum conflict coalitions in* G *is denoted* $W^{mc}(G)$.

It is not difficult to verify that the c-core of a proper social choice game equals the set of minimum conflict coalitions in G.

Theorem 7.4 *Let* $G = (N, W, X, p)$ *be a proper social choice game. Then*

$$Co^c(G) = W^{mc}(G).$$

Proof. *Let* G *be a proper social choice game.*
1). First we prove $Co^c(G) \subseteq W^{mc}(G)$. *Let* $S \in Co^c(G)$. *We have to prove:* $S \in W^{mc}(G)$, *i.e.* $H(R^p(S)) \leq H(R^p(T)$ *for every* $T \in W$.
Since $S \in Co^c(G)$, $\neg\exists T \in W[T \succ_c S]$, *i.e.*
$\neg\exists T \in W[T \succeq_c S \wedge \neg S \succeq_c T]$, *i.e.*
$\neg\exists T \in W[\exists i \in S \cap T[T\rho_i S] \wedge \neg\exists i \in S \cap T[S\rho_i T]]$, *i.e.*
$\forall T \in W[\forall i \in S \cap T[S\pi_i^c T] \vee \exists i \in S \cap T[S\rho_i^c T]$.
Since G *is proper, for every* $T \in W$, $S \cap T \neq \emptyset$. *If* $\forall i \in S \cap T[S\pi_i^c T]$ *for every* $T \in W$, *then* $H(R^p(S)) < H(R^p(T))$ *for every* $T \in W$. *If* $\exists i \in S \cap T[S\rho_i^c T]$ *for every* $T \in W$, *then* $H(R^p(S)) \leq H(R^p(T))$ *for every* $T \in W$. *Thus,* $S \in W^{cm}(G)$.
2). We now prove $W^{cm}(G) \subseteq Co^c(G)$. *Let* $S \in W^{cm}(G)$. *Since* G *is proper, for every* $T \in W$, $S \cap T \neq \emptyset$. *Now, since for every* $T \in W$, $H(R_p(S)) \leq H(R^p(T))$, *we have* $\forall i \in S \cap T[S\rho_i^c T]$ *for every* $T \in W$. *Hence,* $\neg\exists T \in W[T \succ_c S]$, *i.e.* $S \in Co^c(G)$. □.

Clearly, only coalitions from the c-core of a social choice game will be formed. These coalitions all are minimum conflict coalitions.

Hypothesis 7.1 *Let $G = (N, W, X, p)$ be a social choice game and suppose that the c-core of G is not empty. Then only coalitions from the c-core of G will be formed.*

When applying conflict minimization theory to an empirical domain, for example, cabinet formation in multi-party systems, then this hypothesis states that only minimum conflict cabinets will be formed. The following result shows that this hypothesis can be used for any proper social choice game.

Theorem 7.5 *Let G be a proper social choice game. Then $Co^c(G) \neq \emptyset$.*

Proof. *By definition, $Co^c(G) = \mu(W, \succeq_c)$. It suffices to prove that $\succeq_c$ is complete and transitive.*
Completeness. Let $S, T \in W$. Since G is proper, $S \cap T \neq \emptyset$. Let $i \in S \cap T$. Since ρ_i^c is complete, $S\rho_i^c T$ or $T\rho_i^c S$. Therefore, $S \succeq_c T$ or $T \succeq_c S$.
Transitivity. Let $S, T, U \in W$ and suppose $S \succeq_c T$ and $T \succeq_c U$. Then:
$\exists i \in S \cap T : S\rho_i^c T$, hence $H(R^p(S)) \leq H(R^p(T))$.
$\exists i \in T \cap U : T\rho_i^c U$, hence $H(R^p(T)) \leq H(R^p(U))$.
Since $H(R^p(S)) \leq H(R^p(T))$ and $H(R^p(T)) \leq H(R^p(U))$, $H(R^p(S)) \leq H(R^p(U))$. Since G is proper, $S \cap U \neq \emptyset$. Hence, $\exists i \in S \cap U : S\rho_i^c U$ and therefore $S \succeq_c U$. $\square$.

According to theorem 7.5, to compute the c-core of a game, it suffices to determine the Hamming conflict index of every coalition and then to determine the coalitions with the smallest index.

7.4.4 Conflict Minimization Theory: Computation Example

To illustrate the working of conflict minimization theory we apply the theory to an imaginary case. In order to be able to determine the coalitional preferences of the players, we use the concept of conflict range as defined in definition 7.13.

Consider the social choice game $G = (N, W, X, p)$, where

1. $N = \{i, j, k, l\}$,
2. $W = \{\{i, j, k\}, \{i, j, l\}, \{i, k, l\}, \{j, k, l\}, \{i, j, k, l\}\}$,

3. $X = \{x, y, z\}$,

4. p is, in compact notation (cf. chapter 2):

$$\begin{array}{rc} R_i: & xyz \\ R_j: & x(yz) \\ R_k: & yxz \\ R_l: & zxy \end{array}$$

In general, Hamming distances for a preference profile can conveniently be presented in a symmetric matrix whose columns and rows consist of the individual preferences and whose cells contain the Hamming distances between the preferences. Leaving the names of the rows and columns aside, the matrix for the presented situation is:

	R_i	R_j	R_k	R_l
R_i	0	1	2	4
R_j	1	0	3	3
R_k	2	3	0	6
R_l	4	3	6	0

To see how these numbers are calculated, consider cel R_iR_k. We see that the symmetric difference $R_i \oplus R_k = \{xy, yx\}$. Hence, the Hamming distance between R_i and R_k is 2, which we fill in in cel R_iR_k.

Let us investigate the conflict range of each winning coalition (see definition 7.13).

1. Coalition $S_1 = \{i, j, k\}$: $diam(\{R_i, R_j, R_k\}) = 3$ and $mesh(\{R_i, R_j, R_k\}) = 1$. Hence $r(\{R_i, R_j, R_k\}) = 2$.

2. Coalition $S_2 = \{i, j, l\}$: $diam(\{R_i, R_j, R_l\}) = 4$ and $mesh(\{R_i, R_j, R_l\}) = 1$. Hence, $r(\{R_i, R_j, R_l\}) = 3$.

3. Coalition $S_3 = \{j, k, l\}$: $diam(\{R_j, R_k, R_l\}) = 6$ and $mesh(\{R_j, R_k, R_l\}) = 3$. Hence, $r(\{R_j, R_k, R_l\}) = 3$.

4. Coalition $S_4 = \{i, k, l\}$: $diam(\{R_i, R_k, R_l\}) = 6$ and $mesh(\{R_i, R_k, R_l\}) = 2$. Hence, $r(\{R_i, R_k, R_l\}) = 4$.

5. Coalition $S_5 = \{i, j, k, l\}$: $diam(R_i, R_j, R_k, R_l\}) = 6$ and $mesh(\{R_i, R_j, R_k, R_l\}) = 1$. Hence, $r(\{R_i, R_j, R_k, R_l\}) = 5$.

According to theorem 7.4, a c-core coalition is a minimum conflict coalition. Since $\{i, j, k\}$ is the coalition with the smallest conflict range, this coalition is in the c-core. In order to be complete, we construct the individual preferences with respect to the winning coalitions.

Using definition 7.14 we obtain the following preference profile concerning the set of winning coalitions.

$$\begin{array}{ll} i: & S_1 \pi_i S_2 \pi_i S_4 \pi_i S_5 \pi_i S_3 \\ j: & S_1 \pi_j S_2 \iota_j S_3 \pi_j S_5 \pi_i S_4 \\ k: & S_1 \pi_k S_3 \pi_k S_4 \pi_k S_5 \pi_k S_2 \\ l: & S_2 \iota_l S_3 \pi_l S_4 \pi_l S_5 \pi_l S_1 \end{array}$$

To see how the viability relation is constructed, consider coalition S_1 and coalition S_2. We have: $S \cap T = \{i, j\}$. Now i strictly prefers S_1 to S_2 and j also strictly prefers S_1 to S_2. Hence, by definition 3.5, $S_1 \succ_c S_2$. Proceeding in this way, we get:
$S_1 \succ S_2$ via $\{i, j\}$, $S_2 \approx S_3$ via $\{j, l\}$,
$S_3 \succ S_4$ via $\{k, l\}$, $S_4 \succ S_5$ via $\{i, k, l\}$,
$S_1 \succ S_5$ via $\{i, j, k\}$, $S_1 \succ S_4$ via $\{i, k\}$,
$S_1 \succ S_3$ via $\{j, k\}$, $S_2 \succ S_4$ via $\{i, l\}$,
$S_2 \succ S_5$ via $\{i, j, l\}$, $S_3 \succ S_5$ via $\{j, k, l\}$.
As expected, $S_1 = \{i, j, k\}$ is the core coalition since there is no coalition T such that $T \succ_c S_1$. It is predicted that this coalition will be formed.

7.5 Preference Distance Theory of Coalition Formation

In chapter 2, we discussed the concepts of social choice rule, social preference and social choice. These concepts will be used in the preference distance theory of coalition formation presented in this section.

Let $G = (N, W, X, p)$ be a social choice game. Starting point is the idea that in every coalition once formed, a social preference must be produced concerning a social choice problem X. The production of this social preference in a coalition will be done with the aid of a social choice rule.

Now suppose that every player knows the other players' preferences concerning X. In addition, suppose that every player knows the social choice rule of each coalition. Then every player is able to calculate the social preference of every winning coalition and thus he can compare the calculated social preferences with his own individual preference. In comparing these preferences, we assume that each player uses the Hamming distance function and that each player will prefer coalition S to coalition T if the Hamming distance between his preference and the social preference of S is smaller than the Hamming distance between his preference and that of T. The behavioral assumption is that each player strives to minimize the distance between his preference and the social preference of the coalition to be formed.

Preference distance theory as presented in this section is related to and in any case inspired by De Swaan's policy distance theory[10]. Instead of policy positions, however, we use preferences concerning a social choice problem. Further we use the Hamming distance function to calculate the distances between the individual preferences and the expected social preferences of the coalitions with respect to a social choice problem while De Swaan uses a non-metric policy distance notion.

7.5.1 Preference Distance Theory: Descriptive Part

Let $G = (N, W, X, p)$ be a social choice game and consider a coalition $S \in W$. The basic idea is that in each coalition a social choice rule will be used in order to determine a social preference. In chapter 2 we define a social choice rule as a function with domain $\Pi = (O(X))^N$. In the present context, we need social choice rules with domain $(O(X))^S$ where S is a coalition. However, this does not change the basic results of chapter 2. Let us call a member of $(O(X))^S$ a *coalitional preference profile.* A member of $(O(X))^S$ will be denoted by p_S.

Definition 7.17 *Let $G = (N, W, X, p)$ be a social choice game and S be a coalition. . A* social choice rule for *S is a function $F : (O(X))^S \rightarrow B(X)$. $F(p_S)$ is called the* social preference of S under F.

In order to determine a preference concerning W, every player will evaluate the distance between his preference and the social preference under

[10]See chapter 6 for a discussion of De Swaan's theory.

some F for every $S \in W$.. The assumption is that a player $i \in N$ will prefer an $S \in W$ to a $T \in W$ if the distance between R_i^p and $F(p_S)$ is less than the distance between R_i^p and $F(p_T)$. Remember that

$$W_i = \{S \in W \mid i \in S\}.$$

Definition 7.18 *Let $G = (N, W, X, p)$ be a social choice game. Let $i \in N, S, T \in W$ and h be the Hamming distance function. Let F be a social choice rule for S and F' a social choice rule for T.*

1. i strictly prefers S to T, *notation:* $S\pi_i^d T, :=$

   (a) $S \in W_i, T \notin W_i$ *or*

   (b) $S, T \in W_i$ *and* $h(R_i^p, F(p_S)) < h(R_i^p, F'(p_T))$.

2. i *is* indifferent between S and T, *notation:* $S\iota_i^d :=$

   (a) $S, T \notin W_i$ *or*

   (b) $S, T \in W_i$ *and* $h(R_i^p, F(p_S)) = h(R_i^p, F'(p_T))$.

3. i weakly prefers S to T, *notation:* $S\rho_i^d T, := S\pi_i^d T$ *or* $S\iota_i^d T$.

The superscript d in the notation of these preferences remind us that we are dealing with preferences as defined in preference distance theory of coalition formation. It is easy to verify that ρ_i^d is complete and transitive with respect to W for every $i \in N$. Therefore, for every $i \in N$:

$$\mu(W, \rho_i^d) = \{S \in W \mid \neg\exists T \in W[T\pi_i^d S]\}$$

is not empty[11]. A coalition $S \in \mu(W, \rho_i^d)$ is called a *minimum distance coalition for i.*

7.5.2 Preference Distance Theory: Solution Part

The following step is to specify the solution part of preference distance theory. Just like conflict minimization theory we employ the idea that players in the intersection of two coalitions are 'critical' for the survival of these coalitions.

[11]Cf. chapter 4, section 3 for a study of maximal choice sets.

Definition 7.19 *Let $G = (N, W, X, p)$ be a social choice game and let $S, T \in W$.*

1. *S is viable against T, notation: $S \succeq_d T$, := there are $i \in S \cap T$: $S \rho_i^d T$.*
2. *S and T are* viable with respect to each other := $S \succeq_d T$ *and* $T \succeq_d S$.
3. *S is* strictly viable against T, *notation: $S \succ_d T$, := $S \succeq_d T$ but not $T \succeq_d S$.*
4. *The set*
$$Co^d(G) = \mu(W, \succeq_d)$$
*is called the d-*core *of G.*

We assume that if the d-core of a social choice game is not empty, then only coalitions from the d-core of this game will be formed.

Hypothesis 7.2 *Let $G = (N, W, X, p)$ be a social choice game and suppose that the d-core of G is not empty. Then only coalitions from the d-core of G will be formed.*

When is the d-core of a social choice game nonempty? In the next theorem a sufficient condition is given. In this theorem, the notion of linear profile as developed in section 7.3.2 (this chapter), is crucial. Further, the system of majority decision is used[12].

Theorem 7.6 *Let $G = (N, W, X, p)$ be a proper social choice game. Let F be the system of majority decision and suppose F will be used in every $S \in W$. If p is a linear profile, then*

$$Co^d(G) \neq \emptyset.$$

Proof. *Let G be a proper social choice game and let $p = (R_1, R_2, \ldots, R_n)$ be a linear profile. We must analyze two cases, namely, the case that n is odd and the case that n is even.*
<u>Case 1</u>. Let n be odd. Consider the preference $R_{(n+1)/2}$. By theorem 7.3, $M(p) = R_{(n+1)/2}$. Therefore, $N \in \mu(W, \rho_{(n+1)/2})$. We prove that N is in the d-core, i.e., there is no $T \in W$ such that $T \succeq_d N$. Consider a $T \in W$. By theorem 7.3, there are three possibilities:

[12]Cf. chapter 3, section 1 for the definition of this system.

1. $M(p_T)$ *is between* R_1 *and* $R_{(n+1)/2}$,

2. $M(p_T)$ *is between* $R_{(n+1)/2}$ *and* R_n,

3. $M(p_T) = R_{(n+1)/2}$.

Possibility 1. *Let* $M(p_T)$ *be between* R_1 *and* $R_{(n+1)/2}$. *All players in* T *with a preference between* R_1 *and* $M(p_T)$ *prefer* T *to* N. *Also all players with a preference between* $M(p_T)$ *and the majority relation of the preferences of the players between* $M(p_T)$ *and* $R_{(n+1)/2}$ *prefer* T *to* N. *If player* $(n+1)/2$ *is in* T, *then this player will strictly prefer* N *to* T. *Hence* N *is viable against* T *in this case. If player* $(n+1)/2$ *is not in* T, *then* T *must contain a player* k *with* $k > (n+1)/2$ *in order to be winning. Since* $k \in N$ *and since* R_k *with* $k > (n+1)/2$ *is between* $R_{(n+1)/2}$ *and* R_n, k *strictly prefers* N *to* T. *Thus* N *is viable against* T *also in this case.*
Possibility 2. *Let* $M(p_T)$ *be between* $R_{(n+1)/2}$ *and* R_n. *All players with a preference between* $M(p_T)$ *and* R_n *prefer* T *to* N. *Also every player with a preference between* $M(p_T$ *and the majority relation of the preferences of the players with preferences between* $M(p_T)$ *and* $R_{(n+1)/2}$ *prefer* T *to* N. *If player* $(n+1)/2$ *is in* T, *then this player will strictly prefer* N *to* T. *Hence,* T *is not strictly viable against* N *in this case. If* $(n+1)/2$ *is not in* T, *then* T *must contain a player* l *in order to be winning such that* $l < (n+1)/2$. *Since* $l \in N$, *and since* R_l *between* R_1 *and* $M(p) = R_{(n+1)/2}$, l *strictly prefers* N *to* T. *Hence,* T *is neither strictly viable against* T *in this case.*
Possibility 3. *Let* $M(p_T) = R_{(n+1)/2}$. *Then, every player in* $T \cap N$ *will be indifferent between* N *and* T. *Hence,* N *and* T *are viable with resepct to each other.*
Case 2. *Let* n *be even. Then, according to theorem 7.3,*
$M(p) = M(R_{n/2}, R_{(n/2)+1})$. *We again prove that* N *is in the d-core. Consider a coalition* $T \in W$. *Applying theorem 7.3, there are three possibilities:*

1. $M(p_T)$ *is between* R_1 *and* $M(R_{n/2}, R_{(n/2)+1})$,

2. $M(p_T)$ *is between* $M(R_{n/2}, R_{(n/2)+1})$ *and* R_n,

3. $M(p_T) = M(R_{n/2}, R_{(n/2)+1})$.

Possibility 1. *Since G is proper, T must contain at least* $(n/2)+1$ *players in order to be winning. Hence, T must contain a player* k *with* $k \geq (n/2)+1$. *Since* R_k *is between* $M(R_{n/2}, R_{(n/2)+1})$ *and* R_n, *k strictly prefers N to T. Hence T is not strictly viable against N in this case.*
Possibility 2. *Since G is proper, T must contain a player* l *such that* $l \leq n/2$ *in order to be winning. Since* R_l *is between* R_1 *and* $M(R_{n/2}, R_{(n/2)+1})$, *l strictly prefers N to T. Hence T is not strictly viable against N.*
Possibility 3. *In this case, each player in* $T \cap N$ *will be indifferent between T and N. Hence, neither in this case is T strictly viable against N.* □.

Unfortunately, the requirement that a preference profile is a linear profile, is rather strong. It will not frequently be met in reality. The d-core of a proper social choice game may be empty then. If so, then preference distance theory will fail in predicting coalitions. To circumvent this problem we propose to use the theory of generalized stable sets[13].

Definition 7.20 *Let* $G = (N, W, X, p)$ *be a social choice game.*

1. *A nonempty set* $V \subseteq W$ *is a* generalized stable set *of* W :=
   (a) *for all* $S, T \in V$: *not* $S \succ_d^T T$,
   (b) *for every* $T \notin V$ *there is an* $S \in V$ *such that* $S \succ_d^T T$.
2. *A coalition S in a generalized stable set V of W is called a* stable distance coalition *in G.*

Applying theorem 4.16 of chapter 4 yields the following result.

Theorem 7.7 *Let* $G = (N, W, X, p)$ *be a social choice game. Then there is a generalized stable set of W if* $\succ_d$ *is not empty.*

Proof. *See theorem 4.16., chapter 4.* □.

Hence, using the theory of generalized stable sets enables us to make predictions about coalition formation in any social choice game.

[13]Of course, other solution theories may be used, for example, the theory of generalized optimal choice sets. See chapter 4.

Hypothesis 7.3 *Let $G = (N, W, X, p)$ be a social choice game. Then only stable distance coalitions in G will be formed.*

When dealing with an empirical domain, for example, cabinet formation in multi-party systems, we expect that in any case stable distance coalitions will be formed.

Besides coalitions, preference distance theory also predicts the social preference to be produced in a social choice game. That is, if the predicted coalitons are known, then the social preferences these coalitions will produce are known. E.g., if the system of majority decision is used in a winning coalition, then the majority relation for the members' preferences of that coalition is predicted.

7.5.3 Preference Distance Theory: Computation Example

To illustrate the working of the theory, we provide a simple case and compute the relevant sets.

Consider the social choice game $G = (N, W, X, p)$ where

1. $N = \{i, j, k\}$,
2. $W = \{S_1, S_2, S_3, S_4\}$, where
   $S_1 = \{i, j\}$,
   $S_2 = \{i, k, \}$,
   $S_3 = \{j, k\}$,
   $S_4 = \{i, j, k\}$,
3. $X = \{x, y, z, w\}$,
4. p is, in compact notation:

$$\begin{array}{ll} R_i: & xwzy \\ R_j: & yxwz \\ R_k: & zwyx \end{array}$$

In this preference profile no preference is between two other preferences. Hence, it is not a linear profile, so we do not have a guarantee for the existence of a core-solution[14]. The Hamming distance matrix[15] for this preference profile is:

	R_i	R_j	R_k
R_i	0	6	8
R_j	6	0	10
R_k	8	10	0

We now must calculate the social preference for each winning coalition. We do this by using the system of majority decision. We write xy instead of xMy. If x and y are majority indifferent, we write (xy). For convenience, we write $M(S_i)$ instead of the cumbersome $M(p_{S_i})$ where $i = 1, 2, 3, 4$. We then get:

1. $M(S_1) = xw, xz, wz, (xy), (yz), (wy)$,
2. $M(S_2) = wy, zy, (zx), (xw), (wz), (xy)$,
3. $M(S_3) = yx, (xw), (xz), (yw), (yz), (wz)$,
4. $M(S_4) = yx, xw, wz, zy, wy, xz$.

Now the Hamming distance between every individual preference and the social preference of every winning coalition can be calculated. The preferences of i, j and k are denoted, respectively, by R_i, R_j and R_k.

To illustrate, we calculate the distance between R_i and $M(S_4)$.
$R_i = \{xw, xz, xy, wz, wy, zy\}$ and
$M(S_4) = \{yx, xw, wz, zy, wy, xz\}$.
$R_i \cup M(S_4) = \{xw, xz, xy, yx, wz, wy, zy\}$ and
$R_i \cap M(S_4) = \{xw, xz, wz, wy, zy\}$. Hence,
$R_i \oplus M(S_4) = (R_i \cup M(S_4)) - (R_i \cap M(S_4)) = \{xy, yx\}$.
Since $\mid R_i \oplus M(S_4) \mid = 2$, $h(R_i, M(S_4) = 2$.
Proceeding in this way we obtain:
$h(R_i, M(S_1) = 3$,
$h(R_i, M(S_2) = 3$,
$h(R_i, M(S_4) = 2$.

$h(R_j, M(S_1) = 3$,

[14]See theorem 7.6.
[15]See 7.4.4.

$h(R_j, M(S_3) = 5,$
$h(R_j, M(S_4) = 4.$

$h(R_k, M(S_2) = 4,$
$h(R_k, M(S_3) = 5,$
$h(R_k, M(S_4) = 6.$

The preferences of each player with respect to the winning coalitions then are:

$$\begin{array}{ll} i: & S_4 \pi S_2 \iota S_1 \pi S_3 \\ j: & S_1 \pi S_4 \pi S_3 \pi S_2 \\ k: & S_2 \pi S_3 \pi S_4 \pi S_1 \end{array}$$

To illustrate the calculation of the viability relation, we provide an example. Consider coalitions S_1 and S_4. $S_1 \cap S_4 = \{i, j\}$. Since i strictly prefers S_4 to S_1 and j strictly preferes S_1 to S_4, $S_1 \succ_d S_4$. The viability relation can be represented in a matrix in which the rows and columns consist of the coalitions $S_1, \dots S_4$ and in which each cell contains a $\succ$ (strict viability) or an $\approx$ (viability). We note that with $S \prec T$ we mean that T is strictly viable against S.

$$\begin{pmatrix} \approx_{11} & \approx_{12} & \succ_{13} & \approx_{14} \\ \approx_{21} & \approx_{22} & \succ_{23} & \approx_{24} \\ \prec_{31} & \prec_{32} & \approx_{33} & \approx_{34} \\ \approx_{41} & \approx_{42} & \approx_{43} & \approx_{44} \end{pmatrix}$$

Looking at this matrix, we see that coalitions S_1, S_2 and S_4 are undominated. Thus these coalitions are in the d-core and we may expect that one of these coalitions may be formed.

A more complicated example is the proper social choice game $G = (N, W, X, p)$ where

1. $N = \{1, 2, 3, 4, 5\}$,

2. W is any coalition containing at least 3 members,

3. $X = \{x, y, z, w\}$,

4. p is, in compact notation,

$$\begin{array}{rl} R_1: & xyzw \\ R_2: & xwzy \\ R_3: & yxwz \\ R_4: & zwyx \\ R_5: & wzxy \end{array}$$

The social choice rule is the system of majority decision. Calculating this example through yields the d-core $\{\{1,2,3\},\{2,4,5\}\}$. Thus, the prediction is that one of these two coalitions will be formed.

Note that the computation process of this last example already is quite involved. First, since the game is proper, the majority relation of $2^4 = 16$ winning coalitions must be calculated. Then, each player is in 11 winning coalitions. Hence, $5 \times 11 = 55$ distances must be calculated between the players and the majority relations of the coalitions. Then the preferences of the players can be determined. Subsequently, the 16 winning coalitions must be pairwise compared in order to determine the viability relation. Consequently, $(16 \times 15)/2 = 120$ intersections of coalitions must be investigated. For each of these intersections, the preferences of the members must be investigated and with the aid of this the viability must be determined. When putting this in a 16 by 16 viability matrix, the core-coalitions may be detected by inspecting the rows of this matrix. When a row does not contain a $\prec$, the respective coalition is undominated.

7.6 Cabinet Formation Processes in Multi-Party Systems

Conflict minimization theory and preference distance theory are formal theories of coalition formation processes. To embody the theories with empirical content, it is necessary to identify an empirical domain and to relate the theories with this domain. Since the theories can in principle be related to processes of coalition formation in any collective decision-making body, there are several empirical domains possible. We choose here for cabinet formation processes in multi-party systems.

In chapters 5 and 6 we have modelled a multi-party parliamentary system as a weighted majority game in which the players are political parties. The weight of a player is the number of seats of a political party in parliament. A coalition is referred to as a *cabinet*. The quota of the game is the minimum number of seats needed to form a majority cabinet. In order to model a multi-party system as a social choice game, the social choice problem must be specified. As the social choice problem, we may, for example, choose the set of electoral issues about which the parties contested during the electoral competition, extended with additional alternatives that may play a role in the formation process. The preferences of the players then are the party preferences concerning this set of electoral issues and additional alternatives. Of course, the empirical investigation of these party preferences requires a lot of research efforts.

Conflict minimization theory then yields the following hypothesis applied to cabinet formation in multi-party systems:

Hypothesis 7.4 *In proper multi-party parliamentary systems, only minimum conflict cabinets will be formed.*

As conflict minimization theory tells us, a minimum conflict cabinet is a c-core cabinet. This means that it is an undominated cabinet in the sense that for no political party participating in that cabinet there is another cabinet with a lower conflict index.

Preference distance theory yields the following hypothesis with respect to cabinet formation in multi-party systems.

Hypothesis 7.5 *In multi-party parliamentary systems, only d-core coalitions will be formed.*

Since a d-core coalition may not exist, this hypothesis may not work. Therefore we have proposed an alternative hypothesis. Applying this hypothesis to cabinet formation in multi-party systems we get:

Hypothesis 7.6 *In multi-parliamentary systems, only stable distance coalitions will be formed.*

Stable distance coalitions exist when the asymmetric part of the viability relation is not empty[16].

[16] See theorem 7.7.

Testing these hypotheses will be a labor intensive and anyway complicated enterprise. Consider the hypotheses yielded by preference distance theory. For testing these hypotheses data are needed about the social choice problem, that is, about the set of alternatives, policy issues, etc. from which choices have to be made by the parties. Then, data about the preferences concerning this social choice problem must be collected and analyzed. Then, data about the rules determining winning and losing must be collected in order to calculate the set of possible winning coalitions. Also data about the weights of the parties in parliament must be collected in order to calculate the set of possible winning coalitions. Subsequently, data about social choice rules must be collected in order to determine the the social preference of each possible winning coalition concerning the social choice problem. After this, the distance between every party preference and the social preference of each winning coalition must be calculated. On the basis of these distances the preference of each political party with respect to the set of possible cabinets must be calculated. Subsequently, the intersection of each pair of winning coalitions must be investigated in order to construct the viability relation. After the construction of the viability relation, the viability matrix can be constructed. From this matrix the prediction set of cabinets can be derived by detecting the d-core coalitions. Otherwise, we must investigate the viability relation in order to find the stable distance coalitions. Clearly, testing preference distance theory will not be easy.

In testing conflict minimization theory or preference distance theory it may be useful to work with several social problems X_i. Each social choice problem then may be interpreted as an issue set. X may be defined, then, as the cartesian product of the issue sets. This approach will not change our theoretical results.

7.7 Other Fields of Application

The purpose of Hamming distance theory as presented in this chapter is to measure the dissimilarity between preferences in a preference profile. This theory has been used to construct two formal theories of coalition formation, namely, conflict minimization theory and preference distance theory. However, Hamming distance theory together with the notion of

(Hamming) conflict index also can be used for other purposes. In this section we sketch some other fields of application.

1. *Electoral research.* By gathering information by means of survey research about the party preferences of an electorate each time an election must take place, it is possible to construct a series of party preference profiles. The Hamming distance model can be used to analyze the variability of the degree of conflict in elections through time. This research need not be restricted to one specific country. The model allows the comparison of party preference profiles of different countries, eventually at different points of times. In this respect, the model provides a framework for comparative electoral research.

2. *Dynamic decision-making in committees.* The model can be used as a tool to analyze the dynamics of decision-making in committees. This analysis can be done in several ways:

   (a) First, it can be done by observing several preference profiles from the start of a social choice process until the moment of decision-making is arrived at. This research can be coupled with, for example, the hypothesis that the time necessary to arrive at the moment of selecting a social choice is longer, the greater the degree of conflict is at the start.

   (b) The research theme mentioned under (a) has to do with different preference profiles in a committee with respect to the *same* decision-making problem. However, it is also possible to compare preference profiles with respect to *different* decision-making problems in the same committee. Thus the patterns of variability of the degree of conflict in different social choice processes in the same committee can be mutually compared.

   (c) With the model also social choice processes in different committees can be compared. This comparison can be done with respect to the specific moments of voting. It can also be done with respect to the dynamics of the choice processes in the several committees.

3. *Policy-making*. The model has, besides research-technical and theoretical relevance, also a practical relevance. It can produce relevant information in concrete policy-making situations. Knowing the degree of conflict in such a situation is relevant information for politicians and policy-makers. With this information they can make estimations and, therefore, form expectations about the feasability of the possible policy options.

   Further, the model can be used as a tool to produce information about the degree of consensus among the individuals and groups who are the target of the concerned policy. With this information, the policy-makers can form an image of the implementability of the policy option as collectively choosen by the policy-making body.

Surely, there are other fields of application. It appears that the Hamming distance model is flexible enough to be used in several domains. There is even the possiblity to use the model in a policy-making environment. In this chapter, we have used it as a basic tool for the construction of coalition theories.

References

Arrow, K.J. (1963). *Social Choice and Individual Values*. Wiley, New York: Wiley. 2nd ed.

Arrow, K.J. (1967). Values and Collective Decision Making. In: P. Maslett & W.G. Runciman (eds.), *Philosophy, Politics and Society*. Third Series. Oxford: Basil Blackwell.

Arrow, K.J. (1977). Current Developments in the Theory of Social Choice. *Social Research 44*, 607-622.

Axelrod, R. (1970). *Conflict of Interest*. Chicago: Markham.

Banks, J.S. (1985). Sophisticated voting outcomes and the covering relation. *Social Choice and Welfare 1*, 295-306.

Banzhaf, J. (1965). Weighted voting doesn't work: A mathematical analysis. *Rutgers Law Review 19*, 317-343.

Barry, B. (1980). Is it Better to be Powerful or Lucky? *Political Studies 28*, Part 1: 183-194, Part 2: 338-352.

Barthelemy, J.P. & Monjardet, B. (1981). The Median Procedure in Cluster Analysis and Social Choice Theory. *Mathematical Social Sciences 1*, 235-267.

Berge, C. (1985). *Graphs*. Amsterdam: North-Holland.

Behzad, M., Chartrand, G. & Lesniak-Foster, L. (1979). *Graphs and Digraphs*. Prindle: Weber and Schmidt International Series.

Bezembinder, Th. (1987). Problems of content and structure in utilities for social choice. In: E.E. Roskam & R. Suck (Eds.), *Progress in Mathematical Psychology-I*. Amsterdam: Elsevier, 467-483.

Bezembinder, Th. & Van Acker, P. (1985). The Ostrogorski paradox and its relation to nontransitive choice. *Journal of Mathematical Sociology 11*, 131-158.

Black, D. (1958). *The Theory of Committees and Elections*. Cambridge: Cambridge University Press.

Blair, H.D. & Pollak, R.A. (1982). Acyclic collective choice rules. *Econometrica 50*, 931-943.

Blau, J. (1972). Arrow's Theorem with Weak Independence. *Economica 38*, 413-420.

Blau, J. & Deb, R. (1977). Social decision functions and veto. *Econometrica 45*, 871-879.

Bogart, K.P. (1973). Preference Structures I: Distance between Transitive Relations. *Journal of Mathematical Sociology 3*, 49-67.

Bollobas, B. (1986). *Combinatorics*. Cambridge: Cambridge University Press.

Bordes, G. (1976). Consistency, Rationality and Collective Choice. *Review of Economic Studies 43*, 451-457.

Bordes, G. (1979). Some More results on Consistency, Rationality and Collective Choice. In J. Laffont (ed.). *Aggregation and Relevation of Preferences*. Amsterdam: North Holland Publ. Comp.

Boute, S. (1981). On De Swaan's Policy Distance Coalition Theory. *European Journal of Political Research 9*, 313-320.

Boute, S. (1984). Subjective Policy Distance Theory. In: Holler (ed.). (1984).

Braithwaite, R. (1953). *Scientific Exlanation*. Cambridge: Cambridge University Press.

Brams, S. Lucas, W. & Straffin, Ph. (eds.), (1983). *Political and Related Models*. New York, Heidelberg, Berlin: Springer Verlag.

Brown, D.J. (1975). Aggregation of Preferences. *Quarterly Journal of Economics 89*, 456-469.

Browne, E. & Dreijmanis, J. (eds.), (1982). *Government Coalitions in Western Democracies*. New York: Longman.

Browne, E. & Franklin, M.N. (1986). Editor's Introduction: New Directions in Coalition Research. *Legislature Studies Quarterly* 469-493.

Browne, E. Gleiber, D.W. & Mashoba, C.S. (1984). Evaluating Conflict of Interest Theory: Western European Cabinet Coalitions, 1945-1980. *British Journal of Political Science 14*, 1-32.

Condorcet, (1785). Essai sur L'Application de l'Analyse a la Probabilité des Decisions Rendues a la Pluralité des Voix. Paris. Appeared in Condorcet (1986).

Condorcet, (1789). Sur la Forme des Élection. Paris. Appeared in Condorcet (1986).

Condorcet, (1793). Sur les Élections. *Journal d'Instruction Sociale 1*, 25-32. Appeared in Condorcet (1986).

Condorcet, (1986). *Sur les Elections et Autres Textes*. Corpus des Oeuvres de Philosophie en Langue Française. Paris: Librairie Arthème Fayard.

Coombs, C. & Avrunin, (1988). *The Structure of Conflict*. Hillsdale: Lawrence Erlbaum Associates.

Cox, G.W. (1987). The Uncovered Set and the Core'. *American Journal of Political Science*. 31: 408-422.

Curiel, I. (1987). A class of non-normalized power indices for simple games. *Mathematical Social Sciences 13*, 141-152.

Daalder, H. (1984). In search of the center of European multi-party systems. *American Political Science Review 78*, 92-109.

Deegan, J. & Packel, E. (1983). To the (minimal winning) victors go the (equally divided) spoils: A new power index for simple n-person games. In: Brams et.al. (1983). 239-255.

De Swaan, A. (1973). *Coalition Theories and Cabinet Formations*. Amsterdam: North Holland.

De Swaan, A. (1975). A Classification of Parties and Party System According to Coalition Options. *European Journal of Political Research 3*, 361-375.

De Swaan, A. (1982). The Netherlands: Coalitions in a Segmented Society. In: Browne, E. & Dreijmanis, J. (eds.), (1982).

De Waal, F. (1981). *Chimpanzee Politics: Power and Sex among Apes*. London: Jonathan Cape & Harper & Row.

Dodd, L. (1976). *Coalitions in Parliamentary Government*. Princeton: Princeton University Press.

Dodgson, C.L. (1876). A Method of Taking Votes on More than Two Issues. In: Black (1957).

Duchet, P. (1987), A Sufficient Condition for a Graph to be Kernel-Perfect. *Journal of Graph Theory 11*, 81-85.

Einy, E. (1985). On Connected Coalitions in Dominated Simple Games. *International Journal of Game Theory 14*, 103-125.

Elster, J. (1983). *Sour Grapes*. Cambridge: Cambridge University Press.

Elster, J. (1986). The Market and the Forum: Three Varieties of Political Theory. In: Elster, J. & Hylland, A. (eds.), *Foundations of Social Choice Theory*. Cambridge: Cambridge University Press.

Enelow, J.M. & Hinich, M.J. (1984). *The Spatial Theory of Voting*. Cambridge: Cambridge University Press.

Fishburn, P. (1973). *The Theory of Social Choice*. Princeton: Princeton University Press.

Fishburn, P. (1977). Condorcet Social Choice Functions. *SIAM Journal of Applied Mathematics 33*, 469-489.

Galeana-Sanchez, H. & Neumann-Lara, V. (1984). On Kernels and Semikernels of Digraphs. *Discrete Mathematics 48*, 67-76.

Gardenfors, P. (1973). Positionalist Voting Functions. *Theory and Decision 4*, 1-24.

Gardenfors, P. (1985). Rights, Games and Social Choice. *Noûs 25*, 341-356.

Gerhlein, W.V. (1983). Condorcet's Paradox. *Theory and Decision 15*, 161-197.

Gibbard, A. (1969). Social choice and the Arrow conditions. Unpublished

Graaf, J. de (1957). *Theoretical Welfare Economics*. Cambridge: Cambridge University Press. 2nd. ed.

Grofman, B. (1984). The General Irrelvance of the Zero Sum Assumption in the Legislative Context. In: Holler, M. (1984).

Grofman, B. (1989). The Comparative Analysis of Coalition Formation and Duration: Distinguishing Between-Country and Within-Country Effects. *British Journal of Political Science 19*, 291-302.

Grofman, B. Owen G. Noviello N. & Glazer, A. (1987). Stability and Centrality of Legislative Choice in the Spatial Context. *American Political Science Review 81*, 538-53.

Griffin, D.R. (1984). *Animal Thinking*. Cambridge: Harvard University Press.

Hansson, B. (1976). The Existence of Group Preference Functions. *Public Choice 28*, 89-98.

Harary, F. Norman R. & Cartwright D. (1965). *Structural Models: An Introduction to the Theory of Directed Graphs*. New York: John Wiley and Sons, Inc.

Hindess, B. (1988). *Choice, rationality and social theory.* London: Unwin Hyman.

Holler, M. (1982). Forming Coalitions and Measuring Voting Power.. *Political Studies, Volume XXX,* 262-271.

Holler, M. (ed.), (1984). *Coalitions and Collective Actions.* Würburg-Wien: Physica-Verlag.

Holler, M. & Packel, E. (1983). Power, Luck and the Right Index. *Journal of Economics 43,* 21-29.

In 't Veld, R.J. (1975). *Meerderheidsstelsels en Welvaartstheorie.* Leiden: Stenfert Kroese.

Kalai, E. Pazner, E.A. & Schmeidler, D. (1976). Collective Choice Correspondences as Admissible Outcomes of Social Bargaining Processes. *Econometrica 44,* 233-240.

Kalai, E. & Schmeidler, D. (1977). An Admissible Set Occuring in Various Bargaining Situations. *Journal of Economic Theory 14,* 402-411.

Kelly, J.S. (1978). *Arrow Impossibility Theorems.* New York: Academic Press.

Kelly, J.S. (1988). *Social Choice Theory: An Introduction.* Berlin, Heidelberg, New York: Springer-Verlag.

Kemeny, J.G. (1959). Mathematics without numbers. *Daedulus 88,* 577-591.

Kemeny, J.G. & Snell, J.C. (1962). *Mathematical Models in the Social Sciences.* New York.

Krehbiel, K. (1988). Spatial Models of Legislative Choice. *Legislative Studies Quarterly XIII,* 259-319.

Laver, M. (1986). Between Theoretical Elegance and Political Reality. In: Pridham G. (ed.) *Coalition Behavior in Theory and Practice: An Inductive Model for Western Europe.* Cambridge. 32-43.

Lakatos, I. (1978). Popper on demarcation and induction. In: Lakatos, I. *The methodology of scientific research programmes. Philosophical Papers. Volume 1.* Cambridge: Cambridge University Press.

Luce, L.D. & Raiffa, H. (1957). *Games and Decisions.* New York: Wiley.

Lucas, W.F. (1977). The Existence Problem For Solutions. In: Henn, R. & Moeschn, O. (eds.), *Mathematical Economics and Game Theory. Essays in Honor of Oskar Morgenstern.* Berlin: Springer-Verlag.

Lucas, W.F. (1983). Measuring Power in Weighted Voting Games. In: Brams et.al. (1983). 183-238.

Mas-Collel, A. & Sonnenschein, H.F. (1972). General Possibility Theorems for Group Decisions. *Review of Economic Studies 39,* 185-192.

May, K.O. (1952). A Set of Independent, Necessary and Sufficient Conditions for Simple Majority Decision. *Econometrica 20,* 680-684.

McKelvey, R.D. (1986). Covering, Dominance and Institution-Free Properties of Social Choice. *American Journal of Political Science 30,* 283-314.

McKelvey, R. & Ordeshook, P. (1978). Competitive Coalition Theory. In: Ordeshook, P. (ed.). *Game Theory and Political Theory.* New York: New York University Press. 1-39.

McKelvey, R.D., Ordeshook, P.C. & Winer, M.D. (1978). The Competitive Solution Theory for N-Person Games without Transferable Utility with an Application to Committee Games. *American Political Science Review 72,* 599-615.

Mesarovic, M.D. & Takahara Y. (1975). *General Systems Theory: Mathematical Foundations.* New York: Academic Press.

Miller, N.R. (1980). A New Solution Set for Tournaments and Majority Voting. *American Journal of Political Science 24,* 68-96.

Miller, N.R. (1983). The covering relationship in tournaments: Two corrections. *American Journal of Political Science 27,* 382-5.

Moulin, H. (1983). *The Strategy of Social Choice.* Amsterdam: North-Holland Publ. Company.

Mueller, D. (1989). *Public Choice II.* Cambridge: Cambridge University Press.

Murakami, Y. (1968). *Logic and Social Choice.* Dover.

Nakamura, K. (1975). The Core of a Simple Game with Ordinal Preferences. *International Journal of Game Theory 4,* 95-104.

Nakamura, K. (1979). The Vetoers in a Simple Game with Ordinal Preferences. *International Journal of Game Theory 8,* 55-61.

Niemi, R.G. & Weisberg H. (1968). A mathematical solution for the probability of the paradox of voting. *Behavioral Science 13,* 317-323.

Nolte, D. (1988). Ist die Koalitionstheorie am Ende? Eine Bilanz nach 25 Jahren Koalitionsforschung. *Politisches Vierteljahrschrift* 230-251.

Ordeshook, P. (1986). *Game Theory and Political Theory.* Cambridge: Cambridge University Press.

Packel, E. & Deegan, J. (1980). An axiomatic family of power indices for simple n-person games. *Public Choice 35,* 229-239.

Pattanaik, P.K. (1971). *Voting and Collective Choice.* Cambridge: Cambridge University Press.

Pattanaik, P.K. (1978). *Strategy and Group Choice.* Amsterdam: Norh-Holland Publ. Company.

Peleg, B. (1980). A Theory of Coalition Formation in Committees. *Journal of Mathematical Economics 7,* 115-134.

Peleg, B. (1981). Coalition Formation in Simple Games with Dominant Players. *International Journal of Game Theory 10,* 11-33.

Peleg, B. (1983). On Simple Games and Social Choice Correspondence. In: Pattanaik, P. & Salles, M. (eds.). *Social Choice and Welfare.* Amsterdam: North Holland Comp. 251-268.

Peleg, B. (1984). *Game Theoretic Analysis of Voting in Committees*. Cambridge: Cambridge University Press.

Pennock, J.R. (1979). *Democratic Political Theory*. Princeton: Princeton University Press.

Plott, C. (1973). Path independenc, rationality and social choice. *Econometrica 41*, 1075-1091.

Plott, C. (1976). Axiomatic Social Choice Theory: An Overview and Interpretation. *American Journal of Political Science 20*, 511-596.

Popper, K (1959). *The Logic of Scientific Discovery*. London: Hutchinson.

Restle, F. (1959). A Metric and an Ordering on Sets. *Psychometrika 24*, 207-220.

Richardson, M. (1953). Solutions of Irreflexive Relations. *Annals of Mathematics 58*, 573-590.

Riker, W. (1962). *The Theory of Political Coalitions..* New Haven, New York.

Riker, W. (1988). The place of political science in public choice. *Public Choice 57*, 247-257.

Riker, W. & Ordeshook, P. (1973). *Introduction to Positive Political Theory*. New Jersey: Prentice-Hall.

Samuelson, P.A. (1967). Arrow's Mathematical Politics. In: Hook, S. (ed.). *Human Values and Econonmic Policy*. New York: New York University Press.

Samuelson, P.A. (1977). Reaffirming the existence of reasonable Bergson-Samuelson social welfare functions. *Economica 44*, 81-88.

Schofield, N. Grofman B. & Feld, S. (1988). The Core and The Stability of Group Choice in Spatial Voting Games. *American Political Science Review 82*, 195-211.

Schmeidler, D. & Sonnenschein H. (1978). Two Proofs of the Gibbard-Satterthwaite Theorem on the Possibility of a Strategy-proof Social Choice Function. In: Gottinger, H. & Leinfeller, W. *Decision Theory and Social Ethics: Issues in Social Choice*. Dordrecht: Reidel.

Schofield, N. (1982). Bargaining Set Theory and Stability in Coalition Governments. *Mathematical Social Sciences 3*, 9-31.

Schofield, N. & Laver M. (1985). Bargaining Theory and Portfolio Payoffs in European Coalition Governments 1945-1983. *British Journal of Political Science 15*, 143-164.

Schustermann, J. Thomas, J.A. & Wood, F. G. (eds.), (1986). *Dolphin cognition and behavior: a comparative approach*. Hillsdale: L.Erlbaum Associates.

Schwartz, T. (1972). Rationality and the Myth of the Maximum. *Noûs 6*, 97-117.

Schwartz, T. (1986). *The Logic of Collective Choice*. New York: Columbia University Press.

Sen, A. (1970). *Collective Choice and Social Welfare*. San Fransisco: Holden Day Inc.

Sen, A. (1976). Liberty, unanimity and rights. *Economica 43*, 217-245.

Sen, A. (1977). Social Choice Theory: A Re-examination. *Econometrica 45*, 53-89.

Sen, A. (1986). Social Choice Theory. In: Arrow, K.J. & Intriligator, M.D. (eds.). *Handbook of Mathematical Economics*. Volume III, 1073-1182.

Shapley, L. (1962). Simple Games: An Outline of the Descriptive Theory. *Behavioral Science 7*, 59-66.

Shapley, L. (1967). On Committees. In: Zwicky, F. & Wilson, A.G. (eds.). *New Methods of Thought and Procedure* 256-270.

Shapley, L. (1981). Measurement of Power in Political Systems. *Proceedings of Symposia in Applied Mathematics*. Volume 24. 69-81.

Shapley, L. & Shubik, M. (1954). A method for evaluating the distribution of power in a committee system. *American Political Science Review 48*, 787-792.

Shenoy, P.P. (1979). On Coalition Formation: A Game-Theoretical Approach. *International Journal of Game Theory 8*, 133-164.

Shenoy, P.P. (1980). A Dynamic Solution Concept for Abstract Games. *Journal of Optimization Theory and Applications 32*, 151-169.

Shubik, M. (1982). *Game Theory in the Social Sciences: Concepts and Solutions*. Cambridge, Massachusetts, London.

Storcken, T. (1989). *Possibility Theorems for Social Welfare Functions*. Diss. Tilburg.

Straffin, P. (1980). *Topics in the Theory of Voting*. Boston: Birkhäuser.

Straffin, P. (1983). Power indices in politics. In: Brams et. al. (1983). 256-323.

Suppes, P. (1957). *Introduction to Logic*. New York: D. von Nostrand Company.

Suzumura, K. (1983). *Rational Choice, Collective Decisions and Social Welfare*. Cambridge: Cambridge University Press.

Taylor, M. & Laver, M. (1973). Government Coalitions in Western Europe. *European Journal of Political Research 1*, 205-248.

Tops, P. (1990). *Afspiegeling en Afspraak*. Diss. Leiden.

Van Deemen, A. (1987). Dominant Players, Centrum Players and Coalition Formation in Simple Games. Paper delivered at the ECPR Conference. April, Amsterdam.

Van Deemen, A. (1988). Collective Rationality and Social Choice. Paper delivered at the Politicologen Etmaal, Workshop *Filosofie van het Openbaar Bestuur* and workshop *Public Choice*. May. Technical University Twente, Enschede.

Van Deemen, A. (1989). Dominant Players and Minimum Size Coalitions. *European Journal of Political Research 17*, 313-332.

Van Deemen, A. (1990). Theories of Center Players and Cabinet Formations: With an Application to the Dutch Parliamentary System. *Acta Politica 25*, 187-208.

Van Deemen, A. (1990a). *A Note on Generalized Stable Sets*. Accepted for publication in *Sociál Choice and Welfare*.

Van Deemen, A. (1991). Coalition Formation in Centralized Policy Games. *Journal of Theoretical Politics 3*, 139-161.

Van Roozendaal, P. (1990). Centre parties and coalition cabinet formations: a game theoretic approach. *European Journal of Political Research 18*, 325-348.

Von Neumann, J. & Morgenstern, O. (1947). *Theory of Games and Economic Behavior*. Princeton: Princeton University Press. 2nd. ed.

Wriglesworth, J.L. (1985). *Libertarian Conflicts in Social Choice*. Cambridge: Cambridge University Press.

Wuffle, A. Feld, S.F. Owen, G. & Grofman, B. (1989). Finagle's Law and the Finagle Point, a New Solution Concept for Two-Candidate Competition in Spatial Voting Games without a Core. *American Journal of Political Science 33*, 348-375.

Young, H.P. (1974). An Axiomatization of Borda's Rule.' *Journal of Economic Theory 9*, 43-52.

Young, H.P. (1975). Social Choice Scoring Functions. *SIAM Journal of Applied Mathematics 28*, 824-838.

Young, H.P. (1988). Condorcet's Theory of Voting. *American Political Science Review 82*, 1231-1244.

Young, H.P & Balinsky, M.L. (1982). *Fair Representation.* New Haven: Yale University Press.

Samenvatting

De essentie van politiek is winnen. Wat telt is het afdwingen van een collectieve beslissing en de naleving daarvan, het aannemen van een wet, het winnen van een verkiezing, het omverwerpen van een dictatoriaal regime, het doorzetten van beleid, etc. Echter, voor een overwinning is de formatie van een coalitie die voldoende macht heeft om deze overwining af te dwingen noodzakelijk. Derhalve zijn coalitieformatieprocessen veelvuldig waar te nemen in het politieke leven. In dit boek staat het formeren van politieke coalities centraal.

Er bestaat binnen de politieke wetenschap een traditie om coalitieformatieprocessen vanuit een speltheoretische invalshoek te bestuderen. Deze traditie is ontstaan omdat de n-persoons coöperatieve speltheorie zoals ontwikkeld in Von Neumann en Morgenstern (1953) en zoals toegankelijk gemaakt voor sociale en politieke wetenschappers in het excellente werk van Luce en Raiffa (1957), als zodanig niet gericht is op de formatie van coalities maar op uitbetalingen aan spelers. De n-persoons coöperatieve speltheorie is als het ware blind voor de formatie van coalities. Omdat coalitieformatie nu juist zo belangrijk is in de politiek, dient de speltheorie derhalve uitgebreid te worden met additionele concepten en assumpties. Ons werk sluit aan op deze uitgebreide speltheoretische benadering van coalitieformatieprocessen. We zullen binnen dit kader een aantal reeds geformuleerde theorieën kritisch bestuderen. Daarnaast presenteren en bestuderen we een aantal nieuwe theorieën. Deze nieuwe theorieën zijn geformuleerd teneinde een tweetal theoretische tekortkomingen binnen de speltheoretische benadering van coalitieformatie op te heffen.

De eerste tekortkoming betreft de afwezigheid van een actorgerichte beleidstheorie. De in de loop van de tijd geformuleerde coalitietheorieën binnen de speltheoretische traditie kunnen in een aantal klassen onderverdeeld worden. De eerste klasse is die van de machtstheorieën. Een andere klasse betreft beleidstheorieën. Bij machtstheorieën spelen de machtsposities van de speler een doorslaggevende rol in de formatie van coalities. Bij beleidstheorieën spelen naast de machtposities ook de beleidsposities een verklarende rol. Daarnaast kunnen we de zogenaamde actorgerichte machtstheorieën onderscheiden. Deze proberen aan te geven wanneer en hoe de machtspositie van een specifieke speler (actor) doorslaggevend kan zijn in het formeren van een coalitie. We beschikken momenteel over zo'n actorgerichte machtstheorie, nl. de theorie van dominante spelers van Peleg. Nu is binnen deze classificatie ook een deelklasse van actorgerichte beleidstheorieën te onderscheiden. Zo'n actorgerichte beleidstheorie zou moeten aangeven wanneer en hoe de machtspositie tezamen met de beleidspositie van een specifieke speler de doorslag geven in een coalitieformatieproces. Er bestaat echter geen actorgerichte beleidstheorie, ondanks het feit dat bijna alle relevante empirische onderzoeken het belang van beleid in coalitieformatieprocessen ondersteunen. Een eerste doelstelling van dit werk is dan ook het vormen van een actorgerichte beleidstheorie. Dit doen we in hoofdstuk 5.

De tweede theoretische tekortkoming is dat in bijna alle tot nu toe geformuleerde theorieën de voorkeuren van de spelers met betrekking tot coalities een ondergeschikte rol spelen. Er is één uitzondering, nl. de beleidsafstandstheorie van De Swaan. Deze laatste theorie leidt tot een aantal voorkeuren, één voor elke speler, voor coalities die gebruikt worden voor de voorspelling van coalities. We sluiten aan op het werk van De Swaan door verschillende theorieën te formuleren waarin coalitionele voorkeuren een doorslaggevende rol spelen. Uitgangspunt van deze nieuwe theorieën is dat het formeren van een coalitie gezien wordt als een collectief besluitvormingsprobleem. Binnen dit verband onderscheiden we een tweetal typen van besluitvormingstheorieën van coalitieformatie.

1. In het eerste type is de verzameling van mogelijke winnende coalities zelf het collectief besluitvormingsprobleem. Coalitieformatie wordt dan gezien als het collectief kiezen van een coalitie uit deze verzameling. De spelers worden hierbij geacht voorkeuren te formeren met betrekking tot deze verzameling. Vervolgens worden deze voorkeuren geaggregeerd tot een sociale voorkeur die als zodanig de basis is voor het bepalen van een oplossing van het keuzeprobleem.

2. In het tweede type besluitvormingstheorieën van coalitieformatie is het besluitvormingsprobleem een verzameling van alternatieven. De spelers

staan nu voor het probleem collectief een keuze te bepalen. Dit kan echter alleen gebeuren door een coalitie te formeren die deze collectieve keuze kan afdwingen (een winnende coalitie dus). De belangrijke vraag die dit type theorie tracht te beantwoorden is hoe de voorkeuren van de spelers met betrekking tot winnnende coalities bepaald worden door de voorkeuren van de spelers met betrekking tot de alternatieven in het collectief besluitvormingsprobleem.

Het hoofddoel van dit boek is het formuleren van collectieve besluitvormingstheorieën van coalitieformatieprocessen. We formuleren zowel een theorie die van het eerste type is, namelijk de *machtsoverschottheorie*, als twee theorieën die van het tweede type zijn, nl de *conflictminimalisatietheorie* en de *voorkeurafstandstheorie*.

1. De machtsoverschottheorie start vanuit de gedachte dat een speler een winnende coalitie S prefereert boven een winnende coalitie T indien deze speler een groter machtsoverschot heeft in S dan in T.

2. De conflictminimalisatietheorie start vanuit de assumptie dat iedere speler zal proberen een coalitie te formeren van spelers wiens voorkeuren zo dicht mogelijk bij elkaar liggen. Voor de specificatie van de termen 'zo dicht mogelijk' bebruiken we een afstandsmodel dat ons in staat stelt de afstanden voor elk paar voorkeuren te bepalen. De conflictminimalisatietheorie leidt tot profielen van coalitionele voorkeuren die geaggregeerd kunnen worden en waarmee voorspellingen gedaan kunnen worden over de te vormen coalities.

3. De voorkeurafstandstheorie is gebaseerd op de gedachte dat binnen elke eenmaal geformeerde coalitie een besluitvormingsproces plaats zal vinden dat zal leiden tot een sociale voorkeur van die coalitie met betrekking tot het oorspronkelijke besluitvormingsprobleem. Deze sociale voorkeur is een aggregatie van de voorkeuren (m.b.t. de verzameling van alternatieven) van de leden van deze coalitie. De veronderstelling hierbij is

   (a) dat elke speler de voorkeur van elke andere speler kent en

   (b) dat elke speler van elke winnende coalitie weet welke aggregatieprocedure gebruikt zal worden.

Elke speler kan derhalve de afstanden bepalen tussen haar of zijn voorkeur en de voorkeur van elke mogelijke winnende coalitie. Op basis van deze afstanden zal elke speler vervolgens haar of zijn voorkeur met betrekking tot coalities vaststellen. De assumptie is dat een speler een winnende coalitie met een sociale voorkeur die dichter bij haar of zijn individuele voorkeur ligt zal preferen boven een winnende coalitie wiens sociale voorkeur verder af ligt. Voor het bepalen van de afstanden tussen sociale voorkeuren van coalities en individuele voorkeuren wordt gebruik gemaakt van het afstandsmodel dat ook gebruikt wordt voor conflictminimalisatietheorie. Ook de voorkeurafstandstheorie leidt tot profielen van coalitionele voorkeuren die geaggregeerd kunnen worden en waarmee voorspellingen over de te vormen coalities gemaakt kunnen worden.

De sociale keuzetheorie bestudeert collectieve besluitvormingsprocessen. De theorie gaat over het aggregeren van individuele voorkeuren tot een sociale voorkeur. Omdat we in dit werk coalitieformatie zien als een collectief besluitvormingsproces is de sociale keuzetheorie van groot belang. We besteden derhalve uitgebreid aandacht aan deze theorie, nl. in hoofdstukken 2, 3 en 4. Een bekend probleem gesignaleerd door de sociale keuzetheorie is dat een best of maximaal element voor een sociale voorkeur niet altijd hoeft te bestaan. Een best alternatief is in ieder geval afwezig in geval van cyclische sociale voorkeuren. Omdat we apriori niet kunnen uitsluiten dat aggregaties van coalitionele voorkeuren cyclisch zijn, dienen we te beschikken over theorieën die aangeven hoe een collectieve keuze te bepalen in geval van cycli. De formulering van deze theorieën is niet alleen nodig om een adequate coalitieformatietheorie te bouwen, maar heeft ook waarde op zichzelf. Het probleem van het 'kraken' van een cyclische sociale voorkeur is één van de centrale problemen in de sociale keuzetheorie zelf. Een doelstelling van dit werk is dan ook de bestudering van methoden die keuzeverzamelingen voor cyclische relaties kunnen specificeren.

In hoofdstuk 2 worden de basisbegrippen van de sociale keuzetheorie besproken. We gaan hierbij in op standaardthema's als rationele keuzes en collectieve rationaliteit. Verder bespreken we een aantal eigenschappen van aggregatieprocedures. Binnen dit kader bespreken we de stelling van Arrow. We geven een eigen bewijs van deze stelling.

De sociale keuzetheorie zoals gepresenteerd in hoofdstuk 2 is abstract. In hoofdstuk 3 daarentegen bestuderen we een aantal concrete keuzestelsels. De meeste aandacht hierbij gaat uit naar het meerderheidsstelsel. Zoals bekend hoeft een alternatief dat een meerderheid heeft over elk ander alternatief niet te bestaan. Dit verschijnsel is reeds ontdekt door de Franse rationalist Marquis de Condorcet en wordt derhalve wel de Condorcet paradox genoemd. Het doet zich voor indien de door het meerderheidsstelsel voortgebrachte sociale voorkeur (de meerderheidsrelatie) cyclisch is. We zullen in dit hoofdstuk de oplossing van Condorcet bespreken. Tevens presenteren we een eigen theorie voor de behandeling van dit probleem. Deze theorie noemen we de theorie van stabiele meerderheidsoplossingen en is gebaseerd op resultaten die gepresenteerd zullen worden in hoofdstuk 4. De theorie gaat uit van de notie van meerderheidspad. Dit is een rijtje van alternatieven $x_1, x_2, \ldots, x_n$ waarvoor geldt dat er een strikte meerderheid is van x_1 over x_2, van x_2 over x_3, ..., van x_{n_1} over x_n. Een stabiele meerderheidsoplossing nu is een deelverzameling V van de verzameling van alternatieven waarvoor geldt :

1. voor elke x en y in V geldt dat er geen meerderheidspad is van x naar y.
2. voor elke y niet in V is er een x in V zodanig dat er een meerderheidspad is van x naar y.

Door gebruik te maken van een resultaat uit hoofdstuk 4 kan bewezen worden dat stabiele meerderheidsoplossingen altijd bestaan. We besteden ook aandacht aan de Bordaregel en de pluraliteitsregel. We zullen zien dat Marquis de Condorcet deze regels eveneens uitgebreid bestudeerd heeft. Hierbij verwijzen we naar teksten van de Condorcet waaraan tot nu toe, voor zover we weten, niet gerefereerd is. De meest verrassende ontdekking hierbij is dat Marquis de Condorcet de Bordaregel verwierp vanwege de schending van deze regel van de conditie van onafhankelijke alternatieven, een conditie die een speciale rol speelt in de stelling van Arrow. In dit hoofdstuk bestuderen we ook het Nederlandse kiesstelstel. We zullen aantonen dat in Nederland een partij die door een meerderheid van kiezers geprefereerd wordt boven elke andere partij toch het kleinste aantal zetels toegewezen kan krijgen.

In hoofdstuk 4 gaan we nader in op het probleem van cyclische sociale voorkeuren. Theorieën die aangeven hoe keuzes bepaald kunnen worden aan de hand van een sociale voorkeur, dus aan de hand van een relatie die verkregen is door een aggregatieproces, noemen we oplossingstheorieën. Een bekende oplossingstheorie is collectieve rationaliteit. Volgens dit concept dienen die alternatieven gekozen te worden die het beste zijn volgens de sociale voorkeur. Echter, beste alternatieven bestaan niet indien een sociale voorkeur cyclisch is. Omdat we, zoals reeds gezegd, cyclische relaties niet kunnen uitsluiten bij de bestudering van coalitieformatie als collectief besluitvormingsprobleem is dit concept weinig bruikbaar. We bestuderen derhalve enkele alternatieven. Één daarvan is de theorie van maximale keuzes. Een keuze is maximaal indien er geen alternatief bestaat dat beter is. We zullen precies aangeven wanneer maximale keuzes bestaan. Hierbij zal blijken dat deze kunnen bestaan in geval van cyclische relaties. Echter, hoewel de theorie volstaat voor een aantal theorieën in latere hoofdstukken (hoofdstukken 6 en 7), hebben we toch theorieën nodig die elke cyclische relatie kunnen 'kraken'. Het eerste thoeretisch alternatief dat hiertoe in staat is, is de theorie van gegeneraliseerde optimale keuzeverzamelingen. Volgens deze theorie dienen keuzeverzamelingen te bestaan uit elementen die zich bevinden in de topcycli van een relatie. Er kan bewezen worden dat gegeneraliseerde optimale keuzeverzamelingen niet leeg zijn voor willekeurige cycli. We presenteren een stelling die deze keuzes karakteriseert en die tevens duidelijk maakt hoe ze gevonden kunnen worden. Een nadeel van gegeneraliseerde optimale keuzeverzamelingen is dat ze intern instabiel zijn. Dit betekent dat voor een alternatief x uit een gegeneraliseerde optimale keuzeverzameling er een alternatief y uit deze verzameling kan bestaan zodanig dat y strikt sociaal geprefereerd wordt boven x. Teneinde deze tekortkoming te ontlopen hebben we een alternatieve theorie geformuleerd die gekenmerkt wordt door interne stabiliteit. Deze theorie sluit aan op de stabiele verzamelingsleer zoals ontwikkeld in Von Neumann and Morgenstern (1953) en heet derhalve de theorie van gegeneraliseerde stabiele oplossingen. Een gegeneraliseerde stabiele oplossing V bestaat uit elementen waartussen geen strikt voorkeurspad bestaat en waarvoor geldt dat voor elke y niet in V er een x in V is zodanig dat een strikt voorkeurspad bestaat van x naar y. We geven een aantal stellingen met betrekking tot deze keuzeverzamelingen.

In hoofdstuk 5 presenteren we de speltheoretische context voor de coalitietheorieën. Omdat winnen de essentie is van politiek gaan we in op spelen waarin winnen en verliezen centraal staan. Deze spelen heten simple games. Vervolgens bestuderen we een aantal reeds geformuleerde coalitieformatietheorieën. Dit zijn de minimum omvang theorie van Riker, de belangenconflict theorie van Axelrod en de actorgerichte machtstheorie van dominante spelers van Peleg. In dit hoofdstuk presenteren we ook onze actorgerichte beleidstheorie teneinde de eerste bovengenoemde theoretische tekortkoming op te heffen. Deze actorgerichte beleidstheorie noemen we de theorie van centrale spelers. De definitie van een centrale speler is als volgt. Zij een ordening van beleidsposities gegeven. Een speler i is centraal indien noch de spelers met een beleidspositie voorafgaande aan die van i in een ordening van posities, noch de spelers met een beleidspositie volgend op die van i in een ordening van posities een winnende coalitie kunnen vormen zonder i. De hypothese is dat alleen coalities met centrale spelers gevormd zullen worden. De theorie van centrale spelers wordt verfijnd door de specificatie van een theorie, de theorie van gebalanceerde coalities, waarin aangenomen wordt dat centrale spelers streven naar de vorming van coalities die zij zo goed mogelijk in balans kunnen houden. We geven ook een rekenvoorbeeld.

In hoofstuk 6 bestuderen we besluitvormingstheorieën van coalitieformatie van het eerste type. De eerste theorie die we bestuderen is de beleidsafstandtheorie van De Swaan. Deze theorie die een belangrijke inspiratiebron voor dit werk is geweest, gaat uit van de gedachte dat elke speler streeft naar een coalitie met een verwachte beleidspositie die zo dicht mogelijk bij haar of zijn beleidspositie ligt. Teneinde de termen 'zo dicht mogelijk' te specificeren, formuleert De Swaan een niet-metrisch concept van 'afstand'. De Swaan formuleert verschillende varianten van zijn theorie. In dit werk bestuderen we voornamelijk de open versie en de gesloten versie. We zullen zien dat beide versies geplaagd worden door ernstige inconsistenties. Vervolgens presenteren we onze machtsoverschottheorie. Zoals gezegd vertrekt deze theorie van de gedachte dat elke speler streeft naar een coalitie waarin zijn of haar machtsoverschot - zijnde het verschil tussen haar of zijn machtsgewicht en de som van de machtsgewichten van de overige leden van die coalitie - zo groot mogelijk is. We formuleren diverse varianten van deze theorie. We zullen eveneens zien dat de theorie van maximale sociale keuzes gwebruikt kan worden als oplossingstheorie voor deze coalitietheorie. Ook voor deze coalitietheorie presenteren we een berekening aan de hand van een 'real-life' voorbeeld (de zetelverdeling in de Tweede Kamer sinds de verkiezingen van 6 september 1989 en een bekende links-rechtsordening van de Nederlandse politieke partijen).

In het laatste hoofdstuk presenteren we de twee besluitvormingstheorieën van het tweede type, t.w. de conflictminimalisatietheorie en de voorkeurafstandstheorie. Eerst formuleren we een type spel dat we sociaal keuzespel noemen. Dit spel bestaat uit een collectief keuzeprobleem (een politieke agenda), een aantal spelers, een voorkeursprofiel- één voorkeur voor elke speler met betrekking tot het collectief keuzeprobleem - en een aantal regels voor winnen en verliezen. We zullen abstraheren van de aard van regels die winnen en dus verliezen bepalen. Hierdoor zijn we in staat coalitietheorieën te formuleren die onafhankelijk zijn van deze regels en derhalve een grotere reikwijdte hebben. Vervolgens presenteren we een afstandsmodel dat gebruikt zal worden in beide theorieën. Centraal in dit afstandsmodel staat het begrip Hammingafstand. De Hammingafstand van twee voorkeuren is het aantal elementen in het symmetrisch verschil van deze voorkeuren. We zullen zien dat dit begrip inderdaad een metriek is. Vervolgens bestuderen we de notie van een lineair profiel. Deze bestudering geeft aanleiding tot de formulering van een voldoende voorwaarde voor het bestaan van een meerderheidskeuze. Lineaire profielen zullen een speciale rol spelen in de voorkeurafstandstheorie. Daarna presenteren we de reeds besproken conflictminimalisatietheorie. We geven een aantal theoretische resultaten en opnieuw een rekenvoorbeeld. Vervolgens presenteren we de reeds besproken voorkeurafstandstheorie. Ook hierbij geven we een aantal resultaten en een rekenvoorbeeld. We gaan vervolgens in op hypothesen die uit deze theorieën afgeleid kunnen worden met betrekking tot kabinetsformaties in meerpartijsystemen en op de empirische toetsbaarheid van deze hypothesen. Ter afsluiting formuleren we een aantal alternatieve toepassingsgebieden voor het Hamming afstandsmodel zoals gepresenteerd in dit hoofdstuk. Hiermee zijn we dan aan het einde van onze reis gekomen.

Curriculum Vitae

De schrijver van dit proefschrift werd geboren op 13 september 1952 in Nijmegen. In 1968 behaalde hij zijn ULO A en B diploma. Van 1971 tot 1974 volgde hij het Atheneum en behaalde hiervoor in juni 1974 zijn diploma. In 1975 begon hij met zijn studie politicologie aan de Katholieke Universiteit van Nijmegen. In juni 1977 behaalde hij zijn kandidaatsexamen. Gedurende zijn doctoraalstudie werd zijn belangstelling voor de sociale keuzetheorie gewekt en gestimuleerd door prof. dr. R.J. in 't Veld, toen hoogleraar bestuurskunde aan de Katholieke Universiteit van Nijmegen. In mei 1983 behaalde hij zijn doctoraalexamen (cum laude). Gedurende zijn doctoraalstudie deed hij onderzoekservaring op in diverse studentassistenschappen en in een onderzoeksstage.

Na zijn studie deed hij tot september 1983 contractonderzoek. In oktober 1984 werd hij onderwijsmedewerker bij de studierichting politicologie van de Katholieke Universiteit van Nijmegen. In januari 1986 begon hij op dezelfde universiteit met de uitvoering van het poolplaatsproject Bureaucratie en Sociale Keuze, codenummer SW/3/86. De resultaten van dit onderzoek leidde tot het schrijven van dit proefschrift. Vanaf 1 november 1990 is hij universitair docent politicologie bij deze universiteit.

Stellingen behorende bij het proefschrift Coalition Formation and Social Choice

Ad van Deemen

1. Bij Tweede Kamerverkiezingen in Nederland is het mogelijk dat een partij die door een meerderheid geprefereerd wordt boven een andere partij toch een kleiner aantal zetels toegewezen krijgt. Bewijs: zie hoofdstuk 3 van dit proefschrift.

2. De lange duur van de kabinetsformatie Van Agt-Wiegel in 1977 is veroorzaakt door het feit dat de PvdA als grootste partij en als de partij met de grootste verkiezingsoverwinning het initiatief tot het formeren van een kabinet op zich nam terwijl het CDA de centrumpositie innam en dus de loop van het formatieproces kon bepalen.
   Zie voor actoren met een centrumpositie de theorie van gecentraliseerde beleidsspelen in hoofdstuk 5 van dit proefschrift.

3. D66 zit niet in het huidige kabinet Lubbers-Kok omdat

   (a) het CDA in de huidige combinatie een positief machtsoverschot heeft maar in de combinatie $\{PvdA, D66, CDA\}$ een negatief machtsoverschot gehad zou hebben,

   (b) het CDA in de huidige combinatie spilpartij is maar dit niet zou zijn in de combinatie $\{PvdA, D66, CDA\}$. In deze combinatie is D66 spilpartij.

   Zie hoofdstuk 6 van dit proefschrift voor de machtsoverschottheorie en hoofdstuk 5 voor het begrip spilpartij.

4. De propositie van Wittgenstein "Die Logik ist kein Lehre, sondern ein Spiegelbild der Welt" dient vervangen te worden door de propositie "Die Logik ist kein Spiegelbild, sondern eine Begrenzung der Welt," en wel omdat de logica de grens van het logisch mogelijke en het logisch onmogelijke beschrijft en daarmee de grenzen aangeeft van het empirisch mogelijke. Zie L. Wittgenstein, (1922). *Tractatus Logico-Philosophicus.* London: Routledge & Kegan. Satz 6.13.

5. Het 2-persoons prisoner's dilemma wordt opgelost in geval van wederzijdse liefde waarbij liefde gedefinieerd is als 'kiezen wat het beste is voor de ander'. Zie voor het prisoner's dilemma R. Luce & H. Raiffa, (1989). *Games and Decisions*. New York: Dover Publication. (Oorspr. ed. 1957). 5.4.

6. Iteratie van het 2-persoons prisoner's dilemma laat zien waarom onbeantwoorde liefde een lijdensweg is. Immers, diegene die steeds kiest wat het beste is voor de ander ontvangt dan in elk constituerend spel de sucker's payoff terwijl de geliefde steeds de volle buit ontvangt.
Voor iteratie van het 2-persoons prisoner's dilemma, zie R. Axelrod (1984), *The Evolution of Cooperation*. New York: Basic Books, Inc., Publishers.

7. De stemparadox verdwijnt in oneindige samenlevingen.
Bewijs: Laat κ een oneindig kardinaalgetal zijn en beschouw het volgende profiel:
$\kappa : x\ y\ z$
$\kappa : z\ x\ y$
$\kappa : y\ z\ x$
Hierbij betekent bijvoorbeeld $\kappa : x\ y\ z$ dat κ kiezers x strikt prefereren boven y en y strikt prefereren boven z.
Omdat er $\kappa + \kappa = \kappa$ kiezers zijn met $x\ y$ en κ kiezers met $y\ x$, $\kappa + \kappa = \kappa$ kiezers met $y\ z$ en κ kiezers met $z\ y$, en $\kappa + \kappa = \kappa$ kiezers met $z\ x$ en κ kiezers met $x\ z$, verkrijgen we totale indifferentie en is daarmee de paradox verdwenen.

8. De stelling van Arrow laat zien dat het *logisch* onmogelijk is dat een collectieve besluitvormingsprocedure voldoet aan de eisen van collectieve rationaliteit, onafhankelijkheid van irrelevante alternatieven, Pareto en verbod op dictatuur. Omdat iets wat logisch onmogelijk is, empirisch niet kan bestaan, sluit de stelling het *empirisch* bestaan van collectieve besluitvormingsprocedures die voldoen aan de Arrow eisen uit. Dit laat zien dat de stelling van Arrow wel degelijk een empirisch aspect heeft.

9. Elk inleidend boek in de sociale wetenschappen waarin de stelling van Arrow niet behandeld wordt, dient herschreven te worden.

NICI
P.O. Box 9104, 6500 HE Nijmegen
Tel.: 080 - 512633