UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE FILOSOFIA E CIÊNCIAS HUMANAS

PEDRO HENRIQUE CARRASQUEIRA ZANEI

CONSISTENT SYSTEMS OF RIGHTS AND THE LOGICAL LIMITS OF LIBERAL DEMOCRACY

SISTEMAS CONSISTENTES DE DIREITOS E OS LIMITES LÓGICOS DA DEMOCRACIA LIBERAL

CAMPINAS
2023

Pedro Henrique Carrasqueira Zanei

Consistent systems of rights and the logical limits of liberal democracy

Sistemas consistentes de direitos e os limites lógicos da democracia liberal

Tese apresentada ao Instituto de Filosofia e Ciências Humanas da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutor em Filosofia.

Thesis presented to the Institute of Philosophy and Human Sciences of the University of Campinas in partial fulfillment of the requirements for the degree of Doctor, in the area of Philosophy.

Supervisor: Walter Alexandre Carnielli

Este trabalho corresponde à versão final da tese defendida pelo aluno Pedro Henrique Carrasqueira Zanei, e orientada pelo Prof. Dr. Walter Alexandre Carnielli.

Campinas
2023

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Filosofia e Ciências Humanas
Cecília Maria Jorge Nicolau - CRB 8/3387

C231c Carrasqueira Zanei, Pedro Henrique, 1983-
Consistent systems of rights and the logical limits of liberal democracy / Pedro Henrique Carrasqueira Zanei. – Campinas, SP : [s.n.], 2023.

Orientador: Walter Alexandre Carnielli.
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Filosofia e Ciências Humanas.

1. Escolha social. 2. Democracia. 3. Liberalismo. 4. Lógica. 5. Inconsistência (Lógica). I. Carnielli, Walter Alexandre, 1952-. II. Universidade Estadual de Campinas. Instituto de Filosofia e Ciências Humanas. III. Título.

Informações Complementares

Título em outro idioma: Sistemas consistentes de direitos e os limites lógicos da democracia liberal
Palavras-chave em inglês:
Social choice
Democracy
Liberalism
Logic
Inconsistency (Logic)
Área de concentração: Filosofia
Titulação: Doutor em Filosofia
Banca examinadora:
Walter Alexandre Carnielli [Orientador]
Ítala Maria Loffredo D'Ottaviano
Otávio Bueno
Álvaro Sandroni
André Luiz de Almeida Lisbôa Neiva
Data de defesa: 14-07-2023
Programa de Pós-Graduação: Filosofia

Identificação e informações acadêmicas do(a) aluno(a)
- ORCID do autor: https://orcid.org/0009-0006-9786-5191
- Currículo Lattes do autor: http://lattes.cnpq.br/8843690349739339

UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE FILOSOFIA E CIÊNCIAS HUMANAS

A Comissão Julgadora dos trabalhos de Defesa de Tese de Doutorado, composta pelos(as) Professores(as) Doutores(as) a seguir descritos, em sessão pública realizada em 14 de setembro de 2023, considerou o candidato Pedro Henrique Carrasqueira Zanei aprovado.

Prof. Dr. Walter Alexandre Carnielli
Profa. Dra. Itala Maria Loffredo D'Ottaviano
Prof. Dr. André Luiz de Almeida Lisbôa Neiva
Prof. Dr. Álvaro Sandroni
Prof. Dr. Otávio Bueno

A Ata de Defesa com as respectivas assinaturas dos membros encontra-se no SIGA/Sistema de Fluxo de Dissertações/Teses e na Coordenadoria do Programa de Pós-Graduação em Filosofia do Instituto de Filosofia e Ciências Humanas.

To my parents,
Sílvia and Osmar.

Acknowledgments

I would like to thank first and foremost professor Carnielli, for believing in me even when I didn't. I'll be glad if I've managed to prove you right.

I also thank professors Coniglio and D'Ottaviano. I suspect that, without you three, this work would simply not have been possible in Brazil.

I would also like to thank my sister Anna and my friends Álvaro, Anderson, Graziella, Daniel, Lucas, Maynara, Pablo, Pedro, Tainan and Tatyane for your never-ending support. I owe all of you a lot more than you'll ever know.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Abstract

This doctoral dissertation is concerned with the matter of the logical consistency of systems of rights and powers. In the first part of the dissertation, a generalization of the concept of a coherent system of individual rights, originally developed by Gibbard, Farrell and Suzumura, is presented. I take this generalization to be an adequate characterization of the coherence of a system of rights and powers in general. In the second part, I tackle the question of whether this concept of coherence can legitimately be seen as a notion of logical consistency.

This investigation makes use of formal models and mathematical techniques. In particular, in order to characterize coherent systems of rights and powers, I resort to the techniques of social choice theory. For the question of the logical consistency of such systems, I rely on techniques essentially due to Keisler and Kaye, and also on some interesting ideas on the nature of logic suggested by the work of Stalnaker.

It is shown that the formal concept of a coherent system of rights and powers indeed corresponds in a precise way to the underlying notion of consistency of certain Tarskian logic, and therefore can legitimately be seen as logical in nature.

An appendix containing (almost) all of the immediately relevant mathematical and choice-theoretical background knowledge is also provided.

Key words: rights; social choice; logic; consistency

Resumo

Esta dissertação de doutorado trata da questão da consistência lógica de sistemas de direitos e poderes. Na primeira parte da dissertação, é apresentada uma generalização do conceito de um sistema coerente de direitos individuais, originalmente desenvolvido por Gibbard, Farrell e Suzumura. Tomo essa generalização como uma caracterização adequada da coerência de um sistema de direitos e poderes em geral. Na segunda parte, abordo a questão de se esse conceito de coerência pode legitimamente ser visto como uma noção de consistência lógica.

Essa investigação faz uso de modelos formais e técnicas matemáticas. Em particular, para caracterizar sistemas coerentes de direitos e poderes, recorre-se às técnicas da teoria da escolha social. Para a questão da consistência lógica desses sistemas, apoia-se em técnicas essencialmente devidas a Keisler e Kaye, bem como em algumas ideias interessantes sobre a natureza da lógica sugeridas pelo trabalho de Stalnaker.

É demonstrado que o conceito formal de um sistema coerente de direitos e poderes de fato corresponde de um jeito preciso à noção subjacente de consistência de certa lógica tarskiana, e, portanto, pode ser legitimamente vista como de natureza lógica.

Um apêndice contendo (quase) todo o conhecimento matemático e de teoria da escolha imediatamente relevante também é fornecido.

Palavras chave: direitos; escolha social; lógica; consistência

Contents

Chapter 1

Introduction

This doctoral dissertation is concerned with the matter of the logical consistency of systems of rights and powers in liberal democracies. More precisely, it attempts to provide answers to two related questions. One is the question of whether and how can liberal rights of certain kind be made compatible with some fairly mild strictures of democratic ruling, and what are some of the consequences of they being made so. By saying that they are to be made compatible, we mean that the resulting system of rights and powers is supposed to be coherent in an intuitively adequate sense of coherence. This is the subject of the first part of the dissertation. The second question this doctoral dissertation tackles, then, is that of whether the concept of a coherent system of rights and powers, explored in the first part, can legitimately be taken to be a notion of logical consistency. This is the subject of its second part.

The investigation makes heavy use of formal models and mathematical techniques. In particular, in order to answer the first question, it resorts to the techniques of social choice theory.

Because it cannot be assumed that every possibly interested reader of this dissertation will be acquainted with the subject (nor, indeed, with the kind of formal approach to philosophical questions taken here), an appendix containing (almost) all of the immediately relevant mathematical and choice-theoretical background is also provided.

The remaining of this introduction paints, in very broad strokes, an informal version of the argument of this dissertation.

1.1 Democracy and liberal rights

In his seminal work on how to model liberal rights in the framework of social choice theory [55], Sen asks us to consider the simple example of the right to choose the color of one's own walls. To have such a right, according to him, is a matter of your preferences socially prevailing regarding states that diverge only on whether your own wall's are painted, say, either white or pink. So if, say, you prefer them to be pink, then that is the color they should be allowed to be in whatever state of affairs that is socially implemented (provided such a state is feasible).

This notion of what it means to have a right is straightforward enough, but, as Sen himself remarks, it logically excludes the possibility of the preferences of a majority always socially prevailing on all issues. This is so because, if a majority prefers your walls to be white, then, if it were to prevail, we would have it that your walls ought to be both pink and white, — which is impossible.

To the extent, therefore, that the rule of the majority is central to democracy, these simple observations pose a serious question about whether liberal rights are compatible with democratic ruling. Clearly, if we are to have liberal rights at all, then the majority cannot rule on all issues. How much, however, is the rule of the majority to be limited in order to give space to liberal rights?

Well, the answer provided by Sen in that work is: a lot. His theorem on the impossibility of the Paretian liberal states that, if at least two individuals are to have even a single right each, then not even unanimity can be respected on all issues, for otherwise there will be situations in which no state of affairs will be socially permissible.

A fruitful way of looking at the problem is to see it as a matter, not of a clash between liberalism and democracy, but as a clash among rights, or powers, broadly conceived. From a formal point of view, to say that the preferences of a majority should prevail on an issue is the same as saying that it has a right on that issue. This observation, however, leads to the more general problem of how to guarantee that the rights and powers of various individuals and groups in a society can peacefully coexist.

The theorems in section 4.1 make precise what it means for the rights of various individuals and groups in a society to peacefully coexist, by means of the notion of a coherent decisiveness assignment. This notion is a generalization of the notion of a coherent system of individual rights, which is to be found in [16], [23], [66].

In order for a general system of rights and powers to be coherent and still respect unanimity, however, it must satisfy a remarkable condition: namely, that, in every situation in which rights would otherwise conflict with each other, there must be a group of individuals in society (which may depend on how the rights would conflict) that has the power to arbitrarily decide which state of affairs is to be implemented. This is shown in sections 4.2 and 4.3.

We interpret this result as meaning that some form of arbitration among rights is inevitable in a liberal democracy, and that this is so even if the system of rights in question is coherent. So, to the extent that arbitration is to be institutionalized in society, this would amount to a proof that some sort of judicial institution is not only a practical but indeed a logical necessity of liberal democracies.

Of course, not every actual societies will have a coherent system of rights. As a matter of fact, given how intricate the systems of rights of actual societies tend to be, it is rather unlikely that any society at all can completely avoid incoherences in its system of rights. That being the case, in section 4.4 we provide a general schema for converting incoherent systems of rights into coherent ones, and to do so in such a way that, in a precise sense, the resulting system preserves the existing rights to the greatest possible extent. This schema is also of theoretical interest as it comprises the schemata presented in [23] and [59] as particular and restricted cases.

1.2 Rights and logical consistency

In philosophy, there is a longstanding tradition of logical analyses of normative concepts, usually in the framework of modal logic. That approach to normative concepts usually starts from some informal notions of interest — such us the notions of obligation and permission — and goes on to provide a formal account of them either as primitive or defined concepts within some logic.

Our approach to the matter of the logical status of the concept of a coherent system of rights and powers differs from that more traditional one in two aspects. The first one is that the notion of interest to our analysis is already a formal and fairly well-established one in the literature of social choice theory. The second aspect in which it differs is that, instead of taking that notion of coherence as either a primitive our defined concept within a logic, we actually present (in chapter 8) a Tarskian logic such that our notion of interest corresponds in a precise sense to one of its inherent features: namely, its concept of consistency. This is the main result of the second part of this dissertation.

The primary import of our logical account of the notion of a coherent system of rights and powers, therefore, is that it reveals that notion as being logical in nature. This, as far as we can tell, is an entirely novel approach and contribution to the subject of deontic logic, broadly conceived.

Part I

Paretian entitlement and the inevitability of arbitration

Chapter 2

Introduction

There are two classical results of social choice theory which may be seem as foundational, since they have historically set much of the agenda for the field. One is Kenneth Arrow's *impossibility theorem* (or *general possibility theorem*, as he himself called it) [2]. Arrow's theorem is commonly interpreted as showing that even arguably minor conditions a democratic preferential voting procedure is expected to satisfy are in general incompatible with the seemingly most natural requirement of rationality of preferences, which is transitivity, *cf.* [56]. The other result is Amartya Sen's *impossibility of a Paretian liberal*, which shows that a requirement of social rationality much weaker than transitivity can't be met by a social choice rule which respects unanimous strict preferences if individuals are to have even a rather limited amount of rights of certain kind [62].

The two results are not exactly logically unrelated, as they explicitly share some assumptions. Most notably, both Arrow and Sen assume that a social choice function ought to aggregate preferences, have an unrestricted domain and satisfy the so-called weak Pareto condition. Both may also be said to assume that social rankings of options must be at least acyclic, given that transitivity implies the acyclicity assumed by Sen. Sen's condition of minimal liberty — the assumption that at least two individuals have rights of certain kind —, however, is actually *stronger* than Arrow's non-dictatorship, and therefore the impossibility of a Paretian liberal is itself not a weakening of Arrow's much more complex impossibility theorem. This, together with the fact that one result seems to pertain primarily to voting procedures and the other to systems of individual rights, may suggest that Arrow and Sen are dealing with formally related but ultimately distinct subject-matters, — an impression not rarely reinforced by textbook presentations of social choice theory, which tend to treat the theorems and their possible applications separately, *e.g.* [20], [65].

Nonetheless, this treatment seems to be at odds with Sen's original assessment of his own result. In his seminal paper [62] on the subject as well as in other texts, *e.g.*, [56], [59], Sen seems always keen to stress that, unlike Arrow's, his theorem doesn't require the controversial condition of independence of irrelevant alternatives, and therefore is not open to the sort of the criticism Arrow's impossibility theorem has received, *cf.* [56]. In a later work, Sen even attempts a discussion of the common logical core of both theorems, in terms of a type of power weaker than decisiveness, which he dubs “potential semidecisiveness” [59] (or “potential veto power”, as it is perhaps best to call it, *cf.* [65], so as to avoid confusion with the concept of almost decisiveness already found in Arrow). He does so grounded on the fact that minimal liberty, but also the existence of a dictator, both trivially imply the existence of an individual in the society who is decisive between at least two options, — from which it then follows that such an individual must also have global potential veto power if the domain of the social choice function is unrestricted, if it satisfies the weak Pareto condition and, moreover, if it yields only complete and acyclic social preferences.

Although logically irreproachable, Sen’s analysis of the root cause of the two results is conceptually unsatisfactory, for the simple reason that what in the impossibility of a Paretian liberal is assumed right from the start, in Arrow’s theorem appears only as a consequence of a conclusion that is nothing short of surprising. To put it bluntly: because Sen’s analysis doesn’t clearly display why their *explicitly* shared assumptions lead to difficulties, it is therefore not explanatory enough.

My aim in this chapter is to bridge this gap between the two classical results of social choice theory. I do so by proving a theorem which essentially shows that the assumptions of unrestricted domain and the weak Pareto condition, together with the demand for acyclic social rankings of options, *already* suffice to induce a social choice function to bestow seemingly arbitrary powers of certain type to groups within society. I also discuss how this result relates to the classical ones by Arrow and Sen. In a sense, then, this chapter is a thorough investigation into the logical foundations of those two foundational theorems.

Apart from its logical relevance as a missing link of social choice theory, the conceptual import of the results of this chapter arguably rests on the alternative interpretation of the weak Pareto condition it suggests. This interpretation appears to not have been seriously considered in the literature so far, and it seems to be of particular interest for a proper diagnosis of what is really at stake in the impossibility of a Paretian liberal. Incidentally, for the sake of making this point clear enough, I present an alternative proof of the theorem by Sen, — one which I believe highlights its significance in a way that the original proof does not.

This chapter is divided as follows.

In section 2 I present the framework of social choice theory strictly required for the results that follow. This presentation is entirely based in [7], [20], [56], [65] (whose content vastly overlaps in the relevant parts), so I mostly refrain from citing them specifically. Subsection 2.4, in particular, is dedicated to a discussion of the impossibility of a Paretian liberal. In it, I present the promised alternative proof of Sen’s theorem.

Section 3 is the core of the chapter. In it, I introduce the concept of coherent decisiveness assignment, which is a generalization of the concept of coherent system of individual rights found in [23], [16], [64]. Subsection 3.3 presents the main result of the chapter, which is essentially an impossibility theorem to the effect that there can be no coherent Paretian system of decisiveness powers without there being some group in society which has the power to arbitrarily choose in some circumstances what social outcome is to be implemented. Subsection 3.4 introduces a resolution scheme to that impossibility, by showing how to make coherent any incoherent decisiveness assignment, and to do so in such a way that arguably minimizes the relative loss of decisiveness powers that necessarily ensues.

Chapter 3

Social choice theory

3.1 Choice and preferences

Let S be a non-empty set, the elements of which are to be variously interpreted either as *options* of some kind, or as *possible outcomes* of a choice procedure, or even as *possible states* of the world, depending on the context. A *choice function on* S is a (possibly partial) $C : \wp(S) \longmapsto \wp(S)$ such that $C(X) \subseteq X$, for all $X \in \mathsf{dom}(C)$. Intuitively, a choice function assigns to any set X of *available*, or *feasible*, options or outcomes (also to be called a *situation* if its elements are possible states) one of its subsets $C(X)$, which is to be interpreted as the set of options deemed *acceptable* from X, — that is, the options which might end up getting *picked* from it under C. (Accordingly, we say $X \setminus C(X)$ are the options definitely *rejected* from X under C.) A choice functions is thus meant to represent the *choice behavior* of an individual, or of a group of individuals taken as a whole.

We say a choice function is *determined for* X if it is defined for X and either X is empty or $C(X) \neq \varnothing$, and that it is *finitely determined* if it is determined for every non-empty, finite X. Throughout this dissertation, we shall assume any choice function must be defined (but not necessarily determined) for every X. $\mathcal{C}(S)$ will stand for the set of all total choice functions on S. Later on we will also assume that every choice function is rationalizable through best options, meaning that for all such C there will be a preference relation on S such that $C(X)$ is the set of best options in X, for all X.

By a *preference relation on* S what is here meant is just a binary relation $\succsim \subseteq S \times S$. A preference relation is intended to represent the *preferences* of an individual (or, again, perhaps a group of individuals taken as whole). In other words, a preference relation is to be interpreted as a *ranking* of options in terms of *better* and *worse*, possibly admitting ties. $x \succsim y$ means, therefore, that x *is at least as good as* y. Given a preference relation $\succsim$, we write $\sim$ for its *symmetric* or *indifference part* — that is, the relation defined by $x \sim y$ iff $x \succsim y$ and $y \succsim x$ —, and $\succ$ for its *asymmetric* or *strict preference part* — which is defined by $x \succ y$ iff $x \succsim y$ but not $y \succsim x$. Thus, $x \sim y$ means that x *is as good as* y, and $x \succ y$ that x *is better than* y. We refer to a pair x, y of options as an *issue*, and to the four possible preference configurations regarding an issue — strict preference for one option, strict preference for the other, indifference between them and indecision (when neither $x \succsim y$ nor $y \succsim x$) — as a *stance* or *position* on the issue.

Given a preference relation $\succsim$ on S and an $X \subseteq S$, the *best elements of* X *under* $\succsim$ are no other than the set $B_{\succsim}(X) = \{x \in X : x \succsim y$ for all $y \in X\}$ of the greatest elements of X under $\succsim$. A choice function C on S is therefore *rationalizable through best options by* $\succsim$ if $C(X) = B_{\succsim}(X)$, for all X. It should however be noted that a choice function being rationalizable by no means imply that the *choice behavior* it represents is supposed to be regarded as rational in some deeper sense, but merely that it can be *rationally explained* as the expression of certain preferences. As a matter of fact, it is the rationality of the *preferences themselves* that is usually at stake in the classical theorems of social choice theory.

On that note, it ought to be said that at least individuals' preference relations are commonly assumed to be *fully rational* in the sense of being *connected quasi-orders* (or simply *orderings*, as they are usually called in the social choice literature). A relation $\succsim$ is a quasi-order if it is *reflexive* and *transitive*. A relation is reflexive if, for all x, y, if $x = y$, then $x \succsim y$, and it is transitive if, for all x, y, z, if $x \succsim y$ and $y \succsim z$, then $x \succsim z$. A relation is connected if, for all x, y, if $x \neq y$, then either $x \succsim y$ or $y \succsim x$. A relation that is both reflexive and connected — that is, such that either $x \succsim y$ or $y \succsim x$, for all x, y — is said to be *complete*. The assumption that certain preference relations ought to be quasi-orders is tantamount to the notion that the preferences represented by them indeed behave as rankings in the natural sense, while completeness means that the individual or group holding those preferences has a decided position on every issue (that is: indecision never occurs).

In what follows, $\mathcal{O}(S)$ will stand for the set of all orderings on S, and we will assume that individuals' preferences are all orderings. This, it should be said, is a *methodological* rather than an *anthropological* assumption. Because in social choice theory we are for the most part interested in testing the limits of *social rationality*, assuming that *at least* the individuals in the society are themselves rational is necessary in order to isolate the causes of failures of rationality of other types. As for group preferences, any substantial prior assumption of rationality would be disputed, and so none will be made from the outset. That is, we shall only assume that group preference relations, if any exists, are in the set $\mathcal{R}(S)$ of all possible complete binary relations on S.

In order to avoid confusion, it may be convenient to observe at this point that the word "preferences" is used in social choice theory in a technical sense, as a catch-all term for *rankings of relative goodness*. This use is noticeably wider than its natural one in English. In the sense it is here used, moral values *too* count as preferences — given that they arguably take the form of some possible states of the world being deemed (morally) better than others —, while in natural language "preferences" would for the most part be restricted to rankings of relative goodness of *morally indifferent* options. In point of fact, because we are also not restricting our attention to *moral* goodness, the word "good" itself should be understood in the vague sense of *adequacy to some given standard* (with the relevant standard usually depending, of course, on the intended application of the theory).

Now for our final (and central) basic concept, let I be a non-empty set of *individuals*, who we assume are to make a collective choice on an equally non-empty S. We refer to such a pair of a set of individuals and a set of options as a *society*. Given a society, a *preferences list* for it (or, as it is more usually called in social choice theory, a *preferences profile*) is a total $p : I \longmapsto \mathcal{O}(S)$, — that is, a function assigning a preference relation to each and every individual. Let $\mathcal{O}(S)^I$ denote the set of all such functions. A *social choice function for* I, S (heretofore SCF, for short) is a (possibly partial) $A : \mathcal{O}(S)^I \longmapsto \mathcal{C}(S)$. An SCF is, thus, a function assigning a choice behavior to the whole of the society (or, more properly, to the group of all the individuals in the society, taken as a whole), given the complete information about the individuals' preferences. Intuitively, $A(p)(X)$ are the options that might get picked (or: the possible outcomes that might obtain; or, still: the possible states of the world which might end up being brought about) if the individuals in the society, each holding their respective preferences as described by p, were to collectively behave in their choices according to the rules set by A when only the options (or outcomes, or states) in X are feasible. (What this rather long-winded explanation of what an SCF is supposed to be means will be made clearer by way of example soon enough.)

If we assume $A(p)$ to be rationalizable for all $p \in \mathsf{dom}(A)$, then we may think of A as essentially *aggregating* the preferences of the individuals into *social* ones, — namely, those which rationalize $A(p)$ for each p. In such cases, for the sake of simplicity we may alternatively take an SCF to be an (again, possibly partial) $A : \mathcal{O}(S)^I \longmapsto \mathcal{R}(S)$. We refer to such an SCF as a *preferences aggregation function* (PAF, for short).

Given an SCF A for I, S, an individual i and a preference relations profile p, and provided the relevant SCF is fixed by the context, we write $\succsim_i^p$ for $p(i)$; we also write C^p for $A(p)$ when $A(p)$ is not a PAF and $\succsim^p$ for $A(p)$ when it is, as this notation proves to be far more conspicuous than the standard one ever could.

3.2 Majority rule, Condorcet's paradox and Arrow's impossibility theorem

The most obvious example of a social choice rule is majority rule. Given a society I, S, majority rule for it can be taken to be the following SCF A.

Majority rule $x \in C^p(X)$ iff, for all y, $|\{i \in I : x \in B_{\succsim_i^p}(X)\}| \geq |\{i \in I : y \in B_{\succsim_i^p}(X)\}|$

That is: an option is socially acceptable under majority rule if, an only if, it is deemed best by at least as many individuals as any other option. If individuals' preferences are expected to be strict orders (that is, if their preferences are not only complete and transitive, but also *antisymmetric*: if $x \sim y$, then $x = y$, — meaning indifference between distinct options is not allowed), then this boils down to an option being uniquely socially acceptable if and only if it is the best according to a majority of individuals (or if it receives a majority of votes).

As a historical matter, what would later become the field of social choice theory arguably begins as a properly scientific endeavor from an observation of the following kind, traditionally attributed in its full import to Condorcet, about the limitations of majority rule as a purportedly democratic electoral procedure. Let I be a population of one hundred voters, and let x, y, z be the candidates running in the election. Suppose 40 individuals vote for x, 35 for y and 25 for z. Then, according to majority rule, x is the winner of the election. Suppose however that, if x were to run solely against y, then all of z's voters would vote for y. Then y would receive a total of 60 votes, thus being the winner instead of x. In a sense, this is to say that x is not the "real" winner (or *Condorcet winner*) of our hypothetical election, as x doesn't have the favor of a majority of voters *whenever* he or she is available as an option.

Of course, this problem can be reiterated for the second (and, indeed, any) place in the election. For suppose now that all of z's voters would vote for x instead if z left the race and vice-versa, and that all of y's votes would go for x if y left. Then x indeed would have the actual majority against both y and z, but z would also have a majority against y.

The point made by examples such as these is that majority rule would be inadequate as a *fully* democratic electoral procedure, because it ignores relevant information about the preferences of the individuals regarding candidates. The examples also suggest that majority rule could perhaps be improved by moving towards *preferential voting*. In this setting, the winners of an election would be those candidates who are at least as good (and possibly better) than all candidates according to the social preferences.

Now, the obvious analogue of majority rule in this setting is *pairwise majority rule*, which can be taken to be the PAF A defined as follows, for any given I, S.

Pairwise majority rule $x \succsim^p y$ iff $|\{i : x \succsim_i^p y\}| \geq |\{i : y \succsim_i^p x\}|$, for all p

Although pairwise majority is certainly an improvement on majority rule, as it takes into account the whole of the wills of the individuals, it also has some undeniable shortcomings. In particular, it cannot assure the social preferences to be an ordering, as is made evident by the so-called *Condorcet's paradox*. Let $I = \{i_0, i_1, i_1\}$, $S = \{x_0, x_1, x_2\}$, and consider the profile p such that $x_0 \succ_{i_0}^p x_1 \succ_{i_0}^p x_2$, $x_1 \succ_{i_1}^p x_2 \succ_{i_1}^p x_0$ and $x_2 \succ_{i_2}^p x_0 \succ_{i_2}^p x_1$. As can easily be checked, under this profile x_0 has the favor of a majority against x_1 — so that $x_0 \succ^p x_1$ according to pairwise majority rule —, x_1 against x_2 — so that $x_1 \succ^p x_2$ —, and x_2 against x_0 — so that $x_2 \succ^p x_0$. Not only are these social preferences not even a quasi-order, but there is no Condorcet winner either, as there is no socially best option in S.

Despite these shortcomings, pairwise majority has many properties which are indisputably desirable to find in a properly democratic preferential voting procedure. For one, it has a so-called *unrestricted domain* (UD), meaning it is defined (that is, p is in its domain) for all

$p \in \mathcal{O}(S)^I$. More importantly insofar as democracy is concerned, it has the properties known in the literature as *neutrality* and *anonymity*, which mean, respectively, that all options are treated the same by the procedure, and that every vote counts the same (so that it doesn't matter *whose* vote it is, — hence the name). Finally, pairwise majority is *positively responsive*, which means, roughly, that if two options are at least tied (that is, if the society as a whole is at least indifferent on the issue), then any voter who were to alone change his or her vote in favor of one of the options would break the eventual tie in its favor; or, to put it briefly: no one in particular holds exclusively the power of the *swing vote*. In fact, as proven by May [41], pairwise majority rule is fully *characterized* by these four conditions, in the sense that, given a society I, S, pairwise majority rule for it is the *only* PAF satisfying *all* of them.

Because pairwise majority rule is characterized by these four properties, and yet it is susceptible to the Condorcet paradox, we know that we can't have a preferences aggregation procedure with all of the desirable democratic properties which also happens to yield an ordering as output for any possible input. Therefore, if we want fully rational social preferences no matter what, something must give.

Arrow's work may thus be seem as an answer to the question: is there a preference aggregation procedure which is both universally defined, in the sense of yielding orderings as social preferences for every possible preferences profile, and which *also* manages to preserve the democratic qualities of pairwise majority rule to at least some reasonable degree?

More precisely, Arrow postulates such a procedure, as modeled by a PAF A, should satisfy the following conditions.

The first one is having an unrestricted domain (UD), which means that the PAF should make no restriction on what preferences the individuals are allowed to have or express regarding the options. Formally, this is to say that $A(p)$ must be defined (that is, $p \in \mathsf{dom}(A)$, — hence the name) for all $p \in \mathcal{O}(S)^I$.

The second condition is the so-called *weak Pareto* condition (WP) (or the *Pareto principle*), which demands that, if everyone in the society strictly prefer x to y, then the social stance on the issue should also be in favor of x against y. As it is easily seem, pairwise majority satisfies this condition as well, and at first glance at least it looks hardly objectionable in this context.

Third, Arrow requires the hypothetical procedure to have a property known as *independence of irrelevant alternatives* (IIA), which is logically weaker than neutrality and means that the social preference on an issue x, y for profile p depends exclusively on how the individuals rank these options relative to each other in p, — which is to say: the social preference yielded by the procedure is invariant, regarding the issue, under substitutions of p by profiles in which the ranking of x, y relative to each other is exactly as it is in p; in short, the function is *pairwise defined*, in the sense of being defined issue by issue, with no reference to other options. (Formally, for the sake of being thorough: for any two profiles p, p', if both $x \succsim_i^p y$ iff $x \succsim_i^{p'} y$ and $y \succsim_i^p x$ iff $y \succsim_i^{p'} x$, for all i, then both $x \succsim^p y$ iff $x \succsim^{p'} y$ and $y \succsim^p x$ iff $y \succsim^{p'} x$.) IIA may be seem as a strictly procedural requirement, in the sense that it makes the application of the aggregation procedure to the resolution of any issue non-contextual, therefore informationally manageable.

Fourth, and finally, he requires the procedure to be *non-dictatorial*, meaning there should be no i such that, for all p, if i strictly prefers x to y, then that is the social stance on the issue. That is: there should be no individual such that, no matter the preferences of the other individuals, that individual's strict preferences *always* socially prevail. This last property is a logical weakening of anonymity, and one must concede that, if there ever was a fairly mild condition a democratic electoral procedure should be required to satisfy, it is this one.

Of course, rigorously speaking there is no *real* preferences aggregation problem if society consists of a single individual. Similarly, if there are only two options available, then *any* preferences aggregation procedure has the property IIA and is, moreover, bound to yield an ordering if anything at all, given that, in particular, there can be no counterexample to transitivity with only two options. Indeed, under these conditions pairwise majority itself *already* is a PAF satisfying all of the Arrovian requirements. For $|I| \geq 2$ finite and $|X| \geq 3$, however, there can be *no* such PAF: the requirements are *logically inconsistent*.

This, in a nutshell, is Arrow's impossibility theorem.

3.3 Decisiveness

Crucial to Arrow's proof of his famous impossibility theorem is the concept of a coalition being decisive on an issue.

Let p be a preferences profile for I, S, and let J be a *coalition* in this society, — that is, a $J \subseteq I$. The *(strong) Pareto relation for J under p* is the relation P_J^p defined by: xP_J^py iff, for all $i \in J$, $x \succ_i^p y$. Intuitively, xP_J^py means that, under p, the individuals in J unanimously strictly prefer x to y, or that they all vote for x against y on the issue. It is easy to see that xP_J^py is transitive for any p, J and asymmetric for any $p, J \neq \varnothing$. (For $J = \varnothing$ it is the universal relation for every p.) Notice that when J is an *individual coalition* — that is, when $J = \{i\}$ for some i — $P_{\{i\}}^p$ boils down to $\succ_i^p$. For this reason, we naturally conflate individuals and their respective individual coalitions, as there is no conceptual gain in making the distinction.

Now let A be a PAF for I, S. We say coalition J is *decisive* for x against y under A if, for all p, if xP_J^py, then $x \succ^p y$. Informally, it is to say that a coalition is decisive under A if the individuals in the coalition together have the power to impose their unanimous stance for x against y as the social one: impose, that is, in the sense that the rules of aggregation represented by A do not allow the social preferences to disagree from those of the individuals in J on the issue x, y when they all agree that x is to be strictly preferred to y.

In what follows, we write xD_J^Ay to mean that J is decisive for x against y under A, and xE_J^Ay to mean that it is decisive *between* x, y — that is, that both xD_J^Ay and yD_J^Ax. (We also say that a coalition decisive between x, y is *entitled* to the issue x, y). When the relevant

A is fixed by the context, however, we drop the superscript. We say coalition J is decisive *on* issue x, y if it is either decisive for x against y or for y against x, that it is *locally* decisive if there is an issue on which it is decisive, and that it is *globally* decisive if it is decisive on all issues (or, what amounts to the same, if it is entitled to all issues).

The concept of decisiveness is an useful analytical tools of social choice theory, as it allows us to explain the inner workings of a social choice rule in terms of certain powers it assigns to groups of individuals. For instance, under pairwise majority rule, every coalition composed of more than half the individuals in the society is globally decisive (and that is precisely what it is to be expected if the majority is to indeed rule).

It is easy to see that decisiveness has the following properties, which stem from the facts that $P^p_\varnothing$ is the universal relation for all p but for no $J \neq \varnothing$ and no p does $xP^p_J x$, and the fact that for all p it is necessarily the case that, if $xP^p_{J'} y$ and $J' \supseteq J$, then $xP^p_J y$.

Vacuousness $xD_J x$ iff $J \neq \varnothing$

Monotonicity If $xD_J y$ and $J' \supseteq J$, then $xD_{J'} y$

More interestingly, it has the following property under the assumption that the relevant PAF satisfies UD. (Here, $\overline{J}$ stands for the complement of J relative to I.)

Exclusiveness For $x \neq y$, if $xD_J y$, then not $yD_{\overline{J}} x$

It is now shown that this is indeed the case.

Proof. Suppose for a contradiction that both $xD_J y$ and $yD_{\overline{J}} x$ for $x \neq y$, and let p be any profile such that $xP^p_J y$ and $yP^p_{\overline{J}} x$. Because $J, \overline{J}$ are disjoint, such a profile exists. But then both $x \succ^p y$ and $y \succ^p x$, which is impossible, thus contradicting the assumption that A is defined for p. □

Notice that it immediately follows from these two last properties, taken together, that it cannot be the case that $xD_J y$ but also $yD_{J'} x$ for $x \neq y$ if J, J' are disjoint, since J' is then a subset of $\overline{J}$, and thus $yD_{J'} x$ implies $yD_{\overline{J}} x$ by monotonicity.

The following conditions on a PAF A for I, S, which are neatly related to the concept of decisiveness, will be particularly important for our purposes.

Weak Pareto (WP) For all x, y, if $xP^p_I y$, then $x \succ^p y$

Weak non-imposedness (WNI) For no x, y it is the case that, for all p, $x \succ^p y$

WP is just the formal statement that I, the coalition of all the individuals, is decisive on all issues. WNI, in turn, essentially states that no issue is socially decided in advance by the rules represented by A.

As it so happens, WNI is equivalent to the empty set not being decisive on any issue whatsoever, as it is now shown.

Proof. If $xD_{\varnothing}y$, then, because vacuously $xP^{p}_{\varnothing}y$ for whatever p, it follows that $x \succ^{p} y$ for all p, and thus that the strict preference for x over y is socially imposed by A. Conversely, if $x \succ^{p} y$ for all p, then, trivially, every coalition is decisive for x against y, and in particular the empty one. □

Regarding WP, the following simple but important fact holds.

Proposition 3.1. *A PAF satisfies WP iff, for every issue, there is a coalition entitled to it.*

Proof. If A for I, S satisfies WP, then, because the coalition I of all individuals is decisive for x against y for every x, y (and, therefore, also entitled to every issue), for every issue there is a coalition entitled to it. Conversely, if for every x, y there is J entitled to it, then, letting x, y be any, because $J \subseteq I$ for whatever J, by the monotonicity of decisiveness also I is entitled to x, y, and therefore decisive both for x against y and for y against x — which is tantamount to saying that A satisfies WP. □

It should also be noted that, under the assumption of UD, condition WP immediately implies WNI, since it then cannot be the case that $x \succ^{p} y$ for a p such that $yP^{p}_{I}x$.

3.4 On the impossibility of a Paretian liberal

Although social choice theory as presented in the previous subsections is maybe most naturally seem as a theoretical toolkit for the analysis of voting procedures, it is just as well suited for dealing with a host of other problems of social choice. Perhaps one of Sen's most brilliant insights (but also one which, for as far as I can tell, he never actually stated in just so many words as I do bellow) was noticing in [62] how the same social choice techniques used by Arrow can be applied to a great extent to the study of law and rights, in the following way.

Law is arguably a social choice rule in the sense that, given any set of socially feasible states, it distinguishes those which may be acceptably brought about by society — brought about *according to law* — from those which ought to be avoided. Now, *which* states are according to law generally hinges on how the individuals and groups in the society exercise their rights and other legal powers. Of course, what rights and powers they happen to exercise depend, in turn, on what the individuals want, — that is, on what states they prefer would be brought about by the coordinated action of society as a whole. Clearly, this description of law fits squarely within the model provided by a social choice function.

What may then be called the *Senian conception of law* involves three further assumptions.

The first is that law conceived as a social choice rule should be finitely determined for all preferences profiles, — or, in other words: that, no matter the preferences of the individuals, it should yield at least one state which is according to law (or at least *tolerated* by it, given what is feasible), for every finite situation. The second assumption is that being according to law is a matter of *degree*, even if in the loosest possible sense of a state being acceptably brought about under the law if, and only if, there is no other state available which is to be legally preferred.

Let us pause for a moment to spell out what these two assumptions mean from a formal point of view. For this we need the concept of a cycle in a preference relation.

A preference relation $\gtrsim$ has an *improper cycle* if there is a sequence $x_0, \ldots, x_n$ of options such that $x_0 = x_n$, $x_k \gtrsim x_{k+1}$ for all $k < n$ and $x_k > x_{k+1}$ for some $k < n$, and it has a *proper cycle* if there is a sequence $x_0, \ldots, x_n$ of options such that $x_0 = x_n$ and $x_k > x_{k+1}$ for all $k < n$. We say that a preference relation is *strongly acyclic* if it has no improper cycles, and that it is *weakly acyclic* if it at least has no proper cycles (or, equivalently, if its strict part is strongly acyclic). Because in this and later chapters we shall mostly care only for proper cycles and weakly acyclic preference relations, unless otherwise stated by "cycle', "acyclic" and "acyclicity" I mean *proper cycle*, *weakly acyclic* and *weak acyclicity*. One should notice right away that transitivity implies strong acyclicity, and therefore acyclicity as well.

An easy lemma, due to Sen himself, states that a necessary and sufficient condition for $B_{\gtrsim}(X)$ to be non-empty for X finite is that $\gtrsim$ be complete and acyclic. So let I, S be a society, and let $\mathcal{A}(S)$ be the set of all complete and acyclic preference relations on S. Following Sen himself, we say a PAF for I, S is a *social decision function* (SDF) if $\mathsf{imag}(A) \subseteq \mathcal{A}(S)$. That is: an SDF is a PAF yielding complete and acyclic social preferences for every preferences profile in its domain. A trivial but important consequence of Sen's aforementioned lemma is that, for A a PAF, C^p is finitely determined for all $p \in \mathsf{dom}(A)$ iff A is an SDF.

So what the first two assumptions of the Senian conception of law amount to is the claim that law is to be modeled by an SDF, — that is, a social choice function whose output choice is to be rationalizable by a complete and acyclic preference relation, for whatever input preferences profile.

The third assumption is that, if not all of them, at least *some* rights (and, in particular, some *individual* rights) take the form of entitlements on issues, that is, of the power to determine according to one's own strict preferences the part of the social preferences regarding those issues. (It is evident, however, that not *all* entitlements ought to be conceived as rights, for then the clash between individual rights and the weak Pareto condition would be a matter of rights alone, and that is clearly *not* how Sen conceives of it.)

One may think of this notion of what some rights are as follows. Let any two possible states be given. If they are distinct, them they must be distinct in at least one aspect, that is, there must be *something* that is the case in one but not the other. So a right in this sense would be a power to determine, between two states diverging in at least that one aspect (but possibly conditional on they not diverging in certain other aspects), which of them is to be socially strictly preferred. To illustrate this notion, Sen offers us the simple examples of the right to paint one's walls either pink or white, and of the right to sleep on one's back or belly.

The trouble Sen uncovers within this conception of law and rights is that, if at least two individuals are to have rights that are entitlements in a society, then the those three assumptions clash with the weak Pareto condition if the PAF representing the rules of the society is to have an unrestricted domain.

More formally, let A be any SDF for I, S satisfying conditions UD, WP, and suppose it also satisfies the condition of *minimal liberty*.

Minimal liberty (ML) There are individuals $i \neq j$ and states $x \neq y$ and $z \neq w$ (not necessarily all distinct, however) such that xE_iy and zE_jw.

Then the so-called *impossibility of a Paretian liberal* holds: there simply is no such A.

In the alternative proof I present bellow, crucial use is made of the following lemma due to Suzumura (proposition 11.33). Given a preference relation $\gtrsim$, we say $\gtrsim'$ is a *compatible extension* of $\gtrsim$ if $\gtrsim' \supseteq \gtrsim$ and, moreover, $\succ' \supseteq \succ$. In words: a compatible extension is any relation which extends the original one while preserving its strict part (that is, no strict preference stance according to the original relation becomes an indifference according to the extension). Suzumura's extension lemma states that, if a relation is strongly acyclic, then it has an *ordering extension* (that is, a compatible extension that is an ordering). An immediate corollary of this lemma is that, if a relation $\gtrsim$ is acyclic and *strict* (that is, if $\gtrsim\, = \,\succ$), then it has an ordering extension as well.

Now for the promised proof of Sen's theorem.

Proof. Given that A is an SDF for I, S satisfying UD and WP and it also satisfies ML by supposition, let $i \neq j$ and $x \neq y$, $z \neq w$ be any and assume that xE_iy and zE_jw. Let X be any finite situation such that $x, y, z, w \in X$. Because X is finite, assume $|X| = n$ and let $X = \{x_0, \dots, x_n\}$, with $x_n = x_0$. Without loss of generality, assume that $x = x_0, y = x_1$ and, given an $\ell < n$, that $z = x_\ell, w = \ell + 1$. There are two cases to be considered.

Supposing $\ell = 0$, let p be any preferences profile such that $x \succ_i^p y$ and $w \succ_j^p z$. Trivially, such a p exists and thus $p \in \mathsf{dom}(A)$, by UD. However, because both i, j are then decisive between the very same pair x, y, it follows that $x \succ^p y$ and $y \succ^p x$, — which is impossible, due to the asymmetry of $\succ^p$. Therefore A is undefined for p, contradicting the assumption that it satisfies UD.

Supposing now $\ell \neq 0$, let p be any profile such that $x \succ_i^p y$ and $z \succ_j^p w$, and such that, for all $k < n$, $k \neq 0$, $k \neq 1$, $k \neq \ell$, $k \neq \ell + 1$, $x_k \succ_k^p x_{k+1}$ (which, of course, is tantamount to saying that $x_k P_I^p x_{k+1}$ for all $k < n$). That is: i strictly prefers x_0 to x_1, j strictly prefers x_ℓ to $x_{\ell+1}$ and, for every pair of subsequent states in the sequence $x_0, \dots x_n$ *other* than those to which i, j are entitled, every individual in the society strictly prefers the former to the latter. Now, because for every individual there is an issue in the sequence for which his or her preferences are not specified, and every individual's preferences insofar as they *are* specified are therefore strict and acyclic, by Suzumura's compatible extension lemma such a p indeed exists and thus again $p \in \mathsf{dom}(A)$, by UD. However, because $x_0 \succ_i^p x_1$ and $x_0 D_i x_1$; because $x_\ell \succ_j^p x_{\ell+1}$ and $x_\ell D_j x_{\ell+1}$, and because, by WP, $x_k P_I^p x_{k+1}$ and $x_k D_I x_{k+1}$ for all $k < n$, $k \neq 0$, $k \neq 1$, $k \neq \ell$, $k \neq \ell + 1$, it follows that $x_k \succ^p x_{k+1}$ for all $k < n$. Since $x_n = x_0$, this is a preferences cycle, — which means $\succsim^p$ is not acyclic, and thus that A is not an SDF. What is more, because those are *all* the options in X, $C^p = B_{\succsim^p}(X)$ is therefore empty, — which means C^p is *not* determined for X either. □

Sen illustrates the conflict between the exercise of individual rights and the Pareto principle with the much repeated example of a simple society with two individuals, Mr. Lewd and Mr. Prude, and a dispute over a copy of a controversial book, *Lady Chatterley's lover*. On the one hand, Mr. Lewd would very much like to read the book, but he would like it even better if the close-minded Mr. Prude were to read it. On the other hand, Mr. Prude thinks the book should not be read at all, but also that, if it were to be read, he should be the one to do so, as this would at least prevent Mr. Lewd from engaging once more in morally reprehensible actions. Now, because they live in a liberal society, it seems natural that, between either no one reading the book and Mr. Lewd reading it, Mr. Lewd's preference regarding the issue should be the only one to count. Similarly, between either no one reading the book and Mr. Prude reading it, Mr. Prude's preference should prevail. Both, however, think Mr. Prude having the copy is better than Mr. Lewd having it. But then Mr. Lewd reading the book should be socially preferred to no one reading it, no one reading it should be preferred to Mr. Prude reading it, and Mr. Prude reading it should be preferred to Mr. Lewd reading it. This is a social preferences cycle, and so there is an impasse about whether someone, and who, should have the copy of *Lady Chatterley's lover*.

Gibbard offers [23] a different but just as famous example, which is also of some interest because it seems to elicit distinct intuitions on the matter. It is the case of love triangle involving Edwin, Angelina and the judge. Angelina is in love with Edwin and can barely hide that she wishes to marry him; only the judge has proposed to her, however, and she would gladly settle for him so as to avoid becoming a spinster. Now, Edwin is a confirmed bachelor, but he also happens to foster a profound enmity for the judge and, because of that, Edwin would hate to see him married to Angelina. The judge himself, being a sober man of few strong passions, is happy with whatever Angelina decides. Clearly, in a liberal society, whether in these circumstances Angelina marries the judge or remains single should be entirely up to her, and, if she were to reject the judge's proposal, whether Edwin accepts her love for him and marries her, or whether he stays forever a bachelor instead, should be his business alone. Given their preferences, however, this may lead to a social preferences cycle in much the same way as in Sen's example.

But let us get back to technical matters. The gist of the alternative proof of Sen's theorem presented above is that it not only shows that the presence of two entitled individuals in the society renders the social preferences cyclical for some profiles if the aggregation function satisfies UD and WP, but also that *every* finite situation involving those entitlements may be rendered undetermined in the right (or wrong, as it is perhaps more proper to say in this context) circumstances.

Batra and Pattanaik have presented [4] a natural generalization of the impossibility of Paretian liberal, known from the literature as *the impossibility of Paretian federalism*. Instead of two individuals as in Sen's ML, their condition of *minimal federalism* assumes the existence of two disjoint entitled coalition instead. From a logical point of view, an alternative proof for it would read essentially the same as the one above. (Indeed, Sen's theorem is just the particular case in which the two disjoint coalitions are individual ones, so that the Pareto relation for each coincides with the strict preferences of those very individuals). For this reason, we may well treat the two theorems as if they were the same.

Many ways of escaping the puzzle presented by Sen have been proposed. From among these, I believe only two are really worth mentioning here, the one by Gibbard and the one by Sen himself. Obviously enough, in order to escape the conclusion of the theorem some of its assumptions must be weakened. The question is: which ones, and how?

Sen's [59] own proposal is essentially to limit the scope of application of WP by imposing a "virtual" domain restriction. The basic idea appeals to the character of the individuals. In Sen's own terms, an individual is *rights respecting* (or simply *liberal*, in Suzumura's expression) if he or she is so disposed as to refrain from expressing his or her preferences regarding issues to which he or she is not personally entitled, whenever doing so would avoid a social preferences cycles. Sen proves that, if at least one individual in the society is liberal, then the impossibility does not arise.

Gibbard [23] takes the cue from his own example for Sen's theorem by observing that, while Angelina has every incentive to exercise her right to exclude the possibility of never marrying, as this is the worst possible outcome according to her preferences, and therefore she should do so irrespective of what Edwin chooses, Edwin himself, given that his only actual choice are then between either seeing Angelina wedded to the judge or taking her as his wife, has an incentive to *waive* (that is: deliberately not exercise) his right to remain a bachelor in order to avoid an outcome that he (despicable as his motives may be) considers worse then ending up as Angelina's husband. Gibbard's resolution scheme thus preserves WP by weakening the strength of individual rights, as it recommends at least one individual waiving his or her rights whenever a social preferences cycle would otherwise come about.

Chapter 4

Coherent decisiveness assignments and the common source of impossibilities

Given this short but more than thorough review of social choice theory, we can finally move towards the seemingly simple question I'd like to tackle in this chapter, which is this: *why* do puzzles of social choice such as the Sen's impossibility theorem and Condorcet's paradox occur whenever an SDF is supposed to satisfy UD and WP?

In order to properly answer that question, I need to introduce a host of concepts and prove some technical lemmata. So let us begin by the concepts which will allow us to make the question more precise (and therefore manageable).

Definition 4.1 (Potential litigiousness)**.** Let A be a PAF for I, S, and let X be a situation. We say X is *potentially litigious under* A if: X is finite, with $|X| \geq 2$; assuming $|X| \leq n$, there is a sequence $x_0, \ldots, x_n$, with $x_k \neq x_{k+1}$ for all $k < n$ but $x_n = x_0$, such that $X = \{x_0, \ldots, x_n\}$ and a collection $V \subseteq \wp(I)$ such that, for every $k < n$, there is a coalition $J_k \in V$ such that $x_k D_{J_k} x_{k+1}$.

We refer to such a collection V, given such a sequence of states in X, as an *arrangement of powers regarding* X. For the most part, though, the relevant sequence of options will be clearly fixed by the context and won't be explicitly specified.

In words, a situation is potentially litigious if it is finite, *non-trivial* (that is: there are at least two distinct states in it) and if it is possible for the individuals to arrange themselves into coalitions so that the issues occurring in the situation end up being the matter of decisiveness in a way that could, at least *in principle*, lead to social preferences cycles. In point of fact, notice that, because for each $k < n$ there is a J_k decisive for x_k against x_{k+1}, if it were the case that $x_k P^p_{J_k} x_{k+1}$ for every $k < n$ for some $p \in \mathsf{dom}(A)$, then the resulting social preferences $\succsim^p$ restricted to X would have an *overarching* cycle (that is, a cycle involving every state in X) and, therefore, the social choice for X would be undetermined. Of course, there may be no $p \in \mathsf{dom}(A)$ which would actually generate such a cycle. Nevertheless we know from the proof of Sen's theorem presented here that some preferences profiles *do*, at least for *some* PAFs.

Our original question can then be spelled out thus: what additional condition an

SDF satisfying UD and WP has got to satisfy so as to prevent any potentially litigious situation from ever becomes *actually* so, with no state being socially acceptable?

4.1 Canonicity and coherence

Before introducing the one concept that is crucial to answering our question, we need to prove a technical lemma about how the concept of decisiveness relates to the existence of certain PAFs. The results presented in this section are simpler versions of the results in [3]. It is nonetheless worth mentioning that we have arrived at them independently, as generalizations of the results in [64].

Given PAF A for a society I, S, let us think of the decisiveness powers it implements *independently*, as an arbitrary *decisiveness assignment* (or DA, for short), — by which I mean, as a total function $D : \wp(I) \longmapsto \wp(S \times S)$ assigning binary relations on the set of options to coalitions. (Alternatively, one may think of a decisiveness assignment as a total function $D : S \times S \longmapsto \wp(\wp(I))$ assigning sets of coalitions to ordered pairs of options. Both ways work just fine, so the reader should settle for whichever feels more natural to him or her.) Of course, if this is to actually *be* a proper DA — the decisiveness implemented by a PAF satisfying UD at least —, it must have the properties of vacuousness, monotonicity and, additionally, exclusiveness. One might then wonder if there is some procedure for performing the following task: given a DA D (any function D of that type and satisfying those three properties, that is), to find a PAF satisfying UD and yielding complete social preferences such that the DA it implements is precisely D. More precisely, let us say D is *implemented* by A if xDy implies xD^Ay, for all x, y, and that it is *faithfully implemented* by A if the converse is also the case, for all x, y. Then we may put the description of the task in the form of a question: is *any* DA faithfully implementable?

The answer is a straightforward *yes*. We prove this means of a canonical PAF in the simple lemma bellow. But let us first definite what is meant by a PAF being the canonical one for a DA.

Definition 4.2 (Canonical PAF)**.** Let D be a DA for I, S. The *canonical PAF for D* is the PAF defined by: $x \succ^p y$ if xD_Jy and xP^p_Jy for some J, and $y \succsim^p x$ otherwise, for all $p \in \mathcal{O}(S)^I$.

The basic idea of the canonical PAF (which is a generalization of the concept of the *Pareto extension* found in Sen) is *overkill*: if either xD_Jy for no J such that xP^p_Jy, or if not xP^p_Jy for any J such that xD_Jy, then it yields $y \succsim^p x$; in simple words: in all instances in which either no coalition is decisive for x against y, or some are but, in all those coalitions, not everyone strictly prefers x to y, the canonical PAF for D provides a possible counterexample to the decisiveness of every coalition that, according D, indeed is not supposed to be decisive on that issue.

Proposition 4.1 (Implementation lemma)**.** *Any DA is faithfully implementable.*

Proof. Let I, S be a society, and let D be a DA for it (with the properties of vacuousness, monotonicity and exclusiveness, that is), and let A be the canonical PAF for D. It can easily be seen that, provided A is well-defined, it satisfies UD, $\succsim^p$ is complete for every $p \in \mathsf{dom}(A)$ and, moreover, A implements D. (Incidentally, A also happens to satisfy condition IIA, as it is pairwise defined by design.) As can just as easily be checked by inspection of the definition of A, however, for it to be ill-defined one of two things would have to happen, for some p: either the definition of A would entail $x \succ^p x$ for some x, or it would entail instead both $x \succ^p y$ and $y \succsim^p x$ for some x, y. The former could only happen if xD_Jx and xP^p_Jx, and this is impossible because D satisfies vacuousness; the latter could only happen if, for some J, J' both xD_Jy and xP^p_Jy and $yD_{J'}x$ and $yP^p_{J'}x$. This, however, implies J, J' would have to be disjoint, which is ruled out by exclusiveness together with monotonicity.

It remains to be checked whether xD^Ay implies xDy for all x, y, —- that is: if the implementation of D by A is indeed faithful. So let J, x, y be any and suppose xD^A_Jy. Because A satisfies UD, there is a $p \in \mathsf{dom}(A)$ such that xP^p_Jy, — from which it follows, since xD^A_Jy, that $x \succ^p y$. But it can't be the case that $x \not{D}_J\, y$, for then we would have $y \succsim^p x$ by the definition of A, thus contradicting $x \succ^p y$. Therefore, xD_Jy. □

Now we can finally introduce the concept of coherence of a DA, which is a straightforward generalization of one elaborated by Gibbard [23], Farrell [16] and Suzumura [64] regarding systems of individual rights.

Definition 4.3 (Coherent DA)**.** Let D be a DA for I, S. An *incoherence* in D is a non-empty sequence $x_0, \ldots, x_n$ of options, with $x_k \neq x_\ell$ for all $k < \ell < n$ but $x_n = x_0$, such that, for every $k < n$, there is a coalition J_k such that $x_kD_Jx_{k+1}$ and $\bigcap\{J_k\}_{k<n} = \varnothing$. A DA is *coherent* if there is no incoherence in it.

Alternatively, we could have define an incoherence to be a non-empty sequence of *ordered pairs* of options instead, — one of the form $x_0, y_0, \ldots, x_{n-1}, y_{n-1}$, with $x_k \neq x_\ell$ for all $k < \ell < n$, and with $x_{k+1} = y_k$ for all $k < n-1$ and $y_{n-1} = x_0$, such that, for every $k < n$, there is a coalition J_k such that $x_kD_Jy_k$ and $\bigcap\{J_k\}_{k<n} = \varnothing$. We will stick to the first definition for the rest of this chapter, but the second, more roundabout one will also find its use later on.

Allow me to try to make this somewhat esoteric definition a bit more clear.

Given a preference relation $\succsim$, let us call an (ordered) pair x, y such that $x \succsim y$ a *link* in that relation; and, given a PAF for a society, let us say individual i is (partially, at least) *responsible* for the link x, y in $\succsim^p$ if $i \in J$, xD_Jy and xP^p_Jy. That is: i is responsible for the link x, y if he or she is in a coalition decisive for x against y and he or she strictly prefers x to y together with everyone in else in that coalition, thus helping it exercise that particular decisiveness power. Now recall that a cycle in a preference relation is a sequence of options, beginning and ending in the same option, such that, for every pair of subsequent options in the sequence, the former is strictly preferred to the latter. It is easy to see that an individual cannot be responsible for every single link in a cyclic social preference relation, for this would imply his or her preferences are cyclic as well, which contradicts the assumption that individual preferences are transitive. Indeed, if *any* option occurs twice in the sequence even before it ends, then no single individual can be responsible for all the links in between those two occurrences, for this is simply a shorter cycle. So, if there is a social preferences cycle and it is to result *from the exercise of decisiveness powers alone*, no individual can be in every coalition decisive on some pair of subsequent options involved in the cycle.

This informal argument shows coherence, insofar as it excludes such patterns in a DA, is a necessary condition for a DA to never enable individuals to bring about a social preferences cycle. Is it also *sufficient*? — Well, because cycles are by definition finite sequences, within every such sequence there must be a shortest cycle, — one involving the least amount of options, and thus also the least amount of links. The shortest cycle within a cycle is, of course, *non-redundant* (that is: no option occurs twice in it before the repetition of its first option) for, if it weren't, then by definition it wouldn't be the shortest. So, if there's a social preferences cycle, then there's also a non-redundant one, — from which it follows that, if no non-redundant cycle can be brought about solely by the exercise of decisiveness powers, then no redundant ones can be either. Now, if there *is* an incoherence in the DA implemented by a PAF, then there will be profiles (provided the PAF satisfies UD) that will produce non-redundant social preferences cycles in that fashion, as all that is required from the individuals for it to happen is that each be responsible for at most all but one link in the sequence of links involved in a non-redundant cycle. This is so because, if for each individual the part of a non-redundant cycle for which he or she can be responsible is, of necessity, an acyclic and strict relation, then, by Suzumura's compatible extension lemma, it is also a relation having at least one compatible ordering extension. Therefore, for each individual it suffices that we take that ordering to be his or her preferences for us to have a profile that will generate a social preferences cycle.

Notice that in this discussion we have ignored the case of an incoherence with no options in between the first and the last, for this is equivalent to the empty coalition being decisive between an option and itself. Because this implies the DA violates vacuousness, the only PAF implementing it is the empty one (that is, the PAF defined for no preferences profile at all).

Proposition 4.2 (Coherence lemma)**.** *An SDF satisfying UD implements a coherent DA; and, given a coherent DA, the canonical PAF for it is an SDF satisfying UD.*

Proof. Let A be an SDF for I, S satisfying UD and suppose for a contradiction that D is not coherent, which means there is an incoherence in it. Let $x_0, \dots, x_n$ and $J_0, \dots, J_{k-1}$ be as in the definition of incoherence. (As already observed, it can't be that $n < 2$, for this would imply that $\mathsf{dom}(A) = \varnothing$, contradicting UD). Notice that the last part of that definition is equivalent to saying that, for every individual i, there is at least one coalition from among $J_0, \dots, J_{k-1}$ to which he or she doesn't belong. So let p be any profile such that, for each i, if $i \in J_k$, then $x_k \succ_i^p x_{k+1}$. Because his or her preferences as so partially specified are logically consistent, acyclic and strict for the reasons discussed above, as usual by Suzumura's compatible extension lemma such a p indeed exists and $p \in \mathsf{dom}(A)$, by UD. This implies that $x_k P^p_J x_{k+1}$ for each $k < n$; and, because also $x_k D_J x_{k+1}$ for each $k < n$, it follows that $x_k \succ^p x_{k+1}$ for each $k < n$. But then, because $x_n = x_0$, the preference relation $\succsim^p$ is not acyclic, thus contracting the assumption that A is an SDF.

Now let D be a coherent DA for I, S. By proposition 4.1 we know that the canonical implementation of D is a PAF A satisfying UD which faithfully implements D. It remains to be shown that A is an SDF. So suppose not. Then there is a $p \in \mathsf{dom}(A)$ generating a social preferences cycle under A. Let such a p be given, and let $x_0, \ldots, x_n$ be that cycle. Then $x_k \succ^p x_{k+1}$ for each $k < n$, — which means, by the definition of A, that there is at least one J_k such that $x_k D_{J_k} x_{k+1}$ and such that $x_k P^p_{J_k} x_{k+1}$. For each k, let one such J_k be fixed. Now let $x_\ell, \ldots, x_m$ be the shortest cycle within the given cycle (which we know exists and is of necessity non-redundant, as already pointed out), and let i be any individual. Obviously enough, it can't be that $i \in \bigcap\{J_k\}_{\ell<k<m}$, for then $\succsim^p_i$ would be cyclic, — which is impossible so long as individual's preferences are assumed to be transitive. It follows that $\bigcap\{J_k\}_{\ell<k<m} = \varnothing$; and, because $x_\ell, \ldots, x_m$ is a non-redundant cycle, this then implies that D has an incoherence, which contradicts the assumption that it is coherent. □

As perhaps a matter of curiosity if nothing else, it might be noted that one can deduce Condorcet's paradox from this proposition as a corollary.

Corollary 4.2.1 (Condorcet's paradox)**.** *If the society has at least three individuals and at least three options, then, provided the number of individuals is finite, pairwise majority rule for that society is not an SDF.*

Proof. Let I, S such that $|I| = n \geq 3$ finite and $|S| \geq 3$ a society be any, and let X such that $|X| = 3$ be any. Assume $X = \{x_0, x_1, x_2, x_3\}$ with $x_3 = x_0$, and let $I = \{i_0, \ldots, i_{n-1}\}$. Taking π to be the permutation of I such that $\pi(i_k) = i_{k+\lfloor \frac{n}{2} \rfloor \text{ (mod } n)}$, let coalitions J_0, J_1, J_2 be such that $J_0 = \{i_0, \ldots, i_{\lfloor \frac{n}{2} \rfloor}\}$ and, for both $\ell \in \{1, 2\}$, $\pi(i_k) \in J_\ell$ iff $i_k \in J_{\ell-1}$. (For example, if $n = 3$, then $I = \{i_0, i_1, i_2\}$, $J_0 = \{i_0, i_1\}$, $J_1 = \{i_1, i_2\}$ and $J_2 = \{i_2, i_0\}$). Because all such coalitions are majoritarian, all of them are decisive for every issue under pairwise majority rule; in particular, for each $k < 3$, J_k is entitled to x_k, x_{k+1}, — from which it follows that, for every $k < 3$, there is a J_k such that $x_k D_{J_k} x_{k+1}$. As can easily be seen, however, $\bigcap\{J_0, J_1, J_2\} = \varnothing$ by design, as any two such coalitions have at most two individuals in common and there are three coalitions. Therefore, the DA implemented by pairwise majority rule for I, S has an incoherence; and thus, because pairwise majority rule satisfies UD, it can't be an SDF. □

4.2 Litigiousness and the Pareto principle

Now we're finally in a position so as to be able to properly answer the question with which we began this section: namely, how to avoid that an SDF satisfying conditions UD and

WP ever turns a potentially litigious situation into an actually litigious one.

So let X be potentially litigious under some A for I, S, and let V be an arrangement of powers regarding it. We call $\bigcap V$, the set of individuals who are in every coalition in the arrangement, the *arbitral coalition* of V. We explain why call it that shortly. Before doing so, we should note that two simple but important facts follow from what has been proven so far. The first one is that, if an PAF satisfies WP — which is equivalent to saying that every possible entitlement is implemented by it (proposition 3.1) —, then every non-trivial, finite situation must be potentially litigious under it (so that a situation being non-trivial and finite and it being potentially litigious become equivalent). The second is that, because the DA an SDF satisfying UD implements is of necessity coherent (proposition 4.2), the arbitral coalition of any arrangement of powers, regarding any such situation, is non-empty. These two facts together immediately entail the next proposition (and therefore I state it without proof).

Proposition 4.3. *If an SDF satisfies UD and WP, then the arbitral coalition of any arrangement of powers is non-empty.*

This proposition is of the utmost importance, because both Sen's theorem and Condorcet's paradox follow from it as easy corollaries.

Corollary 4.3.1 (Batra's and Pattanaik's impossibility of federalism theorem – 1972 [4])**.** *There is no SDF satisfying UD and WP with two disjoint coalitions being each entitled under it to an issue involving distinct states.*

Proof. Let A for I, S be an SDF satisfying UD and WP be any, and suppose for a contradiction there are disjoint J, J' and $x \neq y, z \neq w$ such that $xE_J y$ and $zE_{J'}w$. Then, for every finite situation X such that $x, y, z, w \in X$ and every arrangement of powers V for it such that $J, J' \in V$, the arbitral coalition $\bigcap V$ is empty, which is impossible because A is an SDF satisfying UD and WP. □

Corollary 4.3.2 (Sen's impossibility of a Paretian liberal theorem – Sen 1970)**.** *There is no SDF satisfying conditions UD, WP and ML.*

Proof. Follows immediately from the previous corollary and from the fact that any two distinct individual coalitions are disjoint. □

4.3 The inevitability of arbitration

As for Arrow's impossibility theorem, it is well known that, if a PAF satisfying UD and WP is to also satisfy IIA and yield at least complete and *quasi-transitive* preferences (that is, if the strict part of the social preference relation for each profile in its domain is to be transitive — a condition stronger than being an SDF), then every coalition decisive on even a single issue is globally decisive, and moreover, if there are finitely many individuals in the society, then there is coalition globally decisive and contained in every other decisive coalition, — an *oligarchy*. Clearly, then, the oligarchy is not only the smallest possible arbitral coalition, but also its singleton is the smallest possible arrangement of powers for whatever finite situation, and in any case it is contained in every possible arbitral coalition. Of course, if the PAF is furthermore required to yield fully transitive preferences for all profiles in its domain, then the oligarchy will consist of a single individual and the PAF will thus be dictatorial, just like Arrow has famously shown.

It is undeniable that oligarchies and dictators hold vast powers, as they are able to choose what the social preferences will be, at least regarding issues on which they happen to have strict preferences; and, since both are essentially "overgrown" arbitral coalitions, one might wonder what kind of power, if any, arbitral coalitions themselves enjoy. This is the question I now answer. As we will see, even a PAF satisfying conditions much weaker than those imposed by Arrow must already exhibits signs of the phenomenon of concentration of power first recognized by him.

Definition 4.4. Let I, S be a society. A coalition J in that society has *potential settling power in situation X under A* if there is a *partial preferences profile* (or a *context*: let us call it so) $q \in \mathcal{O}(S)^{\overline{J}}$ such that, for every $x \in X$, there is a preferences profile $p \in \mathcal{O}(S)^{I}$ extending q such that $x \in B_{\succsim^p}(X)$.

That is, a coalition has potential settling power if, depending on the preferences of all the individuals *not* in it, the coalition may end up being granted the power to freely choose which state is to be implemented.

Now, finally, for our long overdue explanation: arbitral coalitions deserve to be so called on account of to the following fact, which neatly relates all the concepts introduced in this chapter.

Proposition 4.4. *If a situation is potentially litigious under an SDF satisfying UD, then, for every arrangement of powers regarding that situation, if the arbitral coalition of the arrangement is non-empty, it has potential settling power.*

Proof. Let A be an SDF for I, S satisfying UD, let X be potentially litigious and let $V \neq \varnothing$ be an arrangement of powers regarding it. Let $x \in X$ be any. Assuming $x_0, \ldots, x_n$ with $x_n = x_0$ is the relevant sequence, it follows that for each $k < n$ there is a coalition $J_k \in V$ entitled to x_k, x_{k+1}. Without loss of generality, assume $x = x_0$. For each $i \in \overline{\bigcap V}$ (that is: for each individual *not* in the arbitral coalition), let q be any context such that, for all $k < n$, if $i \in J_k$, then $x_k \succ_i^q x_{k+1}$. Because no such i is in every coalition in the arrangement of powers, all such partially specified preference relations are strict and acyclic, and therefore, by Suzumura's compatible extension lemma, all of them have ordering extensions — which implies such a q indeed exists. Now let p be any preferences profile such that $\succsim_i^p = \succsim_i^q$ for all $i \in \overline{\bigcap V}$ and such that, for every $i \in \bigcap V$, $x_k \succ_i^p x_{k+1}$ for every $k < n-1$. Again by Suzumura's aforementioned lemma, such a profile also exists, and it happens to extends q by design.

But then we have it that $x_k P^p_{J_k} x_{k+1}$ for each $k < n - 1$, — and therefore, because J_k is entitled to x_k, x_{k+1}, that $x_k >^p x_{k+1}$ for all $k < n - 1$. Now, since A is an SDF, it can't be that $x_k >^p x_n$ for any $k < n$, for then $>^p$ would have a cycle (not to mention that $x_0 >^p x_n$ is logically impossible, as $x_n = x_0$ by assumption). But then $x_n = x_0 = x$ must be the only state in situation X not dominated by some other state under the social preference relation for the profile, — which in turn entails, given that A is an SDF and $\gtrsim^p$ is therefore complete and acyclic, that $x \in B_{\gtrsim^p}(X) = C^p(X)$. — In sum: the arbitral coalition of V has potential settling power in situation X under A. □

But why care for this rather weak kind of power, one might ask? — At first glance, there may seem to be nothing necessarily wrong with a coalition having potential settling power (or *actual*, even: potential settling power with an empty context), at least when it happens to be the smallest one entitled to every issue in a feasible set or situation. (If the situation involves only an individual's own rights, then this may in fact be desirable.) But what if there are two or more coalitions *concerned* in the situation (that is: decisive on issues in the situation) and, for at least two of them, neither is a part of the other (even if not necessarily disjoint, as in Sen's theorem)?

As our next proposition shows, if one wishes the preferences aggregation rule for a society to always yield minimally rational social preferences with individuals being free to express whatever preferences they happen to have through the exercise of their allotted rights and powers, then the rule cannot obey the Pareto principle without also granting some group the power to arbitrate, *no matter the situation or who is concerned in it* (so long as it is finite, that is).

Proposition 4.5 (The inevitability of arbitration). *If an SDF satisfies UD and WP, then, for every finite situation, there is a coalition with potential settling power in it.*

Proof. If a situation is trivial, then every coalition (vacuously) has potential settling power in it under any SDF. If a situation is non-trivial and finite, then the conclusion follows at once from proposition 4.3 and proposition 4.4. □

This theorem can be cast in an "impossibility form", so to speak, by adopting the following additional condition, which seems rather mild and a plausible thing to expect from a minimally *peaceful* society:

Self-mediation (SM) There is at least one non-trivial, finite situation in which the society can surely do without arbitration, — that is, an X with $|X| \geq 2$ such that no coalition in society has potential settling power in it.

Then it might as well read:

Proposition 4.6 (The impossibility of self-mediation). *There is no SDF jointly satisfying conditions UD, WP and SM.*

How troublesome is this conclusion of the inevitability of arbitration, or of the impossibility of self-mediation? — In some settings, it may well be said to be not at all. Consider, for instance, a choice between two options to be made by three people under pairwise majority rule (which *is* an SDF satisfying UD and WP for simple choices), and let us say that the first of them would choose one option, and the second the other. Then, and naturally enough, it would remain for the third person to join either the first or the second and settle the matter. This is a clear case of arbitration, but it is hard to think there is anything wrong with it. (There *might* be, however, if, for example, the arbitrator were to vote already informed of the votes of the other individuals. The Arrovian framework does not account for informational aspects of collective choice of this sort.) But pairwise majority rule is special in this sense because, if we abstract from certain informational aspects, then no one in particular can be said to be *the* arbitrator in our example. For suppose the third person were to choose the one option, the second the other, and let us assume we are now uninformed of the first person's choice. (Formally, this is tantamount to considering a different arrangement of powers for the same feasible set.) Now it is the first person who turns out to be the arbitrator. In any event, because all of them have potential settling power (and indeed, due to pairwise majority rule satisfying anonymity, because all individuals necessarily have the same amount of power), there is no inequality involved: arbitral power is not *arbitrary* in this case.

The trouble begins when a PAF is expected to *not* satisfy anonymity, — which is *always* the case in any setting in which individual rights conceived as entitlements are supposed to exist. In that regard, perhaps the most important conceptual implication of what has been shown here is that, to make individual rights compatible with the Pareto principle, one must eschew the notion that such rights can be strictly *individual* entitlements (unless, of course, no more than one individual is to have rights); but then one may not be fully able to exercise one's rights without first securing the support of a 'third party", so to speak: whether Angelina and Edwin marry might hinge on what a judge wants, after all.

4.4 A resolution scheme for the impossibility

But *how much*, exactly, *should* what the judge wants matter?

Suppose the social rules were to *actually* assign the role of arbitrator to the judge. More formally, let $I = \{i_a, i_e, i_j\}$ and $S = \{x_e, x_j, x_s\}$ be that society, with i_a, i_e, i_j standing, respectively, for Angelina, Edwin and the judge, and with x_e, x_j, x_s standing for Angelina marrying Edwin, she marrying the judge and everybody remaining single. Let A be an SDF for this society, assume it satisfies conditions UD and WP, and assume, moreover, that the DA it implements makes the judge an arbitrator in any dispute between Angelina and Edwin regarding marriage, otherwise assigning no decisiveness powers; that is, it makes I entitled to every issue, it makes coalition $\{i_a, i_j\}$ decisive between x_j, x_s, coalition $\{i_e, i_j\}$ decisive between x_e, x_s, and that is all. It is easy to see that this DA is coherent. Now let p be any profile that gives rise to a social preferences cycle under the DA originally intended for this society, with Angelina strictly preferring x_e to x_j and x_j to x_s, Edwin preferring x_s to x_e and x_e to x_j, and the judge strictly preferring x_e to x_j. Under the DA implemented by A, however, the cycle is avoided simply because, although I's decisiveness for x_e against x_j is exercised, the decisiveness of $\{i_a, i_j\}$ for x_j against x_s and the decisiveness of $\{i_e, i_j\}$ for x_s against x_e cannot be exercised together, no matter what the preferences of the judge on the other issues may be, due to the transitivity of his preferences.

Now let p' be a different profile, — one in which Angelina's preferences are the same as before, but now Edwin actually wants to marry her: that is, he both strictly prefers x_e to x_s and x_e to x_j in p'; and let the judge's preferences regarding who should Angelina marry remain unchanged. Then, as in p, by condition WP it follows that $x_e \succ^{p'} x_j$. However, let us also say that in p' the judge (perhaps somewhat out of spite) is indifferent between Angelina marrying Edwin and not marrying at all. Then neither the decisiveness of coalition $\{i_a, i_j\}$ between x_j, x_s is exercised (because Angelina and the judge have opposite stances on the issue), nor is the decisiveness of coalition $\{i_e, i_j\}$ between x_e, x_s (because only Edwin has a strict position on the issue). Nothing, however, prevents A from ruling both x_e and x_j better than x_s anyway, since these social preferences neither are cyclical nor do they disrespect any entitlement; and, because both Angelina and Edwin now prefer marrying each other to staying single, there is also no good reason why it shouldn't rule x_e best — no matter what the judge wants.

The point I'm trying to make by way of this example is that there is no good reason for a social choice rule to not operate *as if* implementing rights conceived as entitlements, exactly as Sen wants, *except* in response to profiles that would induce a social preferences cycle if it did.

To make this suggestion more concrete, let A for I, S be a PAF (not necessarily an SDF) satisfying UD such that, if it is cyclic for some profile, the cycles all come from the exercise of decisiveness alone. That is: if $x_0, \dots, x_n$ is a non-redundant cycle of $\succsim^p$, then there are $J_0, \dots, J_{n-1}$ such that $x_k D_{J_k} x_{k+1}$ and $x_k P^p_{J_k} x_{k+1}$, for all $k < n$. Now let V a non-empty coalition in that society be given, and let $\overset{\Delta}{A}$ be the PAF defined as follows: for all x, y, if there is no cycle $x_0, \dots, x_n$ in $\succsim^p$, with $x_k \neq x_\ell$ for all $k < \ell < n$, such that $x = x_k$ and $y = x_{k+1}$ for some $k < n$, then $x \overset{\Delta}{\succsim}{}^p y$ iff $x \succsim^p y$; however, if there is such a cycle, then: if $x P^p_V y$, then $x \overset{\Delta}{\succ}{}^p y$; otherwise, $x \overset{\Delta}{\sim}{}^p y$. That is: if the issue x, y isn't involved in any non-redundant social preferences cycle for profile p, then $\overset{\Delta}{\succsim}{}^p$ agrees with $\succsim^p$ on x, y; if it is, then $\overset{\Delta}{\succsim}{}^p$ agrees with $\succsim^p$ if P^p_V agrees too, and it settles for social indifference otherwise. Because by assumption the link x, y occurs in a cycle of $\succsim^p$ only if there is a coalition J decisive for x against y such that everyone in J strictly prefer x to y in p, one may interpret this construction as meaning that, in the advent of such a social preferences cycle, the "claim" of coalition J to a "right" for x against y is "confirmed" by $\overset{\Delta}{A}$ in the case of that issue being involved in a social conflict if, and only if, coalition V unanimously "rules in favor" of J on the issue, so to speak. For this reason, we may refer to $\overset{\Delta}{A}$ as the *adjudication of A through P_V* (where P_V is to be taken as a function assigning Pareto relations to profiles), and to V as the *judicial coalition* for the adjudication of A through P_V.

Clearly, $\overset{\Delta}{A}$ satisfies UD and yields complete social preferences for any profile by design. Furthermore, decisiveness powers exercised under A remain so under $\overset{\Delta}{A}$, provided the holder of that power is not responsible for any link in a social preferences cycles.

It follows immediately from the definition that the adjudication of an SDF is simply that SDF itself. We now prove that the adjudication of a PAF is an SDF even if the original PAF is not one.

Proof. Let A a PAF for I, S be any, and let $V \neq \varnothing$ a judicial coalition for the adjudication of A through P_V be fixed. Let any preferences profile p be given.

First of all, notice that no cycle of $\gtrsim^p$ can be a cycle of $\overset{\Delta}{\gtrsim}{}^p$. For suppose for a contradiction that they share a cycle, and let $x_0, \ldots, x_n$ be any non-redundant cycle that is part of that cycle. (As we know, there must be at least one such part.) By the definition of $\overset{\Delta}{A}$, this can only be a cycle of $\overset{V}{\gtrsim}{}^p$ if $x_k P^p_\Delta x_{k+1}$ for all $k < n$, — which is impossible because P^p_V is transitive, thus acyclic. Therefore, no such cycle of $\overset{\Delta}{\gtrsim}{}^p$ exists, and consequently neither does any cycle that it might share with $\gtrsim^p$.

Nevertheless suppose, again for a contradiction, that $\overset{\Delta}{\gtrsim}{}^p$ has a cycle $x_0, \ldots, x_n$ anyway. Then it must be because $x \overset{\Delta}{\succ}{}^p y$ for at least one issue x, y such that $x \not\succ^p y$, since they can't share a cycle. But then $\overset{\Delta}{\gtrsim}{}^p$ disagrees with $\gtrsim^p$ on that issue, and by the definition of $\overset{\Delta}{A}$ this can only be because $x \overset{\Delta}{\sim}{}^p y$ while $x \succ^p y$, — a contradiction. Therefore, $\overset{\Delta}{\gtrsim}{}^p$ has no cycles; and, because p is any, the adjudication of A through Δ is an SDF. $\square$

Perhaps the most salient effect of adjudication is that, for every arrangement of powers with an empty arbitral coalition regarding any potentially litigious situation under A, every coalition J in that arrangement, will under $\overset{\Delta}{A}$, stop being decisive on at least one issue in that situation. In simpler terms, it is to say that adjudication necessarily induces a redistribution of decisiveness powers. For every such situation, substituting $J \cup V$ for J in that arrangement of powers will still make an arrangement of powers for it, with the exact same configuration of options; but, of course, now the situation can never become actually litigious because V is the arbitral coalition for that arrangement, and V is by assumption non-empty. Nonetheless, as already observed, $\overset{\Delta}{A}$ will only diverge from A in regard to actually litigious situations, so that the eventual redistribution of powers induced by adjudication won't be as radical as one might fear. Indeed, it is easy to see that adjudication also doesn't happen to accidentally confer any new decisiveness powers. For suppose for a contradiction that it did. Then there would be an issue x, y and a coalition J which is now decisive for x against y (decisive under $\overset{\Delta}{A}$, that is), although it wasn't decisive before (under A). By definition, this means that there is a profile p

such that $y \succsim^p x$ even though $xP_J^p y$, and such that $x \overset{\Delta}{\succ^p} y$. Because, however, an adjudication can only diverge from A on an issue under a profile by making a strict social preference into an indifference, this is impossible, so that no coalition can gain decisiveness powers by means of adjudication alone.

By way of conclusion, two more facts are worth pointing out about the resolution scheme of adjudication.

One is that, as inspection of the proof that $\overset{\Delta}{A}$ is an SDF reveals, there is no need for adjudication to be made through P_V. *Any* SDF Δ for I, S will do the trick, or even any function Δ (perhaps depending on other parameters in addition to preferences profiles), provided it only yields acyclic relations on S as outputs. Indeed, the interested reader can easily check for him or herself that both Sen's [59] and Gibbard's [23] resolution schemes for the impossibility of a Paretian liberal describe functions that fit this scheme under certain conditions (*cf.* section 13.4), so that they are themselves just very particular examples of adjudication. This observation, I believe, opens up the door to a whole field of investigation of possible particular adjudication schemes.

The other fact worth pointing out is that the assumption that cycles result exclusively from decisiveness exercising is somewhat gratuitous and can be dropped, as it plays no role in how adjudication operates. Be that as it may, conflicts of rights are what motivated the introduction of the concept in this dissertation. Other possible applications certainly require further research.

Chapter 5

Discussion and conclusions

As we have seem, the role the Pareto principle plays in the inevitability of arbitration is that it makes *every* issue in society the matter of some entitlement, so that rights must sometimes outcompete other entitlements in order to be effective. Because this can't be done in a coherent fashion if more than one individual in the society is to have rights, then arbitration (and therefore some form of limitation of individual rights) is inescapable. But why care for the Pareto principle in the first place? — This is the question I would like to briefly address in this closing chapter of the first part of this dissertation, as it seems to me to be at the core of what we have investigated up until this far.

Originally, as revealed by the name it bears, the Pareto principle was conceived as a more or less obvious principle of efficiency in welfare economics [56]. In this context, it seems very natural to think that, if everyone in society judges x better then y, then, for all intents and purposes, x being implemented indeed is socially better than y being implemented. Let us refer to this interpretation of the weak Pareto condition as the *welfare interpretation*.

Another possible interpretation of the principle is what we might call the *unanimity interpretation*. It is easy to see that the definition of pairwise majority rule, *cf.* section 3.2, implies the weak Pareto condition, as the coalition of all individuals in society is always a majority. Under this "democratic" interpretation, the Pareto principle prescribes that unanimous strict preferences be respected by the aggregation rule.

It is hard to deny that the Pareto principle seems highly commendable as a principle of social choice, no matter which of those two interpretations makes more sense in the given context. Nevertheless, in the face of the difficulty posed by the impossibility of the Paretian liberal, the Pareto principle has been extensively criticized by Sen [55], [56]. His most poignant criticisms center on the issue of the *informational restriction* imposed on the aggregation function by the principle. Notice that, in saying that x is to be socially strictly preferred to y whenever x is strictly preferred to y by everyone, the principle implicitly imposes the restriction that this must be so *irrespective* of how x and y fare when compared to other options. Just as the condition of independence of irrelevant alternatives is usually pointed out as the main responsible for the dreary conclusion of Arrow's impossibility theorem, and precisely to the extent that it imposes an informational restriction of the same kind, Sen believes it is this aspect of the weak Pareto condition that is mostly to blame for the impossibility of the Paretian liberal.

As announced in the introduction (chapter 2) to this part of the dissertation, I'd like to propose that there is a third possible interpretation of the Pareto principle, — one which, as far as I can tell, hasn't been seriously considered in the literature so far, and one which also suggests an entirely new understanding of what is really at stake in the impossibility of the Paretian liberal.

One of the immediate logical consequences of the weak Pareto condition, besides the already discussed condition of weak non-imposedness, *cf.* section 3.3, is the stronger condition of *citizen's sovereignty* (CS), so named by Arrow himself: for every x, y, there is a profile p in the domain of the PAF such that x ends up being socially strictly preferred to y. Intuitively, this conveys a notion of sovereignty in the sense that, not only is no social preference regarding the issue x, y imposed, but also there is always a way for any option to actually socially prevail over any other, depending only on the preferences of the individuals.

As we have seen in section 3.3, the weak Pareto condition is equivalent to there being in society, for every possible issue, a coalition entitled to it. In this formulation, and insofar as it logically implies condition CS, the Pareto principle can be understood as conveying an even stronger notion of sovereignty: the notion that, ultimately, the decision on every issue hinges on the power or authority of some group within society.

According to the Senian conception of law and rights, *cf.* section 3.4, individual rights are a particular kind of what we have been generically calling (mostly for the sake of not begging any questions) "entitlements". Conversely, then, one may reasonably think of entitlements as something akin to rights, — perhaps as "coalitional" rights, one might say. However one thinks of these notions, the point is that, formally at least, there is no important distinction between rights and entitlements of other possible sorts (such as, for example, entitlements stemming from legal authority or sheer political power). By saying this I don't mean to deny that there may be relevant *moral* distinctions to me made between rights and other kinds of entitlements, nor that those distinctions shouldn't weight in when one is trying to devise appropriate aggregation rules (and above all, perhaps, when trying to devise reasonable adjudication rules for PAFs that implement incoherent DAs). What I do mean is that, so long as every possible entitlement is to be implemented — which means: so long as the Pareto principle holds —, those distinctions can no longer play any substantive role. If the Senian conception of law and rights is sound, then the Pareto principle expresses a conception of sovereignty which puts political power and rights on equal footing; therefore, it also leaves no room for individual rights (at least if we assume that a society in which only a single individual has rights is unacceptable, that is).

Under this interpretation, however, the weak Pareto condition also presents itself as somewhat less plausible than the staple of democratic ruling it appears to be under the unanimity interpretation, or the seemingly neutral principle of economic efficiency commended by the economic interpretation. But this is not to say that it is implausible either.

In modern political philosophy, where the concept has its roots, sovereignty is usually understood as the power of the sovereign. So, who is the sovereign? — The commonsense notion, whose first systematic analysis can be found in Bodin [5], but which was made popular by Hobbes [30], is that the sovereign is a person or group of people holding "absolute power" in society. Admittedly, this is not very informative definitions in the present context, but I guess it does convey the intended notion, which can then be naturally cashed out in terms of decisiveness: the sovereign (if one exists) is the individual or coalition that is decisive on every issue. Now, given that the Pareto principle says precisely that the coalition of all individuals is a sovereign in this sense, and given that, if there is a sovereign in this sense, then the monotonicity of decisiveness implies that the Pareto principle holds, then it turns out that the Pareto principle is also equivalent to the existence of a sovereign.

Of course, we have come a long way since the likes of Bodin and Hobbes, and contemporary liberal democracies are far from being the kind of political body they envisaged as being the best possible arrangement of power. But I believe that this basic idea of sovereignty is in essence still present in the usual arrangement of power that is to be found in contemporary societies, *cf.* [32], [52], — and, thus, even if liberal democracy seems at first sight a notion antagonistic to absolute sovereignty.

Generally speaking, Arrow's theorem has been received as a serious blow to the aspiration of rationality of democratic processes. But it turns out that this may not be of much relevance in contemporary societies, as the kind of decision processes with which the theorem is primarily concerned, under that interpretation (namely, electoral processes), albeit being of great social significance, are not the bread and butter and the day-to-day of decision-making in such societies, with their complex economies and borderline impenetrable bureaucracies. Actually, the decisions which in such societies are usually subjected to all-encompassing electoral processes are the ones that have to do with *who* is to be assigned the power to ascribe decisiveness powers (i.e., rulers and lawmakers in general), with average citizens otherwise exercising decisiveness almost exclusive through their rights.

This, I believe, is an observation of fundamental importance about contemporary societies, in regards to the relevance of social choice theory for understanding them. We are called upon to vote mostly, if ever, for leaders and legislators, whose primary function (at least in functioning liberal democracies, that is) is precisely to produce a coherent and effective rights system for all; and a rights system is nothing more, from the point of view of social choice theory, than a decisiveness assignment of some sort. It seems to me, therefore, that the matter of the rationality of the distribution of powers is a far more pressing one.

Part II

Tarskian logic and abstract notions of consistency

Chapter 6

Introduction

Central to the argument of the first part of this dissertation is the concept of a coherent decisiveness assignment, which is, in a sense, the counterpart of the concept of an SDF, — that is, of a PAF universally yielding acyclic social preferences. Because acyclicity of social preferences is arguably the bottom line of social rationality (or, at the very least, of rationalizable social choice), it seems this concept might be of some interest in and of itself.

As briefly mentioned in the first part of this dissertation, the concept of coherence of a decisiveness assignment is a generalization of the concept of a coherent system of rights, as developed by Gibbard, Farrell and, more rigorously, Suzumura. In the literature, coherent systems of rights are sometimes alternatively called *consistent*. As it stands, however, this is a rather loose occurrence of the word "consistency", at least insofar as it is supposed to be charged with *logical import*.

Sen himself wrote a good deal on the somewhat liberal use of the word "consistency" in social choice theory [54], [57]. In those writings, he is primarily concerned with what he dubs the matter of *internal consistency* of choice — by which he means, of what the choice from a feasible set should imply for the choice from another feasible set with which it shares options —, as opposed to the matter of *external consistency*, — of how a choice function fits some given standard (such as rationalizability by best options through some preference relation). Nonetheless, as we have observed in the first part of this dissertation, there is also a matter of the rationality of standards of choice themselves, as most of Arrovian social choice theory is concerned with conditions of rationality of social preferences. Whether in the case of internal or of external consistency of choice, therefore, the question of the logical import of the purported concept of consistency stands.

This seems to me to be a foundational issue of social choice, and thus one much deserving of careful investigation. But how, exactly, is one to approach it?

From the point of view of *classical* logic, consistency has to do, first and foremost, with sets of *sentences* (or *propositions*; I will stick to the latter for reasons that will become clearer as we advance into the subject). A set of propositions is consistent if it does not entail an absurdity (which is classically equivalent to a contradiction), and it entails an absurdity if and only if it is *explosive*, that is, if it entails *every* proposition. In other words, inconsistency is classically equivalent to *triviality*. So, from this point of view, a set of propositions is inconsistent if and only if it leads to triviality; by duality, then, it is consistent if, and only if, there is a proposition it does not entail, that is, if and only if its deductive closure is not the set of all propositions. Moreover, because classical logic is extensive (meaning a set entails every proposition in it) and monotonic (meaning its logical consequences are preserved under extensions of a set of propositions), it follows that the set of all propositions is inconsistent and that, if a set is inconsistent, then so is any of its supersets; again by duality, we must have that the empty set is consistent and that, if a set is consistent, then so are its subsets.

Even this brief overview of the classical concept of consistency makes it clear that, given a logic with at least *some* of the properties of classical logic, one is also implicitly given a notion of consistency suited for that particular universe of propositions. We, however, are coming from the opposite direction. We already have something we suspect is a concept of consistency, but don't have the logic that would make us sure it has the logical import we expect it to have.

To deal with this issue, in this part of the dissertation an alternative approach to logic is introduced which takes consistency — instead of, say, truth or deduction — to be its primitive concept. The basic idea is to take the structural properties of the concept of consistency we have abstracted from classical logic in the discussion above and use them, together with some suitably intuitive definitions, to show how, given such an abstract notion of consistency, we can always find a Tarskian logic whose notion of consistency will be precisely the one from which we have started.

Unusual as this approach might seem at first, it is by no means without antecedents. According to [29], Keisler was the first to adopt such a perspective about logic, doing so in [34] in order to study the properties of infinitary logics. The same fundamental assumptions about consistency Keisler makes can more recently be found again, for example, in Dietrich's works on *judgment aggregation theory* [13] (which is, by the way, itself a spin-off of social choice, *cf.* [25], [38]). The direct inspiration for what I present here is, however, Stalnaker's theory of propositions [63]. The fact that I borrow from him almost all of the auxiliary definitions notwithstanding, Stalnaker himself takes his theory in a completely different direction, as he is mostly concern with the ontological issues surrounding modal notions. Indeed, as far as I can tell, at least in print Stalnaker never developed his theory of propositions into a theory of logic in the way that I do here; so, it seems, this part of the dissertation may be of interest not only to those interested in social choice theory, but to philosophers and logicians in general as well.

Let us nonetheless recall that our main objective is to provide an answer to the question of whether coherence of a decisiveness assignment can be taken to be a logical concept of consistency. Not all of the theory developed in this part will be needed to properly answer that question. Despite that, mostly for the sake of generality, we shall take a moment from social choice theory to delve deeper into such foundational issues of logic.

Chapter 7

Consistency, deduction and truth

Let us begin by the defining the concept of an abstract notion of consistency, *cf.* [13], [29], [63], which is simply a straightforward formalization of what has been discussed in the introduction of this chapter.

Definition 7.1 (Abstract notion of consistency)**.** An *abstract notion of consistency* (ANC, for short) is a pair Λ, Σ such that $\Lambda \neq \varnothing$ and Σ is a *subset space*, *cf.* section 10.2, on Λ (that is, a collection of subsets of Λ) satisfying the following condition, for all Φ, Ψ.

Closure under subsets If $\Phi \in \Sigma$ and $\Psi \subseteq \Phi$, then $\Psi \in \Sigma$

Given such an ANC, we refer to a $\varphi \in \Lambda$ as a *proposition*. (This is here intended just as a generic name for the universe of objects of an ANC. We do not need to make upfront any ontological commitment about the kind of objects they are supposed to be.) Then, given a set Φ of propositions, we say it is *consistent* $\Phi \in \Sigma$.

Alternatively, we could have defined an ANC as a pair Λ, Σ such that Λ is as above and $\overline{\Sigma}$ satisfies the following condition for all Φ, Ψ.

Closure under supersets If $\Phi \in \overline{\Sigma}$ and $\Psi \supseteq \Phi$, then $\Psi \in \overline{\Sigma}$

It is a basic algebraic fact that the two definitions amount to the same thing. Nonetheless it pays off to always have both in mind, as it is sometimes easier to work with one than the other.

Intuitively, an ANC is comprised of a set of propositions (which we take to be *all* the propositions, or at least all the propositions of the relevant kind) together with an specification of the sets of propositions that are consistent. In principle, this specification could be entirely arbitrary, but, as we've already discussed, closure under subsets seem quite natural and acceptable (at least from a classical point of view, that is). The assumption that the subsets of a consistent set are also consistent properly conveys the intuition that inconsistency is a matter of there being *information* that is *conflicting* in some sense. Therefore, if the information available is not conflicting, then neither should be the information that remains if some information is lost.

By stating this I don't mean to claim that there are no other plausible conceptions of consistency. One sure can think of situations in which, contrary to what has been stated, at first glance it seems plausible that consistency is *gained* instead of lost by the acquisition of new information. Actually, there is in the literature on *non-monotonic logic*, *cf.* [26], a very popular example of just this sort of thing, — the case of the propositions expressed by "Tweet is a bird" and "Tweet doesn't fly", which may be taken to be inconsistent given that, by *default*, birds fly; and yet, consistency is immediately recovered once one is informed that Tweet happens to be a penguin (a bird of a somewhat unusual kind). It may be of some interest to observe that, in the setting of ANCs, this example would then also have to be framed as a case of non-monotonicity (but non-monotonicity of the underlying inconsistency space, not necessarily of some logic). Since, however, this subject is not of our immediate concern for the purposes of this dissertation, I'll drop it.

What *is* of our concern is that familiar logical notions can be directly *defined* in ANCs. This is where I take direct inspiration from Stalnaker's theory of propositions [63] and borrow the most from him.

Definition 7.2 (Logical relations and logical concepts in ANCs)**.** Let Λ, Σ be an ANC, and Φ, Ψ be sets of propositions. Φ *is compatible with* Ψ if $\Phi \cup \Psi$ is consistent. Φ *is equiconsistent to* Ψ if, for any Ξ, the set Φ is compatible with Ξ iff so is Ψ. Φ *implies* Ψ if $\Phi \cup \Psi$ is equiconsistent to Φ. Φ *compatibly extends* Ψ if $\Phi \supseteq \Psi$ and Φ is consistent. Φ *is maximally consistent* if Φ is consistent and there is no $\Psi \supset \Phi$ which compatibly extends Φ.

Let φ be a proposition. Φ *entails* φ if Φ implies $\{\varphi\}$, and φ *is equivalent to* ψ if both $\{\varphi\}$ entails ψ and $\{\psi\}$ entails φ. φ *is a tautology* if $\Phi \cup \{\varphi\}$ is consistent for every consistent Φ, it is an *antilogy* $\{\varphi\}$ is inconsistent, and it is *contingent* if it is neither a tautology nor an antilogy.

In what follows, I write $\Phi \Leftrightarrow \Psi$ to mean Φ is equiconsistent to Ψ, and $\Phi \Rightarrow \Psi$ to mean Φ implies Ψ. I also write $\Phi \Vdash \varphi$ to mean Φ entail φ.

It is worthy to point out that, because a subset space of consistent sets is closed under subsets, the definition of the negation of implication can be somewhat simplified to this: $\Phi \not\Rightarrow \Psi$ iff there is a Ξ such that $\Phi \cup \Psi \in \Sigma$ but $\Phi \cup \Psi \cup \Xi \notin \Sigma$, for all Φ, Ψ. To see that this is the case observe that, by definition, $\Phi \not\Rightarrow \Psi$ iff $\Phi \cup \Psi \not\Leftrightarrow \Phi$, which means there is a Ξ such that either $\Phi \cup \Psi \cup \Xi \in \Sigma$ but $\Phi \cup \Psi \notin \Sigma$, or $\Phi \cup \Psi \in \Sigma$ but $\Phi \cup \Psi \cup \Xi \notin \Sigma$; and the former can never be the case, because $\Phi \cup \Psi \subseteq \Phi \cup \Psi \cup \Xi$ and Σ is closed under subsets. Noticing this is quite useful, as this simpler definition will allows us to shorten some of the proofs that follow.

I guess the intuition about information motivating these definitions is quite clear, but a little discussion can do no harm. A set is compatible with another if the information each of them conveys does not conflict with that conveyed by the other, and they are equiconsistent if they are compatible with exactly the same sets, — that is, if the same information, when added to them, either preserves the consistency of both or of none. A set implies another if the information conveyed by the second "adds nothing" to the one conveyed by the first, in the sense that the resulting set remains compatible and incompatible with exactly the same sets of propositions. The other definitions are familiar from logic, and I suppose they are exactly what is to be expected of them in the present setting.

7.1 Logic in ANCs

We now go on to prove a whole bunch of facts about our (purportedly, for now) logical relations and logical concepts definable in ANCs, which will in turn allow us to define a concept of *deductive inference* in terms of consistency. The basic idea we are going for here, as the reader is sure to have anticipated already, is that we shall use the relation of entailment to build a (Tarskian) consequence relation out from any given ANC.

So let an ANC Λ, Σ be fixed in advance for all propositions bellow.

Proposition 7.1. *A set of propositions is consistent iff it is compatible with itself.*

Proof. Follows immediately from the fact that $\Phi = \Phi \cup \Phi$ for any Φ. □

Proposition 7.2. *A set of propositions is consistent iff it has a compatible extension.*

Proof. Let Φ a set of propositions be any.

First assume Φ is consistent, that is, $\Phi \in \Sigma$. Then, because trivially Φ is a compatible extension of Φ, it follows Φ has a compatible extension.

Now assume there's a Ψ such that Ψ compatibly extends Φ. Then, by definition, $\Phi \subseteq \Psi$ and $\Psi \in \Sigma$, so that, because Σ is closed under subsets, $\Phi \in \Sigma$ as well. Therefore, Φ is consistent. □

Proposition 7.3. *Only consistent sets have compatible extensions.*

Proof. Immediate from the closure under supersets of the subset space of inconsistent sets. □

Proposition 7.4. *Compatibility is a symmetric relation, equiconsistency is an equivalence relation and implication is a quasi-ordering.*

Proof. I prove only that implication is transitive, as the rest follows trivially from the definitions.

Let Φ, Ψ, Ξ be any, and assume $\Phi \Rightarrow \Psi$ and $\Psi \Rightarrow \Xi$. Then $\Phi \cup \Psi \Leftrightarrow \Phi$ and $\Psi \cup \Xi \Leftrightarrow \Psi$, — that is: $\Phi \cup \Psi \cup A$ is consistent iff $\Phi \cup A$ is consistent as well, for whatever A, and similarly, $\Psi \cup \Xi \cup B$ is consistent iff $\Psi \cup B$, too, is, for whatever B. We need to show that $\Phi \Rightarrow \Xi$, that is, that, $\Phi \cup \Gamma$ is consistent iff so is $\Phi \cup \Xi \cup \Gamma$, for any Γ. So let Γ be given.

Suppose, first, $\Phi \cup \Xi \cup \Gamma$ is consistent. Then, because $\Phi \cup \Gamma \subseteq \Phi \cup \Xi \cup \Gamma$ and Σ is closed under subsets, $\Phi \cup \Gamma \in \Sigma$. Therefore, if $\Phi \cup \Xi \cup \Gamma$ is consistent, then so is $\Phi \cup \Gamma$.

Suppose, now, $\Phi \cup \Gamma \in \Sigma$. Then, taking $A = \Gamma$, we have it that $\Phi \cup \Psi \cup \Gamma$ is consistent and, furthermore, taking $B = \Phi \cup \Gamma$, that $\Psi \cup \Xi \cup \Phi \cup \Gamma$, too, is. Again because Σ is closed under subsets, we conclude $\Phi \cup \Xi \cup \Gamma \in \Sigma$. Therefore, if $\Phi \cup \Gamma$ is consistent, then so is $\Phi \cup \Xi \cup \Gamma$.

It follows $\Phi \cup \Xi \cup \Gamma$ is consistent iff $\Phi \cup \Gamma$ is: from which it follows by definition that $\Phi \cup \Xi \Leftrightarrow \Phi$, and therefore, also by definition, that $\Phi \Rightarrow \Xi$. □

The next proposition justifies, at least in part, the name of the relation of equiconsistency.

Proposition 7.5. *If a set is equiconsistent to another, then one is consistent iff so is the other.*

Proof. Let Φ, Ψ be any, and assume $\Phi \Leftrightarrow \Psi$. Then $\Phi \cup \Xi$ is consistent iff $\Psi \cup \Xi$ is, for all Ξ. In particular, $\Phi \cup \Psi$ is consistent iff so is $\Psi \cup \Psi$ (that is, iff Ψ itself is consistent), and similarly, $\Psi \cup \Phi$ is consistent iff $\Phi \cup \Phi$ is (which, similarly, amounts to: iff Φ is). Therefore, Φ is consistent iff Ψ is. □

Proposition 7.6. *Two sets are equiconsistent iff they imply each other; that is: equiconsistency is the symmetric part of implication.*

Proof. Let Φ, Ψ be any.

Suppose, first, $\Phi \Leftrightarrow \Psi$. Then Φ is consistent iff Ψ is, by the previous proposition. So suppose for a contradiction $\Phi \not\Rightarrow \Psi$. Then there's a Ξ such that $\Phi \cup \Psi \cup \Xi$ is not consistent although $\Psi \cup \Xi$ is. Let such a Ξ be given. Because $\Phi \Leftrightarrow \Psi$, so that $\Phi \cup \Gamma$ is consistent iff $\Psi \cup \Gamma$ is, for whatever Γ, by taking $\Gamma = \Psi \cup \Xi$ it follows $\Phi \cup \Psi \cup \Xi$ is consistent iff $\Psi \cup \Psi \cup \Xi$ is (that is, iff $\Psi \cup \Xi$ is). Therefore, because $\Psi \cup \Xi$ is consistent by supposition, it also follows $\Phi \cup \Psi \cup \Xi$ is consistent, thus contradicting the supposition that $\Phi \cup \Psi \cup \Xi$ isn't. Therefore, $\Phi \Rightarrow \Psi$. By a similar reasoning, we conclude $\Psi \Rightarrow \Phi$ as well. Thus, if $\Phi \Leftrightarrow \Psi$, then $\Phi \Rightarrow \Psi$ and $\Psi \Rightarrow \Phi$.

Suppose, now, $\Phi \Rightarrow \Psi$ and $\Psi \Rightarrow \Phi$. Then $\Phi \cup \Psi \Leftrightarrow \Phi$ and $\Psi \cup \Phi \Leftrightarrow \Psi$ —that is: $\Phi \Leftrightarrow \Phi \cup \Psi$ (for $\Leftrightarrow$ is an equivalence relation, therefore symmetric) and $\Phi \cup \Psi \Leftrightarrow \Psi$ —, from which it follows by the transitivity of $\Leftrightarrow$ that $\Phi \Leftrightarrow \Psi$. Therefore, if $\Phi \Rightarrow \Psi$ and $\Psi \Rightarrow \Phi$, then $\Phi \Leftrightarrow \Psi$. □

Proposition 7.7. *A set implies its subsets.*

Proof. Let $\Phi \supseteq \Psi$ be any. Because $\Phi \cup \Psi = \Phi$, trivially $\Phi \cup \Psi \cup \Xi$ is consistent iff $\Phi \cup \Xi$, for all Ξ, — from which it follows by definition that $\Phi \Rightarrow \Psi$. □

Proposition 7.8. *A set implies the subsets of the sets it implies.*

Proof. Let Φ, Ψ, Ξ such that $\Phi \Rightarrow \Psi$ and $\Psi \supseteq \Xi$ be any. Then $\Psi \Rightarrow \Xi$ by the previous proposition, from which it follows by the transitivity of $\Rightarrow$ that $\Phi \Rightarrow \Xi$ as well. □

Proposition 7.9. *Two sets are equiconsistent iff they imply the same sets.*

Proof. Let Φ, Ψ be any. We want to prove that $\Phi \Leftrightarrow \Psi$ iff for all Ξ it is the case that $\Phi \Rightarrow \Xi$ iff $\Psi \Rightarrow \Xi$.

So first suppose $\Phi \Leftrightarrow \Psi$. Let Ξ be any, and suppose $\Phi \Rightarrow \Xi$. Then, $\Phi \cup \Xi \Leftrightarrow \Phi$, which means $\Phi \cup \Xi \cup \Gamma$ is consistent iff $\Phi \cup \Gamma$ is, for all Γ. So let Γ be any. Because $\Phi \Leftrightarrow \Psi$ by supposition, $\Phi \cup \Gamma$ is consistent iff so is $\Psi \cup \Gamma$ and, similarly, $\Psi \cup \Xi \cup \Gamma$ is consistent iff $\Phi \cup \Xi \cup \Gamma$, too, is, — from which it follows $\Psi \cup \Xi \cup \Gamma$ is consistent iff $\Psi \cup \Gamma$ is consistent as well. Because Γ is any, $\Psi \cup \Xi \cup \Gamma$ is consistent iff $\Psi \cup \Gamma$ is, for all Γ, — that is, $\Psi \cup \Xi \Leftrightarrow \Psi$, which means $\Psi \Rightarrow \Xi$. If we now suppose $\Psi \Rightarrow \Xi$, by a similar reasoning we find it that $\Phi \Rightarrow \Xi$, and thus that $\Phi \Rightarrow \Xi$ iff $\Psi \Rightarrow \Xi$. Therefore, because Ξ is any, $\Phi \Rightarrow \Xi$ iff $\Psi \Rightarrow \Xi$, for all Ξ.

Now suppose $\Phi \Rightarrow \Xi$ iff $\Psi \Rightarrow \Xi$, for all Ξ. Then, because by the previous proposition $\Phi \Rightarrow \Phi$, also $\Psi \Rightarrow \Phi$, and similarly, because $\Psi \Rightarrow \Psi$, also $\Phi \Rightarrow \Psi$, — which, as already proved (proposition 7.6), is equivalent to $\Phi \Leftrightarrow \Psi$. □

Proposition 7.10. *A set implies two other sets iff it implies their union.*

Proof. Let Φ, Ψ, Ξ be any.

First suppose $\Phi \Rightarrow \Psi$ and $\Phi \Rightarrow \Xi$. Then $\Phi \cup \Psi \Leftrightarrow \Phi$ and $\Phi \cup \Xi \Leftrightarrow \Phi$. So suppose for a contradiction $\Phi \not\Rightarrow \Psi \cup \Xi$. Then there's a Γ such that $\Phi \cup \Psi \cup \Xi \cup \Gamma$ is not consistent although $\Phi \cup \Gamma$ is. Because $\Phi \cup \Psi \Leftrightarrow \Phi$ by supposition, it follows in particular that $\Phi \cup \Gamma$ is consistent iff $\Phi \cup \Psi \cup \Gamma$, too, is; and, because again by supposition $\Phi \cup \Xi \Leftrightarrow \Phi$, it also follows in particular that $\Phi \cup \Psi \cup \Gamma$ is consistent iff so is $\Phi \cup \Xi \cup \Psi \cup \Gamma$. But, this being so, because $\Phi \cup \Gamma$ is consistent by supposition, it follows $\Phi \cup \Xi \cup \Psi \cup \Gamma$ is consistent as well, thus contradicting the supposition that it is not $\Phi \cup \Psi \cup \Xi \cup \Gamma$. Therefore, $\Phi \Rightarrow \Psi \cup \Xi$ after all.

The converse follows trivially from the fact that a set implies the subsets of the sets it implies (proposition 7.8). □

Corollary 7.10.1. *A set implies certain finitely many sets iff it implies their union.*

Proof. Follows from the proposition by an obvious induction. □

Proposition 7.11. *A proposition is entailed by a set iff it is in some set it implies.*

Proof. Let φ a proposition and Φ a set be any.

First suppose $\Phi \Vdash \varphi$. Then, by definition, $\Phi \Rightarrow \{\varphi\}$, so that $\varphi \in \Psi$ for some Ψ such that $\Phi \Rightarrow \Psi$ (namely, $\Psi = \{\varphi\}$).

Now suppose $\varphi \in \Psi$ for some Ψ such that $\Phi \Rightarrow \Psi$. Then, because $\{\varphi\} \subseteq \Psi$ and because a set implies the subsets of the sets it implies (proposition 7.8), $\Phi \Rightarrow \{\varphi\}$, — so that, by definition, $\Phi \Vdash \varphi$. □

Proposition 7.12. *All inconsistent sets are equiconsistent.*

Proof. Let Φ, Ψ inconsistent be any, and let Ξ be any. Because $\overline{\Sigma}$ is closed under supersets, both $\Phi \cup \Xi$ and $\Psi \cup \Xi$ are inconsistent. Because Ξ is any, $\Phi \cup \Xi$ is consistent for no Ξ, and similarly regarding Ψ. Therefore, $\Phi \cup \Xi$ is consistent iff $\Psi \cup \Xi$ is, for all Ξ, — that is: $\Phi \Leftrightarrow \Psi$. □

Proposition 7.13. *An inconsistent set implies every set.*

Proof. Let Φ inconsistent be any. Because $\overline{\Sigma}$ is closed under supersets, $\Phi \cup \Psi$ is inconsistent for whatever Ψ. So, by the previous proposition, $\Phi \cup \Psi \Leftrightarrow \Phi$, which means $\Phi \Rightarrow \Psi$. Because Ψ is any, $\Phi \Rightarrow \Psi$ for every Ψ. □

Corollary 7.13.1. *If an antilogy is in a set, then the set implies every set.*

Proof. Immediate from the proposition and the definition of antilogy. □

Corollary 7.13.2. *An inconsistent set entails every proposition.*

Proof. Immediate from the proposition and the definition of entailment. □

Proposition 7.14. *A set implies an inconsistent set iff it is itself inconsistent.*

Proof. Let Φ be any and Ψ inconsistent be any. If Φ is inconsistent, then, by the previous proposition, it implies every set, and in particular it implies Ψ. If $\Phi \Rightarrow \Psi$, then $\Phi \cup \Xi \Leftrightarrow \Phi \cup \Psi \cup \Xi$ for all Ξ, and in particular for $\Xi = \varnothing$. Therefore, because $\Phi \cup \Psi$ is inconsistent (as $\overline{\Sigma}$ is closed under supersets), so is Φ. □

Corollary 7.14.1. *A set entails an antilogy iff it is inconsistent.*

Proof. Again, immediate. □

Corollary 7.14.2. *All antilogies are equivalent.*

Proof. Immediate from the previous corollary. □

Proposition 7.15. *All tautologies are entailed by every set.*

Proof. Let φ a tautology and Φ be any, and suppose for a contradiction $\Phi \nVdash \varphi$. Then there is a Ψ such that $\Phi \cup \{\varphi\} \cup \Psi$ is not consistent but $\Phi \cup \Psi$ is. From this, however, it follows by the definition of tautology that $\Phi \cup \{\varphi\} \cup \Psi$ is consistent, — a contradiction. Therefore, $\Phi \Vdash \varphi$. □

Corollary 7.15.1. *All tautologies are equivalent.*

Proof. Let φ, ψ be tautologies. If no set is consistent, then neither are $\{\varphi\}$ nor $\{\psi\}$, — from which it follows, by proposition 7.12, that φ and ψ are equivalent. If some set is consistent, then, by closure under subsets, so is the empty set. By the definition of tautology, however, this entails that both $\{\varphi\}$ and $\{\psi\}$ are consistent, and thus the conclusion immediately follows from the proposition. □

For those who are interested in the subject of *paraconsistent logic*, *cf.* [9], it is perhaps noteworthy that the seemingly thin conception of consistency here adopted happens to be quite strong, as it is provable that, in a sense, inconsistency by itself leads to a form of explosion, — and, more importantly, that is seems to do so even *in the absence of any explicit notion of contradiction whatsoever*. (In effect, we shall later on *define* contradictoriness in terms of consistency.) Of course, this should come as no surprise, since we have start by abstracting our notion of consistency from classical logic itself. In any case, and in accordance with what I have previously already observed, this hints at the hypothesis that the phenomenon of paraconsistency in logic has something to do with a different, non-monotonic notion of consistency. But again, as this is not our present subject, I won't dwell on the matter any further.

7.2 Deduction in ANCs

Certainly the most widely accepted formal concept of logic is the one by Tarski, which takes a logic to consist of a set of propositions (or sentences, or what have you — there's some room for dispute here, but this is immaterial to our purposes, as already pointed out), together with a *consequence relation* specifying what propositions can be inferred from what set. More formally, here is the Tarskian definition of logic, as presented in [15].

Definition 7.3 (Logic)**.** A (Tarskian) *logic* is a pair $\Lambda, \vdash$ such that $\Lambda \neq \varnothing$ and $\vdash \subseteq \wp(\Lambda) \times \Lambda$ satisfies the following conditions, for all $\varphi, \psi, \Phi, \Psi$.

Overlap If $\varphi \in \Psi$, then $\Phi \vdash \varphi$

Dilution If $\Phi \vdash \varphi$ and $\Psi \supseteq \Phi$, then $\Psi \vdash \varphi$

Cut If $\Phi \vdash \varphi$ and $\Phi \cup \{\varphi\} \vdash \psi$, then $\Phi \vdash \psi$

Sometimes a condition stronger then cut is required:

Infinitary cut If $\Phi \vdash \varphi$ for all $\varphi \in \Psi$ and $\Psi \vdash \psi$, then $\Phi \vdash \psi$

The next propositions show how entailment as defined in an ANC fares as a purported consequence relation. Again, let an ANC Λ, Σ be fixed.

Proposition 7.16. *Entailment satisfies overlap.*

Proof. Follows from the definition of $\Vdash$ and the fact that a set of implies its subsets (proposition 7.7). □

Proposition 7.17. *Entailment satisfies dilution.*

Proof. Follows from the definition of $\Vdash$ and the facts that a set of implies its subsets (proposition 7.7) and that a set implies the subsets of the sets it implies (proposition 7.8). □

Proposition 7.18. *Entailment satisfies cut.*

Proof. Let φ, ψ, Φ such that $\Phi \Vdash \varphi$ and $\Phi \cup \{\varphi\} \Vdash \psi$ be any, — which by definition means that $\Phi \Rightarrow \{\varphi\}$ and $\Phi \cup \{\varphi\} \Rightarrow \{\psi\}$. Because $\Phi \Rightarrow \Phi$ (as a set implies itself), and therefore $\Phi \Rightarrow \Phi \cup \{\varphi\}$ (since a set implies the union of two sets it implies), by the transitivity of $\Rightarrow$, it follows that $\Phi \Rightarrow \psi$ as desired. □

Proposition 7.19. *In general, entailment does not satisfy infinitary cut.*

Proof. Let Λ, Σ be the ANC such that $\Lambda = \mathbb{N}$ and Σ is the set of all *finite* $\Phi \subseteq \Lambda$ such that either φ is even for all $\varphi \in \Phi$ or φ is odd for all $\varphi \in \Phi$. It is easy to see that in particular all finite, non-empty sets of even numbers are equiconsistent, so that, for any φ an even and Φ a finite, non-empty set of even numbers, $\Phi \Vdash \varphi$. It follows $\Phi \Vdash \varphi$ for all φ in the set Ψ of all even numbers. However Ψ is an infinite set, and thus inconsistent by the definition of Σ. Being inconsistent, $\Psi \Vdash \psi$ for all ψ, and in particular for all ψ odd; but $\Phi \Vdash \psi$ for no odd ψ. Therefore, $\Vdash$ does not satisfy infinitary cut in this ANC. □

This last proposition shows that it is not necessarily the case that the entailment relation of an ANC is a consequence relation, although it is always the case that it is what we may call a *finitary* consequence relation. (It also provides a great example of why we don't have to commit to proposition being any particular kind of thing for logical relations to obtain among sets of them, by the way.) The question one is then sure to pose is: what else should we assume so as to make entailment *infinitary*, and so that the set of propositions of an ANC, together with its entailment, be a Tarskian logic in the stronger sense?

It is well known that a Tarskian logic can be equivalently presented in terms of a *closure operator* instead of a consequence relation. More precisely, given a Tarskian logic $\Lambda, \vdash$, one can define an operator $K_\vdash : \wp(\Lambda) \longmapsto \wp(\Lambda)$ by setting: $K_\vdash(\Phi) = \{\varphi \in \Lambda : \Phi \vdash \phi\}$. Any such closure operator K will then satisfy the following conditions, for all Φ, Ψ, as can easily be checked: $\Phi \subseteq K(\Phi)$; if $\Phi \supseteq \Psi$, then $K(\Phi) \supseteq K(\Psi)$; if $\Phi \subseteq K(\Psi)$, then $K(\Phi) \subseteq K(\Psi)$. (This third condition is equivalent to this, given dilution: $K(K(\Phi) \subseteq K(\Phi)$.) It is also evident that these three conditions correspond, respectively, to overlap, dilution and infinitary cut and that therefore, given a Tarskian logic Λ, K presented in terms of a closure operator, one can define a consequence relation simply by setting: $\Phi \vdash_K \varphi$ iff $\varphi \in K(\Phi)$. Clearly, given a consequence relation, the consequence relation so associated with its closure operator will be the original consequence relation and similarly if we start with a closure operator instead, so that these are indeed just two ways of conceiving of the same objects.

In order, then, to answer our question, we may cash out on the concept of deductive closure of a set of propositions, by defining such an operator in terms of the entailment of an ANC.

Definition 7.4 (Deductive closure)**.** Let Λ, Σ be an ANC and Φ a set of propositions. The *deductive closure of* Φ is $K(\Phi) = \{\varphi \in \Lambda : \Phi \Vdash \varphi\}$. A set is *deductively closed* if it is its own deductive closure.

Notice that $K(\Phi) = \bigcup\{\Psi \subseteq \Lambda : \Phi \Rightarrow \Psi\}$ is an equivalent definition of deductive closure, for, as already proved (proposition 7.11), a proposition is entailed by a set iff it is in some set it implies. Notice also that the deductive closure of a set is *not* necessarily deductively closed in an ANC, as the counterexample presented in the previous proposition clearly reveals.

Does the deductive closure work as expected? Our next propositions show that it does — up to a point.

Proposition 7.20. *A set of propositions is contained in its deductive closure.*

Proof. Follows immediately from the definition of deductive closure, since entailment satisfies overlap. □

Proposition 7.21. *The deductive closure of an inconsistent set is deductively closed.*

Proof. Follows immediately from the fact that an inconsistent set entails every proposition (corollary 7.13.2), so that its deductive closure is simply the set of all propositions, — which is obviously deductively closed. □

Proposition 7.22. *A maximally consistent set is deductively closed.*

Proof. Let Λ, Σ an ANC and Φ a maximally consistent set be any. (Recall that a set is maximally consistent if it is consistent but none of its proper supersets is.) Let φ be any such that $\Phi \Vdash \varphi$ — which means $\Phi \cup \{\varphi\} \cup \Psi$ is consistent iff $\Phi \cup \Psi$ is, for all Ψ —, and suppose for a contradiction $\varphi \notin \Phi$. Then $\Phi \subset \Phi \cup \{\varphi\}$, — from which it follows, because Φ is maximally consistent by assumption, that $\Phi \cup \{\varphi\}$ is inconsistent. Now, by taking $\Psi = \varnothing$, it also follows $\Phi \cup \{\varphi\}$ is consistent iff so is Φ. But then Φ is consistent even though $\Phi \cup \{\varphi\}$ isn't, — a contradiction. Therefore, $\varphi \in \Phi$. Because φ is any, Φ is deductively closed. □

For an ANC to be deductive in the intended, Tarskian sense, it needs to satisfy an additional condition.

Definition 7.5 (Deductive ANCs)**.** Consider the following condition an ANC Λ, Σ may satisfy, for all Φ.

Deductive implication Φ implies $K(\Phi)$

An ANC is *deductive* if it satisfies the condition of deductive implication.

Let Λ, Σ a deductive ANC be fixed. We prove this condition suffices for an ANC to be essentially a Tarskian logic in disguise.

Proposition 7.23. *In a deductive ANC, a set and its deductive closure are equiconsistent.*

Proof. Immediate from the condition of deductive implication and from the facts that a set of propositions is contained in its deductive closure (proposition 7.20), that a set implies its subsets (proposition 7.7) and that equiconsistency is the symmetric part of implication (proposition 7.6). □

Corollary 7.23.1. *In a deductive ANC, two sets are equiconsistent iff they have the same deductive closure.*

Proof. If two sets have the same deductive closure, then it follows from the proposition and the transitivity of the relation of equiconsistency that they are equiconsistent. The converse follows immediately from the definition of deductive closure and from the fact that two sets are equiconsistent iff they imply the same sets (proposition 7.9). □

Proposition 7.24. *In a deductive ANC, the deductive closure of the deductive closure of a set of propositions is the deductive closure of the set itself; that is: the deductive closure of a set is deductively closed.*

Proof. Let φ, Φ be any.

First suppose $\varphi \in K(\Phi)$. Then $K(\Phi) \Vdash \varphi$ by overlap, so that $\varphi \in K(K(\Phi))$ by the definition of deductive closure. Because φ is any, $K(\Phi) \subseteq K(K(\Phi))$.

Now suppose $\varphi \in K(K(\Phi))$. Then $K(\Phi) \Vdash \varphi$. So suppose for a contradiction $\varphi \notin K(\Phi)$. Then, $\Phi \nVdash \varphi$. Moreover, it follows by the definition of $\Vdash$ that $\Phi \not\Rightarrow \{\varphi\}$, although $K(\Phi) \Rightarrow \{\varphi\}$. However, as already proved (proposition 7.23), in a deductive ANC $\Phi \Leftrightarrow K(\Phi)$ and, as also already proved (proposition 7.9), two sets are equiconsistent iff they imply the same sets; but then there's a set (namely, $\{\varphi\}$) which $K(\Phi)$ implies but Φ does not, — a contradiction. Therefore, $\varphi \in K(\Phi)$ after all; and, because φ is any, $K(K(\Phi)) \subseteq K(\Phi)$. □

Proposition 7.25. *In a deductive ANC, if a superset of a set of propositions is contained in the deductive closure of the latter, then their deductive closures are the same.*

Proof. Let φ, Φ, Ψ such that $\Phi \subseteq \Psi \subseteq K(\Phi)$ be any.

First suppose $\varphi \in K(\Phi)$. Then, $\Phi \Vdash \varphi$, and so $\Psi \vdash \varphi$ by dilution. Because φ is any, $K(\Phi) \subseteq K(\Psi)$.

Now suppose $\varphi \in K(\Psi)$. Then $\Psi \Vdash \varphi$, which means $\Psi \Rightarrow \{\varphi\}$. Moreover, because $\Psi \subseteq K(\Phi)$ by assumption, $K(\Phi) \Rightarrow \Psi$, and, by deductive implication, $\Phi \Rightarrow K(\Phi)$. Therefore, by the transitivity of $\Rightarrow$, it follows $\Phi \Rightarrow \{\varphi\}$, so that $\varphi \in K(\Phi)$. Because φ is any, $K(\Psi) \subseteq K(\Phi)$. □

Proposition 7.26. *If an ANC is deductive, then its entailment satisfies infinitary cut.*

Proof. Immediate from the previous two propositions and from the fact that entailment satisfies dilution (proposition 7.17). □

Corollary 7.26.1. *If an ANC is deductive, then its set of propositions together with its entailment constitute a Tarskian logic.*

Proposition 7.27. *An ANC having finitely many propositions is of necessity deductive.*

Proof. Immediate from the definition of K and from the fact that a set implies the union of every set it implies if it implies only finitely many sets (corollary 7.10.1). □

Although in general an extra assumption is required in order to make a logic out of an ANC, the last proposition remarkably shows that, when only finitely many propositions are involved, the existence of a notion of consistency suffices for deductive inferences to be conduced. This may help explain why, historically, many philosophers (most notable among them, Aristotle) seem to have taken consistency to be essential to logic.

7.3 Negation and truth in ANCs

Of course, no investigation of the concept of consistent can be deemed complete without some discussion of its relation with the concepts of contradiction and negation, truth and falsity.

So let us begin by the concepts of contradiction and negation.

Definition 7.6 (Negation)**.** Let Λ, Σ be an ANC and φ, ψ propositions. *φ, ψ are contraries* if $\{\varphi, \psi\}$ is inconsistent. *φ, ψ are subcontraries* if, for every consistent Φ, either $\Phi \cup \{\varphi\}$ is consistent or $\Phi \cup \{\psi\}$ is. *φ, ψ are contradictories*, *cf.* [63], if they are both contraries and subcontraries. A *contradiction* is a pair φ, ψ such that φ, ψ are contradictories. Finally, *φ is a negation of ψ* if it is equivalent to every contradictory of ψ.

Notice that contradictories are not necessarily unique, as the following example shows.

Example 7.1. Let Λ, Σ be such that $\Lambda = \{\alpha, \beta, \gamma\}$ and $\Sigma = \{\varnothing, \{\alpha\}, \{\beta\}, \{\gamma\}, \{\beta, \gamma\}\}$. It's easy to see that this is indeed an ANC and that, in it, α, β are contradictories, — but then so are α, γ.

That is to say that contradictoriness is not necessarily a *functional relation*. Contradictoriness is, of course, symmetric, and it's easy to see it is irreflexive too.

Notice that, in the example above, β and γ are equivalent, and so both are negations of α (and, consequently, given the symmetry of contradictoriness, α is the sole negation of both β and γ). As it turns out, in ANCs this is always the case: namely, that all contradictories of a proposition are equivalent, — or, in other words, that contradictoriness boils down to negation.

Proposition 7.28. *In an ANC, all contradictories of a proposition are equivalent.*

Proof. Let Λ, Σ be an ANC in which φ has ψ and χ as contradictories, and suppose for a contradiction that ψ and χ are not equivalent. This means that there is a Φ such that either

$\{\psi\} \cup \Phi$ is consistent while $\{\chi\} \cup \Phi$ is not, or $\{\psi\} \cup \Phi$ is not consistent while $\{\chi\} \cup \Phi$ is. So, suppose the former. Then, because φ and χ are contradictories (therefore subcontraries), either $\{\psi\} \cup \Phi \cup \{\varphi\}$ is consistent or $\{\psi\} \cup \Phi \cup \{\chi\}$ is. But the former can't be the case, since φ is a contradictory of ψ (therefore also a contrary), and the latter can't be the case either, since, by supposition, $\{\chi\} \cup \Phi$ is inconsistent and it is contained in $\{\psi\} \cup \Phi \cup \{\chi\}$ (so that, by monotonicity, $\{\psi\} \cup \Phi \cup \{\chi\}$, too, would have to be inconsistent). If we now suppose that $\{\psi\} \cup \Phi$ is not consistent while $\{\chi\} \cup \Phi$ is, then, by a similar reasoning, we again arrive at a contraction. It follows that ψ is equivalent to χ after all; and, because ψ, χ are any contradictories of φ, all contradictories of φ are equivalent. □

Corollary 7.28.1. *In an ANC, if a proposition is a negation of another, then the latter is also a negation of the former. In other words: in an ANC, the law of double negation holds.*

Proof. Immediate from the proposition and from the fact that contradictoriness is a symmetric relation. □

Proposition 7.29. *In an ANC, a proposition is a tautology iff its negations are antilogies.*

Proof. Let φ be any.

First suppose φ is a tautology. Then $\Phi \cup \{\varphi\}$ is consistent for every consistent Φ. So suppose for a contradiction ψ a negation of φ is not an antilogy, so that $\{\psi\}$ is consistent. But then, because $\{\psi\}$ is consistent, $\{\psi\} \cup \{\varphi\}$ would have to be consistent, too, thus contradicting the assumption that ψ is a negation of φ. Therefore, ψ is an antilogy.

Now suppose ψ a negation of φ is an antilogy. Then, because φ, ψ are contradictories (therefore subcontraries), $\Phi \cup \{\varphi\}$ is consistent for every consistent Φ such that $\Phi \cup \{\psi\}$ is inconsistent, — and those are simply all consistent Φ, since $\Phi \cup \{\psi\}$ is inconsistent for whatever Φ, given that ψ is by assumption an antilogy. Therefore, φ is a tautology. □

Proposition 7.30. *In an ANC, a proposition is contingent iff so are its negations.*

Proof. Let φ be any.

Suppose φ is contingent. Then it is neither a tautology nor an antilogy, which means that neither $\Phi \cup \{\varphi\}$ is consistent for every consistent Φ, nor is $\{\varphi\}$ inconsistent. So let ψ a negation of φ be given, and let Φ be any consistent set such that $\Phi \cup \{\varphi\}$ is not consistent. Then, because φ, ψ are contradictories (therefore subcontraries), $\Phi \cup \{\psi\}$ is consistent, which entails (by the closure under subsets of consistency) that $\{\psi\}$ is consistent, and thus that ψ is not an antilogy. However, because $\{\varphi\}$ is consistent but $\{\varphi\} \cup \{\psi\}$ is not — since φ, ψ are contradictories (therefore contraries) —, $\Phi \cup \{\psi\}$ is not consistent for every Φ, which means that ψ is not a tautology either. Therefore, ψ is contingent as well.

The proof of the converse is similar.

Alternatively, the same result follows from the previous proposition and the fact that double negation holds in ANCs (corollary 7.28.1), by supposing for a contradiction that the negation of a contingent proposition is not itself contingent. □

Again, it is noteworthy that we have arrived at a rather classical concept of negation from what at first sight are just quite natural assumptions about consistency and negation. Of course, this is so simply because we are working with the intuitive concepts related to negation *within* an equally classical concept of consistency. But a paraconsistent logician might well ask: what if we kept the concept of negation intact but changed the underlying concept of consistency itself? — This is for sure an interesting line of inquiry (but one which, as usual, we shall not pursue here.)

Now for the concept of truth, we draw from a suggestion by Hartry Field [17]. As Field argues, it is generally agreed (exception made for dialetheists, *e.g.* [49]) that truth implies consistency. That being the case, I believe we may follow Stalnaker's proposal [63] and make use of the notion of *maximally consistent* sets of propositions to try to characterize truth in the framework of ANCs.

Stalnaker's intuition seems to be this. Assumed the propositions we have available suffice for capturing *all* the possibly relevant aspects of the world (what admittedly may be a rather strong assumption, depending on what we mean by "world" *cf.* [63] — but we shall go along with it anyway), if we were to add anything to the set of propositions which contains all that can be truthfully said, then, we would of necessity be adding a false proposition. Since, however, such a "truth set" contains all that can be truthfully said, for each false proposition it should then contain a proposition saying that the false proposition is not true. In Stalnaker's version of what this means — according to what he himself calls his "simple theory of propositions" [63], that is —, we would be thus requiring that each proposition in the ANC has a negation. Indeed, as we shall shortly see (proposition 7.33), provided that each proposition in an ANC has a negation and that each consistent set has a maximally consistent extension, a set is maximally consistent if, and only if, for each proposition it contains, either that proposition (and all equivalent propositions) or all of its negations are in the set, — just as expected by Stalnaker.

The suggestion is, thus, that we simply *identify* truth with some *designated* maximally consistent set of propositions in the ANC, — in the sense of a proposition being true if, and only if, it is in the designated set (or, what amounts to the same, if and only if it is in one of its subsets). Of course, in order for this identification to be possible, there need to *be* maximally consistent sets, and this is not the case in any ANC. (For a counterexample, take an arbitrary ANC in which the consistent sets form an infinitely ascending chain under set inclusion.)

Returning to Field's account of the relation between truth an consistency [17], it is worth noting that also argues that a sufficient condition for truth is *irrefutability*, — in the sense that a set of propositions, all of which are true, cannot imply anything false. Now, this presupposes an independent characterization of the concept of falsehood, and this can be provided in terms of contradiction in ANCs satisfying Stalnaker's requirements. The reasoning is natural enough: if truth is characterized in terms of a maximal set of propositions, then either a proposition or its negation should be in it; therefore, nothing can be added to it without it implying a contradiction (that is: both a proposition and its negation). It follows that ANCs capable of conveying the notion of truth as desired should behave somewhat classically in regards to contradiction. We spell out more precisely what we mean by this bellow.

These considerations lead us to the following "abstract" notion of truth. This concept is something of a weaker version of Stalnaker's aforementioned simple theory of proposition. It is weaker because we do not include in the concept what we feel are the two most controversial postulates of the theory of simple propositions [63]: namely, the postulate that for every set of propositions there is a singleton equiconsistent to it, and the postulate that equivalent propositions are identical.

Definition 7.7 (Abstract notion of truth)**.** An *abstract notion of truth* (ANT, for short) is an ANC Λ, Σ satisfying the following conditions, for all φ, Φ.

Negation existence φ has a negation

Extendibility If Φ is consistent, then it has a maximally consistent extension

It should be noted that ANT is indeed a concept distinct from an ANC, since both negation existence and extendibility can be violated by ANCs, as the next two rather trivial examples show.

Example 7.2. Let Λ, Σ be such that $\Lambda = \{\alpha, \beta, \gamma\}$ and $\Sigma = \{\varnothing, \{\alpha\}, \{\beta\}, \{\gamma\}$. Although, in particular, α, β are contraries in this ANC, they are not subcontraries, as there is a consistent set (namely, $\{\gamma\}$) such that both $\{\gamma\} \cup \{\alpha\}$ and $\{\gamma\} \cup \{\beta\}$ are inconsistent. The same is true, in fact, for every pair of propositions in this ANC. So, no two propositions are contradictories in it, and therefore no proposition has a negation in it.

Example 7.3. Let Λ, Σ be the ANC such that $\Lambda = \mathbb{N}$ and Σ is the set of all *finite* $\Phi \subseteq \Lambda$ such that for no n even both $n \in \Phi$ and $n + 1 \in \Phi$. Clearly, in this ANC there are no maximally consistent sets of propositions.

Observe that example 7.2 is an instance of an ANC satisfying extandibility but not negation existence, while example 7.3 is an instance of an ANC satisfying the latter but not the former. (It is easy to check that, for each n even, $n + 1$ is its negation, — from which it follows, by corollary 7.28.1, that, for each n odd, $n - 1$ is its negation.) The two conditions are, therefore, *logically independent*.

We now go on to show, with the aid of some accessory concepts, that maximally consistent sets in ANTs behave essentially as truth assignments to propositions, in the sense that the truth of a proposition can be identified, for all intents and purposes, with the proposition being in a designated maximally consistent set. Moreover, because negation also behaves as expected in ANTs, maximally consistent sets also properly convey the notion of falsity, as they all include either a proposition or its negations, but not both, for all propositions.

Consider, first, the following conditions an ANT Λ, Σ may satisfy.

Deductive non-contradiction For all φ, if $\Phi \Vdash \varphi$ and $\Phi \Vdash \psi$ for ψ a negation of φ, then Φ is not consistent

***Ex contradictione sequitur quodlibet* (ECSQ)** For all φ, if both $\varphi \in \Phi$ and also $\psi \in \Phi$ for ψ a negation of φ, then $\Phi \Vdash \chi$, where χ is any proposition whatsoever

Let an ANT Λ, Σ be fixed.

Proposition 7.31. *An ANT satisfies deductive non-contradiction.*

Proof. Let φ, Φ be any, and suppose both $\Phi \Vdash \varphi$ and $\Phi \Vdash \psi$ for ψ a negation of φ (whose existence is assured by negation existence). Then, because a set implies two other sets iff it implies their union (proposition 7.10), $\Phi \Rightarrow \{\varphi, \psi\}$, meaning $\Phi \cup \{\varphi, \psi\} \cup \Psi$ is consistent iff $\Phi \cup \Psi$ is, for all Ψ, and in particular for $\Psi = \varnothing$. Thus, because, by the definition of negation, $\Phi \cup \{\varphi, \psi\}$ is inconsistent, it follows that so is Φ. □

Proposition 7.32. *An ANT satisfies ECSQ.*

Proof. Immediate from the definition of negation and from fact that an inconsistent set entails every proposition (corollary 7.13.2). □

Consider, now, the following properties a set Φ of propositions in an ANT Λ, Σ may have.

Fullness Either $\varphi \in \Phi$ or $\psi \in \Phi$ for ψ a negation of φ, for all φ

Full consistency Φ is consistent and full

Proposition 7.33. *In an ANT, a set is maximally consistent iff it is fully consistent.*

Proof. Let Φ be any.

First suppose Φ is maximally consistent, and suppose for a contradiction it is not fully so. Then there's a φ such that neither $\varphi \in \Phi$ nor $\psi \in \Phi$ for ψ a negation of φ. Because, however, Φ is consistent, by the definition of negation (that is, because a proposition and its negation are contradictories, therefore subcontraries) it follows that either $\Phi \cup \{\varphi\}$ is consistent or $\Phi \cup \{\psi\}$ is. Suppose, without loss of generality, the former. Then, because $\varphi \in \Phi \cup \{\varphi\}$ but $\varphi \notin \Phi$, it follows $\Phi \cup \{\varphi\} \supset \Phi$, which contradicts the maximality of Φ. Therefore, Φ is fully consistent.

Now suppose Φ is fully consistent and suppose for a contradiction it is not maximally so. Because it is consistent, by extendibility it has a maximally consistent extension, Ψ; and, because it is not itself maximal, there's some $\varphi \in \Psi$ such that $\varphi \notin \Phi$. Let such φ be given. Now, because Φ is full and $\varphi \notin \Phi$, it follows $\psi \in \Phi$ for all negations of φ (of which there is at least one, by negation existence), so that $\psi \in \Psi$ for all such ψ too, since $\Phi \subseteq \Psi$ by supposition. But then $\varphi, \psi \in \Psi$ for ψ a negation of φ, which implies, by the definition of negation (that is, because a proposition and its negation are contradictories, therefore contraries), that Ψ is inconsistent, — thus contradicting the supposition that it is consistent. Therefore, Φ is maximally consistent. □

Finally, we show that in an ANT the concepts of tautology, antilogy and contingent proposition essentially coincide with how they are usually defined in terms of truth.

Proposition 7.34. *In an ANT, a proposition is a tautology iff it is in every maximally consistent set.*

Proof. Let φ be any.

First suppose φ is a tautology, and let Φ a maximally consistent set be any. Because Φ is consistent, by the definition of tautology, Φ has a compatible extension Ψ such that $\varphi \in \Psi$. Of course, because Φ is maximal, $\Psi = \Phi$, so that $\varphi \in \Phi$. Because Φ is any, φ is in every maximally consistent set.

Now suppose φ is in every maximally consistent set. Then, by extendibility, every consistent set has a compatible extension such that φ is in it, from which it follows φ is a tautology. □

Proposition 7.35. *In an ANT, a proposition is contingent iff it is in some maximally consistent sets but not in others.*

Proof. By definition, a proposition is contingent if it is neither a tautology nor an antilogy. By the previous proposition, then, a contingent proposition cannot be in every maximally consistent set, as it is not a tautology. Now, since it is also not an antilogy, there is at least one consistent set in which it is; therefore, by extendibility, also at least one maximally consistent set. □

Chapter 8

Consistency of DAs as an ANC

Now that we have a sufficiently developed the theory of logic in ANCs, we can get back to our original question of whether coherence of a DA is a properly *logical* concept of consistency in some sense of the word.

The basic idea of how we should proceed on the matter is, in earnest, quite simple: given a society I, S, we take all possible sequences of objects of the form xJy, where $x, y \in S$ and $J \in \wp(I)$, to be our propositions Λ. The intended meaning of such a proposition is the expected one: namely, that J *is decisive for* x *against* y. Sets of propositions such as this will thus describe purported DAs. Thus, given a DA D for I, S, proposition xJy will be *true* of D if xD_Jy, and *false* otherwise.

Now, if we were to follow the most natural course, we would perhaps want to take our set $\overline{\Sigma}$ to be that of all the sets of propositions which do not contain any set of the form: $\{x_kJ_ky_k\}_{k<n}$ for some $n > 0$, with $x_k \neq x_\ell$ for all $k < \ell < n$, and with $x_{k+1} = y_k$ for all $k < n-1$ and $y_{n-1} = x_0$, such that $\bigcap\{J_k\}_{k<n} = \varnothing$, as this is the way of describing an incoherence in a possible DA with our propositions. (We thus refer to such a set as *incoherent* or as an *incoherence* as well.)

Does this work as an ANC? — Well, it is clear that the empty set is not incoherent, so it must be consistent according to the definition. Moreover, if a set is not incoherent, then neither can one of its subsets be, so that the set of all coherent sets is closed under subsets, too. Therefore, this is indeed an ANC as expected.

This is the required answer to our initial question, to be sure; but it also certainly seems to be a very hasty one as it stands, and we would rather have a slightly more elaborate answer, if possible. After all, there seems to be an alternative intuition about how we should find the adequate ANC, and it is not immediately evident that the two coincide.

8.1 The logic of DAs

Suppose I were to try to communicate to you how power is distributed in certain society I, S (perhaps our own society), having as a language to do so only our set of propositions as described above. Suppose, that you *know* what I'm trying to do, having complete understanding of what the propositions I'm communicating to you mean. Because you know I'm trying to describe a DA that is *actually implemented* (therefore implementable), there is a lot you may assume about what I have to say (and what I *should not* say), even before I say anything at all. For one, it is obvious that I might say that xJx, for any $x, J \neq \varnothing$, as this is bound to be true of whatever possible DA I might try to describe, simply because any DA satisfy the condition of vacuousness. By the same token, you surely would be surprised if I were to say that $x\varnothing x$; and perhaps you would think I made a mistake if I happened to say that xJy and, not much later, added that $y\overline{J}x$ for $x \neq y$, as this would mean a violation of exclusiveness. If, however, I were to say that xJy and it were the case that $J' \supseteq J$, then I wouldn't even have to say that $xJ'y$, as this would also have to be true due to the fact that any implementable DA satisfies monotonicity.

The point I'm trying to make is that there seems to be an inherent *logic* to the concept of DA. It is a *material* logic in some sense, not a *formal* one, *cf.* [8], [53] (or, at least, not one directly involving the usual logical operators), but a logic all the same, with a lot of inferences inevitably going on in any speech about DAs. This logic is not hard to describe (indeed, I have just done so informally). Let us present it formally, then. (As far as I can tell, the kind of approach to logic that we are about to implement regarding DAs is essentially due to Kaye [33].)

Definition 8.1 (Logic of DAs)**.** Given a society I, S, let the set Λ of propositions be the set of all sequences of the form xJy such that $x, y \in S$ and $J \in \wp(I)$, and let any $\Phi \subseteq \Lambda$ and $\varphi \in \Lambda$ be given. A (DAs-)*deduction of* φ *from* Φ is a non-empty sequence of propositions, with φ being the last of the sequence, obeying the following rules, for any x, y, J, J', φ

1 If $\varphi \in \Phi$, then ϕ may occur at any point in the sequence

2 If $J \neq \varnothing$, then xJx may occur at any point in the sequence

3 If $x\varnothing x$ occurs at some point in the sequence, then φ may occur at any later point in the sequence

4 If xJy occurs at some point in the sequence and $J' \supseteq J$, then $xJ'y$ may occur at any later point in the sequence

5 If $x \neq y$ and xJy occurs at some point in the sequence and $y\overline{J}x$ occurs at some point in the sequence, then φ may occur at any later point in the sequence

The *logic of DAs* (LDAs, for short) consists of Λ together with the smallest relation $\vdash_{\text{DAs}} \subseteq \wp(\Lambda) \times \Lambda$ defined by: $\Phi \vdash_{\text{DAs}} \varphi$ iff there is a DAs-deduction of ϕ from Φ.

Since we are taking the inconsistent sets associated to a logic to be the *trivial* ones — the ones whose deductive closure is the set of all propositions —, it follows immediately from rules 4 and 5 that inconsistent sets associated to the logic of DAs are those entailing $x\varnothing x$ for some x or, equivalently, the ones entailing both xJy and $y\overline{J}x$ for some $x \neq y$ (or those entailing both xJy and $y\overline{J}x$ for some $x \neq y$ and J, J' disjoint, given rule 3).

But this, of course, does not explain very much. What we want to know is whether the LDAs actually performs as expected regarding DAs.

In order to show this, we now prove that the LDAs adequately captures its intended interpretation. As usual in logic, this is done by means of *soundness* and *completeness theorems* [15], [33].

Given Φ, φ, we write $\Phi \vDash_{\text{DAs}} \varphi$ to mean that φ is true of every DA of which every proposition in Φ is true. To say that the LDAs is *sound* (relative to its intended interpretation) means that $\Phi \vdash_{\text{DAs}} \varphi$ entails $\Phi \vDash_{\text{DAs}} \varphi$ for every Φ, φ, — that is: that, if there is a deduction of φ from Φ, then the inference of φ from Φ is *valid*. To say that the LDAs is *complete* (again, relative to its intended interpretation) means the converse of this.

Proposition 8.1 (Soundness theorem)**.** *The LDAs is sound relative to its intended interpretation.*

Proof. By induction on the length of deductions.

Let Φ, φ be any such that $\Phi \vdash_{\text{DAs}} \varphi$.

If φ has been obtained from Φ by means of the application of a single rule (that is: in a single step), then it must have been either rule 1 or rule 2. If it has been obtained by means of rule 1, then $\varphi \in \Phi$, which immediately entails that $\Phi \vDash_{\text{DAs}} \varphi$; if by means of rule 2, then $\varphi = xJx$ for some x and some $J \neq \varnothing$, and xJx is true of any DA whatsoever, by the vacuousness property of DAs, *cf.* section 3.3.

So suppose that inferences in the LDAs are valid for deductions of length up to k, and let φ have been obtained from Φ by means of a deduction of length $k + 1$.

If φ has been obtained by means of either an application of rule 1 or of rule 2, then $\Phi \vDash_{\text{DAs}} \varphi$ by the same reasons as before.

If it has been obtained by means of an application of rule 3, then $x\varnothing x$ for some x must have occurred earlier in the deduction. Because, again by the vacuousness property of DAs, $x\varnothing x$ is true of no DA whatsoever, if $\Phi \vDash_{\text{DAs}} x\varnothing x$, then, vacuously, $\Phi \vDash_{\text{DAs}} \varphi$.

If it has been obtained by means of an application of rule 4, then it is of the form xJy, and $xJ'y$ for $J' \subseteq J$ must have occurred earlier in the deduction, which entails by supposition that $\Phi \vDash_{\text{DAs}} xJ'y$. Thus, by the monotonicity property of DAs, *cf.* section 3.3, it follows that $\Phi \vDash_{\text{DAs}} \varphi$ as well.

Finally, if it has been obtained by means of an application of rule 5, then both xJy and $y\overline{J}x$ for some x, y, J must have occurred earlier in the deduction, which entails by supposition that both $\Phi \vDash_{\text{DAs}} xJy$ and $\Phi \vDash_{\text{DAs}} y\overline{J}x$. Since, by the exclusiveness property of DAs, *cf.* section 3.3, this is impossible, again, vacuously, $\Phi \vDash_{\text{DAs}} \varphi$. □

Now that we know the LDAs performs as expected, we want to directly characterize the consistent sets associated to it. For this purpose, use will have to be made of the fact that the LDAs is *compact* [15],, or *finitary* [9] — meaning that, if $\Phi \vdash_{\text{DAs}} \varphi$, then there is a finite $\Phi' \subseteq \Phi$ such that $\Phi' \vdash_{\text{DAs}} \varphi$.

Proposition 8.2. *The LDAs is compact.*

Proof. Follows immediately from the facts that DAs-deductions are finite sequences and that its deduction rules use only finitely many *premises* (that is: propositions in the given set.) □

Proposition 8.3. *A set of propositions is non-trivial in the LDAs iff it neither contains a proposition saying that the empty coalition is decisive for an option against that same option, nor does it contain both a proposition saying that a coalition is decisive for an option against another, distinct option, and a proposition saying that a coalition disjoint from that one is decisive for the latter option against the former.*

Proof. Let Φ be a set of propositions.

Let x, y and J, J' disjoint be any. As already observed, it follows immediately from rules 3, 4, and 5 that, if either $x\varnothing x \in \Phi$ for some x or, for x, y distinct and J, J' disjoint, $xJy, yJ'x \in \Phi$, then Φ is trivial (that is: $\Phi \vdash_{\text{DAs}} \varphi$ for all φ).

For the converse, suppose that $x\varnothing x \notin \Phi$, for all x, and that, for all $x \neq y$ and J, J' disjoint, either $xJy \notin \Phi$ or $xJ'y \notin \Phi$; but suppose for a contradiction that, nonetheless, $\Phi \vdash c\varnothing c$ for some fixed c (which is equivalent to Φ being trivial, given rule 3). Then, there is DAs-deduction of $c\varnothing c$ from Φ.

Let such a DAs-deduction be given. Because DAs-deductions are by definition finite sequences, assume without loss of generality that the given DAs-deduction of $c\varnothing c$ from Φ is the shortest (or one of the shortest, if there is more than one of the same length). Close inspection of the deduction rules of the LDAs reveal that either rule 1, rule 3 or rule 5 must have been the last one applied in the deduction leading to $c\varnothing c$; and, by supposition, 1 cannot have been applied. So, either rule 3 or rule 5 must have been applied.

Suppose that rule 3 has been applied, which means that $x\varnothing x$ for some x has occurred at an earlier point in the sequence. Clearly, $x\varnothing x$ cannot have itself been obtained by an application of rule 3, for otherwise there would be a shorter deduction of $c\varnothing c$. Thus, it must have been obtained by an application of either rule 1 or rule 5; and, by supposition, rule 1 cannot have been applied.

So suppose that rule 5 has been applied in order to obtain $x\varnothing x$, which means that yJz and $z\overline{J}y$ for some $y \neq z$ and $J \neq \varnothing$ have occurred at earlier points in the sequence. Without loss of generality, assume yJz has been the latest to occur. Again, inspection of the rules reveal that either rule 1, rule 3, rule 4 or rule 5 must have been the last one applied in the deduction leading to yJz. It is easy to see that we would have a shorter deduction of $x\varnothing x$ from Φ if either rule 3 or rule 5 had been the last one to be applied in the deduction leading to yJz. Therefore, either rule 1 or rule 4 must have been applied.

If rule 1 has been applied, then it follows from our initial supposition that $zJ'y \notin \Phi$, for any J' disjoint from J. Now, because this excludes $z\overline{J}y$ having been obtained by an application of rule 1, and because, by the same reasons why yJz couldn't, neither can it have been obtained by an application of either rule 3 or rule 5, $z\overline{J}y$ can only have been obtained by an application of rule 4.

If instead rule 5 had been originally applied in order to obtain $c\varnothing c$, by a similar line of reasoning we would conclude that rule 4 would have been applied at some point in the deduction. So let us investigate what application of rule 4 at any point in the deduction entails.

For the sake of simplicity, let us assume without loss of generality that $z\overline{J}y$ has been obtained by an application of rule 4, which means that $zJ'y$ for some $J' \subseteq \overline{J}$ must have occurred earlier in the sequence. Now, it can either have occurred later or earlier then yJz in the sequence.

Suppose, first, that $zJ'y$ has occurred later than yJz in the sequence. Then it could only have been itself obtained by application of rule 4, as application of the other rules are all by now excluded. But then $z\overline{J}y$ could have been obtained earlier by an application of rule 4, contradicting the supposition that the given deduction is one of the shortest. So no such DAs-deduction of $c\varnothing c$ from Φ can exist.

If we now suppose that $zJ'y$ has occurred earlier than yJz, then, by a reasoning similar to the one we have conduced regarding $y\overline{J}x$, we discover that we can find a shorter deduction involving an application of rule 5 to $zJ'y$ and $y\overline{J'}z$ instead, — so that, again, no such DAs-deduction of $c\varnothing c$ from Φ can exist.

But then no DAs-deduction of $c\varnothing c$ from Φ exists at all, thus contradicting the supposition that one does. Therefore, if $x\varnothing x \notin \Phi$ for all x and, for all $x \neq y$ and J, J' disjoint, either $xJy \notin \Phi$ or $yJ'x \notin \Phi$, then Φ is non-trivial. □

The last proposition allows us to directly characterize the consistent sets associated to the LDAs as follows. Formally, the ANC associated to the LDAs is the pair $\Lambda, \Sigma_{\text{DA}}$ such that $\Sigma_{\text{DA}} = \{\Phi \in \wp(\Lambda)$: neither $x\varnothing x \in \Phi$ for some x nor both $xJy \in \Phi$ and $yJ'x \in \Phi$ for some $x \neq y$ and J, J' disjoint$\}$.

We now show that the entailment relation $\Vdash_{\text{DAs}}$ of the ANC so defined coincides with the consequence relation of the LDAs. This will be instrumental in proving the completeness of the LDAs relative to its intended interpretation.

Proposition 8.4. *The consequence relation of the LDAs is the same as the entailment relation of the ANC associated to the LDAs.*

Proof. Let Γ, φ be any.

For the implication from left to right, we again proceed by induction on the length of deductions.

So suppose that $\Phi \vdash_{\text{DAs}} \varphi$.

First of all it should be noted that, if Φ is inconsistent, then $\Phi \Vdash_{\text{DAs}} \varphi$ by corollary 7.13.2. Therefore, in general we need only consider the cases in which Φ is consistent. Moreover, notice that, in proving $\Phi \Vdash_{\text{DAs}} \varphi$ (that is: that Φ is equiconsistent to $\Phi \cup \{\varphi\}$), we only need to show that, for all Ψ, if $\Phi \cup \Psi$ is consistent, then so is $\Phi \cup \{\varphi\} \cup \Psi$, since the converse follows immediately from the closure under subsets of the consistent sets.

So, if φ has been obtained from Φ by means of the application of a single rule (that is: in a single step), then it must have been either rule 1 or rule 2. If it has been obtained by means of rule 1, then $\varphi \in \Phi$, which immediately entails that $\Phi \Vdash_{\text{DAs}} \varphi$, since, in this case, $\Phi \cup \{\varphi\} = \Phi$ and Φ is equiconsistent to itself. If it has been obtained by means of rule 2, then $\varphi = xJx$ for some x and some $J \neq \varnothing$, and thus, for whatever Ψ, if $\Phi \cup \Psi$ is consistent, then by definition so is $\Phi \cup \{\varphi\} \cup \Psi$.

If φ has been obtained by means of an application of rule 3, then by supposition $\Phi \Vdash_{\text{DAs}} x\varnothing x$ for some x, which entails that Φ is inconsistent and thus that $\Phi \Vdash_{\text{DAs}} \varphi$, by corollary 7.13.2.

Similarly, if φ has been obtained by means of an application of rule 5, then by supposition $\Phi \Vdash_{\text{DAs}} xJy$ and $x\overline{J}$ for some $x \neq y$ and J, which entails that Φ is equiconsistent to $\Phi \cup \{xJy, y\overline{J}x\}$ (by proposition 7.10), and thus that $\Phi \Vdash_{\text{DAs}} \varphi$, by corollary 7.13.2.

As for rule 4, given x, y, J, it suffices to notice that, if $\Phi \cup \{xJy\} \cup \Psi$ is consistent for a given Ψ, then of necessity so is $\Phi \cup \{xJ'y\} \cup \Psi$ for $J' \supseteq J$.

For the converse, suppose $\Phi \nvdash_{\text{DAs}} \varphi$. This implies that Φ is consistent, since it is non-trivial. So suppose that nevertheless $\Phi \Vdash_{\text{DAs}} \varphi$. Because Φ is consistent, it can't be the case $\varphi = x\varnothing x$ for some x; nor can it be the case that $\varphi = xJx$ for some x and some $J \neq \varnothing$, as $\Phi \nvdash_{\text{DAs}} \varphi$ and, by rule 2, $\Phi \vdash_{\text{DAs}} \varphi$ for all x and all $J \neq \varnothing$. It follows that $\varphi = xJy$ for some $x \neq y$ and J. Let such x, y, J be given.

Because $\Phi \Vdash_{\text{DAs}} \varphi$ (that is, $\Phi \Vdash_{\text{DAs}} xJy$), it follows that $\Phi \cup \Psi$ is equiconsistent to $\Phi \cup \{xJy\} \cup \Psi$ for whatever Ψ; in particular, this must be the case for $\Psi = \{y\overline{J}x\}$. However, $\Phi \cup \{xJy\} \cup \{y\overline{J}x\}$ is inconsistent, which entails that $\Phi \cup \{y\overline{J}x\}$ is inconsistent as well. This, on its turn, entails that $xJ'y \in \Phi \cup \{y\overline{J}x\}$ for some J' disjoint from $\overline{J}$. More precisely, since $x \neq y$ (and, therefore, $\varphi \neq y\overline{J}x$), it must be the case that $xJ'y \in \Phi$ for such J'. But then, since J' is disjoint from $\overline{J}$, of necessity $J' \subseteq J$. Therefore, because $xJ'y \in \Phi$, by rules 1 and 4, it follows that $\Phi \vdash_{\text{DAs}} xJy$ (that is, $\Phi \vdash_{\text{DAs}} \varphi$), thus contradicting the supposition. □

In the next proposition, what we mean by the set of propositions *corresponding* to a DA D is the set $\Delta(D) = \{\varphi \in \Lambda : \varphi = xJy \text{ for some } x, y, J \text{ and } xD_Jy\}$.

Proposition 8.5. *A DA is faithfully implementable by a PAF satisfying UD iff it corresponds to the deductive closure of some non-trivial set in the LDAs.*

Proof. Recall that a DA satisfying UD is faithfully implementable iff it satisfies the conditions of vacuousness, monotonicity and exclusiveness, *cf.* section 3.3. So, let such a DA D be given.

First of all notice that, because D is faithfully implementable, neither $xD_{\varnothing}x$ for some x nor both $xD_{J}y$ and $yD_{J'}x$ for some $x \neq y$ and some J, J' disjoint, so that $\Delta(D)$ must be non-trivial by the previous proposition. Now let φ be any such that $\Delta(D) \vdash_{\text{DAs}} \varphi$. Naturally, $\varphi = xJy$ for some x, y and some J. So let those be given, and let a DAs-deduction of xJy from $\Delta(D)$ be given. Because $\Delta(D)$ is non-trivial, no deduction of xJy from $\Delta(D)$ can involve rules 3 and 5, as we haven seem in the proof of the previous proposition. If the last rule to be applied in the deduction has been rule 1, then $xJy \in \Delta(D)$ follows immediately. If rule 2, then $x = y$ and $J \neq \varnothing$: and, thus, because D satisfies vacuousness, $xD_{J}y$ (that is: $xD_{J}x$), which implies by the definition of $\Delta(D)$ that $xJy \in \Delta(D)$. If rule 4, then there is a $J' \subseteq J$ such that $xJ'y$ has occurred at an earlier point in the sequence. Due to the fact that DAs-deductions are finite sequences, we can't go on forever finding applications of rule 4. So there must be a smallest J' such that $xJ'y$ was obtained by an application of either of rule 1 or of rule 2. Whatever the case, $xJ'y \in \Delta(D)$, from which it follows by the definition of $\Delta(D)$ that $xD_{J'}y$; and, because D satisfies monotonicity, also $xD_{J}y$, which again by the definition of $\Delta(D)$ implies that $xJy \in \Delta(D)$. In any case, if $\Delta(D) \vdash_{\text{DAs}} \varphi$ then $\varphi \in \Delta(D)$, which means that $\Delta(D)$ is deductively closed. Therefore, there is a non-trivial set of propositions to whose deductive closure D corresponds.

Conversely, let Φ non-trivial be given, and assume $\Delta(D)$ is its deductive closure. Because Φ is non-trivial, there's some $\varphi \notin \Delta(D)$. In particular, neither $xD_{\varnothing}x$ for some x, nor both xJy and $yJ'x$ for some $x \neq y$ and J, J' disjoint are in $\Delta(D)$, for otherwise the deductive closure of $\Delta(D)$ would be the whole of Λ, and therefore also $\Delta(D)$ would be trivial (and, consequently, so would Φ), because the LDAs is Tarskian. That being noted, then it's just a simple matter of checking that D satisfies the desired conditions. □

This brief exercise in logic shows that implementability by a PAF with unrestricted domain is a legitimate notion of consistency, in the sense that there is an ANC whose consistent sets correspond to implementable DAs: namely, the ANC associated to the LDAs. In this ANC, all sets of propositions having the same deductive closure can be taken to be distinct representations of the same DA, and every implementable DA has at least one consistent set of propositions representing it. To be more precise, in view of the previous proposition let us say that a set in this ANC *represents* a DA if its deductive closure corresponds to it. Then the following corollary neatly sums up what this truly means.

Corollary 8.5.1. *Two sets of propositions in the ANC associated to the LDAs represent the same DA iff they are equiconsistent.*

Proof. Follows immediately from the previous proposition and from the fact, already proved, that in a deductive ANC two sets are equiconsistent iff they have the same deductive closure. □

In sum: the implementable DAs are the DAs associated to the congruence classes of consistent sets of propositions in this ANC (congruent under the equiconsistency relation, that is). Conversely, two sets are equiconsistent in the ANC associated to the LDAs if, and only if, they represent the same object, — so that, in a sense, equiconsistency boils down to an expression of identity in this case.

An immediate consequence of this fact is the completeness theorem for the LDAs.

Corollary 8.5.2 (Completeness theorem)**.** *The LDAs is complete relative to its intended interpretation.*

Proof. Let Φ, φ be any and suppose that $\Phi \nvdash_{\text{DAs}} \varphi$. Then, by the proposition, there is a DA D such that all and only the propositions in the deductive closure of Φ are true of D. Because $\Phi \nvdash_{\text{DAs}} \varphi$ — so that φ is not in the deductive closure of Φ —, all propositions in Φ are true of D, but φ is not. This entails $\Phi \nvDash_{\text{DAs}} \varphi$, as desired. □

8.2 The logic of coherent DAs

That's all well and good, of course, — but what about our original goal, the logical characterization of coherent DAs?

As it turns out, in order to reach that goal one needs just to add a rule of deduction to the LDAs. More precisely, let us define the logic of coherent DAs for a given society.

Definition 8.2 (Logic of coherent DAs)**.** Given a society I, S, let the set Λ of propositions be, as before, the set of all sequences of the form xJy such that $x, y \in S$ and $J \in \wp(I)$, and let any $\Phi \subseteq \Lambda$ and $\varphi \in \Lambda$ be given. A (CDAs-)*deduction of* φ *from* Φ is a non-empty sequence of propositions, with φ being the last of the sequence, obeying the same rules as DAs-deductions, and one additional rule:

6 If, for some $n > 0$, $x_0J_0y_0, \ldots$ and $x_{n-1}J_{n-1}y_{n-1}$, with $x_k \neq x_\ell$ for all $k < \ell < n$, and with $x_{k+1} = y_k$ for all $k < n-1$ and $y_{n-1} = x_0$, each occur at some point in the sequence, and $\bigcap\{J_k\}_{k<n} = \varnothing$, then φ may occur at any latter point in the sequence, for any φ

The *logic of coherent DAs* (LCDAs, for short) consists of Λ together with the smallest relation $\vdash_{\text{CDAs}} \subseteq \wp(\Lambda) \times \Lambda$ defined by: $\Phi \vdash_{\text{CDAs}} \varphi$ iff there is a CDAs-deduction of ϕ from Φ.

It should be noted that rule 6 actually supersedes rules 3 and 5, as they cover the same cases as rule 6 with $n = 1$ and (in the presence of rule 4) $n = 2$, respectively.

The next four propositions are proven in just the same way as their counterparts for LDAs, except that checking for the cases involving rule 6 under the assumption that the set of premises is not incoherent is actually slightly less laborious (and by no means technically more interesting). For that reason alone, I omit the proofs.

Proposition 8.6. *The LCDAs is compact.*

Proposition 8.7. *A set of propositions is non-trivial in the LCDAs iff it is not incoherent.*

Proposition 8.8. *A coherent DA is faithfully implementable by an SDF satisfying UD iff it corresponds to the deductive closure of some non-trivial set in the LCDAs.*

Proposition 8.9. *Two sets of propositions in the ANC associated to the LCDAs represent the same DA iff they are equiconsistent.*

To close this section, I prove two more propositions, mostly for the sake of completing our exploration of the ANC associated to the LCDAs in the context of the framework developed in the previous chapter.

Proposition 8.10. *In the ANC associated to the LCDAs, every consistent set has a maximally consistent extension.*

Proof. We provide a proof only for the simpler case of Λ being at most countable. The proof for Λ uncountable wouldn't be much different, but it would require the axiom of choice. In any case, the proof is very standard and straightforward.

Let Φ a consistent set of propositions in the ANC associated to the LCDAs be any. Let an enumeration of the propositions in Λ be given. We proceed by constructing a maximally consistent set extending Φ as follows.

Let $\Phi_0 = \Phi$ and, for each $n > 0$, let $\Phi_{n+1} = \Phi_n \cup \{\varphi_n\}$ if $\Phi_n \cup \{\varphi_n\}$ is consistent, and $\Phi_{n+1} = \Phi_n$ otherwise.

Clearly, $\bigcup\{\Phi_n\}_{n\in\mathbb{N}}$ is an extension of Φ by design. We now prove that is consistent.

For suppose there is an incoherence in $\bigcup\{\Phi_n\}_{n\in\mathbb{N}}$. Let $\{x_k J_k y_k\}_{k<m}$, with $m > 0$, be that incoherence. Then, because $\{x_k J_k y_k\}_{k<m}$ is finite, there is a smallest $\ell \in \mathbb{N}$ such that $\{x_k J_k y_k\}_{k<m} \subseteq \Phi_n$, which entails by definition that Φ_n is inconsistent. But that is impossible, because $\Phi_0 = \Phi$ is consistent and, for each $n > 0$, if Φ_{n-1} is consistent, then so is Φ_n, from which it follows by induction that Φ_ℓ must itself be consistent. Therefore, $\bigcup\{\Phi_n\}_{n\in\mathbb{N}}$ is indeed consistent.

It remains to be checked whether $\bigcup\{\Phi_n\}_{n\in\mathbb{N}}$ is maximal.

So let φ be any proposition not in $\bigcup\{\Phi_n\}_{n\in\mathbb{N}}$. By construction, there must an ℓ such that $\varphi = \varphi_\ell$. Let such ℓ be given. Then, because $\varphi \notin \bigcup\{\Phi_n\}_{n\in\mathbb{N}}$, it follows that $\Phi_\ell \cup \{\varphi\}$ is inconsistent, meaning that there is an incoherence $\{x_k J_k y_k\}_{k<m} \subseteq \Phi_\ell \cup \{\varphi\}$. Therefore, because $\Phi_\ell \cup \{\varphi\} \subseteq \bigcup\{\Phi_n\}_{n\in\mathbb{N}} \cup \{\varphi\}$, also $\{x_k J_k y_k\}_{k<m} \subseteq \bigcup\{\Phi_n\}_{n\in\mathbb{N}} \cup \{\varphi\}$, which implies that $\bigcup\{\Phi_n\}_{n\in\mathbb{N}} \cup \{\varphi\}$ is inconsistent as well. Because φ is any proposition not in $\bigcup\{\Phi_n\}_{n\in\mathbb{N}}$, it follows that $\bigcup\{\Phi_n\}_{n\in\mathbb{N}}$ has no proper compatible extensions, thus being maximally consistent after all. □

Proposition 8.11. *If the society has at most two options, then the ANC associated to the LCDAs is an ANT (cf. section 7.3).*

Proof. Extendibility follows immediately from the previous proposition, so that it only remains to be show that the ANC also satisfies negation existence.

Let us first observe that, in the ANC associated to the LCDAs, all proposition of the form xJx are antilogies if $J = \varnothing$ and tautologies if $J \neq \varnothing$. Indeed, it is easy to see that those are also all of the tautologies and antilogies there are. If, therefore, the society has only one option, as a matter of fact those are all the propositions there are, and we know (proposition 7.29) that tautologies and antilogies are each other negations, — from what it follows that the ANC satisfies negation existence.

So let the society have two options. Of course, the same still applies regarding tautologies and antilogies. As for contingent propositions, let xJy, where $x \neq y$, be any contingent one. We show that $y\overline{J}x$ is a contradictory of xJy, from which it follows (by proposition 7.28) that they are each others' negations.

In order to do so, we begin by first observing that there are only two possible types of incoherences involving only two options x, y: incoherences of the form $\{x_k J_k y_k\}_{k<2}$, where $x_0 = x, y_0 = y$ and $\bigcap\{J_k\}_{k<2} = \varnothing$, and incoherences of the form $\{x_k J_k y_k\}_{k<2}$, where $x_0 = y, y_0 = x$ and $\bigcap\{J_k\}_{k<2} = \varnothing$. (Notice that, in both cases, the values of x_1, y_1 are settled by the definition of incoherence.)

So let Φ be any and suppose for a contradiction that it is consistent but neither $\Phi\cup\{xJy\}$ nor $\Phi\cup\{y\overline{J}x\}$ are. Because Φ is consistent but $\Phi\cup\{xJy\}$ is not, there is an H disjoint from J such that $yHx \in \Phi$ and $\{xJy, yHx\}$ is an incoherence. Similarly, because Φ is consistent but $\Phi \cup \{y\overline{J}x\}$ is not, there is an H' disjoint from $\overline{J}$ such that $xHy \in \Phi$ and $\{y\overline{J}x, xH'y\}$ is an incoherence. However, because H is disjoint from J, it follows that $H \subseteq \overline{J}$; and similarly, because H' is disjoint from $\overline{J}$, that $H' \subseteq J$. But then H, H' are themselves disjoint, — from which it follows that $\{yHx, xH'y\}$ is an incoherence. But then $\{yHx, xH'y\} \subseteq \Phi$, thus contradicting the supposition that Φ is consistent. This entails that, if Φ is consistent, then either $\Phi \cup \{xJy\}$ or $\Phi \cup \{y\overline{J}x\}$ is, too.

This, of course, amounts to $xJy, y\overline{J}x$ being subcontraries; and, since they are also contraries, they indeed are contradictories as expected. Therefore, the ANC associated to the LCDAs when society has at most two options satisfies negation existence and is an ANT. □

Corollary 8.11.1. *The ANC associated to the LDAs is an ANT.*

Proof. Essentially the same, except that the assumption that the society has at most two options is not required. □

Proposition 8.12. *If the society has at least three options, then no contingent proposition has a contradictory in the ANC associated to the LCDAs.*

Proof. Let us first observe that, if $x \neq y$, then xJy is a contingent proposition, for whatever x, y, J; and those are all contingent propositions that are, since xJx, for whatever x, is either an antilogy (if $J = \varnothing$) or a tautology (if $J \neq \varnothing$). Let us also observe that all contraries of xJy, for whatever x, y, J, are of necessity of the form $yJ'x$, where J' is disjoint from J. This is so because for no proposition of the form $zJ'w$, where $z \neq y$ or $w \neq x$, is $\{xJy, zJ'w\}$ inconsistent, nor is $\{xJy, yJ'x\}$ if J, J' are not disjoint; but all propositions are of one of those three forms. So suppose xJy has a contradictory, and let $xJ'y$ (for some given J' disjoint from J) be it. Now, let $z \neq x$, $z \neq y$ be any, and consider the set of propositions $\Phi = \{xJz, zJx, y\overline{J}z, z\overline{J}y\}$. It is easy to see that this set is coherent (therefore consistent). However, $\Phi \cup \{xJy\}$ is clearly incoherent (therefore inconsistent), and so is $\Phi \cup \{yJ'x\}$, given that, by assumption, $J' \subseteq \overline{J}$. Therefore, because every extension Ψ of Φ such that $xJy \in \Psi$ is also an extension of $\Phi \cup \{xJy\}$, and, similarly, every extension Ψ of Φ such that $yJ'x \in \Psi$ is also an extension of $\Phi \cup \{yJ'x\}$, any such extension of Φ is inconsistent. This being the case, then for every compatible extension of Φ, neither xJy nor $yJ'x$ is in it. It follows that xJy and $yJ'x$ are not subcontraries, — and, thus, nor are they contradictories. In sum, our supposition that xJy has a contradictory is false; and, because xJy is any contingent proposition, the conclusion follows. □

Chapter 9

Discussion and conclusions

Regarding the concept of ANC, and particularly in its application to the logical nature of decisiveness assignments, there are two issues still worth addressing.

The first issue is that, perhaps contrary to expectations, in all ANCs we have discussed the concrete structures of interest correspond to deductively closed sets, and not necessarily to maximally consistent ones. At first sight, this may seem to be at odds with the concept of an abstract notion of truth, which intuitively takes maximally consistent sets to be the ones corresponding to concrete structures — or, as we have originally put it, to "the world", *cf.* section 7.3. This discrepancy however is, in part at least, a consequence of the fact that none of those ANCs — be it the one related to implementable DAs, coherent DAs, or to strongly transitive relations — has enough expressive resources to capture all the possibly relevant aspects of those structures. In particular, the ANCs related to DAs don't have the resources to *directly* say that a coalition is *not* decisive on an issue x, y.

Of course, those ANCs still have the resources to *exclude* the fact of a coalition being decisive on an issue; but that can be done only at the cost of stating other *positive* facts (such as, in particular, the fact of the complementary coalition being decisive on the issue y, x). In other words, it is to say that the concepts of contradictoriness and negation that we have borrowed from Stalnaker's theory of simple propositions is actually inadequate for the intended purposes of that very same theory (at least if not supplemented by other assumptions; but we won't investigate that possibility here, as we find no natural assumptions in sight).

This realization, however, is perhaps just as it is supposed to be. Perhaps the assumption that truth sets are to be maximally consistent is itself inadequate. It can't be denied that the set of all true propositions ought to be deductively closed, if entailment is to preserve truth; for, otherwise, there would be true propositions not in it. Therefore, the collection of all truth sets in a given ANC ought indeed to be a subcollection of all deductively closed sets. But there is no obvious reason why the collection of truth sets would have to comprise only maximally consistent sets, and our examples of ANCs for concrete structures in fact imply otherwise. Again, we won't pursue this line of inquiry any further, as this is not our subject in this doctoral dissertation. But what we have discussed so far suggests that the postulate of extendibility of ANTs ought to be weakened to the requirement that every consistent set has a deductively closed compatible extension (and not necessarily a maximal one).

The second issue, which may be of some interest to logicians, regards the interpretation of negation in the ANCs related to implementable and coherent DAs. As we have seen, while in the ANC associated to the LDAs negations exist for every proposition (corollary 8.11.1), that is not in general the case in the ANC associated to the LCDAs (proposition 8.12). Thus, alternatively, we could still interpret xJy and $y\overline{J}x$, for every x, y, J as each other's negations in the LCDAs, and instead think of the LCDAs as a *paracomplete* logic, *cf.* [9], — given that, in it, xJy and $y\overline{J}x$ are not subcontraries anymore. From a logical point of view, it is interesting to observe that, under this interpretation, by strengthening the coherence requirement of DAs (and, consequently, the notion of consistency of their related ANCs) we have consequently weakened the notion of negation of a logic.

Part III

Appendices

Chapter 10

Mathematical background

In this chapter I briefly present the mathematical background necessary for understanding this dissertation. In the first section I set some terms and conventions regarding the basic set-theoretic concepts that are constantly required throughout the text. In the second section I present and discuss the concept of subset space, and in particular the kind of subset space known as ultrafilter, which plays a major analytic role in understanding the logical implications of Arrow's impossibility theorem.

Some of the concepts here presented will be discussed again in much greater detail in the next chapter. Because they happen to be presupposed by some of the concepts that are the proper subject of this chapter, this redundancy is unfortunately unavoidable. In any event, I try to keep it to a minimum.

10.1 Set theory

I assume the reader is at least acquainted with some basic mathematical logic [10], [40] and the usual notation and axioms of basic set theory, be they those of ZFC [31], [37], [47] or NBG [19]. In particular, I assume the reader is acquainted with the logical notion of *identity* (notation: $=$) and with the set-theoretic primitive notion of *being in* a class, set or collection (notation: $\in$). In a set theory such as NBG, which distinguishes classes from sets, a class is said to be *proper* if it is not in some other class, and thus by a *set* what is meant is a class which is not proper. I follow that usage here. Moreover, by a *collection* I mean simply a class of sets.

I assume the reader is used to the *class notation* for the subclass of X consisting of all elements of X satisfying sentence φ (that is: the expression $\{x \in X : \varphi(x)\}$, where x occurs free in φ) and its variants. I use the class notation for a set without specifying a class of which it is a subclass (that is: in the form $\{x : \varphi(x)\}$) only when the existence of the set so defined is deducible from the axioms of basic set theory.

I also assume the reader is acquainted with the general theory of ordinal and cardinal numbers. I write $\mathbb{N}$ for the set of *finite ordinals*, or *natural numbers*. Given $n, m \in \mathbb{N}$, I write

$n \leq m$ to mean that $n \in m$ or $n = m$, and $n < m$ to mean that $n \in m$. Given a set X, I write $|X|$ for the *cardinal* of X.

As a matter of notational convention, I routinely make use of the *slash notation* for negation of sentences involving binary predicates. For example, I write $x \neq y$ to mean that not $x = y$, and $x \notin X$ to mean that not $x \in X$. Given a binary predicate P and $t_0, \ldots, t_n$ terms, I routinely write $t_0, \ldots, t_{n-2} P t_n$ to mean that $t_0 P t_n, \ldots$ and $t_{n-1} P t_n$.

It should be noted that, in some occasions, I write symbols for binary predicates inverted (for example, $X \ni x$ to mean the same as $x \in X$, — to be read as "X has x in it", as opposed to "x is in X"). I do so mostly because in such occasions I find the inverted one would be the expected ordering of the corresponding expressions in natural language as well.

10.1.1 Sets and operations on sets

Let X, Y be sets. X *is a subset of* Y (notation: $X \subseteq Y$) if, for all x, if $x \in X$, then $x \in Y$. If $X \subseteq Y$ but $Y \nsubseteq X$, then X is a *proper* subset of Y (notation: $X \subset Y$). The *power set of* X (notation: $\wp(X)$) is the set $\{Y : Y \subseteq X\}$.

The *empty set* (notation: $\varnothing$) is the set $\{x : x \neq x\}$, — that is: the set X such that no $x \in X$.

Let X, Y be sets. The *binary intersection of* X *and* Y (notation: $X \cap Y$) is the set $\{x : x \in X$ and $x \in Y\}$. The *binary union of* X *and* Y (notation: $X \cup Y$) is the set $\{x :$ either $x \in X$ or $x \in Y\}$. The *difference of* X *and* Y (notation: $X \setminus Y$) is the set $\{x : x \in X$ but $x \notin Y\}$. Let $X \subseteq Y$. The *complement of* X *relative to* Y (notation, assumed Y is fixed by the context: $\overline{X}$) is the set $\{x \in Y : x \notin X\}$.

Let U be a non-empty collection of sets. The *arbitrary intersection of* U (notation: $\bigcap U$) is the set $\{x : x \in X$ for all $X \in U\}$. Let U be a (possibly empty) collection of sets, all of which are subsets of some given set S. The *arbitrary intersection of* U *relative to* S (notation: $\bigcap_S U$) is the set $\{x \in S : x \in X$ for all $X \in U\}$. Let U be a (possibly empty) collection of sets. The *arbitrary union of* U (notation: $\bigcup U$) is the set $\{x : x \in X$ for some $X \in U\}$.

10.1.2 Relations, functions, indexed sets and sequences

Let x, y be any. The *ordered pair of* x *and* y (notation: $\langle x, y\rangle$) is the set $\{\{x, \varnothing\}, \{y, \{\varnothing\}\}\}$. For $n \in \mathbb{N}$, $n > 2$ the *ordered* n*-tuple of* $x_0, \ldots, x_{n-1}$ is defined by recursion as the ordered pair of the $n-1$-tuple of $x_0, \ldots, x_{n-2}$ and x_{n-1} (notation: $\langle x_0, \ldots, x_{n-1}\rangle$). For the sake of definiteness, we take the 1-tuple of x to be x itself, and the 0-tuple to be the empty set.

Let X, Y be sets. The *binary product of* X *and* Y (notation: $X \times Y$) is the set $\{\langle x, y\rangle : x \in X$ and $y \in Y\}$. In accordance with our definition of an n-tuple, the n-product of $X_0, \ldots, X_{n-1}$ (notation: $X_0 \times \cdots \times X_{n-1}$) is $\{\langle x_0, \ldots, x_{n-1}\rangle : x_0 \in X_0, \ldots$ and $x_{n-1} \in X_{n-1}\}$. If $X_0 = \cdots = X_{n-1}$, I write X^n for $X_0 \times \cdots \times X_{n-1}$. Thus, X^1 is X itself, and $X^0 = \{\varnothing\}$.

A *binary relation from X to Y* (or simply a *relation from X to Y* — I usually drop the qualification) is an $R \subseteq X \times Y$. Given such a relation, X is called its *source* and Y its *target*. I usually write xRy for $\langle x, y \rangle \in R$.

Let $R \subseteq X \times Y$. The *domain of R* (notation: $\mathsf{dom}(R)$) is $\{x \in X : xRy$ for some $y \in Y\}$. The *image of X in Y under R* (or simply the *image of R*, if X, Y are fixed by the context, — notation: $\mathsf{imag}(R)$) is $\{y \in Y : xRy$ for some $x \in X\}$. *The field of R* (notation: $\mathsf{field}(R)$) is $\mathsf{dom}(X) \cup \mathsf{imag}(R)$.

An $R \subseteq X \times Y$ is *on X* if $Y = X$, and it is *total* if $\mathsf{dom}(R) = X$. A relation is *partial* if it is not necessarily total.

An $R \subseteq X \times Y$ is *functional*, or simply a *function*, if, for all $x \in X$ and $y, z \in Y$, if xRy and xRz, then $y = z$. I write $f : X \longmapsto Y$ to mean f is a functional relation from X to Y, — that is, that $f \subseteq X \times Y$ and f is a function. If xfy and f is function, I usually write $f(x)$ for y.

An $f : X \longmapsto Y$ is *into Y*, *injective* or an *injection* if, for all $x, y \in X$, if $f(x) = f(y)$, then $x = y$. It is *onto Y*, *surjective* or a *surjection* if $\mathsf{imag}(f) = Y$. It is *a one-to-one correspondence between X and Y* (or a simply a *correspondence between X and Y*), *bijective* or a *bijection* if it is both into and onto Y. A function π is a *permutation of X* if it is a one-to-one correspondence between X and X itself.

I write ${}^{X}Y$ for the set of all total functions from X to Y. In this text, all functions are to be assumed total unless otherwise stated.

Let I be a set, and let X be a set such that there is a total function from I onto X. In such cases, we say X *is indexed by I* (by means of the given function, which I always refrain from mentioning), and I write $\{x_i\}_{i \in I}$ for X.

A (finite) *sequence from X* is an $s : n \longmapsto X$ for some $n \in \mathbb{N}$. Usually I denote a sequence $s : n \longmapsto X$ such that $s(k) = x_k$ for each $k < n$ simply by listing the x_k in order, — that is, as $x_0, \ldots, x_{n-1}$.

10.1.3 Orderings and the axiom of choice, equivalences and partitions

Let R be a relation on X. R is *reflexive* if xRx for all $x \in X$; *connected* if, for all $x, y \in X$, $x \neq y$, either xRy or yRx; *complete* if it is both reflexive and connected; *symmetric* if, for all $x, y \in X$, if xRy, then yRx; *antisymmetric* if, for all $x, y \in X$, if xRy and yRx, then $x = y$; *transitive* if, for all x, y, z, if xRy and yRz, then xRz; *well-founded* if for all $Y \subseteq X$ there is a *minimal of Y under R*, — that is, an $x \in Y$ such that, for all $y \in Y$, if yRx, then so does xRy (or, equivalently, such that there is no y for which yRx but $x \not R\, y$).

R is a *quasi-ordering* if it is reflexive and transitive, an *ordering* if it is a complete quasi-ordering, a *partial ordering* if it is an antisymmetric quasi-ordering, a *total ordering* if it is a complete partial ordering, and a *well-ordering* if it is a well-founded total ordering; it is an *equivalence relation* (or simply an *equivalence*) if it is reflexive, symmetric and transitive.

Let Y be a set. The *restriction of R to Y* is $R \cap Y \times Y$. Y is *quasi-ordered by R* (respectively, *partially ordered*, *etc.*) if the restriction of R to Y is a quasi-ordering (respectively, a partial ordering, *etc.*). The *well-ordering theorem* (which is equivalent to the *axiom of choice*) states that every set is well-ordered by some relation.

Let X be non-empty. A *partition of X* is an $M \subseteq \wp(X) \setminus \{\varnothing\}$ such that M is *exhaustive of X* and all elements of M are *pairwise disjoint*, — that is: $\bigcup M = X$, and $P \cap Q = \varnothing$ for all $P, Q \in M$, $P \neq Q$. Let E be an equivalence on X, and let $x \in X$. The *equivalence class of x under E* is the set of all $y \in X$ such that yEx. It is easy to see that, given a partition M of X, there is an equivalence E_M on X defined by $E_M = \{\langle x, y\rangle : x, y \in P \text{ for some } P \in M\}$ and that, conversely, given an equivalence E on X, there is a partition M_E of X defined by $M_E = \{P \in \wp(X) : P \text{ is the equivalence class of } x \text{ under } E \text{ for some } x \in X\}$. It is also easy to check that, given a partition M of X, $M = M_{E_M}$ and that, given an equivalence E on X, $E_{M_E} = E$.

10.2 Subset spaces and filters

The expression "subset space" was coined by Pacuit [44] in the context of modal logic to refer to a collection of subsets of a given set. Many objects of mathematical interest are of this sort, so it seems useful to have an expression to refer to them in general. The definition of the concept is quite simple.

Definition 10.1 (Subset space)**.** Let $S \neq \varnothing$ be a set. A *subset space on S* is a $U \subseteq \wp(S)$.

Of course, most subset spaces of interest will have some structure to them. So let us consider the most relevant (relevant for our purposes, that is) properties a subset space may have.

Definition 10.2 (Properties of a subset space)**.** A subset space U on S

- *contains the top* if $S \in U$;
- *contains the bottom* if $\varnothing \in U$;
- is *closed under supersets* or *monotonic* if, if $X \in U$ and $Y \supseteq X$, then $Y \in U$, for all X, Y;
- is *closed under binary intersection* if $X \cap Y \in U$ for all $X, Y \in U$;
- is *closed under finite intersection* if $\bigcap_S V \in U$ for all finite $V \subseteq U$;
- is *closed under intersection* if $\bigcap_S V \in U$ for all $V \subseteq U$;
- is *exclusive* if not both $X \in U$ and $\overline{X} \in U$, for all X;
- is *full* if either $X \in U$ or $\overline{X} \in U$, for all X;

- is *proper* if $U \neq \wp(S)$.

Some rather trivial relations among there properties are discussed in the following propositions. The reader acquainted with the subject of boolean algebra should already know them well enough.

Proposition 10.1. *A subset space is monotonic iff, if two sets are in the subset space, then so is their binary intersection.*

Proof. Let U on S be any.

First assume U is monotonic, let X, Y be any and suppose $X \cap Y \in U$. Then, because $X, Y \supseteq X \cap Y$ and U is monotonic, $X, Y \in U$. Because X, Y are any, $X, Y \in U$ for all such X, Y.

Now assume that if, for all X, Y, if $X \cap Y \in U$ then $X, Y \in U$. Let X, Y be any such that $X \subseteq Y$, and suppose $X \in U$. Then, because $X \subseteq Y$, $X \cap Y = X$; and, because $X \in U$, also $X \cap Y \in U$. It follows from the assumption that $Y \in U$. Therefore, because X, Y are any, U is monotonic. □

Proposition 10.2. *If a subset space is monotonic and non-empty, then it contains the top.*

Proof. Trivial. □

Proposition 10.3. *A monotonic subset space does not contain the bottom iff it is proper.*

Proof. Again, trivial. □

Proposition 10.4. *A subset space is closed under binary intersection iff it is closed under finite intersection.*

Proof. Obviously, closure under finite intersection implies closure under binary intersection.

For the converse, let U on S be any, and suppose it is closed under binary intersection. Let $n \in \mathbb{N}$ and $X_0, \ldots, X_{n-1} \in U$ be any. We prove by induction that $\bigcap_S \{X_k\}_{k<n} \in U$.

Naturally, $\bigcap_S \{X_0\} \in U$ since $\bigcap_S \{X_0\} = X_0$ and $X_0 \in U$ by assumption. So assume that, for all $k < m$ for some $m < n$, $\bigcap_S \{X_k\}_{k<m} \in U$. Then, because U is closed under binary intersection, $\bigcap_S \{X_k\}_{k<m} \cap X_m \in U$. Let x be any. By definition, $x \in \bigcap_S \{X_k\}_{k<m} \cap X_m$ iff $x \in X_k$ for all $k < m$ and $x \in X_m$. But this is equivalent to $x \in X_k$ for all $k < m+1$, which is the same as saying that $x \in \bigcap_S \{X_k\}_{k<m+1}$, — and, thus, $\bigcap_S \{X_k\}_{k<m} \cap X_m = \bigcap_S \{X_k\}_{k<m+1}$. Therefore, $\bigcap_S \{X_k\}_{k<m+1} \in U$. Because m is any, it follows by induction that $\bigcap_S \{X_k\}_{k<n} \in U$; and, because n is any, it then follows that U is closed under finite intersection. □

Proposition 10.5. *If a subset space is closed under intersection, then it contains the top.*

Proof. This is so because $\varnothing \subseteq U$ for any U and $\bigcap_S \varnothing = S$ (that is, because $\bigcap_S \varnothing$ is, by definition, the set of all $x \in S$ belonging to every $X \in \varnothing$ — and all $x \in S$ do, since there are no $x \in \varnothing$). □

Proposition 10.6. *If a subset space is closed under binary intersection, then, if it does not contain the bottom, then it is exclusive.*

Proof. Let U on S closed under binary intersection be any and suppose it is not exclusive. Then there is an $X \in U$ such that $\overline{X} \in U$: from which it follows, because U is closed under binary intersection, that $X \cap \overline{X} \in U$, — and $X \cap \overline{X} = \varnothing$. □

10.2.1 Filters and ultrafilters

The most important kind of subset spaces, for our purposes at least, are filters and ultrafilters [11], [31], [51]. The following accessory concept shall be helpful in defining those kinds of subset spaces.

Definition 10.3 (Upward closure)**.** Let $V \subseteq \wp(S)$ be any. The *upward closure of* V *relative to* S is the set $\uparrow_S V = \{X \subseteq S : X \supseteq Y \text{ for some } Y \in V\}$.

Convention 10.1. I usually drop the subscript in the symbol for the upward closure of a set, if the set relative to which it is being closed is fixed by the context.

Proposition 10.7. *The upward closure of a set is monotonic.*

Proof. Let S and $V \subseteq \wp(S)$ be any. Let X, Y be any such that $X \subseteq Y$, and suppose $X \in \uparrow V$. Then there's a $Z \in V$ such that $Z \subseteq X$. Because $Y \supseteq X$, it follows $Y \supseteq Z$ as well, — so that, by definition, $Y \in \uparrow V$. Because X, Y are any, $\uparrow V$ is monotonic. □

Definition 10.4 (Filter)**.** A subset space F on S is a *filter on* S if it contains the top, does not contain the bottom, is monotonic and closed under binary intersection.

Proposition 10.8. *A filter is an exclusive subset space.*

Proof. Indirectly already proved (proposition 10.6). □

Proposition 10.9. *A filter is a proper subset space.*

Proof. Again, indirectly already proved (proposition 10.3). □

Let us now consider the possible properties of filters that shall be of interest to us.

Definition 10.5 (Properties of a filter)**.** A filter F on S is

- an *ultrafilter* if it is *maximal on* S, — that is, if there is no filter F' on S such that $F' \supset F$;
- *principal* if $F = \uparrow\{B\}$ for some B;
- *fixed* if $F = \uparrow\{B\}$ for some B and B is a singleton.

Convention 10.2. If $F = \uparrow\{B\}$ for a given B, we refer to B as its *base*. Alternatively, therefore, a fixed filter is a principal filter which has a singleton for a base (or a singleton base, for short).

It should be noted that, unfortunately, the literature is somewhat equivocal regarding the names of the properties of filters presented above. The nomenclature here adopted is simply the one to which I'm used.

We now discuss in some detail how these properties relate to each other. The reader fully acquainted with the theory of filters and ultrafilters may prefer to skim through this part of the text.

Proposition 10.10. *Not every filter is an ultrafilter.*

Proof. Let $S = \{a, b\}$ and let $F = \{S\}$. Such an F is trivially a filter on S, but it's not an ultrafilter because both $\uparrow\{a\}$ and $\uparrow\{b\}$ are filters containing it, as the reader can easily check. □

In order to prove that every filter is contained in an ultrafilter, we shall need to make use of *Zorn's lemma*. Recall from basic set theory that, given a set X and a relation R on X, a *chain in* X *under* R is a $C \subseteq X$ which is totally ordered by R, that an *upper bound of* C *in* X *under* R is an $x \in X$ such that xRy for all $y \in X$, and that a *maximal element of* X *under* R is an $x \in X$ such that, for all $y \in X$, if $y \rotatebox[origin=c]{180}{R} x$, then so does $x \rotatebox[origin=c]{180}{R} y$ (or, equivalently, such that there is no y for which $y \rotatebox[origin=c]{180}{R} x$ but $x \not\rotatebox[origin=c]{180}{R} y$). Given these concepts, Zorn's lemma can be stated as follows. (I state it without proof, as it is well known to be equivalent to the *axiom of choice*.)

Proposition 10.11 (Zorn's lemma)**.** *If, under a certain relation, every chain in a set has an upper bound in it, then the set has a maximal element under the relation.*

Proposition 10.12. *Every filter is contained in an ultrafilter.*

Proof. Let F a filter on S be any, and let $\mathcal{F}$ be the set of all filters containing F. Obviously enough, $\mathcal{F}$ is non-empty, since $F \supseteq F$. So let C a chain in $\mathcal{F}$ under $\subseteq$ be any. If $C = \varnothing$, then F itself is an upper bound of C. Suppose, thus, $C \neq \varnothing$, and consider $\bigcup C$. Because $F \subseteq F'$ for every $F' \in C$ by definition, $F \subseteq \bigcup C$, so that, clearly, if $\bigcup C$ is a filter, then $\bigcup C \in \mathcal{F}$ as well; and therefore, because $F' \subseteq \bigcup C$ for all $F' \in C$, if $\bigcup C \in \mathcal{F}$, then it is an upper bound of C in $\mathcal{F}$ under $\subseteq$. Because C is any, from this and Zorn's lemma it follows $\mathcal{F}$ has a maximal element, which is therefore an ultrafilter. Thus it only remains to be shown $\bigcup C$ is indeed a filter.

That $\bigcup C$ contains the top is trivial, since every element of C does, C being a chain of filters. As for monotonicity, let $X \in \bigcup C$ and $Y \supseteq X$ be any. Then, there's an $F' \in C$ such that $X \in F'$. Because F' is a filter, therefore monotonic, $Y \in F'$, from which it follows $Y \in \bigcup C$ as well. Regarding closure under binary intersection, let $X, Y \in \bigcup C$ be any. Then there are $F', F'' \in C$ such that $X \in F'$ and $Y \in F''$. Because C is a chain, either $F' \subseteq F''$ or $F'' \subseteq F'$. Without loss of generality, suppose the latter is the case. Then $Y \in F'$ as well, so that $X \cap Y \in F'$, — from which it follows $X \cap Y \in \bigcup C$. Therefore, $\bigcup C$ is a filter as desired. □

Proposition 10.13. *A filter is an ultrafilter iff it is full.*

Proof. Let U a filter on S be any.

Assume, first, U is an ultrafilter, and suppose for a contradiction it is not full. Then, there is an $X \subseteq S$ such that neither $X \in U$ nor $\overline{X} \in U$. Let such X be given.

Consider, now, $F = \uparrow\{Y \subseteq S : Y = X \cap Z \text{ for some } Z \in U\}$. Because $X \cap Z \subseteq Z$ for any Z, it follows by the definition of F that $Z \in F$ for any $Z \in U$, so that $U \subseteq F$. In particular, F contains the top because $S \in U$. Moreover, it does not contain the bottom. For suppose for a contradiction that it does. Then of necessity $\varnothing \supseteq X \cap Z$ for some $Z \in U$, so that (because the empty set has no other subset) $X \cap Z = \varnothing$ for some $Z \in U$. This, however, implies there is a $Z \in U$ such that $Z \subseteq \overline{X}$, which then implies (because U is a filter, therefore monotonic) $\overline{X} \in U$, — thus contradicting the supposition that $\overline{X} \notin U$. Of course, F is monotonic by definition. Finally, let P, Q be any and suppose $P, Q \in F$. Then, by the definition of F there's a $Z \in U$ such that $P \supseteq X \cap Z$, and similarly a $W \in U$ such that $Q \supseteq X \cap W$. Because $Z, W \in U$ and U is a filter (therefore closed under binary intersection), also $Z \cap W \in U$, from which it follows that $X \cap (Z \cap W) \in F$. It is now a simple matter of set algebra to see that $(X \cap Z) \cap W \subseteq P$ for whatever W may be, and similarly that $(X \cap W) \cap Z \subseteq Q$ for whatever Z may be, so that $X \cap (Z \cap W) \subseteq P \cap Q$, — from which it follows by the definition of F that $P \cap Q \in F$. Because P, Q are any, F is therefore closed under binary intersection. It follows that F is a filter.

Notice however that $X \cap S = X$, so that $X \in F$ by the definition of F. But then $F \supseteq U$ and $X \in F$, and by supposition $X \notin U$. Therefore, $U \subset F$ for F a filter, thus contradicting the assumption that U is an ultrafilter. So, U is full.

Assume, now, U is full, and suppose for a contradiction it is not an ultrafilter. Then there is a filter $F \supset U$, so that there's an $X \in F$ such that $X \notin U$. Now, because U is full and $X \notin U$, it follows that $\overline{X} \in U$; and, because $U \subseteq F$, also that $\overline{X} \in F$, — thus contradicting the fact that F, being a filter, is exclusive. Therefore, U is an ultrafilter. □

Proposition 10.14. *There are non-principal filters.*

Proof. Let S infinite be any, and let $V = \{X \subseteq S : \overline{X} \text{ is finite}\}$. That is: V is the set of all *cofinite* subsets of S. Clearly, if V is a filter, then it's non-principal. For suppose for a contradiction it is principal. Then, there would have to be a $B \in V$ such that $B \subseteq X$ for all $X \in V$. By the definition of V, however, such a B would have to be cofinite as well, so that, given any non-empty finite $X \subseteq B$, also $B \setminus X$ would be cofinite, and therefore $B \setminus X \in V$ by definition. But $B \setminus X \subset B$ for any such X, which would contradict the supposition that B is a base of V. So it only remains to be shown V is indeed a filter.

But this is also rather trivial. S is cofinite, since it is infinite and $\overline{S} = \varnothing$ is finite; and, obviously enough, $\varnothing$ is not cofinite. Monotonicity is equally obvious. As for closure under binary intersection, it suffices to notice that, if X, Y are cofinite, so that $\overline{X}, \overline{Y}$ are finite, then $\overline{X} \cup \overline{Y}$ is finite as well, and $\overline{X} \cup \overline{Y} = \overline{X \cap Y}$, from which it follows that $X \cap Y$ is cofinite. □

Proposition 10.15. *Every filter on a finite set is principal.*

Proof. Let U on S finite be a filter, and suppose for a contradiction it is not principal. Then, it doesn't have a base, — that is: there's no unique B such that $B \subseteq X$ for all $X \in U$. So, either no $B \subseteq X$ for all $X \in U$, or some $B \in X$ for all $X \in U$ but it's not unique.

Suppose, first, no $B \in U$ is such that $B \subseteq X$ for all $X \in U$, and let $X \in U$ be any. Because $X \subseteq S$ and S is finite, so is X. So, among the subsets of X that are in U, let B be any which is minimal under $\subseteq$. Because X is finite, such a B exists. Now, it can't be that $B \subseteq Y$ for all $Y \in U$, $Y \neq X$, for this would immediately contradict the supposition. So, there is at least one $Y \in U$ such that $Y \nsupseteq B$. Let such Y be given. Because Y itself is finite, among the subsets of Y that are in U there's also a B' minimal under $\subseteq$. By supposition, $B' \neq B$. However, because U is a filter, therefore closed under binary intersection, and $B, B' \in U$, it follows that $B \cap B' \in U$; and, because $B \cap B' \subseteq B'$ and $B' \subseteq Y$, also $B \cap B' \subseteq Y$. But $B \cap B' \subset B'$ by supposition, and, therefore, also $B \cap B' \subset Y$, thus contradicting the assumption that $B' \in U$ is minimal under $\subseteq$ among the subsets of Y that are in U. So it can't be, as assumed, that $B \nsubseteq Y$. Therefore, because $Y \neq X$ is any, $B \subseteq Y$ for all $Y \neq X$; and, because X is any and $B \subseteq X$, it follows that $B \subseteq X$ for every $X \in U$, thus contradicting the supposition that no such B exists.

Suppose, now, there is a $B \in U$ such that $B \subseteq X$ for all $X \in U$ but B is not unique. Let $B' \neq B$ be any $B' \in U$ such that $B' \subseteq X$ for all $X \in U$. Then, because $B \subseteq X$ for all $X \in U$, in particular $B \subseteq B'$; and similarly, because $B' \subseteq X$ for all $X \in U$, in particular $B' \subseteq B$. But then $B = B'$, thus contradicting the supposition that $B' \neq B$. So B is unique after all, which contradicts the supposition that it is not.

In both cases, we find a contradiction. So U indeed has a base, and is therefore principal. □

For our next proposition, we shall need to make use of the following principle of arithmetic. It is provable in set theory, but I state it without proof, assuming the reader is acquainted with it. Recall that a *least element* of an $X \subseteq \mathbb{N}$ is an $x \in X$ such that $x \leq y$ for all $y \in X$.

Proposition 10.16 (Well-ordering principle)**.** *Every non-empty set of natural numbers has a least element.*

Proposition 10.17. *Every set in a non-principal filter is infinite.*

Proof. Let F on S be a non-principal filter, and suppose for a contradiction there's an $X \in F$ finite. Let $G = \{X \in F : X \text{ is finite}\}$ (which, notice, implies $G \subseteq F$). By supposition, $G \neq \varnothing$, and $|X| \in \mathbb{N}$ for every $X \in G$ by the definition of G. So, by the well-ordering principle, there is a least $n \in \mathbb{N}$ for which $|X| = n$ for some $X \in G$. Let such an n and such an X be given. Now, it can't be that $X \subseteq Y$ for all $Y \in F$, for then F would be principal, thus contradicting the assumption that it is not. So there's a $Y \in F$ such that $X \nsubseteq Y$. Let such a Y also be given. Because F is a filter, therefore closed under binary intersection, $X \cap Y \in F$; and, for the same reason, $X \cap Y \neq \varnothing$, since F does not contain the bottom. However, because $X \nsubseteq Y$, of necessity

$X \cap Y \subset X$; and thus, because X is finite, so is $X \cap Y$. But then $X \cap Y \in G$ and $|X \cap Y| < |X|$, which contradicts the assumption that $|X| = n$. Therefore, no such X exists; and, because F is any non-principal filter, every set in a non-principal filter is infinite. □

Proposition 10.18. *Every principal ultrafilter is fixed.*

Proof. Let U on S a principal ultrafilter be any. Because U is principal, it has a base, — B, say. So suppose for a contradiction B is not a singleton. Because U is an ultrafilter, therefore a filter, $B \neq \varnothing$. Now let $x \in B$ be any. Because B is not a singleton, $\{x\} \subset B$. Moreover, because by assumption $\{x\} \notin U$ (for otherwise B wouldn't be a base), and because U is an ultrafilter — therefore full (proposition 10.13) —, $\overline{\{x\}} \in U$. So, again because U is an ultrafilter — therefore closed under binary intersection —, $\overline{\{x\}} \cap B \in U$. But, because $\{x\} \subset B$, it follows that $\overline{\{x\}} \cap B \subset B$, thus contradicting the assumption that B is a base of U. Therefore B must be a singleton, so that U is fixed. □

Proposition 10.19. *There are non-principal ultrafilters.*

Proof. Let S infinite be any, and consider the set V of cofinite subsets of S. This set is a filter, and so, by proposition 10.12, there is an ultrafilter U containing it. Clearly, U can't be principal, for this would imply V has a base as well, and it has also been shown this cannot be the case. □

Proposition 10.20. *Every non-principal ultrafilter on a set contains the filter of its cofinite subsets.*

Proof. Let U a non-principal ultrafilter on S be any, and suppose for a contradiction it does not contain the filter of cofinite subsets of S. Then, there's an $X \subseteq S$ cofinite such that $X \notin F$. Because U is an ultrafilter, therefore full (proposition 10.13), $\overline{X} \in F$. However, as has also been shown, because $\overline{X}$ is finite this is impossible. □

Chapter 11

Choice and preferences

In this chapter we delve deeper into the concepts of choice and preferences first discussed in section 3.1. Our main goal in this chapter is to present a proof of Suzumura's theorem, which shows the equivalence between a relation being strongly acyclic and its having a compatible extension which is an ordering. This central result in the theory of orderings, which extends Szpilrajn's theorem to the effect that every quasi-ordering has an ordering as a compatible extension, is required for most of the results presented in this dissertation.

Just as the aforementioned section 3.1, this chapter, unless otherwise noted, is almost entirely based on [7], [20], [56], [65], so that I mostly refrain from citing them specifically. Both for the sake of ease of reference and of establishing useful notational conventions, some of the content of that section is repeated here (albeit in much greater detail).

11.1 Choices and rationalization

The fundamental concept of choice theory is, of course, that of *choice*. In set theory, the notion of choice is formally articulated by the concept of choice function. A *set-theoretical choice function* c picks, from each non-empty set X of a given collection of sets, an element $c(X)$ of X [19], [31], [47], [37], — a choice on X thus being identified with the element chosen from X by c. Intuitively, c is intended to model the behavior of an agent who makes choices on each and every X in the domain of the function. In choice theory, however, we are interested in a more general concept of choice function which assigns, to each X in the collection, not an element but a *subset* $C(X)$ of X. Given that there is an obvious correspondence between those two types of functions, a set-theoretical choice function can be seem as equivalent to the particular case of a "choice-theoretical" choice function such that $C(X)$ is a singleton for all non-empty X (with $C(\varnothing)$ being the empty set itself for every choice function C by definition).

Intuitively, a choice function in this sense, just as in the set-theoretical one, is intended to formally represent the choice behavior of an agent, but in a somewhat broader, arguably *counterfactual* fashion. It is counterfactual first of all in the sense that it represents how the agent *would* choose if the options were to be presented to him or her in various possible combinations: that is what the assumption that opportunities are subsets of one and same set stands for. This aspect, however, is evidently shared with set-theoretical choice functions having as source the powerset of a given set, and it is not the one we care about here. Second, it is counterfactual in the sense that it represents not only about the option that agent would *actually* pick from each set, but also about all other options that he or she would be willing to alternatively pick, — that is: all those the agent *judges* acceptable. It is in this second sense that this concept of choice function is distinct from the set-theoretical one. This is so not only due to the enlargement it implies of the image of the function to more than just singletons and the empty set, but also, and more importantly, because it purports to model extensionally what is in essence a *subjective* aspect of choices: namely, the *judgments* involved in making them, judgments about what options are *acceptable*. This *dispositional* aspect of choices, although perhaps only apparently absent from the set-theoretical concept of choice function, in any event undeniably comes to the fore and is of primary interest in choice theory, for reasons that will become clear soon enough.

The intuition that choice functions are to somehow represent the dispositions of agents to make certain choices can be made precise by imposing additional constraints on the concept. The most usual and natural away to do this is through the concept of *preferences*.

Preferences are simply binary relations. Although I mostly talk of preferences in what follows (of binary relations *interpreted as* preferences, that is), all theorems I present hold, of course, for binary relations in general. But I talk of preferences nevertheless, instead of relations in general, because there seems to me to be some intuitions that are quite particular to the matter of preferences and choices, and I find that the language used in presenting the theorems can be instrumental in eliciting those intuitions.

Convention 11.1. Let $\succsim$ be a preference relation on S. Informally, when $\succsim$ is fixed by the context, we refer to a $\langle x, y\rangle \in \succsim$ as a *preference* (*for x over y*), and also as a *link in* $\succsim$ (or a $\succsim$-link, for short). Usually I write $x \succsim y$ for $\langle x, y\rangle \in \succsim$. When either $x \succsim y$ or $y \succsim x$, we say *x and y are compared by* $\succsim$. We refer to a pair $x, y \in S$ such that they are not compared by $\succsim$ as an *incompleteness of* $\succsim$. When $x \succsim y$ but not $y \succsim x$, we say *x dominates y under* $\succsim$.

Definition 11.1 (Parts of a relation and the diagonal relation)**.** The *symmetric part of* $\succsim$ (or the *indifference part*, if $\succsim$ is to be interpreted as preferences) is the relation $\sim = \{\langle x, y\rangle \in S \times S : x \succsim y$ and $y \succsim x\}$. The *asymmetric part of* $\succsim$ (or the *strict preference part*, if $\succsim$ is to be interpreted as preferences) is the relation $\succ = \{\langle x, y\rangle \in S \times S : x \succsim y$ but $y \not\succsim x\}$. The *non-compared part of* $\succsim$ is the relation $\wr = \{\langle x, y\rangle \in S \times S :\ x \not\succsim y$ and also $y \not\succsim x\}$. The *diagonal relation on S* is the relation $\Delta = \{\langle x, y\rangle \in S \times S : x = y\}$.

Convention 11.2. In some occasions, and specially when a relation is represented by a letter — R, say —, I write $\tilde{R}$ for its symmetric part, and $\bar{R}$ for its asymmetric part.

Proposition 11.1. *A relation is the union of its symmetric and asymmetric parts.*

Proof. Let $\succsim$ on S and x, y be any, and suppose $x \succsim y$. By excluded middle, either $y \succsim x$ or $y \not\succsim x$. In the former case, $x \sim y$; in the latter, $x \succ y$. Because x, y are any, either $x \sim y$ or $x \succ y$ for all x, y such that $x \succsim y$. Therefore, $\succsim\, = \,\sim \cup \succ$. □

In the theory of rational choice, it is generally assumed that rational agents are *maximizers* (or, as I believe it would be more suitable to say, *optmizers*). That is, rational agents make choices according to what they judge *maximally good*, or *best*, in the sense that only the best options in an opportunity are judged acceptable, if any.

Definition 11.2 (Best and optimal options)**.** Let X be a set. A *greatest* or *best option in X under* $\succsim$ (or a $\succsim$*-best option*, for short) is an $x \in X$ such that, for all $y \in X$, $x \succsim y$. A *maximal* or *optimal option in X under* $\succsim$ (or a $\succsim$*-optimal option*, for short) is an $x \in X$ such that for no $y \in X$ does $y \succ x$.

Convention 11.3. I denote the set of best options in X under $\succsim$ as $B_{\succsim}(X)$, and the set of optimal options by $O_{\succsim}(X)$.

As the definitions suggests, optimization comes in two flavors. It is easy to see (but the proof for it is presented next anyway) that an option being best under a relation implies it being optimal, but not the converse. It is also easy to see that this distinction has to with whether or not all options in the set are compared, as proven latter.

Proposition 11.2. *A best option is optimal. The converse is not true.*

Proof. Let $\succsim$ be a relation, X a set, let x be a $\succsim$-best option in X, and suppose for a contradiction that x is not $\succsim$-optimal. Then there is a $y \in X$ such that $x \succ y$. However, because x is $\succsim$-best, in particular $x \succsim y$, thus contradicting $x \succ y$. Therefore, x is $\succsim$-optimal.

Now let $X = \{x\}$ and $\succsim\, = \varnothing$. Clearly, x is $\succsim$-optimal in X, but not $\succsim$-best. □

It is clear that $B_{\succsim}$, taken as a function from $\wp(S)$ to itself, defines a choice function on S whatever S and $\succsim$ may be, since $B_{\succsim}(X) \subseteq X$ for any X by definition; and the same is true for $O_{\succsim}$. Not every choice function is of this sort, however, as the following example shows.

Example 11.1. Let C on $\{a, b, c\}$ be given by this table:

X	$\varnothing$	$\{a\}$	$\{b\}$	$\{c\}$	$\{a, b\}$	$\{a, c\}$	$\{b, c\}$	$\{a, b, c\}$
$C(X)$	$\varnothing$	$\{a\}$	$\{b\}$	$\{c\}$	$\{a, b\}$	$\{c\}$	$\{c\}$	$\{a, b\}$

There is no $\succsim$ such that $C(X) = B_{\succsim}(X)$ (or that $C(X) = O_{\succsim}(X)$, for that matter) for all X, as can easily be checked by inspection. The issue with C is that $C(\{a, c\}) = C(\{b, c\}) = \{c\}$ is incompatible with $C(\{a, b, c\}) = \{a, b\}$ if there is to be such a $\succsim$.

The intuition that from the point of view of rationality there is something odd with the above choice function can be made precise through the concepts of *rationalization* and *revealed preferences*.

Definition 11.3 (Rationalization)**.** Let C be a choice function on S and $\succsim$ preferences on S. C is *rationalized by* $\succsim$ *through best options* (or $\succsim$*-*B*-rationalized*, for short) if $C(X) = B_{\succsim}(X)$ for all $X \in \mathsf{dom}(C)$; it is *rationalized by* $\succsim$ *through optimal options* (or $\succsim$*-*O*-rationalized*, for short) if $C(X) = O_{\succsim}(X)$ for all $X \in \mathsf{dom}(C)$.

Convention 11.4. We say C is B*-rationalizable* (similarly, O*-rationalizable*) if there is a $\succsim$ such that C is $\succsim$-B-rationalized ($\succsim$-O-rationalized).

Definition 11.4 (Revealed preferences)**.** Let C be a choice function on S. The *preferences revealed by* C (or the C-revealed preferences, for short) is the relation $\succsim_C = \{\langle x, y\rangle \in S \times S :$ for some $X \in \mathsf{dom}(C)$, $x \in C(X)$ and $y \in X\}$.

The next proposition suggests the adequacy of these definitions.

Proposition 11.3. *If a choice function with full domain is rationalized by a preference relation through best options, then the preferences revealed by it are that relation.*

Proof. Let C on S and $\succsim$ on S be any, and suppose C is $\succsim$-B-rationalized. Let x, y be any.

First suppose $x \succsim y$. Then, $x \in B_{\succsim}(\{x, y\})$, from which it follows that $x \succsim_C y$.

Now suppose $x \succsim_C y$. Then there is an X such that $x \in C(X)$ and $y \in X$. Because C is $\succsim$-B-rationalized, $C(X) = B_{\succsim}(X)$. Therefore, $x \in B_{\succsim}(X)$, from which it follows that $x \succsim y$. □

Well, what about the converse?

As the reader can easily check for him or herself, the choice function presented above as an example of an odd one is not rationalizable by any preference relation, but it does reveal one all the same. Only for choice functions having the following property the converse is the case as well.

Definition 11.5 (Extensiveness – Plott 1973 [48])**.** A choice function C is *extensive* if, for all $X \in \mathsf{dom}(C)$ and all x, if $x \in X$ and $x \succsim_C y$ for all $y \in X$, then $x \in C(X)$.

Proposition 11.4. *A choice function is extensive iff it is rationalized through best options by the preferences it reveals.*

Proof. Let C on S be any.

Suppose C is extensive. Let $X \in \mathsf{dom}(C)$ and x be any. First, suppose $x \in C(X)$. Then $x \succsim_C$ for all $y \in X$, from which it follows that $x \in B_{\succsim_C}(X)$. Now suppose $x \in B_{\succsim_C}(X)$. Then $x \succsim_C y$ for all $y \in X$, — from which it follows, because C is extensive, that $x \in C(X)$. Because x and X are any, $C(X) = B_{\succsim_C}(X)$ for all X, — that is: C is $\succsim_C$-B-rationalized.

Now suppose C is $\succsim_C$-B-rationalized. Let $X \in \mathsf{dom}(C)$ and x be any, and suppose $x \in X$ and $x \succsim_C y$ for all $y \in X$. Then $x \in B_{\succsim_C}(X)$ and, therefore, $x \in C(X)$. Because X and x are any, C is extensive. □

Proposition 11.5. *It follows immediately from the previous propositions that an extensive choice function with full domain is rationalizable through best options iff it is rationalized though best options by the preferences it reveals.*

As already observed in section 3.1, it is perhaps unfortunate that the expression "rationalization" is already quite established in the literature, because rationalizability of a choice function does not imply that the *choices* made by the agent are themselves in any sense *rational*. To see what I mean by this, consider the following example.

Example 11.2. Let C on $\{a, b, c\}$ be given by this table:

X	$\varnothing$	$\{a\}$	$\{b\}$	$\{c\}$	$\{a, b\}$	$\{a, c\}$	$\{b, c\}$	$\{a, b, c\}$
$C(X)$	$\varnothing$	$\{a\}$	$\{b\}$	$\{c\}$	$\{a\}$	$\{c\}$	$\{b\}$	$\varnothing$

By definition, $\succsim_C = \{\langle a, a\rangle, \langle b, b\rangle, \langle c, c\rangle, \langle a, b\rangle, \langle b, c\rangle, \langle c, a\rangle\}$.

C *is* $\succsim_C$-B-rationalized (indeed, also $\succsim_C$-O-rationalized, as $B_{\succsim_C}(X) = O_{\succsim_C}(X)$ for all X), but there is nothing really rational about these choices, as they express preferences according to which each option dominates an option which dominates an option which dominates the first one — thus leading to every option in $\{a, b, c\}$ being rejected, even though they all are acceptable from one opportunity each when pairwise compared.

As the example shows, "rationalization" here means *merely* that a choice function may be interpreted as the behavior of an agent who makes at leas optimal choices according to preferences, and so even if the behavior itself doesn't seem to be entirely rational.

This prompts us to the question of what conditions preferences must satisfy in order to lead optimizing agents to actually make rational choices.

11.2 Relations and their properties

Let us begin by recalling from section 3.4 our two concepts of cycle in a relation.

Definition 11.6 (Cycles)**.** An *improper cycle a relation* $\succsim$ *on* S (or a $\succsim$*-improper cycle*, for short) is a sequence $x_0, \ldots, x_n \in S$ such that $x_k \succsim x_{k+1}$ for all $k < n$ but also $x_k \succ x_{k+1}$ for some $k < n$. A *proper cycle of* $\succsim$ (or a $\succsim$*-proper cycle*, or simply a $\succsim$*-cycle* for short), is an improper cycle $x_0, \ldots, x_n$ such that $x_k \succ x_{k+1}$ for all $k < n$.

A cycle $x_0, \ldots, x_n$ (of any sort) is *minimal* if $x_k \neq x_\ell$ for all $k, \ell < n$ such that $k \neq \ell$.

Following Bossert and Suzumura [7], we distinguish two kinds of properties of relations, coherence properties and richness properties. The distinction itself is not essential, but it helps understand the distinct motivations for expecting certain properties from preferences. Intuitively, coherence properties have to do with how the preferences of an agent regarding different pairs of options fit together, while richness properties amount to how much information his or her preferences can provide for the optimization of choice.

Definition 11.7 (Coherence properties of a relation)**.** A relation $\succsim$ on S is

- *symmetric* if, for all x, y, if $x \succsim y$, then $y \succsim x$;
- *asymmetric* if, for all x, y, if $x \succsim y$, then $y \not\succsim x$;
- *antisymmetric* if, for all x, y, if $x \succsim y$ and $y \succsim x$, then $x = y$;
- *strongly acyclic* if there are no $\succsim$-proper cycles;
- *weakly acyclic* (or simply *acyclic*) if there are no $\succsim$-proper cycles;
- *transitive* if, for all x, y, z, if $x \succsim y$ and $y \succsim z$, then $x \succ z$;
- *quasi-transitive* if $\succ$ is transitive.

Definition 11.8 (Richness properties of a relation)**.** A relation $\succsim$ on S is

- *reflexive* if, for all x, $x \succsim x$ or, equivalently, if $\succsim \supseteq \Delta$;
- *irreflexive* if, for all x, $x \not\succsim x$ or, equivalently, if $\succsim \cap \Delta = \varnothing$;
- *connected* if, for all x, y such that $x \neq y$, x and y are compared by it;
- *complete* if it is both reflexive and connected, or, equivalently, if, for all x, y, x and y are compared by it (which, of course, is also tantamount to $\wr = \varnothing$).

Certain classes of relations combining some of those properties shall be of interest, most notably quasi-orderings and orderings.

Definition 11.9 (Orderings and similarities)**.** A relation $\succsim$ on S is

- a *quasi-ordering* if it is reflexive and transitive;
- a *partial ordering* if it is an antisymmetric quasi-ordering;
- an *ordering* if it is a connected quasi-ordering;
- a *strict ordering* if it is asymmetric and transitive;
- a *similarity relation* (or simply a *similarity*, for short) if it is reflexive and symmetric;
- an *equivalence relation* (or simply an *equivalence*, for short) if it is a transitive similarity.

Convention 11.5. Let X be a set. We say $\succsim$ *quasi-orders* X (similarly, that it *partially orders*, *etc.*) if its restriction to X is a quasi-order (similarly, a partial order, *etc.*).

The following series of propositions state, with proofs, the fundamental facts about how all those properties and classes of relations relate to each other, and also what they each imply for the parts of a relation.

Proposition 11.6. *Both the symmetric and the non-compared parts of a relation are symmetric. Indeed, a relation is symmetric iff it is its own symmetric part. Moreover, if a relation is transitive, then so is its symmetric part. From this it follows trivially that if a relation is reflexive, then its symmetric part is a similarity and, if it is also transitive, an equivalence.*

Proof. I prove only the bit about transitivity, as the rest is trivial. Let $\succsim$ on S be transitive, let x, y, z be any, and suppose $x \sim y$ and $y \sim z$. Because $\sim \subseteq \succsim$ and $\succsim$ is transitive, $x \succsim z$. Moreover, because $\sim$ is symmetric, $y \sim z$ and $x \sim y$, it follows that $z \sim y$ and $y \sim x$. Again because $\sim \subseteq \succsim$ and $\succsim$ is transitive, $z \succsim x$. Therefore, $x \sim z$. □

Proposition 11.7. *If a relation is asymmetric, then it is irreflexive. The asymmetric part of a relation is asymmetric, and a relation is asymmetric iff it is its own asymmetric part. Moreover, if it is asymmetric and acyclic, then it is strongly acyclic. If a relation is transitive, then so is its asymmetric part: that is, it is also quasi-transitive. (The converse is not true, however.) From this it follows that, if a relation is an ordering, then its asymmetric part is a strict ordering.*

Proof. Again, I prove only the bit about transitivity. Let $\succsim$ on S be transitive, let x, y, z be any, and suppose $x \succ y$ and $y \succ z$. Because $\succ \subseteq \succsim$ and $\succsim$ is transitive, $x \succsim z$. Suppose now for a contradiction that, nevertheless, $z \succsim x$. Then, because $x \succ y$, $\succ \subseteq \succsim$ and $\succsim$ is transitive, $z \succsim y$, thus contradicting $y \succ z$. Therefore, $x \succ z$.

Now let $S = \{a, b, c\}$ and $\succsim = \{\langle a, b\rangle, \langle b, c\rangle, \langle c, b\rangle\}$. Clearly, $\succsim$ is (vacuously) quasi-transitive, but not transitive. □

Proposition 11.8. *If a relation is transitive, then it is strongly acyclic (and therefore acyclic, as follows immediately from the definition); if it is quasi-transitive, then it is acyclic. The converses are not true. Indeed, quasi-transitivity and strong acyclicity are independent properties.*

Proof. Let $\succsim$ on S be transitive, and suppose for a contradiction that there is an improper cycle $x_0, \ldots, x_n$. Because $x_k \succsim x_{k+1}$ for all k, by repeated applications of transitivity it follows that $x_0 \succsim x_n$, thus contradicting $x_n \succ x_0$. Therefore, $\succsim$ is strongly acyclic. Similarly, let $\succsim$ be quasi-transitive and suppose, again for a contradiction, that there is a cycle $x_0, \ldots, x_n$. Because $x_k \succ x_{k+1}$ for all k, by repeated applications of quasi-transitivity it follows that $x_0 \succ x_n$, thus contradicting $x_n \succ x_0$, from which we conclude $\succsim$ is acyclic.

Now let $S = \{a, b, c\}$ and $\succsim = \{\langle a, b\rangle, \langle b, c\rangle\}$. Clearly, $\succsim$ is acyclic (indeed, strongly acyclic), but not quasi-transitive (and therefore, by the previous proposition, also not transitive). Finally, let $\succsim = \{\langle a, b\rangle, \langle a, c\rangle, \langle b, c\rangle, \langle c, a\rangle, \langle c, b\rangle\}$. This relation is (again, vacuously) quasi-transitive, but not strongly acyclic. □

Proposition 11.9. *If a relation is transitive and, under it, an option relates to another, and the second to a third, and at least one of them relates strictly to the next, then the first relates strictly to the last. Similarly, if a relation is transitive and, under it, an option relates to another, and the second to a third, and none of them relates strictly to the next, then the first does not relate strictly to the last.*

Proof. Let $\succsim$ on S transitive and x, y, z be any.

First assume, without loss of generality, $x \succsim y$ and $y \succ z$. Then, by transitivity, $x \succsim z$. Now suppose for a contradiction $z \succsim x$ as well. Then x, y, z is an improper cycle, which by the previous proposition contradicts the assumption that $\succsim$ is transitive. Therefore, $x \succ z$.

Now assume $x \sim y$ and $y \sim z$. Then, by transitivity, $x \succsim z$. Now suppose for a contradiction $x \succ z$. Then x, y, z is again an improper cycle, thus also contradicting the assumption that $\succsim$ is transitive. Therefore, $x \sim z$. □

Proposition 11.10. *If a relation is strongly acyclic and complete, then it is transitive.*

Proof. Let $\succsim$ on S be strongly acyclic and complete, and suppose for a contradiction that it is not transitive. Then there are $x, y, z \in S$ such that $x \succsim y$, $y \succsim z$ but not $x \succsim z$. Because $\succsim$ is complete and not $x \succsim z$, it follows that $z \succsim x$. So x, y, z is an improper cycle, contradicting the consistency of $\succsim$. Therefore, $\succsim$ is transitive. □

The following three propositions paint a minimal picture of how the concepts of rationalization and preferences relate to each other. These are by absolutely no means the only interesting results on rationalization; they are, however, all that is required for what is to follow.

Proposition 11.11. *If the restriction of a relation to a set is complete, then an optimal option in the set under that relation is also best.*

Proof. Let X be a set, $\succsim$ such that $\succsim \cap (X \times X)$ is complete, suppose x is $\succsim$-optimal in X, and let $y \in X$ be any. Because x is $\succsim$-optimal, in particular not $y \not\succ x$ — that is, either $x \sim y$ or $x \wr y$. However, because $\succsim \cap (X \times X)$ is complete, $x \not\wr y$. Therefore, $x \sim y$, and so, in particular, $x \succsim y$. Because y is any, it follows that x is $\succsim$-best. □

Proposition 11.12. *If the restriction of a relation to a finite non-empty set is acyclic, then it has an optimal option under that relation.*

Proof. Let X a finite non-empty set, let $\succsim$ be such that $\succsim \cap (X \times X)$ is acyclic, and suppose for a contradiction that no option in X is $\succsim$-optimal. Because X is finite and non-empty, $|X| = n + 1$ for some n (where $|X|$ stands for the cardinality of X). Now let x_0 be an arbitrary option of X. Because x_0 is not optimal, there is another, necessarily different option in X — x_1, say — such that $x_1 \succ x_0$. By repeating this same reasoning n times, we conclude that there is an x_{n+1} such that $x_{n+1} \succ x_n$. However, because we have now already exhausted the options of X, we must also conclude that $x_{n+1} = x_k$ for some $k < n$. It follows that $x_k, \ldots, x_n$ is an $R \cap (X \times X)$-cycle, contradicting the acyclicity of $\succsim \cap (X \times X)$. Therefore, X has an $\succsim \cap (X \times X)$-optimal option, — which is also an $\succsim$-optimal option in X, since $\succsim \cap (X \times X) \subseteq \succsim$. □

Proposition 11.13. *It follows immediately from the two previous proposition that, if the restriction of relation to a finite non-empty set is acyclic and complete, then the set has a best option under that relation. More importantly, it also follows that a choice function is finitely decided and rationalizable through best options iff it is rationalized through best options by an acyclic and complete relation.*

11.3 Extensions of a relation

As discussed above, richness properties of preferences are distinguished from coherence properties in that they have to do with how informed an agent's preferences allow his or her choices to be. As we have seen, rationalization by preferences through best and through optimal options only coincided when the preferences are complete. This suggests that an agent can sometimes improve the rationality of his choices by forming preferences about pairs which were previously not compared by his or her preferences. This, however, raises an obvious issue. Because we are taking strict preferences to be represented by the asymmetric part of the agents preference relation, forming additional preferences shouldn't change the compositions of the symmetric and the asymmetric part of the relation regarding pairs that were already in one or the other. Moreover, adding more pairs to a relation may change how those pairs fit together with each other, thus messing with the relation's coherence properties. Naturally, if our concern is rationality, we would usually like to be able to preserve those properties when forming additional preferences, if possible.

These considerations lead to the concept of compatible extension of a relation.

Definition 11.10 (Compatible extension of a relation)**.** Let $\succsim$ be a relation on S, and let $S' \supseteq S$. A *compatible extension of* $\succsim$ *to* S' is a relation $\succsim'$ on S' such that $\succsim' \supseteq \succsim$ and $\succ' \supseteq \succ$.

Convention 11.6. Because we will only be interested in compatible extensions of relations to their own underlying sets, I usually won't explicitly mention the underlying sets of extensions. Moreover, I drop the qualification "compatible" for the sake of brevity, so that "extension" will always mean compatible extension. To avoid confusion, given a $\succsim$ I refer to a $\succsim' \supseteq \succsim$ that is not necessarily an extension of $\succsim$ as a *super-relation of* $\succsim$.

Now for three simple facts about extensions, in order to show that they behave as expected.

Proposition 11.14. *A relation is an extension of itself.*

Proof. Trivial. □

Proposition 11.15. *If a relation is an extension of another, and the latter an extension of the former, then they are the same relation.*

Proof. Again trivial, as $\subseteq$ is an antisymmetric relation. □

Proposition 11.16. *An extension of an extension of a relation is itself an extension of that relation.*

Proof. Let $\succsim$ on S, $\succsim'$ an extension of $\succsim$ to S' and $\succsim''$ an extension of $\succsim'$ to S'' be any. By definition, $S \subseteq S' \subseteq S''$, $\succsim \subseteq \succsim' \subseteq \succsim''$ and $\succ \subseteq \succ' \subseteq \succ''$, so that $S \subseteq S''$, $\succsim \subseteq \succsim''$ and $\succ \subseteq \succsim''$. Therefore, $\succsim''$ is an extension of $\succsim$. □

Trivial as these propositions may be, they show extensions indeed exhibit a mereological behavior, as expected. As for the next proposition, it shows that the concept is not not trivial, in the sense that, again as expected, completeness is the limit to how much a relation can be extended.

Proposition 11.17. *A complete relation has no proper extensions (to its own underlying set, that is).*

Proof. Let $\succsim$ on S be complete, and suppose for a contradiction that $\succsim'$ on S is a proper extension of $\succsim$. Then there are x, y such that $x \succsim' y$ but $x \not\succsim y$. Because $\succsim$ is complete and $x \not\succsim y$, it follows that $y \succsim x$, — indeed, that $y \succ x$. However, because $\succsim'$ is an extension of $\succsim$, from this it follows that $y \succ' x$, thus contradicting $x \succsim' y$. Therefore, $\succsim'$ is not a proper extension of $\succsim$; and, because $\succsim'$ is any, no extension of $\succsim$ for $\succsim$ complete is proper. □

Given that the purpose of introducing the concept of extension is to discuss what happens to the preferences of an agent as he or she forms more preferences, and completeness of his or her preference relation is the limit to extending it, a natural questions seems to me to arise: what happens if we simply throw in the relation *all* the pairs that are not already compared by it? What coherence properties of the relation would be lost, and what preserved?

In order to explore these questions, I introduce the concept of full extension of a relation.

Definition 11.11 (Full extension)**.** The *full extension of* $\succsim$ is the relation $\succsim^\circ = \succsim \cup \wr$.

Although the idea of this device has clearly been in use in the literature on choice and preferences (most notably, in social choice theory, in Sen's concept of the Pareto extension, *cf.* [56]), as far as I can tell, the concept of full extension hasn't been explicitly considered nor named in this literature.

As usual, I prove the following propositions in order to show that the concept of full extension works as desired.

Proposition 11.18. *The full extension of a relation is indeed an extension. Actually, something stronger is the case: their asymmetric parts are the same.*

Proof. Let $\succsim$ be a relation on S, and let x, y be any. Because $\succsim = \succ \cup \sim$, it follows that $\succsim^\circ = \succ \cup \sim \cup \wr$. So it suffices to notice that, if $\langle x, y \rangle$ belongs to $\succ^\circ$, then, because both $\sim$ and $\wr$ are symmetric, it can belong to neither of them, for otherwise we would also have $y \succsim^\circ x$. □

Remark 11.1. Notice that at first if might seem that there is an ambiguity in writing, as I did, $\succ^\circ$ for the asymmetric part of $\succsim^\circ$, since it could mean the full extension of $\succ$ instead. It turns out, however, as shown by the previous proposition, that these things are one and the same, so no ambiguity arises.

Proposition 11.19. *The symmetric part of the full extension of a relation is the union of its symmetric and its non-compared parts.*

Proof. Follows immediately from the previous proposition and the fact that a relation is the union of its symmetric and its asymmetric parts (proposition 11.1). □

Proposition 11.20. *Two relations on the same underlying set have the same full extension iff they have the same asymmetric part.*

Proof. Let $\succsim$ on S and $\succsim' \neq \succsim$ on S be any.

Suppose, first, that $\succsim^\circ = \succsim'^\circ$. Therefore, $\succ^\circ = \succ'^\circ$; and, because $\succ^\circ = \succ$ and $\succ'^\circ = \succ'$, so, too, $\succ = \succ'$.

Now suppose $\succ = \succ'$. Let x, y be any. Of course, if $x \succsim y$ and $x \succsim' y$, or if $x \wr y$ and $x \wr' y$, then $x \succsim^\circ y$ and $x \succsim'^\circ y$. So suppose, without loss of generality, that $x \succsim y$ but $x \not\succsim' y$. Then, because $\succ = \succ'$, of necessity $y \succsim x$ and $y \not\succsim' x$, so that $x \sim y$ and $x \wr' y$; from which it follows that $x \succsim^\circ y$ and $x \succsim' y$. Therefore, because x, y are any, $\succsim^\circ = \succsim'^\circ$. □

Proposition 11.21. *The full extension of a relation is complete.*

Proof. Again, this follows by excluded middle. For let $\succsim$ on S x, y be any. Either $x \succsim y$ or $x \not\succsim y$, and either $y \succsim x$ or $y \not\succsim x$. If $x \succsim y$ and $y \succsim x$, then $x \sim y$. If $x \succsim y$ but $x \not\succsim y$, then $x \succ y$, and, if $x \not\succsim y$ but $y \succsim x$, then $y \succ x$. Finally, if $x \not\succsim y$ and $y \not\succsim x$, then $x \wr y$. As already noticed, $\succsim^\circ = \sim \cup \succ \cup \wr$, so that, in any case, either $x \succsim^\circ y$ or $y \succsim^\circ x$, — that is: $\succsim^\circ$ is complete. □

Proposition 11.22. *A relation is complete iff it is its own full extension.*

Proof. This follows immediately from the fact that completeness amounts to $\wr = \varnothing$, so that $\succsim^\circ = \succsim \cup \varnothing = \succsim$. □

The next proposition explains why the concept of full extension is of interest in the context of the study of extensions: the full extension of a relation contains all of its extensions.

Proposition 11.23. *The full extension is the largest extension of a relation (to its own underlying set, that is).*

Proof. Let $\succsim$ on S and $\succsim'$ an extension of $\succsim$ be any, and suppose for a contradiction that $\succsim^\circ \not\supseteq \succsim'$. Then there are x, y such that $x \succsim' y$ but $x \not\succsim^\circ y$. Because $\succsim^\circ \supseteq \succsim$ and $x \not\succsim^\circ y$, also $x \not\succsim y$, — so that $\succsim' \supset \succsim$. Now, by excluded middle, either $y \succsim x$ or $y \not\succsim x$. The former can't be the case because then we would have $y \succ x$ and, since $\succsim'$ is an extension of $\succsim$, also $y \succ' x$, — thus contradicting $x \succsim' y$. The latter can't be the case either, because then we would have $x \wr y$, but $\wr \subseteq \succsim^\circ$, — thus contradicting $x \not\succsim^\circ y$. It follows that, contrary to supposition, $\succsim^\circ \supseteq \succsim'$ after all. Because $\succsim'$ is any, $\succsim^\circ \supseteq \succsim'$ for all extensions $\succsim'$ of $\succsim$. □

The next two, quite trivial propositions show how the concepts of quasi-transitivity and full-extension relate.

Proposition 11.24. *A relation is quasi-transitive iff so is its full extension.*

Proof. Immediate from the fact that the asymmetric part of a relation and of its full extension are the same (proposition 11.18). □

Proposition 11.25. *If a relation is transitive, then its full extension is quasi-transitive. The converse is not true.*

Proof. Immediate from the previous proposition and from the fact that a transitive relation is quasi-transitive (proposition 11.7).

Being complete (therefore, its own full extension), the same relation used before as example that not every quasi-transitive relation is transitive suffice as a counterexample to the converse. □

The true interest of the concept of full extension lies in the next proposition, which shows that, in a sense, it allows us to ignore the difference between rationalization of choice function through best and through optimal options.

Proposition 11.26. *An option in a subset of the underlying set of a relation is optimal under it iff it is best under the full extension of the relation.*

Proof. Let $\gtrsim$ on S and $X \subseteq S$ be any.

First let $x \in X$ be $\gtrsim$-optimal. Then there are no $y \in X$ such that $y > x$. So, whatever $y \in X$ may be, either $x > y$, $x \sim y$ or $x \wr y$; thus, $x \gtrsim^\circ y$. Therefore, because y is any, $x \gtrsim^\circ y$ for all $y \in X$, and x is $\gtrsim^\circ$-best in X.

Now let $x \in X$ be $\gtrsim^\circ$-best. Then $x \gtrsim^\circ y$ for all $y \in X$. So, whatever $y \in X$ may be, either $x > y$, $x \sim y$ or $x \wr y$; thus, $y \not> x$. Therefore, because y is any, $y \not> x$ for all $y \in X$, and x is $\gtrsim$-optimal in X. □

Corollary 11.26.1. *A choice function is rationalizable through optimal options by a relation iff it is rationalizable through best options by the full extension of that relation.*

This proposition is of some theoretical importance. Because of it, in regard to rationalization of choice, we don't really need two distinct concepts: rationalization through best options suffices. From a practical point of view, an agent making choices through optimal options is tantamount to acting as if he or she were indifferent between every pair of options about which he or she hasn't formed preferences. This fact somewhat justifies why, in general, rational choice theorists haven't really cared much for rationalization through optimal options.

Definition 11.12 (Transitive closure)**.** Let $\gtrsim$ be a relation on S. The *transitive closure of* $\gtrsim$ is the relation $\gtrsim^* = \{\langle x, y\rangle \in S \times S :$ there are $x_0, \dots, x_n \in S$ such that $x_0 = x$, $x_k \gtrsim x_{k+1}$ for all $k < n$ and $x_n = y\}$.

Proposition 11.27. *The transitive closure of a relation includes the relation itself.*

Proof. Simply take $n = 1$ in the definition of transitive closure. □

Proposition 11.28. *The transitive closure of a relation is transitive. Indeed, it is the smallest transitive relation including it.*

Proof. For the first part, let $\succsim$ be a relation on S, let x, y, z be any, and suppose $x \succsim^* y$ and $y \succsim^* z$. Then there are $x_0, \ldots, x_n$ such that $x_0 = x$, $x_i \succsim x_{i+1}$ for all $i < n$ and $x_n = y$, and there are also $y_0, \ldots, y_m$ such that $y_0 = y$, $y_j \succsim y_{j+1}$ for all $j < m$ and $y_m = z$. So, simply by taking $x_{n+j} = y_j$, it follows that there are $x_0, \ldots, x_{n+m}$ such that $x_0 = x$, $x_k \succsim x_{k+1}$ for all $k < n + m$ and $x_{n+m} = z$, — that is, $x \succsim^* z$. Therefore, $\succsim^*$ is transitive.

For the second part, let $\succsim'$ be any transitive relation including $\succsim$, and let x, y be any such that $x \succsim^* y$. Then there are $x_0, \ldots, x_n$ such that $x_0 = x$, $x_k \succsim x_{k+1}$ for all $k < n$ and $x_n = y$. If $n = 1$, we have that $x \succsim y$, from which it follows immediately that $x \succsim' y$ since $\succsim \subseteq \succsim'$ by assumption; so suppose $n > 1$. Then, because $\succsim \subseteq \succsim'$ and $\succsim'$ is transitive, from $x_0 \succsim x_1$ and $x_1 \succsim x_2$ it follows that $x_0 \succsim' x_1$, $x_1 \succsim' x_2$ and, therefore, that $x_0 \succsim' x_2$. By repeating this same reasoning $n - 1$ times, we conclude that $x_0 \succsim' x_n$, that is, $x \succsim' y$. Therefore, $\succsim^* \subseteq \succsim'$. Because $\succsim'$ is any, $\succsim^* \subseteq \succsim'$ for all transitive $\succsim' \supseteq \succsim$. □

Proposition 11.29. *A relation is transitive iff it is its own transitive closure, — from which it follows trivially that the transitive closure of the transitive closure of a relation is the transitive closure itself.*

Proof. Immediate from the two previous propositions. □

Proposition 11.30. *The transitive closure of a super-relation of a relation includes the transitive closure of that relation.*

Proof. Let $\succsim$ on S, $\succsim' \supseteq \succsim$ and x, y be any, and suppose $x \succsim^* y$. Then there are $x_0, \ldots x_n$ such that $x_0 = x$, $x_k \succsim x_{k+1}$ for all $k < n$ and $x_n = y$. Because $\succsim' \supseteq \succsim$, it follows that $x_k \succsim' x_{k+1}$ for all $k < n$ as well, so that $x \succsim'^* y$. Therefore, $\succsim'^* \supseteq \succsim^*$. □

Proposition 11.31. *A relation is strongly acyclic iff its transitive closure is an extension of it.*

Proof. Let $\succsim$ on S be any.

First suppose $\succsim$ is not strongly acyclic. Then, there are $x_0, \ldots, x_n$ such that $x_k \succsim x_{k+1}$ for all $k < n$ but also $x_n \succ x_0$. Because $x_n \succsim x_0$ and $\succsim \subseteq \succsim^*$, it follows $x_n \succsim^* x_0$. However, because $x_k \succsim x_{k+1}$ for all $k < n$, it also follows $x_0 \succsim^* x_n$. Therefore, $\langle x_n, x_0 \rangle$ does not belong to the asymmetric part of $\succsim^*$, and $\succsim^*$ is not an extension of $\succsim$.

Now suppose $\succsim^*$ is not an extension of $\succsim$. Then, because $\succsim^* \supseteq \succsim$, this can only be the case if there are x, y such that $\langle x, y \rangle$ does not belong to the asymmetric part of $\succsim^*$ even though $x \succ y$. However, again because $\succsim \subseteq \succsim^*$, from this it follows that $x \succsim^* y$, and therefore, since $\langle x, y \rangle$ does not belong to the asymmetric part of $\succsim^*$, also that $y \succsim^* x$. This means that there are $x_0, \ldots, x_n$ such that $x_0 = y$, $x_k \succsim x_{k+1}$ for all $k < n$ and $x_n = x$; but $x \succ y$ by supposition, that is, $x_n \succ x_0$, — a $\succsim$-improper cycle. Therefore, $\succsim$ is not strongly acyclic. □

Corollary 11.31.1. *If the asymmetric part of the transitive closure of a relation is the same as the transitive closure of its asymmetric part, then the relation is strongly acyclic.*

Proof. Let $\succsim$ on S be any such that $\bar{\succsim}^* = \succ^*$. Because by definition $\succ^* \supseteq \succ$, also $\bar{\succsim}^* \supseteq \succ$, from which it follows (because $\succsim^* \supseteq \succsim$ by definition as well) that $\succsim^*$ is an extension of $\succsim$, therefore strongly acyclic by the previous proposition. □

Finally we get to the two main theorems of the chapter, Szpilrajn's theorem and Suzumura's theorem. Szpilrajn's theorem, in its original form, shows any partial ordering can be extended to an ordering. In his seminal work on social choice theory Arrow stated, without proof, the same to be the case for quasi-orderings as well. The proof is indeed a simple modification of Szpilrajn's original proof, but it was only to appear in print much later in a paper by Bengt Hansson. Suzumura's theorem, however, generalizes Szpilrajn's result further still by showing the existence of an ordering extension to be actually *equivalent* to the concept of consistency of a relation.

For Szpilrajn's theorem, we need again to make use of Zorn's lemma.

Proposition 11.32 (Szpilrajn 1930 [67], Hansson 1969 [28])**.** *If a relation is a quasi-ordering, then it has an extension that is an ordering.*

Proof. Let $\succsim$ on S be a quasi-ordering, and let Q be the set of all quasi-orderings that are extensions of $\succsim$ to S. Clearly, $Q \neq \varnothing$, as $\succsim$ is an extension of itself and, therefore, $\succsim \in Q$. Given X a chain in Q under $\subseteq$, let $\succsim^X = \bigcup X$. Clearly, $\succsim^X$ is reflexive, for $\succsim \subseteq \succsim^X$ and $\succsim$ is reflexive. As for transitivity, let x, y, z be any, and suppose $x \succsim^X y$ and $y \succsim^X z$. Because $x \succsim^X y$, there is a $\succsim' \in X$ such that $x \succsim' y$ and, similarly, because $y \succsim^X z$, there is a $\succsim'' \in X$ such that $y \succsim'' z$. Now, because X is a chain, either $\succsim' \subseteq \succsim''$ or $\succsim'' \subseteq \succsim'$. Without loss of generality, suppose the latter is the case. Then it follows that also $y \succsim' z$ and, because $\succsim'$ is a quasi-ordering, that $x \succsim' z$. But, by definition, $\succsim' \subseteq \succsim^X$, and so, because x, y, z are any, $\succsim^X$ is transitive. Therefore, $\succsim^X$ is a quasi-ordering. Moreover, $\succsim^X$ is an extension of $\succsim$. For suppose not. Then there are x, y such that $x \succ y$ but $x \not\succ^X y$. Because $\succsim \subseteq \succsim^X$, and therefore $x \succsim^X y$, this is only possible if $x \sim^X y$. But then there must be a $\succsim' \in X$ such that $y \succ x$ — which is impossible because, by assumption, $\succsim'$ is an extension of $\succsim$. So it follows that $\succsim^X$ is an upper bound of X in Q under $\subseteq$; and, because X a chain in Q under $\subseteq$ is any, by Zorn's lemma we have that Q has a maximal element — $\succsim^+$, say. It remains to be shown that $\succsim^+$ is an ordering.

Because $\succsim^+ \in Q$, it is a quasi-ordering, therefore reflexive and transitive. As for completeness, let $s, t \in S$ be any and suppose for a contradiction that $s \wr^+ t$ (from which it follows that $s \neq t$, since $\succsim^+$ is reflexive). Let $\succsim^{++} = \succsim^+ \cup \{\langle x, y\rangle \in S \times S : x \succsim^+ s \text{ and } t \succsim^+ y\}$. We show that $\succsim^{++}$ is a quasi-ordering extending $\succsim$ to S that is greater than $\succsim^+$.

Obviously enough, $\succsim^+ \subseteq \succsim^{++}$ by definition. Moreover, because $s \succsim^+ s$ and $t \succsim^+ t$, in particular we have that $s \succsim^{++} t$. Therefore, because by supposition $s \wr^+ t$, it follows that $\succsim^+ \subset \succsim^{++}$. That $\succsim^{++}$ is reflexive is trivial, as $\succsim^+ \subseteq \succsim^{++}$ and $\succsim^+$ is reflexive. As for transitivity, again let x, y, z

be any and suppose that $x \succsim^{++} y$ and $y \succsim^{++} z$. If $x \succsim^{+} y$ and $x \succsim^{+} z$, then $x \succsim^{+} z$ because $\succsim^{+}$ is transitive, and therefore $x \succsim^{++} z$ because $\succsim^{+} \subset \succsim^{++}$. If $x \succsim^{+} y$ but $y \succsim^{++} z$ only, then we have that $y \succsim^{+} s$ and $t \succsim^{+} z$. Because $\succsim^{+}$ is transitive, $x \succsim^{+} y$ and $y \succsim^{+} s$, it follows that $x \succsim^{+} s$ and, therefore, that $x \succsim^{++} z$ by definition. The argument for $x \succsim^{++} y$ only but $y \succsim^{+} z$ is similar to the one for the previous case. Finally, it is impossible that $x \succsim^{++} y$ only and $y \succsim^{++} z$ only, for then we would have, in particular, that $t \succsim^{+} y$ and $y \succsim^{+} s$, and therefore, by the transitivity $\succsim^{+}$, that $t \succsim^{+} s$, thus contradicting the assumption that $s \wr^{+} t$. Therefore, because x, y, z are any, $\succsim^{++}$ is transitive, and thus a quasi-ordering.

That $\succsim^{++}$ is indeed an extension of $\succsim$ is easy to see. Clearly, $\succsim \subseteq \succsim^{++}$, since $\succsim \subseteq \succsim^{+} \subseteq \succsim^{++}$. So let x, y be any and suppose, again for a contradiction, that $x \succ y$ but $x \not\succ^{++} y$. Then, of necessity, $x \sim y$. Because $\succsim^{+}$ is an extension of $\succsim$, it can't be that $y \succsim^{+} x$, and so $y \succsim^{++} x$ only. Therefore, $y \succsim^{+} s$ and $t \succsim^{+} x$, — from which it would follow, since $x \succ^{+} y$, that $t \succsim^{+} s$, thus contradicting again the assumption that $s \wr^{+} t$. So we have that $x \succ^{++} y$. Because x, y are any, $\succsim^{++}$ is an extension of $\succsim$.

But then $\succsim^{++}$ is quasi-ordering that extends $\succsim$, although also $\succsim^{++} \supset \succsim^{+}$ — which contradicts the maximality of $\succsim^{+}$ is Q. Therefore, $\succsim^{+}$ is complete, and thus an ordering extending $\succsim$. □

We shall routinely make implicit use of Szpilrajn's extension theorem, as follows. Whenever I present a transitive relation on a subset of the set of options, it is to be assumed that I mean some ordering extension of it, which is known to exist due of the theorem. This spares us the trouble of specifying the whole relation, which is quite useful because in most occasions the rest of the relation won't really matter for our purposes, but only that it is an ordering.

Proposition 11.33 (Suzumura 1976 [66]). *A relation has an extension which is an ordering iff it is strongly acyclic.*

Proof. Let $\succsim$ on S be any.

First suppose it is not strongly acyclic. Then, there is a $\succsim$-improper cycle. So let $\succsim'$ an extension of $\succsim$ be any. Because $\succsim' \supseteq \succsim$ and $\succ' \supseteq \succ$, it follows trivially that any $\succsim$-improper cycle is also a $\succsim'$-improper cycle. Therefore, $\succsim'$ is also not strongly acyclic, thus not transitive, and thus also not an ordering.

Now suppose $\succsim$ is strongly acyclic. Let $\succsim^{+} = \succsim^{*} \cup \Delta$. It is easy to see that $\succsim^{+}$ is a quasi-ordering. Reflexivity is given by definition. As for transitivity, let x, y, z be any and suppose that $x \succsim^{+} y$ and $x \succsim^{+} z$. If both $x \succsim^{*} y$ and $y \succsim^{*} z$, then $x \succsim^{+} z$ because $\succsim^{*} \subseteq \succsim^{+}$. If $x \succsim^{*} y$ and $y \Delta z$, then $y = z$ and thus $x \succsim^{+} z$; and similarly if $x \Delta y$ and $y \succsim^{*} z$. Obviously enough, if $x \Delta y$ and $y \Delta z$, then $x = y = z$, and thus $x \succsim^{+} z$ is given by definition. In any case, $x \succsim^{+} z$, and therefore, because x, y, z are any, $\succsim^{+}$ is transitive, and thus also a quasi-ordering. It follows from the previous proposition that it has an extension that is an ordering. So, if we can show that $\succsim^{+}$ is itself an extension of $\succsim$ we are done, as it has already been proven that an extension of an extension of a relation is itself an extension of that relation.

That $\succsim^+ \supseteq \succsim$ is trivial, as $\succsim^+ \supseteq \succsim^* \supseteq \succsim$. So let x, y be any and suppose for a contradiction that $x \succ y$ but $x \not\succsim^+ y$. As usual, because $\succsim^+ \supseteq \succsim$, this is only possible if $x \sim^+ y$, and thus if $y \succsim^+ x$. Clearly, $x \not\Delta y$, since $x \succ y$ and thus $x \neq y$. If, however, $y \succsim x$, then $\succsim$ wouldn't be strongly acyclic as assumed. Therefore, $x \succsim^+ y$; and, because x, y are any, $\succsim^+$ is indeed an extension of $\succsim$, — from which it immediately follows by the previous proposition that $\succsim$ has an extension which is an ordering. □

Corollary 11.33.1. *A relation has an extension which is a strict ordering iff it is asymmetric and acyclic.*

Proof. Follows immediately from the theorem and the fact that a strict ordering is asymmetric, together with the fact that an asymmetric and acyclic relation is strongly acyclic (proposition 11.7). □

Chapter 12

Social choice and Arrow's impossibility theorem

In this chapter, I present with proof Arrow's much celebrated impossibility theorem, which is arguably the foundational result of social choice theory. I also present Hansson's alternative proof of that theorem, which requires the order-theoretic concept of ultrafilter discussed in chapter 10. This is instrumental to a proper understanding of the role played in Arrow's impossibility theorem by the concept of decisiveness, *cf.* section 3.3, which on its turn plays a central role in the results presented in this dissertation.

This chapter presupposes the content of chapter chapter 3.

12.1 Aggregation and welfare

Let us begin by providing a few examples of social choice rules (and classes of such rules) so as to make the notion of a PAF, *cf.* section 3.2, somewhat more concrete to the reader. (For the definition of the unanimity rule, recall the definition of the full extension of a relation from section 11.3.)

Unanimity rule The PAF defined by: $x \succsim^p y$ if $xP_I^{p\circ}y$, for all p

Pairwise simple majority rule The PAF defined by: $x \succsim^p y$ iff $|\{i \in I : x \succsim_i^p y\}| \geq |\{i \in I : y \succsim_i^p x\}|$, for all p

Taboo rules Any PAF defined by: $x \succsim^p y$ iff $x \succsim y$ for some given $\succsim$, for all p

Dictatorial rules Any PAF defined by: $x \succsim^p y$ iff $x \succsim_i^p y$ for some given i, for all p

Unanimity rule (also known as the Pareto rule) settles an option socially preferred to another iff everyone in the society prefers that option to the other. Pairwise simple majority rule, which we have already discussed in full in section 3.2, settles an option socially preferred to another iff the number of individuals who prefer the former to the latter is at least as great as that of those who prefer the latter to the former. Taboo rules simply ignore the preferences of the individuals and settle the social preferences as some previously established preferences. Notice that this is a class of rules, not a single rule: for each given preferences, we have a different taboo rule. Similarly, autocratic rules are those which settle the social preferences as those of some individual in the society, thus ignoring the preferences of every other individual.

These are by no means all the rules that may be of interest, and evidently nor are they the most complex ones. They are, however, quite representative of rules that have been used extensively in political processes through human history. That being the case, it is natural that one major lines of research in social choice be the study of the properties of particular rules or classes of rules. That's the subject to which we turn now.

First, let us present the class of PAFs with which we shall be concerned in Arrow's theorem, which are the so-called "social welfare" functions.

Definition 12.1 (Social welfare functions)**.** Let A be a PAF for I, S. A is a *social welfare function* (or *SWF*, for short) if $\mathsf{imag}(A) \subseteq \mathcal{O}(S)$.

Obviously enough, every SWF is an SDF (*cf.* section 3.4), but not the converse.

SDFs and SWF can be seen as two extremes of rationality requirements on social preferences. As we know from the previous chapter, acyclic and complete preferences are necessary and sufficient for there to be a rationalizable finitely decided choice function. Intuitively, therefore, SDFs merely require that the social preferences be reasonable enough so that the society may reach a decision on every matter involving at most finitely many options. This is a reasonable but weak requirement, if a PAF is to lead to justifiable social decisions at all. At the other extreme we find SWFs, which require *full rationality* of the social preferences, in the sense that it assumes the social preferences must be a complete scale of socially better and worse possible outcomes.

12.2 Conditions on PAFs

In this section, all of the conditions on a PAF discussed in section 3.2 are formally (albeit summarily) presented, both for the sake of ease of reference and so that the reader may appreciate the logical connections among them in preparation for the discussion of Arrow's impossibility theorem.

Definition 12.2 (Conditions on a PAF)**.** Consider the following conditions a PAF A for a society I, S may satisfy, for all $p \in \mathsf{dom}(A)$.

Unrestricted domain (UD) $\mathcal{O}(S)^I \subseteq \mathsf{dom}(A)$

Weak Pareto (WP) If $xP_I^p y$, then $x \succ^p y$

Pareto inclusivity (PI) If $x\breve{P}_I^p y$, then $x \succsim^p y$

Strong Pareto (SP) If $x\breve{P}_I^p y$, then $x \succ^p y$

Independence of irrelevant alternatives (IIA) If both $x \succsim_i^{p'} y$ iff $x \succsim_i^p y$ and $y \succsim_i^{p'} x$ iff $y \succsim_i^p x$, for all i, then both $x \succsim^{p'} y$ iff $x \succsim^p y$ and $y \succsim^{p'} x$ iff $y \succsim^p x$, for all $p' \in \mathsf{dom}(A)$

Neutrality (N) If for some permutation π of the options it is the case that both $\pi(x) \succsim_i^{p'} \pi(y)$ iff $x \succsim_i^p y$ and $\pi(y) \succsim_i^{p'} \pi(x)$ iff $y \succsim_i^p x$, for all i, then both $\pi(x) \succsim^{p'} \pi(y)$ iff $x \succsim^p y$ and $\pi(y) \succsim^{p'} \pi(x)$ iff $y \succsim^p x$, for all $p' \in \mathsf{dom}(A)$

Non-dictatorship (ND) There is no *dictator*, — that is, an i such that, for all p and all x, y, if $x \succ_i^p y$, then $x \succ^p y$

Anonymity (A) If for some permutation π of the individuals it is the case that $\succsim_{\pi(i)}^{p'} = \succsim_i^p$ for all i, then $\succsim^{p'} = \succsim^p$, for all $p' \in \mathsf{dom}(A)$

Non-negative responsiveness (NNR) If x and p' are such that, for all i, y, if $x \succsim_i^p y$, then $x \succsim_i^{p'} y$; if $x \succ_i^p y$, then $x \succ_i^{p'} y$, and, for all z, w, $z \neq x$, $w \neq x$, $z \succsim_i^{p'} w$ iff $z \succsim_i^p w$, then, if $x \succsim^p y$, then $x \succsim^{p'} y$, for all x

Positive responsiveness (PR) If p' is such that, for all i, $x \succ_i^{p'} y$ if $x \succ_i^p y$ and $x \succsim_i^{p'} y$ if $x \sim_i^p y$, and there is a j such that $x \sim_j^p y$ but $x \succ_i^{p'} y$ or $y \succ_j^p x$ but $x \succsim_j^{p'} y$, then, if $x \succsim^p y$, then $x \succ^{p'} y$, for all $p' \in \mathsf{dom}(A)$

Some informal explanation of the four not previously discussed conditions is in order here.

Condition CS, citizen's sovereignty, essentially require that the results of elections aren't rigged or in some sense settled in advance. Any option can end up being socially judged better than any other.

Condition PI, Pareto inclusivity requires that, whenever the individuals unanimously prefer (but not necessarily strictly prefer) an option over another, so does the society as a whole. Just as WP, it can be seen as optimization principle, or principle of welfare, but also as a principle of popular sovereignty in the sense that, under either WP or PI, unanimity is guaranteed to trump any other considerations regarding the relevant options. (Indeed, notice that, under UD, both WP and PI trivially imply CS).

Condition SP, strong Pareto, requires something that is much more contentious for a preferences aggregation, — namely, that strict preferences of individuals be socially met whenever they are unopposed. This principle won't be of any particular interest to us, though, and is here presented mostly for the sake of contrast, and also to give the reader an additional opportunity of developing intuitions about the subject of social choice.

Condition NNR, non-negative responsiveness (also called *positive association*, as Arrow originally named it) requires that, if the preferences of the individuals remain otherwise unchanged except perhaps that the relative ranking of an option in each of their preferences is either the same as before or higher, then, for whatever options to which the social preferences previously judged it superior, the social preferences should still do so. In other words, NNR states that an option shouldn't be socially ranked any lower than before when the only thing that has changed in the preferences of the individuals is that now some of them rank it higher. Evidently, as Arrow himself points out, if one is concern with welfare at all, this is a necessary requirement a PAF should satisfy. Notice that condition NNR coupled with CS implies WP.

12.3 Arrow's impossibility theorem

As already discussed in section 3.2, conditions WP and ND are perhaps the most natural ones to impose on a PAF that is supposed to be democratic, — unless, of course, the "society" consists of a *single* individual. Indeed, if a PAF for such a society satisfies WP, then that individual is of necessity a dictator. Moreover, if there's only one possible outcome, then, in a sense, there's no real choice to be socially made; and, if only two, then conditions such as IIA, as well as requirements such as transitivity of the social preferences, are trivially satisfied by any PAF yielding complete social preferences. It is easy but tedious to check that pairwise simple majority rule satisfy *all* of the conditions discussed above, *and* it's also an SWF for a society with two or more individuals and only two possible outcomes. As has been already pointed out, *cf.* section 3.2, pairwise simple majority rule satisfies conditions UD, WP, ND and IIA. However, due to being subjected to Condorcet's paradox, it is not an SWF.

In view of these observations, let us call a society with at least two individuals and at least three options a *proper* one. The question, then, this: is there any good alternative to pairwise simple majority rule for a democratic but welfaristic society? More precisely: is there a PAF satisfying at least some of the conditions that characterize pairwise simple majority rule but yields fully rational social preferences?

As Arrow's justly celebrated theorem shows, for proper finite societies, even the arguably mild conditions UD, WP and IIA cannot be jointly satisfied for SWFs. Simply put: the requirement that the social preferences be fully rational is an exceedingly strong one, as far as the requirements of a democratic aggregation procedure are concerned. Democracy and welfarism seem to make irreconcilable demands on social choice procedures.

In order to present a proof of Arrow's impossibility theorem, let us first recall from section 3.4 the concept concept of decisiveness, and also present the concept of almost decisiveness, which plays a central role in the proof of the Arrow's theorem.

Definition 12.3 (Decisiveness [2])**.** Let A be a PAF for I, S, J a coalition and x, y options. *J is almost decisive for x over y under A* if, for all p, if $xP^p_J y$ and $yP^p_{\overline{J}} x$, then $x >^p y$. *J is decisive for x over y under A* (or *decisive*, for short) if, for all p, if $xP^p_J y$, then $x >^p y$.

Informally, it is to say a coalition is almost decisive for a pair of options if, whenever every individual in the coalition holds the same strict preference about the pair, and albeit every individual not in the coalition holds the opposite preference, the social preference agrees with former.

Remark 12.1. Notice, by the way, that conditions WP and ND can be restated in terms of decisiveness.

Weak Pareto (WP) The full coalition is decisive for x over y under A, for all x, y

Non-dictatorship (NP) There is no i such that coalition i is decisive for x over y under A, for all x, y

The heavy lifting, so to speak, of Arrow's impossibility theorem is done by the following lemma, which shows that, for an SWF satisfying conditions UD, WP and IIA, decisiveness is an "all or nothing" matter: if a coalition is decisive for even a single pair, it must be decisive for all of them. It is not too complicated a lemma, as the reasoning involved is quite straightforward; but it is a laborious one, as it requires checking many similar but distinct cases.

Proposition 12.1 (Contagion lemma - Arrow 1951 [2])**.** *If an SWF for a proper society satisfies conditions UD, WP and IIA, then, if a coalition is almost decisive on a single issue, then it is decisive on every issue.*

Proof. Let A an SWF for I, S a proper society be any satisfying conditions UD, WP and IIA. Let $x \neq y$ and J be any, and suppose J is almost decisive for x against y under A.

Let z be any such that $z \neq x, z \neq y$.

By condition UD, let p be any such that $xP^p_J y$ and $yP^p_I z$, but $yP^p_{\overline{J}} x$. That is: everyone in J strictly prefers x to y, everyone in the society strictly prefers z to y, but everyone not in J strictly prefers y to x. Also let the preferences of the $j \in \overline{J}$ on x, z be any. Now, because J is almost decisive for x against y, because $xP^p_J y$ and because $yP^p_{\overline{J}} x$, we have it that $x \succ^p y$; because $yP^p_I z$, by condition WP we have it that $y \succ^p_I z$; and, because $x \succ^p y$, $y \succ^p_I z$ and A is an SWF, with $\succsim^p$ being therefore transitive, we have it that $x \succ^p z$.

But $xP^p_J y$; and, because $yP^p_I z$, in particular $yP^p_J z$. Therefore, because P^p_J is transitive, we also have it that $xP^p_J z$. Notice however that, regarding x, z, profile p is such that $xP^p_J z$ and also, by assumption, such that the preferences of the $j \in \overline{J}$ on x, z are any. But $x \succ^p z$. So, by condition IIA, for any profile p' that is just like p regarding x, z, it must be the case that $x \succ^{p'} z$, — from which it immediately follows that J is decisive for x against z under A. Because z is any distinct from x and y, if J is almost decisive for x against y under A, then it is decisive for x against z for any z (any $z \neq x$, $z \neq y$ that is). Moreover, because decisiveness implies almost decisiveness, it also follows that, if J is almost decisive for x against y under A, then it is almost decisive for x against z as well.

Now, by condition UD, let p be any such that $xP^p_J y$ and $zP^p_I x$, but $yP^p_{\overline{J}} x$. Just like before, let the preferences of the $j \in \overline{J}$ on y, z be any. Because J is almost decisive for x against

y, because xP^p_Jy and because $yP^p_{\bar{J}}x$, we have it again that $x >^p y$; because zP^p_Ix, by condition WP we have it that $z >^p_I x$; and, because $z >^p x$, $x >^p_I y$ and A is an SWF, we have it that $z >^p y$.

But xP^p_Jy; and, because zP^p_Ix, in particular zP^p_Jx. Therefore, because P^p_J is transitive, we also have it that zP^p_Jy. From this, by the same reasoning as in the previous case, it follows that J is decisive for z against y under A, and therefore that if it is almost decisive for x against y under A, then it is decisive for z against y for any $z \neq x$, $z \neq y$. Again, it is almost decisive for z against y as well.

Now that we have found out that J is almost decisive not only for x against y under A but for x against z and for z against y as well, we proceed by reasoning from these new pair exactly as in the two cases above: that is, we place the third option either on the top or or the bottom of the strict rank of everyone in J, assume that everyone not in J has opposite preferences regarding the pair on which J is decisive, and also assume that everyone in the society has the same preferences regarding the third option and the next one in the rank of everyone in J, letting the other preferences of everyone not in J be any. This leads to the following implications, the first two summing up what we have shown so far.

- Almost decisive for x against $y \Longrightarrow$ decisive for x against z
- Almost decisive for x against $y \Longrightarrow$ decisive for z against y
- Almost decisive for x against $z \Longrightarrow$ decisive for x against y
- Almost decisive for x against $z \Longrightarrow$ decisive for y against z
- Almost decisive for z against $y \Longrightarrow$ decisive for z against x
- Almost decisive for y against $z \Longrightarrow$ decisive for y against x

If the society has only three options, then we are done. Otherwise, now that that J turns out to be decisive on every issue comprising x, y, z for whatever $z \neq x$, $z \neq y$, it is a simple matter of repeating the same reasoning in order to extend this conclusion to any option whatsoever. More precisely, let w be any such that $w \neq x$, $w \neq z$. Then we have the following implications.

- Almost decisive for z against $x \Longrightarrow$ decisive for z against w
- Almost decisive for x against $z \Longrightarrow$ decisive for w against z

In order to see that we have indeed covered all possible pairs, let u, v be any. If $u = x$ and $v = y$, or if $u = y$ and $v = x$, then, because J is decisive both for x against y and for y against x, we have it that J is decisive for u against v. If, say, $u = x$ but $v \neq y$, then, because J is decisive for x against z and z is any such that, in particular, $z \neq y$, again we have it that J is decisive for u against v. The reasoning for $u = y$ and $v \neq x$ is similar, given that J is decisive also for y against z and z is any such that, in particular, $z \neq x$. Finally, if $u \neq x$ and $v \neq y$ or $u \neq y$ and $v \neq x$, then, because J is decisive both for w against z and z against w, because w is any such that $w \neq x$ and z is any such that, in particular, $z \neq y$, once more we have it that J is decisive for u against v. So, whatever u, v may be, indeed we have it that J is decisive for u against v. □

Finally, we get to the theorem itself. I present here two proofs for it: Arrow's original proof, and the proof by Hansson, which highlights the role played by the concept of decisiveness in establishing the theorem by showing that, if an SWF satisfies conditions UD, WP and IIA, then the set of the coalitions that are globally decisive under it constitutes an ultrafilter.

Proposition 12.2 (Arrow's impossibility theorem – Arrow 1951 [2])**.** *For a society with finitely many individuals and at least three possible outcomes, there is no SWF satisfying conditions UD, WP, IIA and ND.*

Proof. Let A an SWF for I, S be any satisfying conditions UD, WP and IIA, and suppose for a contradiction that it also satisfies ND. Because A satisfies condition WP, the coalition of all individuals is globally decisive. Because $I \neq \varnothing$ and I is finite, by the well-ordering principle there is a smallest $n \leq |I|$ for which there is at least one coalition $J \neq \varnothing$, with $|J| = n$, such that J is globally decisive. Because by supposition A satisfies ND, of necessity $n > 1$, for otherwise the single $j \in J$ would be a dictator and a contradiction would immediately ensue.

So let J be the smallest of such coalitions (or one of the smallest, if there is more than one), and let $J' \subseteq J$ be a singleton. Clearly, because $|J| > 1$, both J' and $J \smallsetminus J'$ are non-empty. By condition UD, let p be any such that: $xP^p_{J'}y$ and $yP^p_{J'}z$; that $zP^p_{J \smallsetminus J'}x$ and $xP^p_{J \smallsetminus J'}y$, and that, moreover, $yP^p_{\overline{J}}z$ and $zP^p_{\overline{J}}x$ (provided $\overline{J}$ is non-empty, that is; but the proof goes through just the same if it is).

Because J is decisive for x against y and both $xP^p_{J'}y$ and $xP^p_{J \smallsetminus J'}y$ (so that, therefore, $xP^p_J y$ as well), $x \succ^p y$. Now, it can't be the case that $z \succ^p y$ for, if it were, then, because (by transitivity) $zP^p_{J \smallsetminus J'}y$ and both $yP^p_{J'}z$ and $yP^p_{\overline{J}}z$ (so that, therefore, $yP^p_{J' \cup \overline{J}}z$ as well), it would follow by condition IIA that $J \smallsetminus J'$ is almost decisive for y against z, thus contradicting the assumption that there are no almost decisive coalitions smaller than J. This being so, because A is an SWF, and therefore $\succsim^p$ complete, it must be the case that $y \succsim^p z$; but then, by the transitivity of $\succsim^p$, it must be the case that $x \succsim^p z$. Notice, however, that $xP^p_{J'}y$ and $yP^p_{J'}z$, and therefore $xP^p_{J'}z$ by transitivity, while both $zP^p_{J \smallsetminus J'}x$ and $zP^p_{\overline{J}}x$ (so that $zP^p_{\overline{J'}}x$), — from which it follows, again by condition IIA, that J' is almost decisive for z against x. But then, by the previous proposition, the single $j \in J'$ is a dictator, thus contradicting the assumption that A satisfies ND. □

Alternative proof – Hansson 1972 [27]. Again, let A an SWF for I, S be any satisfying conditions UD, WP and IIA, and suppose for a contradiction it also satisfies ND. Let U denote the set of all coalitions in the society that are decisive for all pairs. We show U is a proper ultrafilter, *cf.* section 10.2.1.

Clearly, $I \in U$, by WP, so that U contains the top. Moreover, it follows trivially from the definitions of decisiveness and of U that it is monotonic.

As for not containing the bottom, notice that, if $\varnothing \in U$, then, because both $xP^p_{\varnothing}y$ and $yP^p_{\varnothing}x$ for any x, y, p, we would have both $x \succ^p y$ and $y \succ^p x$ for any x, y, p as well, thus contradicting the asymmetry of $\succ^p$.

To see that U is closed under binary intersection, let $J, J' \in U$ be any, and let x, y, p be any such that $xP^p_{J\cap J'}y$. Now let $z \neq x$, $z \neq y$ be any, and let p' be any such that $\succsim^{p'}_i \cap (\{x,y\} \times \{x,y\}) = \succsim^p_i \cap (\{x,y\} \times \{x,y\})$ for all i and such that: $xP^{p'}_{J\smallsetminus J'}z$ and $yP^{p'}_{J\smallsetminus J'}z$; $xP^{p'}_{J\cap J'}z$ and $zP^{p'}_{J\cap J'}y$ (so that, by the transitivity of $P^{p'}_{J\cap J'}$, also $xP^{p'}_{J\cap J'}y$, as expected), and $zP^{p'}_{J'\smallsetminus J}x$ and $zP^{p'}_{J'\smallsetminus J}y$. Because $J \in U$ and $xP^{p'}_J z$, we have it that $x \succ^{p'} z$; and similarly, because $J' \in U$ and $zP^{p'}_{J'}y$, we have it that $z \succ^{p'} y$. So, because $x \succ^{p'} z$ and $z \succ^{p'} y$, being A an SWF, by the transitivity of $\succsim^p$ it follows that $x \succsim^{p'} y$. Therefore, because $z \neq x$, $z \neq y, p'$ are any, $x \succ^p y$ by IIA; and, because x, y, p are any, $J \cap J' \in U$ indeed.

Regarding the fullness of U, let J be any and suppose $J \notin U$. Then there are x, y such that J is not decisive for x against y, which implies that there is a p such that $x \not\succ^p y$ (and, thus — because $\succsim^p$ is complete —, a p such that $y \succsim^p x$), even though $xP^p_J y$. So let $z \neq x$, $z \neq y$ be any, and let p' be any such that $\succsim^{p'}_i \cap (\{x,y\} \times \{x,y\}) = \succsim^p_i \cap (\{x,y\} \times \{x,y\})$ for all i and such that $xP^{p'}_J z$ and $zP^{p'}_J y$ (so that, by the transitivity of $P^{p'}_J$, also $xP^{p'}_J y$, as expected), and $zP^{p'}_{\overline{J}}x$ and $zP^{p'}_{\overline{J}}y$. Then, by WP, $z \succ^{p'} y$ and, by IIA, $y \succsim^{p'} x$, so that, by the transitivity of $\succsim^{p'}$, also $z \succ^{p'} x$. Notice however that $xP^{p'}_J z$ and $zP^{p'}_{\overline{J}}x$. Therefore, again by IIA, $z \succ^{p'} x$ for any p'' such that $xP^{p''}_J z$ and $zP^{p''}_{\overline{J}}x$, — which means $\overline{J}$ is almost decisive for z against x. Therefore, by the contagion lemma, $\overline{J}$ is decisive for x against y for every x, y, and thus $\overline{J} \in U$ by the definition of U.

So U is a subset space on I which contains the top, does not contain the bottom, is closed under supersets and binary intersection and is full, thus an ultrafilter on I. But I is finite, and so U is fixed. Therefore, there is an i such that $\{i\} \in U$, — which means, contradicting ND, that i is a dictator. □

The construction of the ultrafilter of decisive coalitions is of importance because it allows us to understand that Arrow's conditions on an SWF are not *logically* inconsistent, since they *can* be jointly satisfied at least for infinite societies, as the next theorem reveals. In its proof, given a preferences profiles $p \in \mathcal{O}(S)^I$, R^p stands for the relation defined by: xR^py iff $x \succsim^p_i y$ for all i.

Proposition 12.3 (Fishburn 1970 [18], Kirman and Sondermann 1972 [35], Hansson 1972 [27])**.** *There are SWFs satisfying conditions UD, WP, IIA and ND for proper infinite societies.*

Proof. Let I, S be a proper infinite society, and let U be a non-principal ultrafilter on I. Let $A_U : \mathcal{O}(S)^I \longmapsto \mathcal{O}(S)$ be defined by $x \succsim^p y$ iff $xR^p_J y$ and $J \in U$, for any x, y and $p \in \mathcal{O}(S)^I$.

We first show $\succsim^p$ is an ordering for any $p \in \mathsf{dom}(A_U)$. So, let p be any.

Clearly, $x \succsim^p x$ for any x, since $xR^p_I x$ for any x and $I \in U$ by the definition of ultrafilter, which means $\succsim^p$ is reflexive.

As for transitivity, let x, y, z be any and suppose $x \succsim^p y$ and $y \succsim^p z$. Then, because $x \succsim^p y$, there's a $J \in U$ such that $xR^p_J y$, and similarly a $J' \in U$ such that $yR^p_{J'} z$. Because U is an ultrafilter, $J \cap J' \in U$. Therefore, by the definition of R^p, we have that $xR^p_{J \cap J'} y$ and $yR^p_{J \cap J'} z$, from which it follows by the transitivity of R that $xR^p_{J \cap J'} z$ as well. Thus, because $J \cap J' \in U$, also $x \succsim^p z$. Because x, y, z are any, $\succsim^p$ is transitive.

Regarding connectedness, let x, y, $x \neq y$ and p be any. Let $J = \{i \in I : x \succsim^p_i y\}$. By excluded middle, either $J \in U$ or not. If $J \in U$, then, because $xR^p_J y$, it follows $x \succsim^p y$ by the definition of A_U. If $J \notin U$, then $\overline{J} \in U$ because U is an ultrafilter. Let $i \in \overline{J}$ be any. By the definition of J, if $i \in \overline{J}$, then $x \not\succsim^p_i y$, which implies $y \succ^p_i x$ — and, thus, $y \succsim^p_i x$ — by the connectedness of $\succ^p_i$. Because i is any, $y \succsim^p_i x$ for all $i \in \overline{J}$, — that is, $yR^p_{\overline{J}} x$. Therefore, because $\overline{J} \in U$, it follows $y \succsim^p x$; and thus, because x, y are any, $\succsim^p$ is connected.

As for the conditions satisfied by A_U, UD is satisfied by definition, IIA follows trivially from the fact that by definition $x \succsim^p y$ depends only on the individuals' preferences about x, y in p for whatever x, y, p, and ND is equally trivial, given that U is a non-principal ultrafilter. Regarding WP, let x, y, p be any, and suppose $xP^p_I y$. Then, because R^p_I is an extension of P^p_I, also $xR^p_I y$, — from which it follows by the definition of A_U that $x \succsim^p y$. Now, also by the definition of A_U, $y \succsim^p x$ as well iff there's a $J \in U$ such that $yR^p_J x$. But, because $xR^p_I y$, the only such J is the empty set, — and $\varnothing \notin U$. It follows $y \not\succsim^p x$, and, so, $x \succ^p x$. Therefore, if $xP^p_I y$, then $x \succ^p y$. Because x, y, p are any, A_U satisfies WP. □

Remark 12.2. It is noteworthy that A_U so defined satisfies conditions other than (and some of them stronger than) those of Arrow's theorem [27]. Indeed, it satisfies PI by definition, and it satisfies N trivially because $x \succsim^p y$ is defined for all x, y, p uniformly. It is also easy to see it satisfies NNR. No such A_U can satisfy A, however, for the following reason. Because I is assumed to be infinite, it has a proper subset — J, say — such that $|J| = |I|$, and such that $|\overline{J}| = |I|$ as well, *cf.* [19]. Because U is an ultrafilter, either $J \in U$ or $\overline{J} \in U$. Suppose without loss of generality the former, and let $f : J \longmapsto \overline{J}$ be one-to-one. Clearly, $\pi = f \cup f^{-1}$ is a permutation of I. Now let x, y, p be any such that $xP^p_J y$ but $yP^p_{\overline{J}} x$, and let p' be such that $yP^{p'}_J x$ but $xP^{p'}_{\overline{J}} y$. Then, as the reader can easily check for him or herself, $x \succ^p y$ but $y \succ^{p'} x$ by the definition of A_U, even though A_U satisfies the antecedent of condition A.

Actually, it's not merely the case that no such A_U for a proper infinite society satisfies A: indeed, no SWF A for a proper society (whether infinite or not) satisfying WP and IIA *can* satisfy it. The reason for this is relatively simple.

If the society is finite, then A can't satisfy A because there is a dictator. If the society is infinite, then, as we know from Hansson's proof of Arrow's theorem, the set of coalitions decisive under A is an ultrafilter, U. Of course, if U is principal, then it is also fixed, and so there's a dictator, from which it follows again A doesn't satisfy A. If U is non-principal, however, then we can define a PAF A_U as in the theorem which is itself an SWF satisfying WP and IIA. Now this implies, by Hansson's proof of Arrow's theorem, that the set of coalitions decisive under A_U is *also* an ultrafilter, U'.

Well, it so happens $U' = U$. To prove this, it suffices to show every $J \in U$ is decisive under A_U, — that is, to show $U \subseteq U'$; for, if this is the case, because both U and U' are ultrafilters (therefore maximal), $U' = U$. So let $J \in U$ and x, y, p be any, and suppose $xP^p_J y$. Then $x \gtrsim^p y$ by the definition of A_U because R_J is an extension of P^p_J. Now let J' be any such that $yR^p_{J'}x$. Because $x \succ^p_i y$ for all $i \in J$ but $y \gtrsim^p_i x$ for all $i \in J'$, it follows $J \cap J' = \varnothing$, so that $J' \notin U$ since U is an ultrafilter and $J \in U$. Because J' is any, there's no $J' \in U$ such that $yR^p_{J'}x$, — from which it follows, by the definition of A_U, that $y \not\gtrsim^p x$, and so that $x \succ^p y$. Therefore, because p is any, J is decisive for x, y and thus, because x, y are any, also for any x, y, so that $J \in U'$. Finally, because J is any, $U \subseteq U'$, and so $U' = U$ as expected.

In conclusion, if A were to satisfy condition A, then so would A_U; but, as pointed out, no A_U does, and so no SWF satisfying conditions WP and IIA satisfies A, — *even if* it satisfies ND.

The upshot of this discussion is this. Even though full social rationality and welfarism are not logically incompatible with even some strong democratic requirements, no such form of preference aggregation can be carried out, even in principle, by means of a fully *impersonal* procedure.

12.4 On the matter of decisiveness

As we have seem, although by no means simpler than Arrow's original proof of the theorem, Hansson's proof has the virtue of pinpointing where the conditions imposed by Arrow on an SWF become problematic. In this section, I'd like to discuss the issue a little more, by reviewing the properties of ultrafilters in the light of its social choice applications.

Recall from section 4.1 that an (arbitrary) decisiveness assignment (DA) for a society I, S is an assignment of a binary relation on S to each coalition from I.

Definition 12.4 (Arrovian conditions on a DA and Arrovian PAFs)**.** Consider the following conditions a DA D for I, S may satisfy (some of which have already been discussed in section 3.3).

Totality If xD_Jy for some x, y, then xD_Jy for all x, y, for all J

Weak Pareto xD_Iy for all x, y

Weak non-imposedness $xD_\varnothing y$ for no x, y

Monotonicity If xD_Jy and $J' \supseteq J$, then $xD_{J'}y$, for all J, x, y

Closure under binary intersection If xD_Jy and $xD_{J'}y$, then $xD_{J \cap J'}y$, for all J, J', x, y

Exclusiveness If xD_Jy, then $y \not\!\!D_{\overline{J}}\, x$, for all J, x, y

Fullness If $x \not\!\!D_J\, y$, then $xD_{\overline{J}}y$, for all J, x, y

Let us call a DA *Arrovian* if it satisfies all of the conditions above, and let us equally call a PAF *Arrovian* if it is an SWF implementing (*cf.* section 4.1) an Arrovian DA.

It should be immediately clear to the reader what we are implicitly doing here. Given a subset space U on I, we define from it a DA D by setting, for all x, y, xD_Jy iff $J \in U$, so that we may then impose conditions on an SWF implementing D by means of the properties of U. It is trivial to prove that, if U is an ultrafilter, then D so defined satisfies the seven conditions above, and that, conversely, D satisfying them implies that U is an ultrafilter. It follows that resorting to a DA D satisfying those conditions is just another way, albeit more conspicuous for our purposes, of presenting the conditions on a SWF that lead to the negative conclusion of Arrow's theorem.

Now, at least for proper societies, as the contagion lemma together with Hansson's proof of Arrow's impossibility theorem show, accepting Arrow's conditions on a PAF forces us to commit to all these conditions on the DA implemented by it. Intuitively, however, not all of them are on a par, as far as justification for their adoption goes.

Totality is an incredibly strong condition, as it means that decisiveness is an "all or nothing" matter. It entirely excludes aggregation procedures under which certain groups or individuals in the society may have a voice only regarding specific issues. Evidently, this would not be desirable for a rule designed to implement individual rights.

As for weak non-imposedness, it excludes the possibility that some social preferences be settled in advance, independent of the preferences of the individuals (hence the name). Weak non-imposedness implies some aspects of condition CS, citizen's sovereignty, in the sense that it excludes at least one way in which a taboo rule could be built. It should be noted, however, that taboos in this technical sense are not *always* undesirable, even in full-blown democratic societies. Much to the contrary, it is quite usual that in contemporary democracies some issues are settled in advance by constitutional law, and thus are not open for any sort of discussion short of an actual coup.

Closure under binary intersection is the truly questionable condition of the bunch. For consider pairwise simple majority rule. It is easy to see that, given a finite society I, S, a coalition J in it is decisive for an x over a y under this rule iff $|J| > |\overline{J}|$. So, as the reader may just as easily check, for such a society pairwise simple majority rule satisfies all of the seven conditions above except closure under binary intersections and fullness, and also fullness if the number of individuals is odd. It only satisfies closure under binary intersection, however, if the society has *at most* two individuals, and simply because, in this case, the only numerical majority is the full coalition itself, so that majority and unanimity coincide. Therefore, pairwise simple majority rule satisfies all seven conditions solely in the trivial case of there being a single individual in the society, who will then of necessity be a dictator.

Regarding fullness, it is perhaps best to start by taking a look again at the concrete example of pairwise simple majority rule. For any society I, S, no coalition J with $|J| \leq |I|/2$ is decisive under this rule, — so that, if $|I|$ is even, no coalition J with $J = |I|/2$ is decisive, and nor is $\overline{J}$, as $|\overline{J}| = |J|$ in this case. So consider what this means for the society in an actual situation in which, regarding x, y, half of the individuals strictly prefer x to y, and half strictly prefer x to y, and these are the only two possible outcomes eligible for implementation, — so that, whichever is implemented, half of the individuals will possibly be upset. This is a tough spot for a society to be in, to say the least. Of course, insofar as it is assumed that only majorities have the *legitimacy* for social decision — as majority rules implicitly do —, this is also the right spot for a society to be, in this situation. But it is nevertheless odd that the social preference regarding x and y should be ruled as indifference when, in actuality, *no one* is indifferent between those two possible outcomes. Such situations lead to impasses under majority rule. Fullness, however, avoid impasses by requiring that, no matter how the individuals happen to divide about an issue, one of the coalitions so formed has a legitimate claim to deciding it. In other words, fullness conveys the notion that a reasonable decisiveness assignment should distribute powers in such a way that, no matter the possible outcomes or the (strict) preferences of the individuals, one could always point to an individual or coalition that is supposed to be able to decide the issue, if necessary.

Chapter 13

Social choice and systems of individual rights

In this chapter, we continue our review of social choice theory by discussing in greater depth Sen's theorem on the impossibility of a Paretian liberal, with particular attention to the concept of coherence of a system of individual rights. We also present both Gibbard's and Sen's resolution schemata for the impossibility.

13.1 The impossibility of a Paretian liberal

In this section, the original proof of Sen's theorem on the impossibility of a Paretian liberal is presented.

Recall that, for Sen, individual rights (or some of them, at least) are entitlements on non-trivial issues, — meaning that a PAF A for I, S implements a right on the issue x, y for individual i if $x \neq y$ and i is decisive between x and y under A.

Given this notion of individual rights, Sen claims that an SDF A representing the rules of a liberal society I, S should satisfy the following condition.

Liberty (L) For each i there are $x \neq y$ such that xE_iy

That is: under condition L, every individual in the society has at least one right that is an entitlement. This seems to be a mild condition on an SDF for a liberal society.

Sen, however, settles for an even milder one, the condition of minimal liberty (ML), *cf.* section 3.4. I repeat it here for ease of reference.

Minimal liberty (ML) There are $i, j, i \neq j$, for which there are $x \neq y$ and $z \neq w$ (not necessarily all distinct) such that xE_iy and zE_jw

Obviously enough, L implies ML if $|I| > 1$.

As we know, this condition together with UD conflicts with WP if the PAF is required to be an SDF.

Proposition 13.1 (Sen's impossibility of a Paretian liberal theorem – Sen 1970 [62])**.** *There is no SDF satisfying conditions UD, WP and ML (and therefore, for a society with at least two individuals, also L).*

Proof. Let A be an SDF for I, S and suppose for contradictions that it satisfies UD, WP and ML. Let i, j and x, y, z, w be as described in the statement of condition ML.

If $x = z$ and $y = w$, or if $x = w$ and $y = z$, then by UD, letting p be such that i and j have opposite preferences regarding x, y, a contradiction follows immediately from ML.

So suppose, without loss of generality, that $x \neq z$ but $y = w$. By UD, let p be any such that $x \succ_i^p y$ and $y \succ_j^p z$, with $z \succ_k^p x$ for all k (i, j included). Because xD_iy by ML and $x \succ_i^p y$, it follows that $x \succ^p y$; because wD_jz (that is, yD_jz) by ML and $y \succ_j^p z$, it follows that $y \succ^p z$; and, because $zP_I^p x$, by WP it follows that $z \succ^p x$. But then we have that $x \succ^p y$, $y \succ^p z$ and $z \succ^p x$, — a $\succsim^p$-cycle, thus contradicting the assumption that A is an SDF.

For a $\succsim^p$-cycle in the case with four distinct socially possible outcomes, take p to be any such that $x \succ_i^p y$ and $z \succ_j^p w$, with $w \succ_k^p x$ and $y \succ_k^p z$ for all k.

In any case, the assumption that A is an SDF is contradicted. Therefore no SDF satisfying, UD, WP and ML exists. □

13.2 On the significance of the result

Perhaps somewhat surprisingly, given its simplicity, and much to Sen's own surprise [56], the theorem on the impossibility of a Paretian liberal has elicited an overwhelming amount of responses over the five decades since its publication. Of course, this is far from unusual. Much like so many other paradoxes known from the history of philosophy, and quite unlike Arrow's theorem, the enduring interest of Sen's puzzle is due precisely to its simplicity, which leaves much less room for debate than Arrow's much more complex result. Because the literature it has spawned is nothing short of insurmountable, I have no hope of covering it here in its entirety, or even of doing it justice; nor do I wish to do so. Most of this literature consists of critiques of Sen's framework for the representation of rights, and in this regard I believe I can do no better than defer to Sen's own replies [58], [60], [61].

This is by no means to say, however, that such criticisms are of no importance. Much to the contrary, this is no minor issue, as it pertains not only to the adequacy but also to the *stability* of the social choice model of rights. Clearly, Sen's theorem wouldn't be very convincing if changing the form of representation of rights was to lead to wildly different conclusions. Nevertheless, and important as it may be, this is a whole other subject, and there's simply no way I can address those criticisms here without taking us far away from our subject.

I'd like to mention, however, at least Nozick's often quoted response to Sen's theorem, as I believe briefly discussing it may help dispel one possible misconception which, being recurrent in the literature itself, may just as well occur to the reader.

Nozick claims [43, pp.165-66]:

> A more appropriate view of individual rights is as follows. Individual rights are co-possible; each person may exercise his rights as he chooses. The exercise of these rights fixes some features of the world. Within the constraints of these fixed features, a choice may be made by a social choice mechanism based upon a social ordering; if there are any choices left to make! (...) *How else can one cope with Sen's result?*

This passage has been of great historical importance for the later developments on the representation of rights in social choice, as it has motivated two approaches alternative to the one proposed by Sen.

One is the approach by means of *effectivity functions* [1], [46], first proposed by Gärdenfors [22]. Recent developments [45], [24] have shown it to be equivalent to the approach, primarily championed by Suzumura and others [12], [21], by means of *game forms*. These two approaches are, at least on the face of it, quite different from the social choice one, and also quite different from one another, and discussing them here in any detail is simply impossible; so I'll just drop the subject. In any event, it is generally agreed that impossibilities analogous to Sen's puzzle occur in the game-theoretic approach to rights, with the Pareto-inefficient Nash equilibrium of the *prisoner's dilemma* being an example of the same phenomenon in this alternative framework.

The other, essentially due to Gibbard [23], assumes the possible social states to have internal structure. More precisely, it assumes them to be tuples consisting of one entry associated to each individual in the society, and perhaps one additional entry relating to strictly public matters. Between two options which are mere variants of each other insofar as a certain individual is concerned — that is: options such that they are the same for every entry other than the one associated with that individual —, his or her preference regarding the options should count for the composition of the social preferences; but not necessarily otherwise. Intuitively, this is just a more formal way of representing the notion that the existence of a right on an issue is conditional on all other relevant aspects of the world remaining the same no matter how the right is exercised. This notion seems to be already implicit in Sen's own examples of one's painting the walls either pink or white or sleeping either on one's back or belly. In Nozick's terms, it would be tantamount to taking the relevance of the individuals preferences for the determination of the social outcome to be restricted to the features of the world which are assigned to that individual.

There is no doubt this is a richer representation of possible social outcomes and of the rights pertaining to them. As Sen points out, however, claiming that rights should be so represented amounts to no real objection to the approach to rights by means of the usual social choice framework, for it is quite evident that the social choice, being more abstract, is also more general. In fact, if an individual is decisive between two possible outcomes that are mere variants of each other, then, trivially, he or she is decisive between two possible outcomes, and thus has a right in Sen's sense. But the converse is obviously not true in general. So it can easily be seen that having a right in Sen's sense is a necessary condition for having a right in the other, apparently more intuitive sense; and, as is therefore to be expected and Gibbard indeed shows, the impossibility of the Paretian liberal stands just the same for this richer form of representation.

On a more technical note, it is interesting to notice that, contrary to what its name may suggest, minimal liberty is by no means a weak condition on a PAF, as the following propositions clearly reveal.

Proposition 13.2. *If a PAF satisfying UD and ML, then it in non-dictatorial.*

Proof. Let A for I, S satisfying UD and ML be any, and suppose for a contradiction it is dictatorial. Because it satisfies ML, there are $i \neq j$ for whom there are $x \neq y$ and $z \neq w$ such that xE_iy and zE_jw. Let such i, j, x, y, z, w be given. Let d be the dictator. Because $i \neq j$, either $d \neq i$ or $d \neq j$. Without loss of generality, suppose the former. Now, by UD, let p be any such that $x \succ_i^p y$ and $y \succ_d^p x$. Then, because xD_iy, it follows that $x \succ^p y$; but also $y \succ^p x$, because d is a dictator, — a contradiction. Therefore, A is non-dictatorial. □

Proposition 13.3. *If a PAF satisfying UD and ML, then it does not satisfy condition A, anonimity.*

Proof. Let A for I, S satisfying UD and ML be any. As in the previous proposition, let the i, j, x, y, z, w to which the statement of condition ML refers be given. By UD, let p be any such that i, j have opposite preferences regarding x, y, and similarly, opposite preferences regarding z, w. For the sake of concreteness, without loss of generality assume that $x \succ_i^p y$ and $z \succ_i^p w$ but $y \succ_j^p x$ and $w \succ_j^p z$. Then, by ML, $x \succ^p y$ and $w \succ^p z$. (Notice that, if x, y and z, w are the same issue, then there is no such PAF because it is supposed to satisfy UD.) Now let π be any permutation of I such that $\pi(i) = j$ and $\pi(j) = i$, and, by UD again, let p' be any such that $\succsim_{\pi(k)}^{p'} = \succsim_k^p$ for all k. Then, in particular, $y \succ_i^{p'} x$ and $z \succ_j^p w$, from which it follows by ML that $y \succ^{p'} x$ and $z \succ^{p'} w$, so that $\succsim^{p'} \neq \succsim^p$. Therefore, A does not satisfy condition A. □

These propositions show ML, at least under UD, excludes both the existence of a dictator and of a fully impersonal aggregation procedure. This, of course, should be of no surprise, as it is precisely what is to be expected of a society in which more than one individual has *personal* rights: power is not entirely concentrated in the hands of a single individual, and the social preferences cannot be settled without some knowledge of who has exercised which right.

13.3 On the coherence of systems of individual rights

Going back to Sen's theorem itself, it should be noted that, for the case with only two options, conditions UD and ML are inconsistent by themselves under the assumption that the preference aggregation is an SDF (and so independently of condition WP). This is due to the fact that in this case ML implies that the DA D implemented by A must assign the same pair to at least two individuals. At first, this may seem to suggest that we should adopt, as a coherence condition, something such as the following.

Exclusive rights (ER) If xD_iy, then xD_jy for no $j \neq i$, for all i, x, y

In and of itself, this is a reasonable principle for reasonable systems of rights, as it conveys the notion that the rights assigned to an individual really are in some sense *personal*, — that is, exclusive to that individual. Of course, it does the job of avoiding the inconsistency of UD and ML (but not their inconsistency with WP though). By itself, however, it won't do to avoid further logical inconsistencies involving those very same conditions. Indeed, it is easy to see (as it follows from basic combinatorics) that, if $|I| = \frac{|S|\times(|S|-1)}{2}$ and condition ER is satisfied, then each individual is entitled to exactly one non-trivial issue under condition L — in which case, given the obvious profile p, a $\gtrsim^p$-cycle can easily be construed if $|I| > 1$. Moreover, if $|I| > \frac{|S|\times(|S|-1)}{2}$, then at least two individuals must be entitled to one and the same non-trivial issue under condition L — in which case ER cannot be satisfied. In general, if there is more than one individual and the system of individual rights satisfying L is otherwise *unrestricted*, then UD and L are already inconsistent under the assumption that the preference aggregation is an SDF.

Considerations such as these led to the introduction of the concept of a *coherent* system of individual rights. A coherent system of individual rights is just a DA, *cf.* section 4.1 from which only incoherences involving individual coalitions are excluded. Formally, this means that, given a DA D for I, S, there are no non-empty sequence $x_0, \ldots, x_n$, with $x_k \neq x_\ell$ for all $k < \ell < n$ but $x_n = x_0$, such that, each $k < n$, there is an individual i_k such that $x_k D_{i_k} x_{k+1}$ but, for some $k, \ell < n$, also $i_k \neq i_\ell$. The proof that a coherent system of individual rights can never lead to a social preferences cycle by the exercise of rights alone is an immediate corollary of proposition 4.2.

13.4 Resolution schemata for the impossibility of the Paretian liberal

Let us say that a DA D for I, S is a *Paretian system of individual rights* if it satisfies the following four conditions, for all x, y, J.

- D_J is symmetric

- If xD_Jy, then either $J = I$ or $J = \{i\}$ for some i
- xD_Iy for all x, y
- The restriction of D to singletons is a coherent DA (or, equivalently, a coherent system of individual rights)

Notice that it is not required that the whole DA be coherent, but only its restriction to singletons (that is, when we consider only the rights of individuals).

Given D a Paretian system of individual rights, let A be any PAF for I, S such that, for all $p \in \mathsf{dom}(A)$, if $\succsim^p$ has a cycle, then it results from the exercise of decisiveness alone. Gibbard's as well as Sen's resolution schemata for the impossibility of a Paretian liberal apply to PAFs of this kind.

So let D a Paretian system of individual rights and A a PAF for I, S implementing D as described be given.

For each i and $p \in \mathsf{dom}(A)$, let $W_i(p)$ be the set defined as follows. $\langle x, y\rangle \in W_i(p)$ iff there are $x_0, \ldots, x_n$ such that:

- $x_0 \neq y$;
- $x_n = x$;
- $y \succsim_i^p x_0$, and
- for all $k < n$, at least one of the following holds:
  - $x_k P_I^p x_{k+1}$, or
  - for some $j \neq i$, $x_k D_j x_{k+1}$ and $x_k \succ_j^p x_{k+1}$.

We say that individual's i right to x over y is *waived* under profile p if $\langle x, y\rangle \in W_i(p)$.

Gibbard [23] proves that, under those conditions, if individuals exercise all and only rights that aren't waived, then the resulting PAF will be an SDF, even if A itself is not.

Sen's resolution schema involves the introduction of a distinction between an individual's preferences, and the part of those preferences that the individual wants to count for the purpose of implementing the decisiveness powers of the coalition of all individuals (that is: the Pareto principle). Formally, for each profile $p \in \mathsf{dom}(A)$, we assume there is a second profile, $\overline{p}$, such that, for all i, $\succsim_i^{\overline{p}}$ is compatibly extended by $\succsim_i^p$. It should be noted that such $\overline{p}$ is not a preferences profiles in the usual sense, since we are not requiring that the preferences they express be either transitive or complete, but merely that they can be compatibly extended by an ordering (which implies, by proposition 11.33, however, that they must be strongly acyclic).

Now, for each $p \in \mathsf{dom}(A)$, let R^p be the relation defined by: xR^py iff there is an i such that xD_iy and $x \succ_i^p y$. We say individual i *respects individual rights* if, for all $p \in \mathsf{dom}(A)$, there is an ordering $\succsim$ compatibly extending both $\succsim_i^{\overline{p}}$ and R^p.

Sen [59] proves that, under those same conditions, if the Pareto principle is so restricted and at least one individual in the society respects individual rights, then, again, the resulting PAF will be an SDF even if A itself is not.

Bibliography

[1] ABDOU, J; KEIDING, H. *Effectivity functions in social choice*. Springer, 1991

[2] ARROW, K. J. *Collective choice and individual values – Second edition*. John Willey & Sons, 1963

[3] BANKS, J. S. *Acyclic social choice from finite sets*. Social Choice and Welfare, 1995, pp. 293-310

[4] BATRA, R. N.; PATTANAIK, P. K. *On some suggestions for having non-binary social choice functions*. Theory and Decision, v. 3, 1972, pp. 1-11

[5] BODIN, J. *On sovereignty: Four chapters from 'The six books of the commonwealth'*. Cambridge University, 1992

[6] BOSSERT, W. *Suzumura consistency*. In: PATTANAIK, P. K. *et al.* (eds.). *Rational choice and social welfare – Theory and applications*. Springer, 2008

[7] BOSSERT, W.; SUZUMURA, K. *Consistency, choice, and rationality*. Harvard University, 2010

[8] BRANDOM, R. *Articulating reasons – An introduction to inferentialism*. Harvard University, 2000

[9] CARNIELLI, W.; CONIGLIO, M. E. *Paraconsistent logic: Consistency, contradiction and negation*. Springer, 2016

[10] CHISWELL, I; HODGES, W. Mathematical logic. Oxford University, 2007

[11] DAVEY, B. A.; PRIESTLEY, H. A. *Introduction to lattices and orders – Second edition*. Cambridge University, 2002

[12] DEB, R. *Rights as alternative game forms*. Social Choice and Welfare, 2004, pp. 83-111

[13] DIETRICH, F. *A generalized model of judgment aggregation*. Social Choice and Welfare, 2007, v. 28, pp. 529-565

[14] DIETRICH, F.; LIST, C. *A liberal paradox for judgment aggregation*. Social Choice and Welfare, 2008, v. 31, pp. 59-78

[15] DUNN, J. M.; HARDEGREE, G. M. *Algebraic methods in philosophical logic*. Oxford University, 2001

[16] FARRELL, M. J. *Liberalism in the theory of social choice*. The Review of Economic Studies, 1976, v. 43, pp. 3-10

[17] FIELD, H. *Metalogic and modality*. Philosophical Studies, 1991, v. 62, pp. 1-22

[18] FISHBURN, P. C. *Arrow's impossibility theorem – Concise proof and infinite voters*. Journal of Economic Theory, 1970, v. 2, pp. 103-106

[19] FITTING, M.; SMULLYAN, R. M. *Set theory and the continuum problem*. Dover, 2010

[20] GAERTNER, W. *A primer in social choice theory – Revised edition*. Oxford University, 2009

[21] GAERTNER, W.; PATTANAIK, P. K.; SUZUMURA, K. *Individual rights revisited*, 1992

[22] GÄRDENFORS, P. *Rights, games and social choice*. Noûs, 1981, v. 15, pp. 341-356

[23] GIBBARD, A. *A Pareto-consistent libertarian claim*. Journal of Economic Theory, 1974, v. 7, pp. 388-410

[24] GORANKO, V.; JAMROGA, W.; TURRINI, P. *Strategic games and truly playable effectivity functions*. Autonomous Agents and Multi-Agent Systems, 2013, v. 26, pp. 288-314

[25] GROSSI, D.; PIGOZZI, G. *Judgment aggregation – A primer*. Morgan & Claypool, 2014

[26] HALPERN, J. Y. *Reasoning about uncertainty*. MIT, 2003

[27] HANSSON, B. *The existence of group preference functions*. Public choice, 1976, v. 28, pp. 89-98

[28] HANSSON, B. *Voting and group decision functions*. Synthese, 1969, v. 20, pp. 526-537

[29] HINMAN, P. G. *Fundamentals of mathematical logic*. Taylor & Francis, 2005

[30] HOBBES, T. *Leviathan*. Hackett, 1994

[31] HRBACEK, K.; JECH, T. *Introduction to set theory – Third edition, revised and expanded*. Marcel Dekker, 1999

[32] JACKSON, R. *Sovereignty – Evolution of an idea*. Polity, 2007

[33] KAYE, R. *The mathematics of logic – A guide to completeness theorems and their applications*. Cambridge University, 2007

[34] KEISLER, H. J. *Model theory for infinitary logic – Logic with countable conjunctions and finite quantifiers*. North-Holland, 1971

[35] KIRMAN, A. P.; SONDERMANN, D. *Arrow's theorem, many agents, and invisible dictators*. Journal of Economic Theory, v. 5, pp. 267-277

[36] LAGERSPETZ, E. *Social choice and democratic values*. Springer, 2016

[37] LEVY, A. *Basic set theory*. Dover, 1979

[38] LIST, K. *The theory of judgment aggregation: An introdutory review*. Synthese, 2012, v. 187, pp. 179-207

[39] LOCKE, J. *Second treatise of government*. Hackett, 1980

[40] MAKINSON, D. *Sets, logic and maths for computing – Third edition*. Springer, 2020

[41] MAY, K. O. *A set of independent necessary and sufficient conditions for simple majority decision*. Econometrica, 1952, v. 20, pp. 680-684

[42] MONTESQUIEU. *The spirit of the laws*. Cambridge University, 1989

[43] NOZICK, R. *Anarchy, state and utopia*. Blackwell, 1974

[44] PACUIT, E. *Neighborhood semantics for modal logic*. Springer, 2017

[45] PAULY, M. *A modal logic for coalitional power in games*. Journal of Logic and Computation, 2002, v. 12, n. 1, pp. 149-166

[46] PELEG, B. *Effectivity functions, game forms, games, and rights*. Social choice and welfare, 1998, v. 15, pp. 67-80

[47] PINTER, C. *A book of set theory*. Dover, 2014

[48] PLOTT, C. R. *Path independence, rationality and social choice*. Econometrica, 1973, v. 41, pp. 1075-1091

[49] PRIEST, G. *In contradiction – A study of the transconsistent*. Oxford University, 2006

[50] RISSE, M. *What to make of the liberal paradox?*. Theory and Decision, 2001, v. 50, pp. 169-196

[51] ROMAN, S. *Lattices and ordered sets*. Springer, 2008

[52] SALTER, A. W. *Sovereignty as exchange of political property rights*. Public Choice, 2015, v. 165, pp. 79-96

[53] SELLARS, W. *Inference and meaning*. Mind, 1953, v. 62, pp. 313-338

[54] SEN, A. K. *Behaviour and the concept of preference*. Economica, 1973, v. 40, pp. 241-259

[55] SEN, A. K. *Choice functions and revealed preference*. Review of Economic Studies, 1971, v. 38, pp. 307-317

[56] SEN, A. K. *Collective choice and social welfare - An expanded edition*. Harvard University, 2017

[57] SEN, A. *Internal consistency of choice*. Econometrica, 1993, v. 61, pp. 495-521

[58] SEN, A. *Liberty and social choice*. The Journal of Philosophy, 1983, v. 80, pp. 5-28

[59] SEN, A. *Liberty, unanimity and rights*. Economica, 1976, v. 43, pp. 217-245

[60] SEN, A. *Minimal liberty*. Economica, 1992, v. 59, pp. 139-159

[61] SEN, A. *Rights: Formulations and consequences*. Analyse & Kritik, 1996, v. 18, pp. 153-170

[62] SEN, A. K. *The impossibility of a Paretian liberal*. Journal of Political Economy, 1970, v. 78, pp. 152-157

[63] STALNAKER, R. *Mere possibilities – Metaphysical foundations of modal semantics*. Princeton University, 2012

[64] SUZUMURA, K. *On the consistency of libertarian claims*. Review of Economic Studies, 1978, v. 45, pp. 329-342

[65] SUZUMURA, K. *Rational choice, collective decisions, and social welfare*. Cambridge University, 1983

[66] SUZUMURA, K. *Remarks on the theory of collective choice*. Economica, 1976, v. 43, pp. 381-390

[67] SZPILRAJN, E. *Sur l'extension de l'ordre partiel*. Fundamenta Mathematicae, 1930, v. 16, pp. 386-389

UNIVERSIDADE ESTADUAL DE CAMPINAS
Instituto de Filosofia e Ciências Humanas
Comissão de Programa de Pós-Graduação em Filosofia

Diretoria Acadêmica

Ata da Sessão pública de defesa de tese para obtenção do título de Doutor em Filosofia, a que se submeteu o aluno Pedro Henrique Carrasqueira Zanei - RA 154029, orientado pelo Prof. Dr. Walter Alexandre Carnielli.

Aos quatorze dias do mês de setembro do ano de dois mil e vinte e três, às 14:00 horas, CLE/UNICAMP, de forma híbrida, reuniu-se a Comissão Examinadora da defesa em epígrafe indicada pela Comissão de Pós-Graduação do(a) Instituto de Filosofia e Ciências Humanas, composta pelo Presidente e Orientador Dr. Walter Alexandre Carnielli (Universidade Estadual de Campinas - IFCH) e pelos membros Profa. Dra. Itala Maria Loffredo D'Ottaviano (IFCH/ UNICAMP), Prof. Dr. André Luiz de Almeida Lisbôa Neiva (Universidade Federal de Alagoas) (por videoconferência), Prof. Dr. Álvaro Sandroni (Kellogg School of Management) (por videoconferência) e Prof. Dr. Otávio Bueno (University of Miami) (por videoconferência), para analisar o trabalho do candidato Pedro Henrique Carrasqueira Zanei, apresentado sob o título "Consistent systems of rights and the logical limits of liberal democracy" (Sistemas consistentes de direitos e os limites lógicos da democracia liberal).

O Presidente declarou abertos os trabalhos, a seguir o candidato dissertou sobre o seu trabalho e foi arguido pela Comissão Examinadora. Terminada a exposição e a arguição, a Comissão reuniu-se e deliberou pelo seguinte resultado:

[X] **APROVADO**
[ ] **APROVADO CONDICIONALMENTE** (ao atendimento das alterações sugeridas pela Comissão Examinadora especificadas no parecer anexo)
[ ] **REPROVADO** (anexar parecer circunstanciado elaborado pela Comissão Examinadora).

Para fazer jus ao título de Doutor, a versão final da tese, considerada Aprovada ou Aprovada Condicionalmente, deverá ser entregue à CPG dentro do prazo de 60 dias, a partir da data da defesa. De acordo com o previsto na Deliberação CONSU-A-10/2015, Artigo 42, parágrafo 1°, inciso II e parágrafo 2°, o aluno Aprovado Condicionalmente que não atender a este prazo será considerado Reprovado. Após a entrega do exemplar definitivo e a sua conferência pela CPG, o resultado será homologado pela Comissão Central de Pós-Graduação da UNICAMP, conferindo título de validade nacional aos aprovados.

Nada mais havendo a tratar, o Senhor Presidente declara a sessão encerrada, sendo a ata lavrada por mim, que segue assinada pelos Senhores Membros da Comissão Examinadora, pelo Coordenador da Comissão de Pós-graduação, com ciência do aluno.

Prof. Dr. Walter Alexandre Carnielli
Presidente da Comissão Examinadora

Profa. Dra. Itala Maria Loffredo D'Ottaviano

Prof. Dr. André Luiz de Almeida Lisbôa Neiva

Prof. Dr. Álvaro Sandroni

Prof. Dr. Otávio Bueno

Pedro Henrique Carrasqueira Zanei
Aluno(a)

Secretaria de Pós Graduação

Profa. Dra. Nashieli Cecilia Rangel Loera
Coordenador(a) da CPG

CÓDIGO DE AUTENTICIDADE

Verifique a autenticidade deste documento na página www.dac.unicamp.br
Código: ea398ed5c41a0654889ff186a80d794ed25c55ee

UNIVERSIDADE ESTADUAL DE CAMPINAS
Instituto de Filosofia e Ciências Humanas
Comissão de Programa de Pós-Graduação em Filosofia

UNICAMP

Diretoria Acadêmica

Parecer da Comissão

O candidato foi aprovado com distinção pela Banca Examinadora que considerou seu trabalho de excelente nível, com alto grau de interdisciplinaridade, transitando entre as áreas de filosofia, lógica, direito e economia. Após responder com proprie-dade e segurança às questões apresentadas, o candidato recebeu de um dos membros da Banca o Prof. Otávio Bueno, editor da coleção Synthese Library - Studies in Epistemology, Logic, Methodology and Philosophy of Science da Editora Springer, um convite para transformar sua tese de doutorado em um volume da série.

CÓDIGO DE AUTENTICIDADE

Verifique a autenticidade deste documento na página www.dac.unicamp.br
Código: ea398ed5c41a0654889ff186a80d794ed25c55ee