Alma Mater Studiorum · Università di Bologna

SCUOLA DI SCIENZE
Corso di Laurea Magistrale in Matematica

LATTICES AND DISCRETE METHODS IN COOPERATIVE GAMES AND DECISIONS

Tesi di Laurea in
Teoria dei Giochi

Relatore:
Chiar.mo Prof.
Giovanni Rossi

Correlatore:
Chiar.mo Prof.
Massimo Ferri

Presentata da:
Enrico Gandolfo

I Sessione
Anno Accademico 2014-2015

Abstract

Questa tesi si pone l'obiettivo di presentare la teoria dei giochi, in particolare di quelli cooperativi, insieme alla teoria delle decisioni, inquadrandole formalmente in termini di matematica discreta. Si tratta di due campi dove l'indagine si origina idealmente da questioni applicative, e dove tuttavia sono sorti e sorgono problemi più tipicamente teorici che hanno interessato e interessano gli ambienti matematico e informatico. Anche se i contributi iniziali sono stati spesso formulati in ambito continuo e utilizzando strumenti tipici di teoria della misura, tuttavia oggi la scelta di modelli e metodi discreti appare la più idonea.

L'idea generale è quindi quella di guardare fin da subito al complesso dei modelli e dei risultati che si intendono presentare attraverso la lente della teoria dei reticoli. Ciò consente di avere una visione globale più nitida e di riuscire agilmente ad intrecciare il discorso considerando congiuntamente la teoria dei giochi e quella delle decisioni. Quindi, dopo avere introdotto gli strumenti necessari, si considerano modelli e problemi con il fine preciso di analizzare dapprima risultati storici e solidi, proseguendo poi verso situazioni più recenti, più complesse e nelle quali i risultati raggiunti possono suscitare perplessità. Da ultimo, vengono presentate alcune questioni aperte ed associati spunti per la ricerca.

Introduzione

La teoria dei giochi nasce a metà degli anni '40 con il libro "Theory of Games e Economic Behaviour" [59] di von Neumann e Morgenstern. Anche se con il termine "giochi" di solito ci si riferisce a quelli non cooperativi, ovvero quelli in cui i giocatori scelgono tra varie strategie, fin da subito von Neumann e Morgenstern considerarono giochi cooperativi non appena il numero di giocatori eccede 2. In questi giochi i giocatori devono scegliere se cooperare (in qualche modo) o meno. Più precisamente, i giochi cooperativi sono funzioni che assumono valori reali su strutture ordinate che hanno come base un insieme di giocatori. In particolare, anche se si possono contemplare modelli di restrizione alla cooperazione tali che il gioco cooperativo che ne risulta è un funzione su un poset (partially ordered set), tuttavia in questo lavoro consideriamo solo funzioni sui reticoli, escludendo quindi qualsiasi restrizione alla cooperazione. Più precisamente, ci concentreremo sulle funzioni a valori reali definite sul reticolo dei sottoinsiemi o su quello delle partizioni. Infine affronteremo anche il caso delle funzioni reali sul reticolo dei sottoinsiemi *embedded*, incluso nel prodotto tra il reticolo dei sottoinsiemi e quello delle partizioni.

Da un punto di vista matematico, i giochi, in particolar modo quelli cooperativi, condividono molti strumenti e concetti con la teoria delle decisioni. Infatti, al momento di scegliere una strategia in un gioco non cooperativo un giocatore è a tutti gli effetti un decisore in condizione di incertezza. Tale problema decisionale ha infatti guidato molte ricerche nel campo della decisione in condizioni di incertezza. In questo lavoro, per probabilità si intenderà una *valuation* di un reticolo di sottoinsiemi, e ci concentreremo sulle probabilità generalizzate o *fuzzy measures* che non sono valuations ma semplicemente monotone e normalizzate. Inoltre, in teoria delle decisioni il concetto fondamentale di informazione è formalizzato in termini di partizioni di un insieme di stati di natura. Di conseguenza, ci concentreremo sulle funzioni su partizioni (*partition functions*) che quantificano il valore di tali informazioni per un decisore. In particolare, quest'ultimo sceglie un'azione massimizzando la propria utilità attesa, la quale assume valori reali su coppie formate da uno stato di natura e un'azione. Partizioni più fini hanno quindi un valore maggiore rispetto a quelle meno fini poichè consento di sceglire azioni più specifiche: una diversa per ciascun blocco della partizione.

Questo lavoro è diviso in quattro capitoli:

Nel Capitolo 1 si introduce il necessario background combinatorio. In primo luogo, si definisce la relazione d'ordine parziale e il concetto associato di insieme parzialmente ordinato (poset), includendo alcuni importanti risultati utili in seguito. Successivamente, consideriamo le relazioni binarie di "meet" e "join" che caratterizzano i reticoli. Questi ultimi costituiranno la base di tutta la successiva analisi. I reticoli vengono poi classificati come distributivi, modulari, semimodulari e geometrici, consentendo in tal modo di introdurre due esempi principali, vale a dire il reticolo dei sottoinsiemi e il reticolo delle partizioni, molto importanti sia per la teoria dei giochi che per la teoria delle decisioni. Infine, descriviamo nel dettaglio alcune proprietà delle funzioni sui reticoli (e più in generale sui poset), con particolare attenzione al rango e all'inversione di Möbius (di *poset functions*).

Nel Capitolo 2 si mostrerà come gli strumenti e i risultati forniti nel Capitolo 1 siano utili sia nella teoria dei giochi che nella teoria delle decisioni. Per quanto riguarda la prima, consideriamo in dettaglio lo sviluppo della teoria dei giochi cooperativi fin dalla sua nascita, avvenuta con la pubblicazione del già citato libro "Theory of Games and Economic Behavoiur" di von Neumann e Morgenstern. Illustriamo poi vari giochi cooperativi, come i simple games, gli unanimity games e i voting quota games, e consideriamo le principali soluzioni puntuali (*point-valued solutions*), ovvero lo Shapley value e il Banzhaf value. Per quanto riguarda la teoria delle decisioni, consideriamo le probabilità non additive, cioè le fuzzy measures, e il valore atteso (di variabili aleatorie) rispetto ad esse calcolato secondo l'integrale discreto di Choquet. Infine, si mostrerà che l'integrale di Choquet definisce una relazione binaria di preferenza razionale (cioè completa e transitiva) su un tipo particolare di fuzzy measure note come *necessity measures*.

Nel Capitolo 3 considereremo modelli più complessi. Per quanto riguarda la teoria dei giochi, l'attenzione è posta sui global games, che mappano coalition structures o partizioni dell'insieme di giocatori in numeri reali. Per questi giochi, insieme alle soluzioni puntuali già menzionate (cioè lo Shapley value e il Banzhaf value), si considera anche la principale soluzione (*set-valued*), vale a dire il *core* (o nucleo). Per quanto riguarda la teoria delle decisioni, esamineremo le rappresentazioni additive dell'integrale discreto di Choquet, con particolare attenzione alle fuzzy measures supermodulari. Successivamente, ci concentreremo sulle *information functions* che assegnano ad ogni partizione degli stati di natura il valore reale dell'informazione che essa incorpora, come detto sopra. In particolare, queste funzioni sono caratterizzate dall'esistenza di una *set function* tale che il valore di ogni partizione è dato dalla somma dei valori dei suoi blocchi come quantificato dalla set function.

Infine, nel Capitolo 4, descriviamo un ulteriore tipo di gioco cooperativo, definito *in partition function form*, e discutiamo il concetto di soluzione da un punto di vista generale, cioè interpretando i giochi come funzioni su reticoli. Consideriamo successivamente un problema con una storia piuttosto importante in teoria delle decisioni, ossia come definire il valore atteso condizionato rispetto ad una probabilità non additiva. Infine, quest'ultimo problema (ovvero come condizionare nel caso non additivo) è ulteriormente studiato in relazione ad un problema più complesso, cioè come definire l'equilibrio di Nash (di un gioco non cooperativo) quando i giocatori randomizzano e calcolano la loro utilità attesa rispetto a distribuzioni non additive su insiemi di strategie.

Introduction

Game theory was born in the mid-40s with the book "Theory of Games and Economic Behaviour" [59] by von Neumann and Morgenstern. Although "games" usually refers to non-cooperative ones, where players choose strategies, still since the very beginning von Neumann and Morgenstern already considered cooperative games as soon as the number of players exceeded 2. In these latter games, players basically choose whether to cooperate (in some form) or not. More precisely, cooperative games essentially are functions taking real values on ordered structures grounded on a player set. In particular, although one may want to consider cooperation restrictions modelled in a way such that the resulting cooperative game is a poset function, in this work we only deal with lattice functions, without any cooperation restrictions. Specifically, we shall focus on real-valued functions defined on subset or partition lattices. We shall finally also deal with real-valued functions on the lattice of embedded subsets, where this latter is included in the product of the subset and partition lattices.

Mathematically speaking, games and especially cooperative ones share many settings and tools with decision theory. In fact, when choosing a strategy in a non-cooperative game a player is typically a decision maker facing uncertainty. Such a decision problem has indeed driven much investigation in the field of decision under uncertainty. In this work a probability is treated as a valuation of a subset lattice, and the focus is on generalized probabilities or fuzzy measures which are not valuations but only monotone and normalized. Also, in decision theory the fundamental concept of information is formalized in terms of partitions of a set of states of nature. Accordingly, we shall focus on partition functions quantifying the worth of such information for a decision maker. In particular, this latter chooses an action maximizing expected utility, where this latter takes real values on pairs of a state (of nature) and an action. Finer partitions are thus more valuable than coarser ones.

This work is divided into four Chapters:

In Chapter 1 we introduce the needed combinatorial background. Firstly, we define the partial order relation and the associated concept of partially ordered set, together with some important results useful in the sequel. Secondly, we consider the "meet" and "join" binary relations that characterize lattices, where these latter constitute the basis for all subsequent analysis. Lattices are then classified as distributive, modular, semimodular and geometric, thereby allowing for two main examples, namely the subset lattice and the partition lattice, both very important for game theory and decision theory. Lastly, we detail certain properties of lattice (and more generally poset) functions, with special attention on the rank (function) and the Möbius inversion of poset functions.

In Chapter 2 we consider how tools and results provided in Chapter 1 are useful both in game theory and decision theory. As for the former, we consider in detail the development of cooperative game theory since its foundation, i.e. von Neumann and Morgestern's already mentioned book "Theory of Games and Economic Behaviour". We illustrate various coalitional games such as simple games, unanimity games and voting quota games, and consider the principal point-valued solutions: the Shapley and Banzhaf values. Concerning decision theory, we discuss non-additive probabilities, i.e. fuzzy measures, and the expectation (of random variables) with respect to them computed according to discrete Choquet integral. Finally, Choquet integration is shown to provide a ranking criterion over necessity measures.

In Chapter 3 the focus turns on more complex settings. As for game theory, attention is placed on global games, mapping coalition structures or partitions of players into real numbers. For these games, together with the point-valued solutions mentioned above (i.e. the Shapley and Banzhaf values), we also consider the main set-valued solution, namely the core. Concerning decisions, we examine alternative (i.e. additive) representations of the discrete Choquet integral, with special attention on supermodular fuzzy measures. Next, the focus is placed on information functions, assigning to every partition of states the real-valued worth of the information it encodes, as mentioned above. In particular, these functions shall be characterized by the existence of a set function such the worth of every partition is given by the sum over its blocks of these latter's worth as quantified by the given set function.

Finally, in Chapter 4, we begin by describing a further type of cooperative games, termed $in\ partition\ function\ form$, and discuss the solution concept from a general perspective, that is while looking at games as lattice functions. We next consider an issue with a quite long history in decision theory, namely how to define the conditional expectation with respect to a non-additive probability. Finally, this latter issue (i.e. how to condition in the non-additive case) is further studied in conjunction with a more complex problem: how to define the Nash equilibrium (of a non-cooperative game) when players randomize and compute their expected utilities with respect to non-additive distributions over strategy sets.

Contents

Chapter 1

Combinatorial background on ordered structures

This first Chapter aims to briefly present ordered structures (i.e. poset and lattices) within the standard combinatorial theory framework.
Developing from this, the following Chapters shall detail how these structures are fundamental in cooperative games and decision making.
Definitions, notations and results appearing hereafter are presented following mainly [1], as well as [57].

1.1 Partially ordered sets

The basic ordered structure to start with is a partially ordered set or poset.

Definition 1.1. A partially ordered set $(S, \leqslant)$ is a set S endowed with a binary relation $\leqslant$ satisfying the following properties for all a, b and c in S

- $a \leqslant a$ reflexivity,
- if $a \leqslant b$ and $b \leqslant a$, then $a = b$ antisymmetry,
- if $a \leqslant b$ and $b \leqslant c$, then $a \leqslant c$ transitivity.

Order relation $\leqslant$ is termed "partial" in that, in general, it does not allow to compare any two elements $x, y \in S$; that is to say, it may well be that neither $x \leqslant y$ nor $y \leqslant x$ hold. Historically, a fundamental example of partial order is divisibility among natural numbers $a, b \in \mathbb{N}$, i.e. $a \leqslant b$ if $\frac{b}{a} \in \mathbb{N}$. Of course, given any two strictly positive integers, it is not necessarily true that one is divisible by the other[1]. On the other hand, if any two elements $x, y \in S$ are comparable through order relation $\leqslant$, then $(S, \leqslant)$ is a totally ordered set or a chain. The antisymmetric component of order relation $\leqslant$, i.e. $<$, is defined as

$$x < y \Leftrightarrow x \leqslant y \quad \text{and} \quad x \neq y.$$

Definition 1.2. Let P be an ordered set. Then

$$K = \{a = x_0 < x_1 < \dots x_{n-1} < x_n = b\},$$

is a *chain* if, for all $x_i, x_j \in K$, either $x_i \leqslant x_j$ or $x_j \leqslant x_i$.

The set $\mathbb{R}$ of real numbers ordered by the standard less-than-or-equal relation i.e. $\leq$, is perhaps the simplest example of a totally ordered set.

Definition 1.3. Let P be an ordered set, with $x, y \in P$. The *covering relation* $\lessdot$ is defined as follows: $x \lessdot y$ (which reads "y covers x" or equivalently "x is covered by y") if $x < y$ and there is no $z \in P$ such that $x < z < y$.

In the first example of a poset we gave, namely $\mathbb{N}$ ordered by divisibility, a number x covers the number y if and only if x divided by y is a prime number. Conversely, in the poset of real numbers ordered by the standard less-than-or-equal relation no element covers another. Note that $\mathbb{N}$ is countably infinite while $\mathbb{R}$ is a continuum. More simply, in this thesis we deal only with finite ordered structures (i.e. posets and lattices, see below).

Definition 1.4. A chain $K = \{a = x_0 < x_1 < \dots x_{n-1} < x_n = b\}$ between two elements a and b is *maximal* (or unrefinable) if each element is covered

[1] In fact, this framework yields the *divisor lattice*, where a main tool used in this work, namely the Möbius function, was firstly conceived (see Rota (1964) [49]).

by its successor, i.e. $x_i \lessdot x_{i+1}$ for $0 \leqslant i \leqslant n-1$. The length of a chain K is defined as $l(K)$=$|K|$-1, where $|K|$ is the cardinality of K.

The definitions provided thus far lead to introduce the Jordan-Dedekind JD-condition, which is fundamental in that it yields the *rank* function. In fact, as discussed in the sequel (see last section of this first Chapter), the rank constitutes the first and perhaps simplest example of poset function (as well as of lattice function), and functions mapping poset (and lattice) elements into the set of real numbers constitute the main concern of the present work.

Definition 1.5. Let P be an ordered set. We say P has a *bottom element*, denoted by 0, if $0 \in P$ and $0 \leqslant x$ for all $x \in P$. Dually, P has a *top element*, denoted by 1, if $1 \in P$ and $x \leqslant 1$ for all $x \in P$.

Definition 1.6. A poset P satisfies the JD-condition if for any given $x, y \in P$ with $x < y$, all maximal chains between x and y have the same length. If all maximal chains from the bottom element 0 to x have the same finite length, this common length is called the rank of x and denoted by $r(x)$.

Proposition 1.1.1. *Let P be a poset with 0. If P satisfies the JD-condition, then for the rank function it holds:*

- $r(0) = 0$,
- $a \lessdot b$ *implies* $r(b) = r(a) + 1$ *for all* $a, b \in P$.

The rank basically measures the *height* of poset elements. This can be visualized through the *Hasse diagram*, which is essentially a graph where nodes (or vertices) are poset elements and any two nodes are linked whenever one covers the other (in terms of the covering relation, see Definition 1.3 above). Nodes are grouped into levels so that the bottom element is the only node in level 0, next all elements covering the bottom element are in level 1, and in general all elements with rank k ($k = 0, 1, 2, \ldots$) are in the k-th level. Hence, within the Hasse diagram, the rank of poset elements is seen to be the length of any shortest path connecting them with the bottom element.

Definition 1.7. An ordered set $(P, \leqslant)$ is an *antichain* if any two distinct elements are incomparable, that is $x \leqslant y$ implies $x = y$ for all $x, y \in P$.

A *segment* $[x, y]$, $x, y \in P$, is the set of all elements z between x and y, that is $[x, y] = \{z \in P : x \leqslant z \leqslant y\}$. A partially ordered set is *locally finite* if every segment is finite. As already clarified, this thesis only deals with finite, and thus a fortiori locally finite, structures.

A fundamental poset we introduce in this Chapter is the power set $P(X)$ of X. Take X to be a finite set; the power set $P(X)$ consists of all subsets of X ordered by set inclusion, i.e. for $A, B \in P(X)$, define $A \leqslant B$ if and only if $A \subseteq B$. The power set of a finite set $N = \{1, \ldots, n\}$ is usually denoted by 2^N. In fact, if $|N| = n$, then $|P(N)| = 2^n$. In game theory N is a set of players and its subsets are coalitions. In decision theory, N is either a set states of nature, or else a set of criteria for evaluating different alternatives. The former case corresponds to decision making under uncertainty, while the latter is commonly referred to as multicriteria decision making.

Also note that $2^n = \sum_{0 \leqslant k \leqslant n} \binom{n}{k}$, where binomial coefficients $\binom{n}{k} = \frac{n!}{k!(n-k)!} = \binom{n}{n-k}$ provides the cardinality of the k-th and $n-k$-th levels of poset $(2^N, \subseteq)$.

The next fundamental ordered structure to be considered is the lattice one, whose definition relies upon the notion of upper and lower bound (the same often used in calculus), detailed hereafter following Stern [57]. Let $Q \subseteq P$ be an arbitrary subset of poset P; then, an element $u \in P$ is said to be an *upper bound* (maximal element) of Q if $x \leqslant u$ for all $x \in Q$. An upper bound u of Q is said to be its *least upper bound*, or *join*, or *supremum*, if $u \leqslant x$ for all upper bounds x of Q. A generic poset needs not have a least upper bound, but it cannot have more than one. Dually, $l \in P$ is said to be a *lower bound* of Q if $l \leqslant x$ for all $x \in Q$. A lower bound l of Q is said to be its *greatest lower bound*, or *meet*, or *infimum*, if $x \leqslant l$ for all lower bounds x of S.

Definition 1.8. A *lattice* L is a partially ordered set in which every two-element set has a supremum and a infimum. In this case, for all $x, y \in P$ the supremum of x and y, $\sup\{x, y\}$, will be denoted as $x \vee y$ and named the

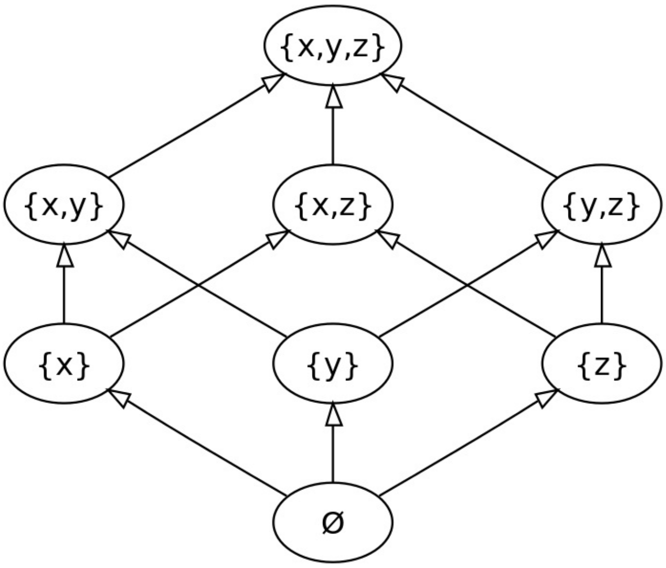

Figure 1.1: Power set $P(X)$ of X where $X=\{x, y, z\}$.

join of x e y. Analogously the infimum $\inf\{x, y\}$ will be denoted by $x \wedge y$ and called the *meet* of x and y.

A lattice $(L, \vee, \wedge)$ is *complete* if $\sup Y$ and $\inf Y$ exist for all $Y \subseteq L$. A simple example of lattice is provided by any subset $L \subseteq P(X)$ which is closed under intersection and union. In this case L is a *lattice of subsets*. Here $A \vee B = A \cup B$ and $A \wedge B = A \cap B$.

A *bounded lattice* is a lattice that, in addition, has a greatest element 1 (the top element) and a least element 0 (the bottom element) satisfying

$$0 \leqslant x \leqslant 1 \quad \text{for all} \quad x \in L.$$

This thesis deals exclusively with bounded lattices.

1.2 Join-irreducible elements and distributive lattices

This and the following Sections contain definitions enabling to introduce two specific lattices, namely the *distributive* and the *geometric* ones. The aim is to outline some main differences and similarities between the two, in order to best detail the key role played by the lattice of subsets (distributive) and that of partitions (geometric) in games and decisions.

Definition 1.9. Let L be a lattice. An element $x \in L$ is *join-irreducible* if

- $x \neq 0$ (in case L has a bottom element),
- $x = a \vee b$ implies $x = a$ or $x = b$ for all $a, b \in L$.

This means that a join-irreducible element x cannot be represented as a join of two lattice elements unless one of them is x itself. A *meet-irreducible* element y is defined dually: $y \neq 1$ (in case L has a top element) and $y = a \wedge b$ entails $y = a$ or $y = b$ for all $a, b \in L$.

Denote the set of join-irreducible elements of lattice L by $J(L)$ and the set of meet-irreducible elements by $M(L)$. Both $J(L)$ and $M(L)$ inherit the order relation $\leqslant$ and thus are ordered set themselves.

In a finite lattice L, an element is join-irreducible if and only if it covers just one element (see Davey-Priestley [9, p.53]). This makes $J(L)$ extremely easy to identify in the Hasse diagram. In subset lattice $P(X)$, join-irreducible elements are singletons, i.e. $J(P(X)) = \{\{x\} : x \in X\}$.

Definition 1.10. A lattice is *distributive* if the following identities hold for all $x, y, z \in L$,

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z),$$

$$x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z).$$

Weakening this definition enables to define modular lattices.

Definition 1.11. A lattice is *modular* if for all elements $x, y, z \in L$

$$z \leqslant x \ \text{ implies } \ x \wedge (y \vee z) = (x \wedge y) \vee z.$$

Every distributive lattice is also modular, but the converse is not true. Any power set lattice $P(X)$ is distributive. More generally, any lattice of sets is distributive. In Davey-Priesley [9] it is proved that every distributive lattice is isomorphic to a lattice of subsets. In Markowski [38] the following is proved:

Theorem 1.2.1. *A finite lattice is distributive if and only if the number of its join-irreducible elements is equal to the length of the lattice itself, where the length of a lattice is the length of a longest maximal chain in the lattice.*

For example, in the lattice 2^N of subsets of $N = \{1, ..., n\}$, there are n join-irreducible elements (i.e. the n sigletons) and n is indeed the length of the lattice. In fact, singletons $\{x\}$, $x \in N$, are the atoms of the lattice.

Definition 1.12. Let L be a lattice with bottom element 0 and top element 1. An *atom* in L is an element that covers the bottom element 0. Dually, a *coatom* is an element covered by the top element 1.

Definition 1.13. A lattice L with bottom element 0 is called *atomic* if for every $x \in L$, $x \neq 0$, there exists an atom $p \in L$ such that $p \leqslant x$.

Definition 1.14. A lattice with bottom element 0 is called *atomistic* if every element ($\neq 0$) is a join of atoms.

A simple example of atomic lattice which is not also atomistic is $\{1, 2, 4\}$, the set of divisors of 4, ordered by the "divisor of" relation (see above): it is atomic, with $\{2\}$ being the only atom, but it is not atomistic, since 4 cannot be obtained as least common multiple of atoms. In fact, 4 is a join-irreducible element which is not also an atom.

The difference between atoms and join-irreducible elements, as well as between atomic and atomistic lattices is further detailed and studied in the sequel, when dealing with the so-called lattice of embedded subsets (see Grabish 2010 [23]), which is relevant in cooperative game theory insofar as games in partition function form are concerned (see Thrall and Lucas 1963 [58]).

1.3 Closure operator: semimodular and geometric lattices

This Section introduces the partition lattice, used next in Chapter 2 when dealing with global games and information functions [18]. It is a lattice which is semimodular and geometric.

Definition 1.15. A lattice is called *semimodular* if for all $a, b \in L$

$$a \wedge b \lessdot a \quad \text{implies} \quad b \lessdot a \vee b.$$

Any semimodular lattice always satisfies the JD-condition (see Aigner [1, p. 47]). The following important theorem also holds:

Theorem 1.3.1. *Let L be a lattice with 0. L is semimodular if and only if L possesses a rank function such that for all $x, y \in L$*

$$r(x \vee y) + r(x \wedge y) \leq r(x) + r(y).$$

L is modular if and only if for all $x, y \in L$

$$r(x \vee y) + r(x \wedge y) = r(x) + r(y).$$

In our setting, the most important semimodular lattice is that of partitions. A partition of a set S, denoted by $\pi = \{A_1, \ldots A_{b(\pi)}\}$ with $|\pi| = b(\pi)$, is a family of non-empty disjoint subsets $A_1, \ldots A_{b(\pi)} \subseteq S$ called *blocks*, whose union is S. Hence $A_k \bigcap A_{k'} = \emptyset$ for $1 \leq k < k' \leq b(\pi)$, $\bigcup_{1 \leq k \leq b(\pi)} A_k = S$. Denote by $\mathcal{P}^N$ the set of all partitions of N, and consider the *coarsening* orded relation defined as follows: given $b(\pi) \leq b(\sigma)$, $\pi = \{A_1, \ldots A_{b(\pi)}\}$ is coarser than $\sigma = \{B_1, \ldots B_{b(\sigma)}\}$, $\sigma \leqslant \pi$, if for each $i \in \{1, \ldots, b(\sigma)\}$ there is $j \in \{1, \ldots, b(\pi)\}$ such that $B_i \subseteq A_j$.

Here $\sigma \leqslant \pi$ reads "π is coarser than σ" or, equivalently, "σ is finer than π". The rank function is defined by $r(\pi) = n - b(\pi)$ for all $\pi \in \mathcal{P}^N$ $(n < \infty)$. Hence, the $(n-k)$-level of $\mathcal{P}^N$, denoted by $\mathcal{P}^N(n-k)$, consists of all partitions with exactly k blocks. For $0 < k \leq n$, numbers $|\mathcal{P}^N(n-k)| =: S_{n,k}$ are the *Stirling numbers of the second kind*, while numbers $|\mathcal{P}^N| =: B_n$ the *Bell numbers* [27, Chapter 6, p. 257]. Thus we have

$$B_n = \sum_{k=1}^{n} S_{n,k} \text{ for all } k, n \in \mathbb{N}.$$

The top element of $\mathcal{P}^N$ is the coarsest partition, with a single block, i.e. $P_\top = \{N\}$, while the bottom element has n blocks, each being a singleton: $P_\perp = \{\{1\}, \ldots, \{n\}\}$. The meet $\pi \wedge \sigma$ is the coarsest partition finer than both π, σ and, analogously, $\pi \vee \sigma$ is the finest partition coarser than both π, σ. With these operations, $(\mathcal{P}^N, \wedge, \vee)$ is a lattice.

In particular, it is semimodular but not modular (see [1, Chapter 2]). Semimodular lattices also obtain, in general, by means of a *closure operator*, together with the *Steinitz exchange axiom*.

Definition 1.16. Let P be a poset. A map $cl : P \to P$ is called a *closure operator* (on P) if for all $x, y \in P$ the following properties are satisfied:

- $x \leqslant cl(x)$ (extensive),
- $x \leqslant y$ implies $cl(x) \leqslant cl(y)$ (increasing),
- $cl(cl(x)) = cl(x)$ (idempotent).

Definition 1.17. A map $cl : x \mapsto cl(x)$ satisfies the *Steinitz exchange axiom* if for all $A \subseteq S$ and $p, q \in S$,

$$p \notin cl(A) \text{ and } p \in cl(A \vee q) \quad \text{implies} \quad q \in cl(A \vee p).$$

The complete lattice of closed subsets (i.e. all subsets $A \subseteq S$ such that $A{=}cl(A)$) is semimodular and geometric.

The first and most important example (from which the term "Steinitz exchange axiom" derives) is that of a vector space over a division ring. Let

V be the set of vectors and $A \longrightarrow \overline{A}$ the linear closure, i.e., $v \in \overline{A}$ if and only if v is linearly dependent on A. That is v can be expressed as linear combination $v = \alpha_1 u_1 + \cdots + \alpha_k u_k$ of elements $u_1, ..., u_k \in A$.

Definition 1.18. A lattice is *geometric* if it is finite, atomistic and semimodular.

The partition lattice is geometric because it is finite and all its elements can be obtained as a join of atoms, where these latter are the $\binom{n}{2}$ (distinct) partitions consisting of $n-1$ blocks, $n-2$ of which are singletons while the remaining one is a pair (i.e. a 2-cardinal subset of N; see Definition 1.12 above).

Definition 1.19. A lattice is *complemented* if it has bottom and top elements 0 and 1, and if every element $x \in L$ has a complement x', that is to say an element $x' \in L$ such that $x \vee x' = 1$ and $x \wedge x' = 0$.

All geometric lattices are relatively complemented, meaning that every segment is complemented. In 2^N every subset has a unique complement, while in $\mathcal{P}^N$ a generic partition may well have several complements.
Another important difference between the subset and partition lattices concerns *modular elements* (see Stern [57, p. 74]).

Definition 1.20. An ordered pair (a, b) of elements of a lattice L is a *modular pair* and we write $a\ M\ b$, if for all $c \in L$

$$c \leqslant b \text{ implies } c \vee (a \wedge b) = (c \vee a) \wedge b.$$

Definition 1.21. In a lattice L, an element b is called a *modular element* if $x\ M\ b$ holds for every $x \in L$.

In the subset lattice 2^N every element is modular, while in the partition lattice $\mathcal{P}^N$ there are $2^n - n$ modular elements, namely all partitions having at most only one block whose cardinality is greater than 1.

1.4 Lattice functions and Möbius inversion

As cooperative games are in fact lattice (or, more generally, poset) functions, this section introduces lattice functions and their basic properties.

A lattice function is any map $f : L \to \mathbb{R}$, where L is a lattice. Accordingly, the rank is one of the simplest such functions. In addition, it is also monotone and bottom-normalized.

Definition 1.22. A lattice function $f : L \to \mathbb{R}$ is *bottom-normalized* if $f(0) = 0$.

Definition 1.23. A lattice function is *monotone* if $x \leqslant y$ implies $f(x) \leqslant f(y)$ for all $x, y \in L$.

Definition 1.24. A lattice function is

- *supermodular* if $f(x \vee y) + f(x \wedge y) \geq f(x) + f(y)$ for all $x, y \in L$,
- *submodular* if $f(x \vee y) + f(x \wedge y) \leq f(x) + f(y)$ for all $x, y \in L$,
- a *valuation* of L if it is both supermodular and submodular.

In cooperative game theory bottom normalization and monotonicity are both standard assumptions (see below).

Let P be a locally finite poset with bottom element, and F a field of characteristic 0, usually $\mathbb{R}$. A function $f : P \times P \to F$ is called *incidence function* on P with values in F if $x \nleqslant y$ implies $f(x, y) = 0$.
The set $\mathbb{A}$ of all incidence functions of P over F, forms a vector space over F that is to say, for $f, g \in \mathbb{A}, k \in F$,

$$(f + g)(x, y) := f(x, y) + g(x, y),$$

$$(kf)(x, y) := kf(x, y).$$

The product $f * g$ of $f, g \in \mathbb{A}$ is defined by convolution:

$$(f * g)(x, y) := \sum_{x \leqslant z \leqslant y} f(x, z)g(z, y).$$

Since every segment in P is finite, the sum is finite, so that the above expression is always well defined. With these operations (namely $+$, $\cdot$, $*$), the space of all incidence functions forms the so-called *incidence algebra* of P over F. Perhaps the most important incidence function (especially in this thesis) is the so-called *zeta function* defined as follows:

$$\zeta(x,y) := \begin{cases} 1 & \text{if } x \leqslant y, \\ 0 & \text{otherwise.} \end{cases}$$

The elements in $\mathbb{A}$ which possess an inverse with respect to the convolution are called the *units* of $\mathbb{A}$.

Theorem 1.4.1. *A function $f \in \mathbb{A}$ is a unit if and only if $f(x,x) \neq 0$ for all $x \in P$. A unit f possesses a unique inverse f^{-1}.*

Therefore ζ is invertible in $\mathbb{A}$ and its inverse is the Möbius function $\mu := \zeta^{-1}$. In particular, the Möbius function $\mu \in \mathbb{A}$ is obtained recursively as follows (see Aigner [1, p. 139]):

$$\mu(x,x) = 1 \quad \text{and} \quad \mu(x,y) = -\sum_{x \leqslant z < y} \mu(z,y) \quad \text{for all } x,y \in P,\ x \leqslant y.$$

Hence, in particular, if $x \lessdot y$ then $\mu(x,y) = -1$.

Definition 1.25. The *Möbius inversion* of a poset function $f : P \to \mathbb{R}$ is $\mu^f(x) = \sum_{y \leqslant x} \mu(y,x) f(y)$.

Möbius inversion was described as the combinatorial analogue of the Fundamental Theorem of calculus thinking of μ^f as the derivative of f, for one in fact has $f(x) = \sum_{y \leqslant x} \mu^f(x)$ (see Rota [49]).

An important result concerns the Möbius inversion of valuations of distributive lattices [1, Theorem 4.63, p. 190 (Davis-Rota)].

Theorem 1.4.2. *If L is a locally finite distributive lattice with bottom element, then every valuation is uniquely determined by its values on the set of join-irreducible elements.*

In particular, valuations of subset lattice 2^N (which is distributive and finite, and thus a fortiori locally finite), have Möbius inversions living only on $J(L) = N$, i.e. on the atoms $\{i\}$ of 2^N. As we shall see, this has important implications for the solution concept in cooperative game theory, as it will be discussed in the next Chapter.

Chapter 2

Games and decisions in discrete settings

In this Chapter we use previous definitions and results of lattice theory with the intent to provide the basics of cooperative game theory and decision theory. More complex settings shall be treated in the following Chapters. In the first part, after a short historical presentation allowing to better delimit our framework, we introduce coalitional games, with special emphasis on some special ones (unanimity, simple and supermodular or convex games [54]) and on the solution concept. In the second part, attention turns on decision theory, focusing on von Neumann-Morgenstern expected utility model, discrete fuzzy measures and Choquet expectation with respect to these latter.

Game theory starts in the late twenties with John von Neumann's analysis of two-person zero-sum games, where one player's gain equals the other's loss (as in poker, chess and, possibly, war). In collaboration with Oskar Morgenstern he extended his research, leading to games which have more players or are not zero-sum (for example, in the well-known prisoners' dilemma, where both players attain maximum payoff when they both do not confess). In their book "*Theory of Games Economic Behavior*" [59] von Neumann and Morgenstern proposed methods for analysing games in more general settings, and their pioneering contribution is probably the most important milestone in the history of game theory. Subsequent work can be classified into two main categories. The first is non-cooperative game theory, concerned with games in normal or extensive form, where players have each to choose a strategy. The other approach is cooperative game theory, where strategic details are ignored. Rather, the focus is on coalitions (= subsets) $A \subseteq N$ of players and what they can achieve when all members cooperate with each other, and utilities are transferable across players. The transferable-utility assumption entails that how to share the fruits of cooperation is the central issue.

A common interpretation is that cooperative game theory studies the outcome of join (i.e. coordinated) actions in a situation with external commitment. Von Neumann and Morgenstern's approach is non-cooperative as long as the concern is on two-player zero-sum games. But the approach is also cooperative for other games. In particular, they provide the first analysis in history for simple games that we will discuss later in this Chapter.

2.1 Coalitional games

Let $N = \{1, ..., n\}$ be a finite player set. Ideally, in *coalitional games*, players can make binding agreements about the distribution of payoff or the choice of strategies even if these agreements are not specified by the rules of the game. Usually, an agreement or a contract is binding if its violation entails high monetary penalties which deter the players from breaking it. There are several real-life situations (e.g. contract law) that allow for this modelling. In such an applicative scenario, coalitional games may be divided into two categories: games with transferable utilities (TU) and games with non-transferable utilities (NTU). In this work we only deal with TU games.

Definition 2.1. A *coalitional game* (with transferable utility) on a set N of players is a function $v : 2^N \to \mathbb{R}_+$ that associates a real number $v(A)$ with each subset A of N. We always assume that $v(\emptyset) = 0$ (bottom-normalized) and that for all $A, B \in 2^N$, $B \subseteq A \Longrightarrow v(A) \geq v(B)$ (monotone).

Let v be a coalitional game and let A be a subcoalition of N. If A is formed, then its members get the amount $v(A)$, called *worth* of the coalition A. In real-life situations, the worth of a coalition is typically represented by a certain amount of money. In a coalitional game, players may be assumed to choose what coalition to join, according to their estimate of the way the payment will be divided among coalition members.

If a coalition A forms, then it can divide its worth, $v(A)$, in any possible way among its members. That is, A can achieve every payoff vector $x \in \mathbb{R}_+^{|A|}$ which is feasible, that is, which satisfies

$$\sum_{i \in A} x_i \leq v(A).$$

Real-life games, where rewards are in money, give examples of TU games. Von Neumann and Morgenstern derive TU coalition functions from the strategic form of games with transferable utilities (i.e. utilities which are linear in money).

The simplest but perhaps most important coalitional game is the unanimity game, first introduced by Shapley in its 1953 paper "Value of n-person games" [55].

Definition 2.2. For all $A, B \in 2^N$, $A \neq \emptyset$ the *unanimity game* U_A associated with the coalition A is defined by:

$$U_A(B) = \begin{cases} 1 & \text{if} \quad B \supseteq A, \\ 0 & \text{otherwise.} \end{cases}$$

As the name suggests, in unanimity games U_A a unit of (transferable) utility is produced if all members $i \in A$ cooperate with each other. Put it differently, there must be unanimous agreement within coalition A. Unanimity games are important because the set $\{U_A : \emptyset \neq A \in 2^N\}$ is a linear basis of the vectorial space of coalitional games (see Peleg and Sudhölter [43, p. 153]). It is easily recognized that unanimity game $U_A(\cdot) = \zeta(A, \cdot)$ is another name for the zeta function in the incidence algebra of subset lattice

$(2^N, \cap, \cup)$. Hence its inverse is Möbius function $\mu(A, \cdot)$. In particular, for subset lattices the Möbius function takes form $\mu(A, B) = (-1)^{|B \setminus A|}$ for all $A \subseteq B \in 2^N$.

2.1.1 Probabilistic and random-order solutions

Let us now turn to the solution concept, i.e. how to share the fruits of cooperation. The solution of a coalitional game v is an additive set function $\phi(v)$ or, equivalently, a valuation of subset lattice $(2^N, \cap, \cup)$.

That is, $\phi(v) = (\phi_1(v), \phi_2(v), ..., \phi_n(v)) \in \mathbb{R}^n$ and $\phi(v)(A) = \sum_{i \in A} \phi_i(v)$. Equivalently, $\phi(v) : 2^N \to \mathbb{R}$ satisfies

$$\phi(v)(A \cup B) + \phi(v)(A \cap B) = \phi(v)(A) + \phi(v)(B),$$

for all $A, B \in 2^N$. The most known solutions of coalitional games are the Shapley and the Banzhaf values, detailed in the next Section. In order to give a proper axiomatization of these values, we first introduce probabilistic and random-order solutions. For definitions and results in this Section we refer to Weber [60]. Fix a player i and let $p^i = \{p^i_A : A \subseteq N \setminus i\}$ be a probability distribution over the collection of coalitions not containing i, hence $\sum_{A \subseteq N \setminus i} p^i_A = 1$ for all $i \in N$ and $p^i_A \geq 0$ for all $A \subseteq N \setminus i$.

Definition 2.3. A mapping $v \xrightarrow{\phi} \phi(v)$ is a *probabilistic value* if for all $i \in N$ there is a probabilistic distribution p^i as above such that

$$\phi_i(v) = \sum_{A \subseteq N \setminus i} p^i_A [v(A \cup i) - v(A)]. \tag{2.1}$$

Geometrically, $\phi : \mathbb{R}^{2^n} \to \mathbb{R}^n$. This is an expectation: any player $i \in N$ shall eventually cooperate by joining some coalition $A \subseteq N \setminus i$ (that has already formed). Then, i will receive *marginal contribuition* $v(A \cup i) - v(A)$. Accordingly, p^i_A is the subjective probability that i joins coalition A, and $\phi_i(v)$ is thus the expected payoff from the game for this player. Note that, for $A \subseteq N \setminus i$, marginal contribution can be expressed in terms of Möbius inversion as

$$v(A \cup i) - v(A) = \sum_{B \in 2^A} \mu^v(B \cup i).$$

We now introduce three important axioms:

- Linearity axiom: ϕ is a linear function, i.e. $\phi(v + w) = \phi(v) + \phi(w)$ and $\phi(\alpha v) = \alpha\phi(v)$, with v, w coalitional games and $\alpha > 0$.
- Dummy axiom: if i is a dummy player in v, i.e. $v(A \cup i) = v(A) + v(i)$ for all $A \subseteq N \setminus i$, then $\phi_i(v) = v(i)$.
- Monotonicity axiom: if v is monotonic, i.e. $v(A) \geq v(B)$, for all $A \supseteq B$, then $\phi_i(v) \geq 0$.

These axioms characterize the class of probabilistic values (see Weber [60]).

Theorem 2.1.1. *A value ϕ is probabilistic if and only if satisfies the linearity, dummy and monotonicity axioms.*

Let $\pi : N \to N$ be a permutation or ordering of the players, with $\pi(i)$ denoting the position occupied by $i \in N$ in ordering π. Also let $\Pi(N)$ be the set (or symmetric group) of all $n!$ such permutations on N.

Definition 2.4. A mapping $v \overset{\phi}{\to} \phi(v)$ is a *random-order value* if there is a probability distribution $p : \Pi(N) \to [0, 1]$ such that for all $i \in N$

$$\phi_i(v) = \sum_{\pi \in \Pi(N)} p(\pi)[v(\{j \in N : \pi(j) \leq \pi(i)\}) - v(\{j \in N : \pi(j) < \pi(i)\})],$$

with $p(\pi) \geq 0$ for all $\pi \in \Pi(N)$, and $\sum_{\pi \in \Pi(N)} p(\pi) = 1$.

To interpret this definition, note that monotonicity of v entails that the players have as their goal the eventual formation of the grand coalition N. Further assume that they see coalition formation as a sequential process: given any ordering π of the players, each player i joins with his predecessors in π gaining a marginal contribution in the game v. Then if the players share a common perception $p(\pi)$ of the likelihood of the various ordering π, the expected marginal contribution of a player is precisely his component of the random-order value. In order to understand the last theorem of this Section and the characterization of the Shapley value, we enunciate two more important axioms:

- Efficiency axiom: solution (or value) ϕ is efficient if $\sum_{i \in N} \phi_i(v) = v(N)$ for all games v.
- Symmetry axiom: if in any game v there are players $i \in N, j \in N \setminus i$ such that $v(A \cup i) = v(A \cup j)$ for all $A \subseteq N \setminus \{i, j\}$, then $\phi_i(v) = \phi_j(v)$.

The following result is a main one contained in Weber (1988) [60].

Theorem 2.1.2. *Any random-order value is a probabilistic value that also satisfies the efficiency axiom.*

2.1.2 Shapley and Banzhaf values

The Shapley value is definitely the key solution concept in cooperative game theory, while the Banzhaf value is the main index quantifying power of voters in voting games. It also has applications in collective coin flipping for distributed randomized computation.

Theorem 2.1.3. *The Shapley value ϕ^{Sh} is the unique probabilistic solution satisfying, in addition, symmetry and efficiency (see [55, 60]). In particular, for all $i \in N$,*

$$\phi_i^{Sh}(v) = \sum_{A \subseteq N \setminus i} \frac{a!(n-a-1)!}{n!} [v(A \cup i) - v(A)],$$

where $a = |A|$.

Note that for $A \subseteq N \setminus i$, the number of maximal chains which include A and $A \cup i$ is $a!(n-a-1)!$, as the number of chains from $A \cup i$ to N is $(n-a-1)!$, and there are further $a!$ chains from $\emptyset$ to A. The denominator $n!$ is the total number of maximal chains in 2^N.

Let us observe that the intuitive interpretation of the formula, as provided in [42], is the following: suppose the players agree to meet at a specific place. Most likely, each will arrive at a different time and it is assumed that all orders of arrival are equally likely, with probability $\frac{1}{n!}$. Finally suppose that if a player i arrives and finds all members $j \in A \subseteq N \setminus i$ (and no others) already there, he receives the amount $v(A \cup i) - v(A)$, that is his marginal contribution to that coalition A. The Shapley value is thus seen to be the random-order value placing the uniform distribution over permutations.

$$\phi_i^{Sh}(v) = \sum_{\pi \in \Pi(N)} \frac{v(\{j \in N : \pi(j) \leq \pi(i)\}) - v(\{j \in N : \pi(j) < \pi(i)\})}{n!}.$$

It seems interesting to observe that, although termed differently, both the Shapley value and Möbius inversion of coalitional games were found (somehow indipendently) by Harsányi in 1963 [30].

The approach yielding the so-called Harsányi dividends is based on the observation that the worth $v(A)$ of (cooperation within) a coalition A consists of 3 parts: (i) the intrinsic value of its members, as singletons, (ii) the added value of cooperation among subsets of these members, and finally (iii) the added value of forming A as an improvement over all existing forms of cooperation. When put this way, we can see that only item (iii) is a merit of forming A, while the rest of the value is generated in strict subsets of A. This item (iii) is referred to as the Harsányi dividend of coalition A. Turning the formula around, expressing the dividends from the payoffs and extending it to the empty set, we obtain the following definition.

Definition 2.5. Let v be a coalitional game. For each coalition A, the dividend $\Delta^v(A)$ is defined, recursively, as follows:

$$\Delta^v(\emptyset) = 0,$$

$$\Delta^v(B) = v(B) - \sum_{A \subset B} \Delta^v(A) \quad \text{if } |B| \geq 1.$$

Moreover, the previous definition leads to the explicit formula

$$\Delta^v(B) = \sum_{A \subseteq B} (-1)^{|B|-|A|} v(A) \quad \text{for all } B \in 2^N.$$

Actually, $\Delta^v(B)$ measures the pure contribution of cooperation within B, since one can interpret it as the contribution of cooperation within the coalition B in addition to what cooperation brings about in all possible subcoalitions that could have formed before the coalition B is determined. This representation or mapping $B \mapsto \Delta^v(B)$ is clearly the specification for the current setting of the more general Möbius inversion of poset functions (see Chapter 1). The mapping is bijective and the two forms are equivalent: given Möbius inversion μ^v, it is possible to retrieve the underlying game v.
Indeed, we have, for any coalition B,

$$v(B) = \sum_{A \subseteq B} \mu^v(A) = \sum_{A \in 2^N} \mu^v(A) U_A(B).$$

That is to say, as already observed, any coalitional game v is a linear combination of unanimity games. The family of unanimity games thus forms a basis of the vector space of coalitional games with dividends as coordinates. Now, let $\sum_{A \in 2^N} \mu^v(A) U_A$ be the (unique) representation of the game v as a linear combination of unanimity games. Calculating the Shapley value using this decomposition is actually quite simple. In particular, in view of the efficiency, dummy and symmetry axioms, for every $i \in N$ we have:

$$\phi_i^{Sh}(U_A) = \begin{cases} \frac{1}{|A|} & \text{if } i \in A, \\ 0 & \text{otherwise,} \end{cases}$$

for each element U_A of the unanimity basis. Indeed, this is precisely how uniqueness of the Shapley value is proved: since its behaviour is determined on each element of the basis, adding the linearity axiom characterizes a unique solution. For generic game v, the Shapley value thus finds the following neat expression:

$$\phi_i^{Sh}(v) = \sum_{A \subseteq N : i \in A} \frac{\mu^v(A)}{|A|} \text{ for all } i \in N.$$

Although the Shapley value mostly applies to generic or real-valued coalitional games, its behaviour on the special class of $\{0,1\}$-valued (or simple) games is now detailed.

Definition 2.6. A *simple game* is a coalitional game where $v : 2^N \to \{0,1\}$.

In particular we have here two sets: $W = \{A \in 2^N : v(A) = 1\}$ the set of *winning coalitions* and $W^C = \{A \in 2^N : v(A) = 0\}$ the set of *losing coalitions*. These sets are complementary, hence $|W| + |W^C| = 2^n$. For the monotonicity of v, each subset including a winning coalition is also winning and, dually, each subset included in a losing coalition is also losing.

A classical example of simple games is provided by *voting quota games*, the generic being denoted by v_w hereafter. Players $i \in N$ are voters, each with an associated strictly positive weight $w_i > 0$, and there is a threshold $w^* > 0$ such that any coalition $A \in 2^N$ is winning if $\sum_{i\in A} w_i > w^*$. Hence,

$$v_w(A) = \begin{cases} 1 & \text{if } \sum_{i\in A} w_i > w^*, \\ 0 & \text{otherwise.} \end{cases}$$

In case $w_i = 1$ for all i and $w^* = \frac{n}{2} + 1$, then v_w is the *voting majority game* [56].

The Banzhaf value Ba is a further probabilistic solution satisfying, in addition, symmetry but not efficiency. This power index of a player in a monotone simple game counts the number of coalitions that are losing but become winning when that player joins them. In voting quota games the sum of all players' power indexes does not necessarily equal the worth of the grand coalition, which is 1, of course.

Definition 2.7. The Banhaf value $v \overset{Ba}{\to} Ba(v)$ is a probabilistic value also satisfying symmetry where all players have the uniform probability distribution over the 2^{n-1} coaltitions they may join:

$$Ba_i^v = \sum_{A\subseteq N\setminus i} \frac{v(A\cup i) - v(A)}{2^{n-1}} \text{ for all } i \in N.$$

In [56] it is shown that its l_1 norm, $||Ba^v||_1 = \sum_i Ba^v_i$, takes its maximum on the voting majority game (see above), where it is about $\sqrt{\frac{2n}{\pi}}$.

Over the years, interest arose in computer science in finding methods for collective coin flipping by n processors which take part in a distributed normalized computation. It is the goal to generate coin flips which are unbiased as possible despite malfunctioning of some of the participating processors. This problem turns out to be equivalent to the quest of simple games with small l_∞ norm for $||Ba^v||_\infty = \max Ba^v_i$. Examples were given of games where exactly half of the coalitions win and $||Ba^v||_\infty = O(\frac{\log n}{n})$ (see [4]).

2.2 Decision Theory

Generally speaking, decision theory deals with formalizing how a decision maker should optimally chose within a set of available alternatives. If such alternatives are in fact strategies available to a player in a non-cooperative game (see above), then the decision is characterized by uncertainty about what other players will do. Although this is a somehow special case, still it has greatly contributed to the main literature on the subject. Alternatively, in this Chapter we consider a decision maker who has to choose an action while nature chooses a state. This is the standard setting for decision under uncertainty [50]. Note that decision theory has more ancient origins than game theory, in that the problem of decision under uncertainty was already studied in the 17th century by Blaise Pascal.

A main model of decision under uncertainty is in terms of choice among lotteries. These latter are probability distributions for a random variable taking finitely many real values. Formally, let $X = \{X_1, X_2, ..., X_n\}$ be a vector, with X_i seen as a quantity of money. Consider the natural ascending order, i.e. $X_1 < X_2 < X_3 < .. < X_n$, and two probability distributions $p = (p_1, p_2, ..., p_n)$ and $q = (q_1, q_2, ..., q_n)$. Fixed X, the decision problem is how to choose between p and q. More generally, the issue is how to characterize a rational [39, Chapters 1 and 6] preference (binary) relation $\gtrsim$ between any two probability distributions $p, q \in \Delta_n$ over X, where Δ_n is the simplex containing all such distributions.

Definitely, the main and perhaps most natural ranking of lotteries obtains through expectation:

$$p \gtrsim q \Longleftrightarrow E_p[x] \geq E_q[x],$$

where $E_p[x] = \sum_{k=1}^{n} p_k X_k$ is the expected (money) value of the lottery. This is indeed the von Neumann-Morgenstern expected utility model. It relies on the following two axioms:

- *Continuity*:
  For all probability distributions $p, q, r \in \Delta_n$,
  $\{\alpha \in [0,1], \alpha p + (1-\alpha)q \gtrsim r\}$ and $\{\alpha \in [0,1], r \gtrsim \alpha p + (1-\alpha)q\}$ are closed intervals of [0,1].

- *Indipendence*:
  For all probability distributions $p, q, r \in \Delta_n$
  $p \gtrsim q \Longleftrightarrow \alpha p + (1-\alpha)r \gtrsim \alpha q + (1-\alpha)r$.

Theorem 2.2.1. *(von Neumann-Morgenstern 1944 [59]) If preference relation $\gtrsim$ satisfies the independence and continuity axioms, then there exists a utility function (over money values) $u : X \to \mathbb{R}$ such that*

$$p \gtrsim q \Longleftrightarrow E_p[u(X)] \geq E_q[u(X)].$$

The independence axiom is however violated in many real-world decision problems. Main examples are the Allais and Ellsberg paradoxes. In particular, in order to introduce the Choquet expected utility model presented in the next Section, the Ellsberg paradox is detailed hereafter in a simplified version.

An urn contains three balls: one is blue, while the others two can be each either yellow or red. One ball is drawn with uniform distribution. Define four lotteries as follows:

- lottery g1: 100 dollars if the ball is blue, 0 if yellow or red.

- lottery g2: 100 dollars if the ball is red, 0 if yellow or blue.
- lottery g3: 100 dollars if the ball is either blue or yellow, 0 if red.
- lottery g4: 100 dollars if the ball is either red or yellow, 0 if blue.

The decision maker DM has to choose between lotteries g1 and g2. Next, the same DM chooses between g3 and g4. When proposed as real-life decisions, the DM commonly prefers g1 over g2 and g4 over g3. This contradicts the von Neumann-Morgenstern expected utility model.

Denote by p_i, $i \in \{0, 1, 2\}$ the subjective probability that (precisely) i red balls are in the urn. Now compare the expected utility of the first two lotteries. Given that g1 is preferred over g2, i.e g1 $>$ g2, it must be:

$$E_p(g1) = \frac{1}{3}u(100) > E_p(g2) = u(100)\left(\frac{1}{3}p_1 + \frac{2}{3}p_2\right),$$

$$\frac{1}{3}u(100) > u(100)\left(\frac{1}{3}p_1 + \frac{2}{3}p_2\right),$$

$$1 > p_1 + 2p_2,$$

where, of course, $p_0 + p_1 + p_2 = 1$. Comparing now g3 and g4 with the latter preferred over the former, we obtain

$$E_p(g4) = \frac{2}{3}u(100) > E_p(g3) = u(100)\left(\frac{1}{3}p_0 + \frac{2}{3}p_1 + p_2\right),$$

$$\frac{2}{3}u(100) > u(100)\left(\frac{1 + p_0 + 2p_2}{3}\right),$$

$$1 > p_1 + 2p_0 \Leftrightarrow -1 > -p_1 - 2p_2,$$

$$1 < p_1 + 2p_2.$$

This contradiction (or impossibility result), together with Allais paradox (which is perhaps even more known) led to search for alternative models of decision under uncertainty. Among these latter, the Choquet expected utility model, where expectation is taken with respect to fuzzy or non-additive probabilities, is a main one. It is detailed in the next Section. In view of this, recall that probability distributions over finite sets are, in fact, set (and thus lattice) functions, precisely the same as coalitional games, but with states of nature instead of players.

As already mentioned, the general setting applying decision under uncertainty is one where the DM has to choose an optimal action, given a bijective probability over states, with a utility function taking real values on pairs of an action and a state [50, 52]. Formally, let $\Omega = \{\omega_1, ..., \omega_n\}$ be the finite set of states of nature. Also let $p : 2^\Omega \to [0, 1]$ be a probability distribution. For any event (or subset of states) $A \in 2^\Omega$, probability $p(A) = Prob[\omega \in A]$ is interpreted as the subjective belief (of the DM) that the true state of nature, i.e. the one that will actually realize, shall be some $\omega \in A$. Clearly, this is a lattice function; in particular, it is a valuation of subset lattice $(2^\Omega, \cap, \cup)$. That is to say,

$$p(A \cap B) + p(A \cup B) = p(A) + p(B) \text{ for all } A, B \in 2^\Omega.$$

As already observed from a different perspective, such valuations have Möbius inversion living only on atoms $\{\omega\} \in 2^\Omega$, i.e. $p(A) = \sum_{\omega \in A} p(\omega)$.

2.2.1 Discrete Choquet expected utility

Although the Choquet integral was originally conceived in terms of measure theory, an extensive literature now deals with the discrete case. In particular, in addition to the finite set Ω of states of nature and probability p as above, consider a set $\mathbb{A} = \{a_1, ..., a_m\}$ of actions available to the DM, and a utility function $u : \mathbb{A} \times \Omega \to \mathbb{R}_+$ taking positive real values on pairs (a, ω) of an action and a state. The utility is supposed to take positive values because when the integrand also takes negative values (discrete) Choquet integration may be symmetric or asymmetric (see [25]). For reasons of space, this double possibility is not addressed here.

Definition 2.8. A *non-additive probability* (or fuzzy measure) γ is a function $\gamma : 2^\Omega \to [0, 1]$ such that $\gamma(\emptyset) = 0$, $\gamma(\Omega) = 1$ and $A \supseteq B \Rightarrow \gamma(A) \geq \gamma(B)$.

Here the decision problem amounts to choose an action $a \in \mathbb{A}$. This is achieved through the Choquet integral by associating with every action a an expectation of random variable $u_a(\omega)$, $\omega \in \Omega$. In particular, expectation is taken with respect to a fuzzy measure γ. Therefore, such a Choquet expected utility of actions provides a criterion for ranking them: the higher their expected value, the better.

Adopting a standard notation in this field, consider a permutation $(\cdot)$ of the indexes $i = 1, \ldots, n$ in Ω, i.e. $(\cdot) : \{1, ..., n\} \to \{1, ..., n\}$, such that $u_a(\omega_{(1)}) \leq u_a(\omega_{(2)}) \leq \cdots \leq u_a(\omega_{(n)})$, and fix $u(\omega_{(0)}) := 0$. Then, the *discrete Choquet integral* of $u_a : \Omega \to \mathbb{R}_+$ with respect to γ is

$$E_\gamma(u_a) := \sum_{k=0}^{n} [u_a(\omega_{(k)}) - u_a(\omega_{(k-1)})] \cdot \gamma(\{\omega_{(k)}, \omega_{(k+1)}, ..., \omega_{(n)}\}).$$

This integral now has a variety of important applications not only in decision under uncertainty, but also in multicriteria decision making, where actions are alternative options in real-life problems, while states are criteria assigning each a score to every alternative. In this case, the fuzzy measure quantifies how these criteria interact with each other, and Choquet integration enables to determine what options get an higher overall (i.e. expected) score [24, 8].

The following calculations are intended to determine a fuzzy measure enabling to overcome the Ellsberg paradox for the simplified version presented in the previous section. Here $\Omega = \{\omega_1, \omega_2, \omega_3\}$ and ω_i is the state of nature where the number of red balls in the urn is exactly $i - 1$ (for $i = 1, 2, 3$). For g1, all 3! permutations are equivalent, because the payoff is indipendent from which state is the "true" one:

$$E_\gamma(g1) = \frac{1}{3} u(100).$$

For g2, the only permutation is the identity $(i) = i$ for $i = 1, 2, 3$:

$$E_\gamma(g2) = u(100) \left[(0 - 0)\gamma(\Omega) + \left(\frac{1}{3} - 0 \right) \gamma(\{\omega_{(2)}, \omega_{(3)}\}) + \left(\frac{2}{3} - \frac{1}{3} \right) \gamma(\{\omega_{(3)}\}) \right].$$

For g3, the only permutation is $\omega_{(1)} = \omega_3,\ \omega_{(2)} = \omega_2,\ \omega_{(3)} = \omega_1$:

$$E_\gamma(g3) = u(100)\left[\frac{1}{3} + (0-0)\gamma(\Omega) + (\frac{1}{3} - 0)\gamma(\{\omega_{(1)}, \omega_{(2)}\}) + (\frac{2}{3} - \frac{1}{3})\gamma(\{\omega_{(1)}\})\right].$$

For g4, as for g1, any of the 3! permutations is fine because the probability of receiving 100 is $\frac{2}{3}$ independently from states:

$$E_\gamma(g4) = \frac{2}{3}u(100).$$

Now, observed choices g1 over g2 and (by the same DM) g4 over g3 entail:

$$E_\gamma(g1) = \frac{1}{3}u(100) > E_\gamma(g2) = u(100)\frac{1}{3}[\gamma(\{\omega_{(2)}, \omega_{(3)}\}) + \gamma(\{\omega_{(3)})\}],$$

Hence,

$$1 > \gamma(\{\omega_{(2)}, \omega_{(3)}\}) + \gamma(\{\omega_{(3)}\}).$$

On the other hand,

$$E_\gamma(g4) > E_\gamma(g3) \Leftrightarrow \frac{2}{3} > \frac{1}{3} + \frac{1}{3}\gamma(\{\omega_{(1)}, \omega_{(2)}\}) + \frac{1}{3}\gamma(\{\omega_{(1)}\}),$$

Hence,

$$1 > \gamma(\{\omega_{(1)}, \omega_{(2)}\}) + \gamma(\{\omega_{(1)}\}).$$

These two inequalities no longer constitute a contradiction, in that even the simple fuzzy measure below is one example where they both hold:

$$\gamma(\emptyset) = 0, \gamma(\Omega) = 1, \gamma(A) = \frac{1}{3} \text{ for all } A \in 2^\Omega \text{ such that } \emptyset \subset A \subset \Omega.$$

We now consider the discrete Choquet integral as a criterion for ranking a special type of non-additive probabilities, namely necessity measures.

2.2.2 Ranking necessity measures via Choquet integration

The same problem considered by von Neumann and Morgenstern, namely how to rank probability distributions over a fixed set of (positive) values (of money), has been more recently addressed for the case where probability distributions are replaced with necessity measures [45]. These latter are peculiar fuzzy measures introduced below.

A lottery has form $\sum_{\omega\in\Omega} p(\{\omega\})\delta_\omega$ with $p(\{\omega\}) \geq 0, \sum_{\omega\in\Omega} p(\{\omega\}) = 1$, where $\delta_\omega : 2^\Omega \to \{0,1\}$ is the Dirac measure at ω. That is,

$$\delta_\omega(A) = \begin{cases} 1 & \text{if } \omega \in A, \\ 0 & \text{if } \omega \in A^c = \Omega \setminus A. \end{cases}$$

It may be recognized that in terms of coalitional games (looking at Ω as containing players) δ_ω is the unanimity game $U_{\{\omega\}}$ defined on 1-cardinal coalition $\{\omega\} \in 2^\Omega$. Conversely, in the present setting, δ_ω is the elementary lottery giving a unitary amount of money whenever state ω occurs.

Let $Prob(\Omega)$ denote the set of lotteries on Ω. If the DM follows the already cited von Neumann-Morgenstern axioms, then there exists a *utility* function $u : \Omega \to [0,1]$ such that for all probabilities $p, q \in Prob(\Omega)$

$$p \succcurlyeq q \Leftrightarrow \int u \; dp \geq \int u \; dq.$$

where $\int u \; dp := \sum_{\omega\in\Omega} p(\{\omega\})u(\omega)$ denotes discrete integration (i.e. summation) for notational convenience. Moreover there are $\omega_1, \omega_0 \in \Omega$ with $u(\omega_1) = 1, u(\omega_0) = 0$ such that for all $p \in Prob(\Omega)$ the following equivalence $\sim$ holds:

$$p \sim \left(\int u \; dp\right)\delta_{\omega_1} + \left(1 - \int u \; dp\right)\delta_{\omega_0}.$$

The interpretation is that for the DM any lottery p can be reduced to a bet on head versus tail, that is to a lottery having for support the best and the worst states, where expected utility $\int u \; dp$ is interpreted as the probability that he wins the bet [35].

Following [45], in this Section the generic fuzzy measure is denoted by v, the same as for coalitional games (rather than γ as in the previous Section). As we shall see, this choice is also useful for representing the Choquet integral in the terms of the core $\mathcal{C}(v)$ of a supermodular fuzzy measure v [21].

Definition 2.9. A fuzzy measure v on Ω is a *necessity measure* if for all $A, B \subset \Omega$, $v(A \cap B) = \min\{v(A), v(B)\}$.

An important class of non-additive probabilities which contains necessity measures is that of belief functions. These latter received some attention for modelling the DM's knowledge [41]. They can also be considered as objects that a DM would have to rank. The aim of Rébillé's paper [45] is to rank different necessity measures without any recourse to von Neumann-Morgenstern's techniques. It turns out that the Choquet expectation will be the criterion for ranking necessity measures.

For short we will denote the set of necessity measures on Ω by $Nec(\Omega)$.

Now we state some axioms that the binary relation $\succsim$ may fulfil.

- (WO) $\succsim$ is a weak order (i.e. complete and transitive).

- (MON) Monotonicity: for all $v, w \in Nec(\Omega)$, $[v \geq w] \Rightarrow [v \succsim w]$, where

  $$[v \geq w] \Leftrightarrow \text{ for all } A \subset \Omega,\ [v(A), v^d(A)] \subset [w(A), w^d(A)],$$

  while v^d is the *dual* of v, i.e. $v^d(A) = 1 - v(A^c)$ and the same for w^d.

- (AGR) Agreement: for all $u, v, w \in Nec(\Omega)$, for all $\alpha \in (0,1)$ if u, w agree and v, w agree then $[u \sim v] \Rightarrow [\alpha u + (1-\alpha)w \sim \alpha v + (1 - \alpha w)]$. Any two $v, w \in Nec(\Omega)$ are defined to agree if for all $A, B \in 2^\Omega$, $(v(A) - v(B))(w(A) - w(B)) \geq 0$.

- (ARCH) $\succsim$ is Archimedean: for all $v, w \in Nec(\Omega)$,

  $$[v \prec w] \Rightarrow [\exists \alpha \in (0,1) \text{ such that } v \prec \alpha w + (1-\alpha)u_\Omega]$$

  and $[\exists \alpha \in (0,1)$ such that $\alpha w + (1-\alpha)u_\Omega \prec v \precsim w] \Rightarrow$
  $\Rightarrow [\exists \alpha' \in (\alpha, 1)$ such that $\alpha' w + (1-\alpha')u_\Omega \precsim v]$.

  The (ARCH) axiom can be understood in the following manner in conjunction with (MON). Let $v \prec w$ and $\alpha \in (0,1)$. Since $U_\Omega \leq w$, we have that $U_\Omega \leq \alpha w + (1-\alpha)U_\Omega \leq w$, under (MON) we get $U_\Omega \preceq$

$\alpha w + (1-\alpha)U_\Omega \preceq w$, the (ARCH) axiom tells us that if α is close enough to 1, then one should obtain also $v \prec \alpha w + (1-\alpha)U_\Omega$. The last axiom ensures that the preference relation is not trivial.

- (NDEG) $\succsim$ is not degenerate: $\exists v, w \in Nec(\Omega)$ such that $v \succ w$. This axiom can be further specified, under (WO) and (MON), (NDEG) is equivalent to: $\exists v \in Nec(\Omega)$ such that $v \succ U_\Omega$.

Consider $F(\Omega) := \{A^u : A \neq \emptyset, A \subset \Omega\}$, where $A^u = \{B : A \subset B \subset \Omega\}$ (A^u stands for the upset generated by A). These sets of subsets of Ω are known as *principal filters*, and $F(\Omega)$ contains all of them. Also let v be a necessity measure on 2^Ω.

There is a unique decomposition of v as a linear combination of unanimity games defined on elements of a maximal chain $\mathcal{K} = \{\emptyset = A_0, A_1, \ldots, A_n = \Omega\}$ of subsets, by means of Möbius inversion:

$$v = \sum_{i=1}^{n} \mu^v(A_i) U_{A_i}$$

where $\mu^v(A_1), ..., \mu^v(A_n) > 0$ and $\sum_{i=1}^{n} \mu^v(A_i) = 1$, while U_A is the usual unanimity game. Note that $\mathcal{K}$ is a poset, and indeed μ^v is the Möbius inversion of the restriction of v on poset $(\mathcal{K}, \supseteq)$. Such a (unique) decomposition, in turn, provides a fundamental tool used in [45], namely the Choquet integral $\int v \, d\beta$ of necessity measure v with respect to monotone set function $\beta : F(\Omega) \to [0,1]$. In particular,

$$\int v \, d\beta := \sum_{i=1}^{n} \mu^v(A_i)\beta(A_i^u).$$

We are now able to state the main non-additive preference representation theorem:

Theorem 2.2.2. *Let $\succsim$ be a binary relation on $Nec(\Omega)$; if $\succsim$ satisfies (WO), (MON), (AGR), (ARCH), (NDEG), then there exists a monotone set function $\beta : F(\Omega) \to [0,1]$ such that for all $v, w \in Nec(\Omega)$*

$$v \succcurlyeq w \Leftrightarrow \int v\ d\beta \geq \int w\ d\beta.$$

Moreover, there is a $\omega_1 \in \Omega$ *such that for all* v *in* $Nec(\Omega)$

$$v \sim \left(\int v\ d\beta\right) \delta_{\omega_1} + \left(1 - \int v\ d\beta\right) U_\Omega,$$

and $\beta(\{\omega_1\}^u) = 1$ *as well as* $\beta(\{\Omega\}) = 0$*. Conversely, if the binary relation is represented by a Choquet integral with respect to a monotone set function* $\beta : F(\Omega) \to [0,1]$ *such that* $\beta(\{\omega_1\}^u) = 1$ *as well as* $\beta(\{\Omega\}) = 0$ *for some* $\omega_1 \in \Omega$*, then* $\succsim$ *satisfies (WO), (MON), (AGR), (ARCH), (NDEG).*

An interpretation of the last equivalence is that for the DM any necessity measure can be reduced to a bet on being perfectly informed of the state which occurs or being totally ignorant. The Choquet expectation value of a necessity measure is interpreted as the degree of information it encodes.

With the following Chapter, attention turns on more complex modelling of cooperative games and decisions. In particular, we now consider settings where the lattice of partitions also enters the picture.

Chapter 3

More complex settings: partitions

Developing from the standard setting introduced in the previous Chapter, the aim now turns at presenting more complex situations both for games and decisions. Concerning the former, this Chapter deals with global games, mapping partitions of players (or coalition structures) into real numbers. Concerning decisions, the focus is mainly placed on information functions, assigning to every partition of states (of nature) a real-valued worth, when partitions encode information. In this respect, note that the entropy of partitions is typically a measure of how informative these latter are.

3.1 Game Theory

Let $N = \{1, \ldots, n\}$ be a finite set and denote by $(2^N, \cap, \cup)$ and $(\mathcal{P}^N, \wedge, \vee)$ the associated subset and partition lattices, respectively [1, 57]. As already observed, in cooperative game theory, N contains players, and set functions $v : 2^N \to \mathbb{R}$ are coalitional games [55], while partition functions $h : \mathcal{P}^N \to \mathbb{R}$ are global games [17].

For a coalition $A \in 2^N$, worth $v(A)$ attains when all and *only* members $i \in A$ cooperate. In a partition $P = \{A_1, \ldots, A_{|P|}\} \in \mathcal{P}^N$, where $A_k \cap A_l = \emptyset$

for $1 \leq k < l \leq |P|$ and $A_1 \cup \cdots \cup A_{|P|} = N$, there are *all* players $i \in N$, although distributed over different blocks $A_i \in P$. Hence, worth $h(P)$ may be interpreted as that achieved when all players $i \in N$ cooperate, and do so in a way yielding P as the prevailing coalition structure. Furthermore, $h(P)$ is a global utility level, common to all players, attained when P is the outcome of cooperation. Accordingly, h is conceived to model interaction on global issues, possibly among nations or other organizations, where any P-cooperation provides an amount of a public good. Examples of such global issues provided in [17] are environmental clean-up and preservation, medical research, water scarcity and pollution, etc.

We now introduce the following notation, enabling to deal simultaneously with both coalitional and global games in a seemingly comprehensive manner. Let $(L, \wedge, \vee)$ denote a lattice with order relation $\geqslant$ such that $L \in \{2^N, \mathcal{P}^N\}$. In particular, if $L = 2^N$, then $\geqslant$ is set inclusion $\supseteq$, while $\wedge$ is intersection $\cap$ as well as $\vee$ is union $\cup$. On the other hand, if $L = \mathcal{P}^N$, then $\geqslant$ is coarsening, while $\wedge$ and $\vee$ denote respectively the 'coarsest-finer-than' or *meet* and the 'finest-coarser-than' or *join* between (any two) partitions.

Cooperative games are lattice functions $f : L \to \mathbb{R}$. Their Möbius inversion [49, p. 344] is $\mu^f : L \to \mathbb{R}$ given by $\mu^f(x) = \sum_{x_\perp \leqslant y \leqslant x} \mu_L(y, x) f(y)$, where $x_\perp$ is the bottom element and μ_L is the Möbius function, defined recursively on ordered pairs $(y, x) \in L \times L$ by $\mu_L(y, x) = -\sum_{y \leqslant z < x} \mu_L(z, x)$ if $y < x$ (i.e. $y \leqslant x$ and $y \neq x$) as well as $\mu_L(y, x) = 1$ if $y = x$, while $\mu_L(y, x) = 0$ if $y \not\leqslant x$ (see Chapter 1). Bottom elements are $x_\perp = \emptyset$ for $L = 2^N$ and $x_\perp = P_\perp = \{\{1\}, \ldots, \{n\}\}$ for $L = \mathcal{P}^N$. Concerning partitions $P, Q \in \mathcal{P}^N$, if $Q < P = \{A_1, \ldots, A_{|P|}\}$, then for every block $A \in P$ there are blocks $B_1, \ldots, B_{k_A} \in Q$ such that $A = B_1 \cup \cdots \cup B_{k_A}$, with $k_A > 1$ for at least one $A \in P$. Segment $[Q, P] = \{P' : Q \leqslant P' \leqslant P\}$ is thus isomorphic to product $\times_{A \in P} \mathcal{P}(k_A)$, where $\mathcal{P}(k)$ denotes the lattice of partitions of a k-set. Accordingly, let $l_k = |\{A : k_A = k\}|$ for $k = 1, \ldots, n$. Then [49, pp. 359-360],

$$\mu_{\mathcal{P}^N}(Q, P) = (-1)^{-n+\sum_{1 \leq k \leq n} l_k} \prod_{1<k<n} (k!)^{l_{k+1}}.$$

In cooperative game theory, Möbius inversion is important primarily because it provides a very useful basis for the vector space $\mathbb{R}^{|L|}$ of real-valued functions on L. In fact, in the incidence algebra of L, Möbius function μ_L is an inverse of zeta function $\zeta_L : L \times L \to \{0,1\}$ defined by $\zeta_L(y,x) = 1$ if $y \leqslant x$ and 0 otherwise. Hence $f(\cdot) = \sum_{x \in L} \mu^f(x)\zeta_L(x,\cdot)$ for all $f(\cdot) \in \mathbb{R}^{|L|}$. When $L = 2^N$ zeta function $\zeta(A,\cdot)$ corresponds to traditional [55] unanimity game $U_A(\cdot)$.

3.1.1 Global and coalitional games

Definition 3.1. A *global game* is a partition function $h : \mathcal{P}^N \to \mathbb{R}$. We assume $h(P_\perp) = 0$, where $P_\perp$ is the finest partition in $\mathcal{P}^N$ i.e. the bottom element of the partition lattice.

Let P^A be a partition of A, P^B be a partition of B and $A \cap B = \emptyset$, then $P^A \cup P^B$ is a well-defined partition of $A \cup B$. Let $P_\perp^A$ and $P_\top^A$ denote the finest and coarsest partition of A, respectively, i.e., $P_\perp^A = \{\{i\} : i \in A\}$ and $P_\top^A = \{A\}$. Let $F_0(L)$ the subspace of 0-normalized lattice functions, namely those $f : L \to \mathbb{R}$ such that $f(x_\perp) = 0$.

Definition 3.2. Let $h \in F_0(\mathcal{P}^N)$ be a global game. Gilboa e Lehrer [17] define the induced coalitional game $v_h \in F_0(2^N)$ by

$$v_h(A) = h(\{A\} \cup P_\perp^{A^c}),$$

for $\emptyset \neq A \in 2^N$ and $v_h(\emptyset) = 0$.

That is, the worth of a coalition A is the worth of the partition where A is the only non-singleton block. Note that this definition entails

$$v_h(\{i\}) = h(P_\perp) = \sum_{j \in N} v(\{j\}) = 0$$

for all $i \in N$.

Definition 3.3. Let $v \in F_0(2^N)$. Gilboa and Lehrer [17] define the induced (additively separable or partially additive) global game $h_v \in F_0(\mathcal{P}^N)$ by

$$h_v(P) = \sum_{A \in P} v(A).$$

In fact, Gilboa and Lehrer state that $h_v \in F_0(\mathcal{P}^N)$, but this is the case only if $v \in F_0(2^N)$ satisfies $\sum_{i \in N} v(\{i\}) = 0$.

In Gilboa-Lehrer [17] is proposed a Shapley value for global games. Their axiomatization is presented below.

Definition 3.4. An operator $\psi : F_0(\mathcal{P}^N) \longrightarrow \mathbb{R}^N$ is a Shapley value for global games if satisfies these four axioms:

- Linearity: for all $h, h' \in F_0(\mathcal{P}^N)$, $\psi(h+h') = \psi(h)+\psi(h')$ and $\psi(\beta h) = \beta(\psi h)$, for all $\beta > 0$.

- Dummy player: if i is a dummy in h, then $\psi_i(h) = 0$, where $i \in N$ is a dummy player in h if for all $P \in \mathcal{P}^N$

$$h(P) = h(P \wedge \{\{i\}, N \setminus i\}).$$

- Interchangeable players: for all $h \in F_0(\mathcal{P}^N)$ and $i, j \in N$, if i and j are interchangeable in h, then $\psi_i(h) = \psi_j(h)$, where i and j are interchangeable in h if for all $P \in \mathcal{P}^N$

$$h(P \wedge \{\{i\}, N \setminus i\}) = h(P \wedge \{\{j\}, N \setminus j\}).$$

- Efficiency: $\sum_{i \in N} \psi_i(h)(i) = h(\{N\})$.

Let us comment briefly on the interpretation of these axioms. Linearity has its usual meaning: suppose that the players in the global game h (say, environmental clean-up) are also involved in a different global game g (e.g., art treasures preservation). It is desirable that one will be able to solve each game separately and obtain the same outcome that would result from considering the two global issues together $(h + g)$. Similarly, homogeneity (that is, $\psi(\beta h) = \beta\psi(h)$ simply means scale invariance.

Next consider the dummy axiom. A player i is a "dummy" in a global game h if the payoff is independent of i's cooperative behaviour. As formulated, it is only required that for every partition P, $h(P)$ will equal the payoff of the partition obtained from P by player i's desertion. Obviously, this also means that player i may decide to join another set in P but will still not affect the payoff. It seems reasonable that such a player will have no share in the surplus of cooperation $h(\{N\})$. As for the third axiom, two players i and j are "interchangeable" if for every partition P the desertion of i from his/her current coalition to form a separate coalition $\{i\}$ has the same impact on $h(P)$ as j would have (notice that in the formulation given above the term $h(P)$ was cancelled on both sides of the equality). The requirement that i and j will get the same payoff according to $\psi(h)$ has a flavour of "symmetry" or "fairness". Finally, the efficiency axiom simply requires that the overall surplus of cooperation, $h(\{N\})$, will be shared among the players.

The following theorem is central in [17].

Theorem 3.1.1. *There is a unique Shapley value ψ for the space of global games and it is equal to the Shapley value of the induced game, i.e.*
$\psi(h) = \phi^{Sh}(v_h)$ *for all* $h \in F_0(\mathcal{P}^N)$.

Evidently, this is because global games h are in fact dealt with in terms of the associated coalitional game v_h. The kind of issues arising from such an approach are perhaps best introduced by quoting directly Gilboa and Lehrer [17, Remark 5.1.2, p. 144]:

"It may seem surprising that the Shapley value of h does not depend on all of the numbers $\{h(P)\}_{P\in\mathcal{P}}$. As a matter of fact, the (small) subset $\{h(\{A\} \cup P_{\perp}^{A^c})\}_{A\subseteq N}$, i.e. the value of h on "all-or-none" partitions alone determines $\phi(h)$, while the value of h on partitions which are not of this form is immaterial. An attempt to understand this phenomenon may be the following. The axiom which should be held responsible for it is the interchangeability axiom: it focuses on the damage that a player may cause by deserting his/her coalition, and should two such players have the same "threat"

power, they are given the same payoff. In a way, this axiom simply distinguishes between those players who do cooperate in some way (i.e. in some non-trivial coalition) and those who do not (singletons). The former have a viable threat, the latter do not. The precise way in which the "cooperative" players cooperate (i.e. via which coalition) does not matter; it only matters that they do. Hence, the payoff depends only on the best that the "cooperative" players may obtain $h(\{A\} \cup P_{\perp}^{A^c})$ where A is the set of "cooperative" ones. Whether this property is desirable or not is debatable. We believe that in some situations it will be quite intuitive and will capture the essence of the cooperative global game, while in others it may well be inappropriate. Since the interchangeability axiom seems innocent, yet guarantees uniqueness, we chose it to define "the Shapley value." However, one may certainly wish to consider other solution concepts".

The Shapley value in Gilboa and Lehrer's axiomatization [17] rises doubts from distinct points of view. Firstly, because Möbius inversion $\mu^{h_v}(P)$, in their notation $\alpha_P(h_v)$, is not properly defined on the bottom element $P_{\perp}$. In particular, their finding is:

$$\alpha_P(h_v) = \begin{cases} \alpha_A(v) = \mu^v(A) & \text{if } P = (\{A\} \cup P_{\perp}^{A^c}), \\ 0 & \text{otherwise.} \end{cases}$$

This means that if a global game is additively separable, then its Möbius inversion lives only on the $2^n - n$ modular elements of $\mathcal{P}^N$, i.e. on those partitions of the form $\{A\} \cup P_{\perp}^{A^c}$ for $1 < |A| < n$, in addition to $P_{\perp}$ and $P_{\top}$ [1, Ex. 13, p. 71]. However, for $|A| = 1$ coefficient $\alpha_{P_{\perp}}(h_v)$ is not defined because there shall be, in general, n distinct coefficients $\alpha_A(v)$ such that $|A| = 1$. In [17], coalitional games $v \in F_0(2^N)$ only satisfy $v(\emptyset) = 0$; therefore, the general case is of course $0 < \alpha_{\{i\}} \neq \alpha_{\{j\}} > 0$ for $i \in N$, $j \in N \setminus i$. Then, any further coalitional game $w \neq v$ satisfying $\sum_{i \in N} v(\{i\}) = \sum_{i \in N} w(\{i\})$ as well as $\mu^v(A) = \mu^w(A)$ for all A such that $|A| > 1$ also additively separates h, i.e. $h_v = h_w$ or equivalently

$$\sum_{A \in P} v(A) = \sum_{A \in P} w(A).$$

Hence, unless $v(\{i\}) = 0$ for all i, there is a continuum of such $w \neq v$. It is evident that the Shapley value $\phi^{Sh}(v_h) = \psi(h)$ also crucially suffers this non-uniqueness issue. In fact, if $w \neq v$ and $h_w = h_v$ (as above), then $\phi^{Sh}(w) \neq \phi^{Sh}(v)$ and still $\psi(h_w) = \psi(h_v)$.

3.1.2 The core

In Chapter 2 we defined the Shapley and Banzhaf values. These are point-valued solutions, while the core, detailed below, is a set-valued solution, that is a possibly empty convex subset of $\mathbb{R}^n$. Historically, the core was conceived in n-player strategic or non-cooperative games (briefly mentioned in the previous Chapter 2), as the set of outcomes (or n-vector of payoffs) such that no coalition can make all its members better off by deviating in a coordinated (i.e. correlated [31]) manner.

In cooperative game theory, the core $\mathcal{C}(v)$ of a coalitional game v is a (possibly empty) set of valuations ϕ of 2^N, i.e. $\phi(A) = \sum_{i \in A} \phi(\{i\})$.

Definition 3.5. The core $\mathcal{C}(v)$ of v is the set

$$\mathcal{C}(v) = \{\phi \text{ valuation of } 2^N : \ \phi(A) \geq v(A) \ \text{ for all } A \in 2^N, \phi(N) = v(N)\}.$$

Hence, the core is in fact the collection of point-valued solutions that assign to each coalition A a worth $\phi(A)$ which exceeds or equals the worth $v(A)$ achieved through cooperation within A only. The interpretation is that if players are rewarded according to a value (if any) $\phi \in \mathcal{C}(v)$, then cooperation is promoted toward the formation of the grand coalition N. Note that $\mathcal{C}(v) \subseteq \mathbb{R}^n$ is a convex polyhedron in $\mathbb{R}^n$. For example, if $v(N) = 1$ and $v(A) = 0$ for all $A \subset N$, then the core of v is in fact the $n-1$-dimensional unit simplex. To see that the core is convex, let $\phi, \phi' \in \mathcal{C}(v)$. Then, $\psi := (\alpha\phi + (1-\alpha)\phi')$ also belongs to $\mathcal{C}(v)$ for all $\alpha \in [0, 1]$, in that

$$\psi(A) = \alpha\phi(A) + (1-\alpha)\phi'(A) \geq \alpha v(A) + (1-\alpha)v(A) = v(A).$$

The main result concerning the core of coalitional games is due to Shapley [54, Theorem 4, p. 21].

Theorem 3.1.2. *If v is a supermodular coalitional game , then $\mathcal{C}(v) \neq \emptyset$.*

The proof is a constructive one: if v is supermodular, then the set $ex(\mathcal{C}(v))$ of core extreme points is easily determined in terms of maximal chains in subset lattice $(2^N, \cap, \cup)$ or, equivalently, permutations $\pi : \{1, ..., n\} \rightarrow \{1, ..., n\}$. Let v be a coalitional game, for all permutations $\pi \in \Pi(N)$ and all $i \in N$, define

$$\phi_i^\pi(v) = v(\{j \in N : \pi(j) \leq \pi(i)\}) - v(\{j \in N : \pi(j) < \pi(i)\}).$$

This expression appears in the definition of random-order values, and is the marginal contribution of player i along the maximal chain

$$\mathcal{K}^\pi = \{A_0^\pi, A_1^\pi, ..., A_n^\pi\}$$

(of coalitions) identified by permutation π, where

$$A_k^\pi = \{j \in N : \pi(j) \leq k\}.$$

Now, for any $A \in 2^N$, either $A \in \mathcal{K}^\pi$ and thus $\sum_{i \in A} \phi_i^\pi(v) = v(A)$, or else

$$v(A) + v(B) \leq v(A \cup B) + v(A \cap B) \text{ for all } A, B \in 2^N,$$

entails

$$v(A \cup i) - v(A) \geq v(B \cup i) - v(B) \text{ for all } B \subseteq A \subseteq N \setminus i.$$

In turn, this guarantees that $\sum_{i \in A} \phi_i^\pi(v) \geq v(A)$ for all $A \in 2^N \setminus \mathcal{K}^\pi$.

In [54], the vertices of the core are shown to correspond to such vectors or points $\phi^\pi \in \mathbb{R}^n, \pi \in \Pi(N)$. Hence, the core has at most $n!$ (distinct) extreme points. Since the Shapley value has been shown in Chapter 2 to be a convex combination of extreme points $\phi^\pi, \pi \in \Pi(N)$, as long as v is supermodular such a value is in the core. That is to say, as long as v is supermodular,

$$\mathcal{C}(v) \ni \phi_i^{Sh}(v) = \sum_{\pi \in \Pi(N)} \frac{\phi^\pi(v)}{n!}.$$

Moreover, from a geometrical perspective, it can be seen as the center of gravity of the extreme points of the core.

Among other things, in the reminder of this Chapter it is shown how the core of a supermodular non-additive probability (or belief function) provides useful results in terms of additive representations of the discrete Choquet integral.

3.2 Decision Theory

We begin dealing with decisions by following Gilboa and Schmeidler [21] with the aim to examine alternative representations of the discrete Choquet integral. Next, attention is placed on information functions, assigning to every partition of states the real-value worth of the information it encodes.

3.2.1 Additive representation of non-additive measure

The representation of beliefs by real-valued set functions which do not necessarily satisfy additivity has a long history. "Belief functions" were introduced by Dempster [10, 11] and Shafer [52]. Their theory is not directly related to decision making under uncertainty, nor is their concept of "probability" derived from preferences. Rather, they assume that "weight of evidence" for events is a primitive, and study the "belief functions" which are generated by summation of such weights. Belief functions are a special class of "non-additive measures" or "capacities", characterized by a condition called "total monotonicity".

In Gilboa and Schmeidler [19] are characterized preferences which may be represented by a utility function and a set of additive measures, in the sense that preferences obey maximization of the minimal expected utility over all measures in the given set. These preferences can also be represented by the non-additive model (with maximization of the Choquet integral) in case the set of measures is the core of a supermodular measure. In particular, supermodular measures correspond to uncertainty aversion and that belief functions are supermodular.

We have already detailed in the previous Chapter a well-known theorem in cooperative game theory, according to which the space of all non-additive measures ("games") is spanned by a natural linear basis (of "unanimity games"). This result may be viewed as suggesting an isomorphism between non-additive set functions on the original space (of states of the world) and additive ones on a larger space (of all events). Using this result, Gilboa and Schmeidler [21] show that the Choquet integral with respect to any non-additive set function v is simply some linear combination of the minima of the integrand (over various events). Furthermore, if v is a belief function, this linear combination reduces to a weighted average. Thus, for such probabilities v, the integral is both mean of minima (over events) and, since they are also convex, minimum of means (where the minimum is taken over additive measures in the core).

Hereafter definitions and theorems are from [21]. Let Ω be a non-empty set of states of nature and let Σ be a finite algebra of events defined on it. We will assume without loss of generality that $\Sigma = 2^{\Omega}$. It will also be useful to define $\Sigma' = \Sigma \setminus \{\emptyset\}$. A function $v : \Sigma \to \mathbb{R}$ with $v(\emptyset) = 0$, is called a non-additive signed measure or a *capacity*. The space of all capacities will be denoted by V and will be considered as a linear space (over $\mathbb{R}$) with the natural operations. We now recall some definitions, adding new ones. For $v \in V$:

- v is *monotone* if $A \subseteq B$ implies $v(A) \leq v(B)$ for all $A, B \in \Sigma$.
- v is *normalized* if $v(\Sigma) = 1$.
- v is *additive* if $v(A \cup B) = v(A) + v(B)$ for all $A, B \in \Sigma$ with $A \cap B = \emptyset$ (that is, a valuation).
- v is *supermodular* if for all $A, B \in \Sigma$
$$v(A \cup B) + v(A \cap B) \geq v(A) + v(B).$$
- v is *non-negative* if $v(A) \geq 0$ for all $A \in \Sigma$.

- v is *totally monotone* if it is non-negative and, for all $A_1, ..., A_n \in \Sigma$, $v(\cup_{i=1}^n A_i) \geq \sum_{I:\emptyset \neq I \subseteq \{1,...,n\}} (-1)^{|I|+1} v(\cap_{i \in I} A_i)$.
- v is a *measure* or additive probability if it is non-negative and additive.
- v is a *belief function* if it is normalized and totally monotone.

We denote the space of real-valued functions on Ω (or random variables) by $F = \{f \text{ such that } f : \Omega \to \mathbb{R}\} = \mathbb{R}^\Omega$. For $v \in V$ and $f \in F$, the Choquet integral of f with respect to v is defined to be (see Chapter 2, section 2.2.1)

$$\int f \, dv := \sum_{k=0}^{n} [f(\omega_{(k)}) - f(\omega_{(k-1)})] \cdot v(\{\omega_{(k)}, \omega_{(k+1)}, ..., \omega_{(n)}\}).$$

such that $f(\omega_{(k)}) \leq f(\omega_{(k+1)})$ for $0 < k \leq n$. Observe that the Choquet integral is linear in the game v, that is for all $v, w \in V$, $\alpha, \beta \in \mathbb{R}$ and $f \in F$,

$$\int f \, d(\alpha v + \beta w) = \alpha \int f \, dv + \beta \int f \, dw.$$

Further important properties of the Choquet integral can be found in [22]. In view of Shapley [54] theorem (see above) stating that if v is supermodular, than $\mathcal{C}(v) \neq \emptyset$, consider the following result.

Theorem 3.2.1. *(Rosenmuller [46, 47]) A monotone game v is supermodular if and only if $\mathcal{C}(v) \neq \emptyset$ and for every $f \in F$,*

$$\int f \, dv = \min_{p \in \mathcal{C}(v)} \int f \, dp.$$

In order to demonstrate the main theorem in this Section (i.e. that if v is a belief function, then the Choquet integral with respect to v is both a minimim of averages and an average of minima), we firstly need the following results.

Lemma 3.2.2. *For $f \in F$ and $A \in \Sigma'$,*

$$\int f \, dU_A = \min\{f(\omega) : \omega \in A\}.$$

The proof of this lemma can be found in [46]. Recalling the (unique) representation $v = \sum_{A\in\Sigma'} \mu^v(A)U_A$ of v, we are now able to state the following theorem:

Theorem 3.2.3. *For every $v \in V$ and $f \in F$,*

$$\int f\ dv = \sum_{A\in\Sigma'} \mu^v(A) \left[\min_{\omega\in A} f(\omega)\right].$$

Given the previous results, this theorem can be proved as follows:

$$\begin{aligned}
\int f\ dv &= \int f\ d\left(\sum_{A\in\Sigma'} \mu^v(A)U_A\right) = \\
&= \sum_{A\in\Sigma'} \mu^v(A) \left(\int f\ dU_A\right) = \\
&= \sum_{A\in\Sigma'} \mu^v(A) \left[\min_{\omega\in A} f(\omega)\right].
\end{aligned}$$

Recall that if v is totally monotone, then $\mu^v(A) \geq 0$ for all $A \in \Sigma'$. If, in addition, v is normalized, i.e. it is a belief function, then

$$\sum_{A\in\Sigma'} \mu^v(A) = v(\Omega) = 1,$$

entailing that the Choquet integral of a function f with respect to v can be expressed as a weighted average over all minima on all non-empty events. In the extreme case where v is additive or a valuation, we find again a specialization of a main theorem presented in Chapter 1 applying to valuations of locally finite distributive lattices, namely

$$\mu^v(A) = 0 \text{ for all } A \text{ such that } |A| > 1.$$

In this case, indeed, the integral of f with respect to v is an average of the values of f or, if you will, of the minima of f over singletons. Another extreme case is where $v = U_\Omega$, and the integral of f with respect to v is simply the minimum of f over the whole of Ω. While both these extremes cases were known to be special cases of the Choquet integral, the last theorem shows that any Choquet integral (to be precise, the integral with respect to any fuzzy measure v) is no more than some average over minima. On the other hand, let us recall that a totally monotone v may be also represented as the minimum of all integrals of f with respect to measures in a certain set (the core of v). If v is also normalized, each of these measures p is simply some weight vector and the integral of f with respect to p is a p-average over f's values. To sum, if v is a belief function, then the Choquet integral with respect to v is both a minimum of averages and an average of minima:

Theorem 3.2.4. *Assume that v is a belief function. Then for every $f \in F$,*

$$\int f \; dv = \sum_{A \in \Sigma'} \mu^v(A) \left[\min_{\omega \in A} f(\omega) \right] = \min_{p \in \mathcal{C}(v)} \sum_{\omega \in \Omega} p(\{\omega\}) f(\omega).$$

3.2.2 Information functions

Let $(\mathcal{P}^\Omega, \wedge, \vee)$ be the lattice of partitions of finite set $\Omega = \{\omega_1, ..., \omega_n\}$. This lattice is described in Chapter 1, Section 1.3 and also already appears in this Chapter in terms of global games (Section 3.1). Information functions assign to every partition $P \in \mathcal{P}^\Omega$ a real value $f(P)$ quantifying the worth of the information it provides. In this respect, perhaps a main example of such a quantification is given by the entropy (of partitions), detailed in the sequel. It seems best to clarify immediately, though, that while global games are monotone partition functions with respect to the coarsening order relation, as long as information is concerned finer partitions are more valuable that coarser ones. In other terms, the Hasse diagram of the partition lattice is turned upside-down, with the finest partition on top.

In information theory, originated by Claude Shannon's 1948 seminal work "*A Mathematical Theory of Communication*" [53], partitions play indeed a central role. According to Shannon's source coding theorem, the number of bits needed to represent the result of an uncertain event is, on average, given by its entropy. Formally, let $p = (p_1, p_2, ..., p_n)$ be a probability distribution over Ω. The *entropy* of p is ($\log = \log_2$)

$$H_p := -\sum_{i=1}^{n} p_i \log p_i.$$

The maximum of this entropy attains on the uniform probability distribution

$$\bar{p} = \left(\frac{1}{n}, \frac{1}{n}, ..., \frac{1}{n}\right) \Rightarrow H_{\bar{p}} = -\sum_{i=1}^{n} \frac{1}{n} \log \frac{1}{n} = -\log \frac{1}{n} = \log n.$$

Developing form this maximaizer $\bar{p}$, the entropy of a partition P obtains by summing over blocks the probability that the "true" state (i.e. the one which will realize) shall be in that block.

Definition 3.6. The entropy of any partition $P \in \mathcal{P}^{\Omega}$ is

$$H_P := -\sum_{A \in P} \frac{|A|}{n} \log \frac{|A|}{n}.$$

Evidently, this entropy measure attains its maximum $\log n$ on the finest partition (which is indeed the most informative, see below), and its minimum 0 on the coarsest one. Hence, the entropy $H : \mathcal{P}^{\Omega} \to [0, \log n]$ of partitions thus provides first example of an information function.

As in the decisional model introduced in Chapter 2, Section 2.2.1, consider a set of action $\mathbb{A}$ and a utility function $u : \mathbb{A} \times \Omega \to \mathbb{R}$, with the DM optimally choosing a *strategy* $s : \Omega \to \mathbb{A}$. Let $\mathcal{S}$ contain all such strategies ($|\mathcal{S}| = |\mathbb{A}| \times |\Omega|$). Now, if the DM is endowed with information P, than this means that any two states ω, ω' can only be distinguished if $\omega \in A, \omega' \in A'$, $A, A' \in P$ and $A \neq A'$. In other terms the DM cannot choose two distinct actions a, a' for distinct states ω, ω' unless it is possible to distinguish between these latter. Formally, one may say that the strategy has to be "measurable" with respect to P, i.e. constant over each block.

For every P in $\mathcal{P}^{\Omega}$, let $\mathcal{S}_P$ denote the set of all P-measurable strategies:

$$\mathcal{S}_P = \{s : s \in \mathcal{S}, \omega \in A \ni \omega' \Rightarrow s(\omega) = s(\omega') \text{ for all } A \in P\}.$$

With this additional ingredients it may be recognized that the traditional [50] decision problem detailed in Chapter 2, namely $\max_{a \in \mathbb{A}} \int_{\Omega} u_a \, dp$ where p is a subjective probability, as usual, and $u_a(\cdot) = u(a, \cdot)$ and , amounts in fact to choose an optimal $P_\top$-measurable strategy ($P_\top = \{\Omega\}$).

Since all our sets are finite ($|\mathbb{A}| = m$), there surely exists $\max_{a\in\mathbb{A}} \int_B u_a \, dp$ for every $B \in 2^\Omega$. For each block of $A \in P$, the DM may *condition* the choice of a distinct optimal action for each block. Accordingly, the worth of (the information provided by) P is

$$f(P) = \sum_{A\in P} \max_{a\in\mathbb{A}} \int_A u_a \, dp.$$

In other terms, the worth of P is the DM's expected utility associated with an optimal P-admissible strategy:

$$E_{(\mathbb{A},u)}(P) = \max_{s\in\mathcal{S}_P} \int_\Omega u_{s(\cdot)} \, dp = f(P).$$

Generally speaking, any $f : \mathcal{P}^\Omega \to \mathbb{R}$ may be termed *information function* as long as there are actions and a utility $(\mathbb{A}, u)$ satisfying $f = E_{(\mathbb{A},u)}$.

Definition 3.7. For a set function v and $B \in 2^\Omega$, the *B-anticore* of v, denoted $\mathcal{AC}_B(v)$, is the set of all valuations λ of subset lattice $(2^B, \cap, \cup)$ satisfying $\lambda(A) \le v(A)$ for $A \in 2^B$ and $\lambda(B) = v(B)$.

The existence characterization of information functions relies on the notion of additive separability introduced in Section 3.1.1 (see Definition 3.3 pag. 38).

Theorem 3.2.5. *Given Ω and p as well as a partition function f (on $\mathcal{P}^\Omega$), the following are equivalent:*

- *there are $(\mathbb{A}, u)$ as above, with $0 \le u \le M$, such that $f = E_{(\mathbb{A},u)}$;*

- *f is additively separated by a set function v such that $\mathcal{AC}_B(v) \ne \emptyset$ and $v(B) \le Mp(B)$ for all $B \in 2^\Omega$.*

Information functions defined in this way (specifically with a traditional or additive probability p as prior) are interesting in their own right. In addition, they are useful for reasoning about the Savage's Sure-Thing Principle, which is in itself quite debated within decision theory.

The Sure-Thing Principle, conceived by Savage in 1954 says that if a decision maker would take a certain action if he knew that an event E obtained, and also if he knew that its negation $\overline{E}$ obtained, then he should take that action even if he knows nothing about E. Savage [50, p. 21] illustrates this as follows: *A businessman contemplates buying a certain piece of property. He considers the outcome of the next presidential election relevant. So, to clarify the matter to himself, he asks whether he would buy if he knew that the Democratic candidate were going to lose, and decides that he would. Similarly, he considers whether he would buy if he knew that the Republican candidate were going to lose, and again finds that he would. Seeing that he would buy in either event, he decides that he should buy, even though he does not know which event will obtain. It is all too seldom that a decision can be arrived at on the basis of this principle, but I know of almost no other extralogical principle governing decisions that finds such ready acceptance.*

In [3], the authors use the concept of conditional probability to address a conceptual puzzle related to Savage's "Sure-Thing Principle". As P represents the DM's information, if the true state of nature is ω, then the DM does not know that, but knows only that the true state is included in the block $A \in P$ to which ω belongs. Probability p is a DM's prior (belief) about the likelihood of states, before information P is available, and in particular before knowing what block $A \in P$ contains the true state. In this view, conditioning appears as follows: as soon as the DM get informed that the true state of nature is some $\omega \in A \in P$, he updates his prior p by assigning null probability to states $j \in A^c$ and probability $\frac{p(\{\omega'\})}{p(A)}$ to states ω' in A. This means that the conditional probability of any event B, given that event A is known to have occurred, is $p(B|A) := \frac{p(B \cap A)}{p(A)}$. This is the common conditional probability of B, given A.

Gilboa and Lehrer's approach [21] allows to investigate what mathematical conditions characterize those partition functions that satisfy the Sure-

Thing Principle. To this end, they introduce the notion of *non-intersecting* partitions. Any two partitions $P, Q \in \mathcal{P}^\Omega$ are said to be non-intersecting if for every $A \in P$ either (i) there is $B \in Q$ such that $A \subseteq B$, or else (ii) there are $B_1, \dots B_k \in Q$ such that $A = B_1 \cup \cdots \cup B_k$. This may be equivalently stated by means of the following notation: for all $A \in 2^\Omega$ such that $\emptyset \subset A \subset \Omega$ and for all $P = \{B_1, \dots B_{|P|}\}$, denote by $P^A = \{A \cap B : B \in P, A \cap B \neq \emptyset\}$ the partition of A induced by P. Then, any two partitions P and Q are non intersecting if (and only if) there is an event $A \in 2^\Omega$ such that $P = P^A \cup P^{A^c}$, $Q = Q^A \cup Q^{A^c}$ (where $P^{A(A^c)}, Q^{A(A^c)}$ are partitions of $A(A^c)$), with $P^A \leqslant Q^A$ and $P^{A^c} \geqslant Q^{A^c}$ (or the opposite). Note that if P and Q are comparable, say $P \geqslant Q$, then they are trivially non-intersecting.

Definition 3.8. (See [18, pp. 447-8].) A partition function f is partially commutative if $f(P) + f(Q) = f(P \wedge Q) + f(P \vee Q)$ for all pairs $P, Q \in \mathcal{P}^\Omega$ of non-intersecting partitions.

We note that a partition function f satisfing such a condition for *all* pairs P, Q of partitions, whether intersecting or not, is in fact a valuation of partition lattice $(\mathcal{P}^\Omega, \wedge, \vee)$. Such valuations are constant functions, that is $f(P) = f(Q)$ for *all* $P, Q \in \mathcal{P}^\Omega$ (see [1, Exercise 12.(ii), p. 195] and [17, Proposition 4.6, p. 140]).

Information functions and the Sure-Thing Principle relate as follows.

Definition 3.9. (See [18, pp. 452-3].) A partition function f is said to satisfy the *Sure-Thing Principle* if the following holds: for all $\emptyset \neq A \in 2^\Omega$ and all $P_1, P_2 \in \mathcal{P}^A$ and all $Q_1, Q_2 \in \mathcal{P}^{A^c}$ with $P_1 \geqslant P_2$,

$$f(P_1 \cup Q_1) - f(P_2 \cup Q_1) = f(P_1 \cup Q_2) - f(P_2 \cup Q_2).$$

Gilboa and Lehrer offer the following interpretation: *"in all four partitions the DM would know whether A has occurred or not. Hence, by the Sure-Thing Principle, the DM should not care about what he/she will know should A not occur in order to evaluate information given A. Thus, the left-hand side, which is the marginal value of P_1 to a DM having P_2, (in case* *he/she has Q_1 for A^c), should be the same as in the case Q_2 is the DM's information on A^c"*. In this view, they conclude with the following observation [18, Observation 3.4, p. 453].

Proposition 3.2.6. *An information function f satisfies the Sure-Thing Principle if and only if it is partially commutative (see above).*

Chapter 4

Recent developments and issues

In this final Chapter we firstly describe a further type of cooperative games and discuss the solution concept in general terms, that is while looking at games as lattice functions. Next, we consider an issue with a quite long history in decision theory, namely how to define the conditional expectation with respect to a non-additive probability. Finally, this latter issue (i.e. how to condition in the non-additive case) is further studied in conjunction with a more complex problem: how to define the Nash equilibrium (of a non-cooperative game) when players randomize and compute their expected utilities with respect to non-additive distributions over strategy sets.

4.1 Games on embedded coalitions

This Section focuses on games in partition function form PFF, firstly introduced by Thrall and Lucas in 1963 [58], and then further studied over the years as poset functions [40, 2, 48, 32, 34, 5, 44, 37] More recently, the ordered structure where these functions take their real values has been endowed with the meet and join operators. Hence PFF games, like coalitional and global games, are now lattice functions [26, 23].

While global games assign a worth to every partition of players (see Chapter 3), PFF games assign a worth to every pair consisting of a coalition and a partition such that this latter embeds the former as one of its block. For this reason, in cooperative game theory such pairs are sometimes called *embedded coalitions* (or embedded subsets).

The product lattice $\times^N := 2^N \times \mathcal{P}^N$ is ordered by relation $\sqsubseteq$ obtained by applying pairwise the order relations of the subset and partition lattices:

$$(A, P) \sqsubseteq (B, Q) \Leftrightarrow A \subseteq B \text{ and } P \leqslant Q.$$

Furthermore, $\times^N$ is a lattice with the following meet and join:

$$(A, P) \wedge_\times (B, Q) = (A \cap B, P \wedge Q),$$

$$(A, P) \vee_\times (B, Q) = (A \cup B, P \vee Q),$$

for all $(A, P), (B, Q) \in \times^N$. Now consider the family $\mathcal{E}^N$ of all embedded coalitions $\mathcal{E}^N := \{(A, P) \in \times^N : A \in P\} \subset \times^N$. Evidently, $(\mathcal{E}^N, \sqsubseteq)$ is a poset inheriting the order of $\times^N$. There is no bottom element, while neither the meet nor the join are easily defined in a way such that $(A, P) \wedge (B, Q), (A, P) \vee (B, Q) \in \mathcal{E}^N$ for all pairs $(A, P), (B, Q) \in \mathcal{E}^N$. In fact, PFF games have been dealt with as poset functions until Grabish 2010 define the lattice of embedded lattice [23]. The top element is, of course, $(N, P_\top)$. Concerning the bottom, all elements of the form $(\{i\}, P_\perp)$ are minimal , but they cover no element. Accordingly, a bottom element denoted by $\perp$ to $\mathcal{E}^N$ in introduced, and the resulting poset is

$$\mathcal{E}^N_* := \mathcal{E}^N \cup \{\perp\}.$$

For $|N| > 2$ (of course), the meet and join defined by Grabish are:

$$(A, P) \vee_\mathcal{E} (A', P') := (B \cup B', Q),$$

$$(A, P) \wedge_\mathcal{E} (A', P') := \begin{cases} (A \cap A', P \wedge P') & \text{if } A \cap A' \neq \emptyset, \\ \perp & \text{otherwise.} \end{cases}$$

where B, B' are blocks of $P \vee P'$ containing respectively A and A' , and Q is the partition obtained by merging B, B' in $P \vee P'$. Note that the meet is the same as for the product lattice $\wedge_{\mathcal{E}} = \wedge_{\times}$, while the join is such that the lattice of embedded subsets $(\mathcal{E}^N_*, \wedge_{\mathcal{E}}, \vee_{\mathcal{E}})$ is atomic but not atomistic (see Chapter 1). In particular, there are n atoms $(\{i\}, P_\perp)$ for $i \in N$, together with join-irreducible elements of the form $(\{i\}, P_\perp^{\{jk\}})$, $i, j, k \in N$, $i \notin \{j, k\}$. These latter are the singleton subsets embedded in those partitions that are atoms of the partition lattice, and whose unique 2-cardinal block does not include the singleton subset. Hence, the total number of such non-atomic join-irreducible element is $n\binom{n-1}{2}$, while the total number of join-irreducible elements, whether atoms or not, is $n\left(1 + \binom{n-1}{2}\right)$. The lattice of embedded subsets is neither distributive nor geometric (precisely because it is not atomistic).

Definition 4.1. A PFF game (on N) is any mapping $h : \mathcal{E}^N_* \longrightarrow \mathbb{R}$ such that $h(\perp) = 0$.

As for lattice functions in general, Möbius inversion enables to represent any PFF game as a linear combination of unanimity PFF games. For notational convenience, embedded coalitions (A, P) in the sequel are denoted by denoted by $AP := A\{A, A_2, ..., A_k\}$. Unanimity PFF games are defined as usual, that is

$$U_{AP}(A', P') = \begin{cases} 1 & \text{if } A'P' \sqsupseteq AP, \\ 0 & \text{otherwise,} \end{cases}$$

for all $AP, A'P' \in \mathcal{E}^N_*$, and $h = \sum_{AP \in \mathcal{E}^N_*} \mu^h(AP) U_{AP}$. Möbius inversion μ^h is defined through Möbius function $\mu_{\mathcal{E}}$ in the incidence algebra of the lattice. Denote the modular partition consisting of block A, with $|A| > 1$, and $n - |A|$ singleton blocks by $P^A_\perp$. Given $AP = A\{A, A_2, ..., A_k\}$, consider $A'P' := A'\{A', A_{12}, ..., A_{1l_1}, A_{21}, ..., A_{2l_2}, ..., A_{k1}, ..., A_{kl_k}\}$ with $A'P' \sqsubset AP$, where blocks $A_{ml_1}, ...A_{ml_m} \in P'$ are those that merge into block $A_m \in P$, for $m = 1, ..., k$, with $A = A_1 = A' \cup A_{12} \cup \cdots \cup A_{1l_1}$. Finally let $a_i := |A_i|$ and $k' := \sum_{i=1}^k l_i$,

Proposition 4.1.1. *(see Proposition 8, p. 486 [23]) The Möbius function of lattice* $(\mathcal{E}_*^N, \wedge_{\mathcal{E}}, \vee_{\mathcal{E}})$ *is*

$$\mu(\bot, AP) = \begin{cases} (-1)^{|A|} & \text{if } P = P_\bot^A, \\ 0 & \text{otherwise.} \end{cases}$$

$$\mu(A'P', AP) = (-1)^{k'-k}(l_1 - 1)! \cdots (l_k - 1)! \quad \text{for } A'P' \sqsubseteq AP.$$

In particular $\mu(iP_\bot, AP) = (-1)^{n-k}(a-1)!(a_2-1)!\cdots(a_k-1)!$. Once the Möbius function is available, Möbius inversion of any game h on $\mathcal{E}_*^N$ is given by

$$\mu^h(AP) = \sum_{A'P' \sqsubseteq AP} \mu(A'P', AP)h(A', P'), \text{ for all } AP \in \mathcal{E}_*^N.$$

Coming to the solution concept, there exists a variety of value functions mapping PFF games into n shares, one for each player, and such n shares are indeed a valuation of subset lattice $(2^N, \cap, \cup)$. Some of this value mappings are proposed as the Shapley value of PFF games, and this name is justified on the ground that the provided n shares obtain as a weighted average or expectation of players' marginal contributions to embedded coalitions. Yet, it should be noted and perhaps emphasized that while the marginal contributions of players are well defined in coalitional games, in more complex games individual players have a limited capability to modify any existing cooperation level. In global games, looking at players' marginal contributions leads to take to account only the worth of modular partitions, thereby disregarding the vast majority (and precisely $\mathcal{B}_n - (2^n - n)$, where $\mathcal{B}_n$ is the n-th Bell number (See Chapter 1)) of the values taken by global games (regarded as lattice functions). In PFF games the same argument applies, in that players' marginal contributions to embedded coalitions may be conceived in alternative ways. Specifically, in [26] the average or expectation is over players' marginal contributions to maximal chains of embedded coalitions, rather than to single embedded coalitions. In any case, it seems not immediate to identify how the Shapley value there provided behaves on unanimity PFF games U_{AP} as above.

Along another line of investigation, an interesting test for value mappings seems to check whether they have fixed points. In this respect, the Shapley value of an additive coalitional game or valuation is the valuation itself. That is, valuations are precisely the fixed points of the Shapley value mapping for coalitional games. Now, more complex games such as global and PFF ones can be induced by a coalitional game, and this latter can be, in particular, a valuation. Then, value mappings for global and PFF games may be tested by checking their behaviour on global and PFF games induced by valuations (of subset lattice 2^N). In fact, this has already been partially considered when describing additively separable global games. Since Gilboa and Leher [17] only consider additively separating coalitional games v such that $v(\{i\}) = 0$ for all $i \in N$, they factually ignore global games additively separated by (non-trivial) valuations of 2^N. Finally note, though, that a global game additively separated by a valuation of 2^N is in fact a valuation of partition lattice $\mathcal{P}^N$, i.e. a constant partition function. The equivalent of additive separability for PFF games may be conceived in alternative ways. For example, one may say that PFF game h is induced by a coalitional game v if for all $AP \in \mathcal{E}^N$,

$$h(AP) = v(A) + \sum_{A' \in P} v(A') = 2v(A) + \sum_{B \in P \setminus A} v(B).$$

Then, existing value mappings for PFF games might be tested on PFF games induced by global games v in this way and, in particular, v could be a valuation of 2^N. That is $h(AP) = v(A) + v(N)$ for all $AP \in \mathcal{E}^N$. Therefore, such a PFF game h is *not* a constant lattice function, i.e. it is not a valuation of the lattice $(\mathcal{E}^N, \wedge_{\mathcal{E}}, \vee_{\mathcal{E}})$ of embedded subsets [23, pp. 484-5].

4.2 Conditional non-additive probabilities and Choquet expectation

The issue of updating non-additive probabilities (Schmeidler (1989) [51]) has been given extensive attention. Several theories have been proposed for the conditional probability in the non-additive case (see [10, 52, 16, 6, 14, 29, 20]). Most suggest that the probability of an event B conditioned on an event A depends not only on the probabilities of A, B and $A \cap B$, as in the traditional Bayes formula, but also on the probabilities of other events, such as $A^c \cap B$ and $(A \cap B) \cup A^c$ (see [19, Section 8, pp. 61-3] and [52, 10]). Once the conditional probability given A is defined, say, $p(\cdot|A)$, one may define the conditional expectation of a function X (e.g., $X = u_a$ a state-of-nature-dependent payoff, derived from a certain action $a \in \mathbb{A}$, as already defined in Chapter 2), given the event A, by simply integrating the restriction of X over A with respect to the conditional probability $p(\cdot|A)$.

This method of calculating the conditional expectation is conceptually inconsistent for the following reason. While the conditional probability of B with respect to A depends on the behaviour of B outside of A, the conditional expectation of X, given A, depends only on the behaviour of X over A. Thus, two functions may be significantly different on the complement of A, and yet, as long as they coincide on A, their conditional expectations are equal. A similar method of calculating the conditional expectation is to restrict the probability and the function to the conditioned event and to consider only the restricted items. More precisely, the conditional expectation is defined as the Choquet integral (see Choquet (1953) [7]) of the restricted function with respect to the normalized restricted probability. This method implies that the derived conditional probability of an event B, given A, depends only on the probability of $A \cap B$ and of A.

It may also imply that the conditional expectation of a function X on A is equal to its conditional expectation over A^c and yet, both differ from the Choquet integral of X. In Lehrer [36], it is presented a geometric approach, inspired by the theory of additive probabilities, which suggests a theory of conditional expectation that does not pass through the conditional probabilities. Rather, the conditional probability is a by-product. The conditional expectation of a function X, given a field of events, say, $\mathcal{F}$, will be defined as the closest (in some formal sense) function, which is $\mathcal{F}$-measurable. This represents a conservative attitude: the conditional expectation of a function X is another function that first, is compatible with the information (modelled by a field of events) and second, is the closest to the original one.

Recall that a field $\mathcal{F}$ of sets is a non-empty subset of the power set 2^{Ω}, closed under the intersection and union of pairs of sets and under complements of individual sets. From the previous Chapter, also recall that information is typically formalized by means of partitions (of state of nature). In fact, any partition P of Ω identifies the field $\mathcal{F}_P = 2^P$. This is the subset lattice whose elements are all and only those subsets of Ω obtained as the union of (zero or one or more) blocks of P. As usual, let us assume that the underlying probability space Ω is finite. Let p be an additive probability. We denote by $\mathcal{D} = 2^N = 2^{P_\perp}$ the field containing all subsets of Ω. A generic subfield of $\mathcal{D}$ is denoted by $\mathcal{F}$. The trivial field (containing $\emptyset$ and Ω only), is denoted as $\mathcal{T}$. On the other hand, the field that consists of $\emptyset, \Omega, A$ and the complement of A, i.e. A^c, is denoted by $\mathcal{F}_A$. Assume that X is a random variable and let $\mathcal{F}$ be a field. It turns out that X can be written as $X = Y + X_\perp$, where Y is $\mathcal{F}$-measurable (i.e. Y is constant on the atoms of $\mathcal{F}$) and $X_\perp$ satisfies

$$\int ZX_\perp \, dp = 0 \text{ for all } \mathcal{F}\text{-measurable variables } Z.$$

The conditional expectation $E(X|\mathcal{F})$ is equal to Y. In other words, $X = E(X|\mathcal{F}) + X_\perp$. In the appropriate space, $E(X|\mathcal{F})$ is the closest $\mathcal{F}$-measurable function to X. More precisely, denote by $\mathcal{M}(\mathcal{F})$ the set of all $\mathcal{F}$-measurable functions or random variables. Then,

$$E(X|\mathcal{F}) = \operatorname*{arg\,min}_{Y \in \mathcal{M}(\mathcal{F})} \int (X - Y)^2 \, dp. \tag{4.1}$$

In other words, Y is the closest, with respect to the l_2 norm variable in $\mathcal{M}(\mathcal{F})$, to X. Stated differently, $E(X|\mathcal{F})$ is the projection of X to the subspace (of variables) $\mathcal{M}(\mathcal{F})$. Let γ be a monotonic non-additive probability (i.e. a fuzzy measure), that is, $\gamma(\emptyset) = 0$, $\gamma(\Omega) = 1$ and if $A \subseteq B$, then $\gamma(A) \leq \gamma(B)$. Such a geometric approach presented above may be interpreted in various ways. More precisely the right side of (4.1) can be written in any of the following ways. Here is a sample:

- $\underset{Y \in \mathcal{M}(\mathcal{F})}{\arg\min} \int X^2 + Y^2 - 2XY \ dp.$
- $\underset{Y \in \mathcal{M}(\mathcal{F})}{\arg\min} \int X^2 + Y^2 \ dp - 2 \int XY \ dp.$
- $\underset{Y \in \mathcal{M}(\mathcal{F})}{\arg\min} \int Y^2 \ dp - 2 \int XY \ dp.$
- $\underset{Y \in \mathcal{M}(\mathcal{F})}{\arg\max} \int 2XY - Y^2 \ dp.$
- $\underset{Y \in \mathcal{M}(\mathcal{F})}{\arg\max} \int 2XY - Y^2 - X^2 \ dp.$

In the case where $p = \gamma$ is fuzzy and the integral is understood as the Choquet integral, no two of these methods are equivalent. Whatever method is adopted, it seems natural to require that the sought conditional expectation minimally satisfies the following two main properties:

(A1) $E(X|\mathcal{F}) = X$ if X is $\mathcal{F}$-measurable,

(A2) $E(X|\mathcal{T}) = \int X \ dp.$

Here (A1) states that if X is already measurable with respect to the field $\mathcal{F}$, then the expectation of X conditional on $\mathcal{F}$ is X itself, while (A2) states that with respect to the trivial field (that is, when no information is available) the conditional expectation coincides with the Choquet integral of X (this is in fact a real number, and therefore the conditional expectation of any random variable with respect to the trivial field is to be interpreted as a constant function). Any definition of the conditional expectation implies a definition of the conditional probability of an event given any field. Consider two events B and A. The conditional probability $p(B|A)$ is the updating of the probability of the event A when the available information is given by field $\mathcal{F}_A$. That is, if any $\omega \in A$ realizes, then what is known is (only) that the whole event A has realized. Similarly, if any $\omega' \in A^c$ realizes, then the whole complement event A^c is observed. Formally, the conditional probability $p(B|A)$ is defined as the value of $E(\mathbf{1}_B|\mathcal{F}_A)$ on A, where $\mathbf{1}_B : \Omega \to \{0, 1\}$ is the characteristic function of B, i.e. $\mathbf{1}_B(\omega) = 1$ if $\omega \in B$ and $\mathbf{1}_B(\omega) = 0$ if $\omega \in B^c$ for all $\omega \in \Omega$ (see [36, Example 4, p. 49]).

The conditional expectation of the function or random variable X, given a field $\mathcal{F}$, is thus defined as a $\mathcal{F}$-measurable function that satisfies, together with (A1) and (A2) above, further properties listed in [36, Section 6, p. 52] some properties. In particular, the conditional expectation of X, given a field $\mathcal{F}$, may be defined as

$$E(X|\mathcal{F}) = \operatorname*{arg\,min}_{Y \in \mathcal{M}(\mathcal{F})} \int (X - Y)^2 \; dp = \operatorname*{arg\,min}_{Y \in \mathcal{M}(\mathcal{F})} \int X^2 + Y^2 - 2XY \; dp.$$

The problem with this definition is that (A2) is not always satisfied, but the flaw is corrected as follows. Denote $\underline{X}(\omega) = \min_{\omega' \in \mathcal{F}(\omega)} X(\omega')$ where $\mathcal{F}(\omega)$ is the atom of $\mathcal{F}$ containing ω. Similarly denote $\overline{X}(\omega) = \max_{\omega' \in \mathcal{F}(\omega)} X(\omega')$. Let $\mathcal{N}(X, \mathcal{F})$ be the subset of those $Y \in \mathcal{M}(\mathcal{F})$ which satisfy $\int (X - Y) \; dp = 0$ and $\underline{X}(\omega) \leq Y(\omega) \leq \overline{X}(\omega)$ for every ω. This set $\mathcal{N}(X, \mathcal{F})$ of $\mathcal{F}$-measurable functions (or random variables) is shown to be non-empty and compact (see Leherer [36, Lemma 1, p. 50]). Accordingly, the definition of (geometric) conditional expectation of a random variable with respect to a field is the following.

Definition 4.2. The conditional expectation of X with respect to $\mathcal{F}$, denoted $E(X|\mathcal{F})$, is a random variable $Y \in \mathcal{N}(X, \mathcal{F})$ that minimizes $\int (X - Y)^2 \; dp$. Formally,

$$E(X|\mathcal{F}) \in \operatorname*{arg\,min}_{Y \in \mathcal{N}(X, \mathcal{F})} \int (X - Y)^2 \; dp.$$

In words, we say that Y is a conditional expectation of X given $\mathcal{F}$ if it is an $\mathcal{F}$-measurable function which minimizes the integral of the difference between X and Y squared, among the functions Y that have two properties: (i) Y is bounded between the minimum and the maximum of X in each atom of $\mathcal{F}$; and (ii) the integral of the difference between X and Y is equal to zero. Note that typically there is no unique solution to the problem,

$$\min_{Y \in \mathcal{N}(X,\mathcal{F})} \int (X-Y)^2 \, dp \text{ subject to } \int (X-Y) \, dp = 0.$$

We say that Y is $E(X|\mathcal{F})$ if Y solves this minimization problem, where this latter always admits a solution [36, Theorem 1, p. 52].

4.3 Nash equilibrium with Choquet expected utility

Under some existing updating schemes it may turn out that the conditional probabilities of B, given A, and of B, given A^c, are both less than some constant, and yet, the probability of B is greater than this constant. Under the updating scheme proposed in [36] this cannot occur. This feature extends to the conditional expectation. The fact that the conditional expectation of a function is uniformly greater than a certain constant implies that the integral of this function is greater than the same constant. In particular, if, given any event in the informational partition, an act is valued, say, 7, then this act is unconditionally valued 7. This approach may be used to define Nash equilibrium (of non-cooperative games) when players randomize their action or strategies according to non-additive probabilities.

The traditional definition of Nash equilibrium with randomized or mixed strategies involves two conditions. First, the players play independently and thus their play induces independent probabilities over the product of their action spaces. Second, each player plays his or her best response, given his or her choice and given other players' actions. In case the mixed actions of the players are non-additive, the first condition calls for a definition of independence of non-additive probabilities defined on a product space. This issue alone has recently been paid a big deal of attention [13, 12, 61, 28, 15].

The geometric approach suggested in the previous Section leads to the following notion of independence: the mixed actions of the players are independent if there is a measure over the set of all joint actions such that (i) the marginal probability over every player's actions coincides with the players' mixed action; and (ii) the players can induce nothing about other players' actions from their own. Only through conditional probability can players learn about others' actions from their own. Therefore, condition (ii) of independence can be conveyed more formally as follows. There exists a probability over the product space (typically, not the product probability) such that the probability of player i playing an action in a set B coincides with the conditional probability of B, given the partition induced by what player i knows (i.e. his or her actions). The second condition of Nash equilibrium refers to incentive compatibility. It states that each player plays his or her best response to other players' actions. However, the payoff given to a player when he or she plays an action, is nothing but the conditional payoff, with respect to the independent probability (over the product space), given that action. Therefore, both conditions of Nash equilibrium require the concept of conditional expectation provided in the previous Section.

In order to formalize these ideas, we need to briefly introduce the notation applying to non-cooperative games. A non-cooperative or strategic game with a finite player set $N = \{1, ..., n\}$ is a triple $\Gamma = (N, \mathbb{S}, u)$ where $\mathbb{S} = \underset{i \in N}{\times} \mathbb{S}_i$ is the product space of all players' strategy sets $\mathbb{S}_i = \{s_i^1, s_i^2, ..., s_i^{k_i}\}$ with $|\mathbb{S}_i| = k_i \geq 2$, and $u : \mathbb{S} \to \mathbb{R}^n$ is a utility function which assigns a n-dimensional payoff vector $u(s) = (u_1(s), u_2(s), ..., u_n(s))$ to all generic strategy profile $s = (s_1, s_2, ..., s_n) \in \mathbb{S}$. Let $\mathbb{S}_{-i} := \underset{j \in N \setminus i}{\times} \mathbb{S}_j$ the product space of all non-i players' strategies, such that for all $s \in \mathbb{S}$, $s = (s_i, s_{-i})$ with $s_{-i} = (s_1, ..., s_{i-1}, s_{i+1}, ..., s_n)$.

Definition 4.3. A pure (i.e. non-randomized) strategy Nash equilibrium is a n-tuple $s \in \mathbb{S}$ such that

$$u_i(s) = u_i(s_i, s_{-i}) \geq u_i(s'_i, s_{-i}) \text{ for all } s'_i \in \mathbb{S}_i \text{ and all } i \in N.$$

The set of pure strategy equilibria of game Γ can be empty or, conversely, can contain multiple strategy profiles. If players may randomize then there always exists al least one equilibrium, as detailed below.

Let $\Delta_{\mathbb{S}_i}$ denote the $k_i - 1$ dimensional simplex whose extreme points correspond to non-random strategies $s_i \in \mathbb{S}_i$, that is

$$\Delta_{\mathbb{S}_i} = \left\{ (\sigma_i^1, ..., \sigma_i^{k_i}) \in [0,1]^{k_i} : \sum_{1 \leq k \leq k_i} \sigma_i^k = 1 \right\}.$$

For player $i \in N$, a *mixed strategy* $\sigma_i \in \Delta_{\mathbb{S}_i}$ is a point in this simplex or, equivalently, a probability over the set $\mathbb{S}_i$ of pure strategies; in fact, $\sigma_i^k = \sigma_i(s_i^k)$ is interpreted as the probability (or the frequency in repeated games) according to which player i plays pure strategy $s_i^k \in \mathbb{S}_i$, $1 \leq k \leq k_i$ when choosing mixed strategy σ_i.

Any profile $\sigma = (\sigma_1, ..., \sigma_n) \in \Delta_{\mathbb{S}_1} \times \cdots \times \Delta_{\mathbb{S}_n}$ of mixed strategies chosen by the n players induces the unique probability distribution $p_\sigma \in \Delta_{\mathbb{S}}$ on the product space $\mathbb{S}$ given by

$$p_\sigma(s_1, ..., s_n) := \prod_{1 \leq i \leq n} \sigma_i(s_i) \text{ for all } s = (s_1, ..., s_n) \in \mathbb{S}.$$

Evidently, $p_\sigma(s) \in [0,1]$ since $\sigma_i(s_i) \in [0,1]$ for all $i \in N$, $s_i \in \mathbb{S}_i$, while $\sum_{s \in \mathbb{S}} p_\sigma(s) = 1$ can be easily checked by induction on the number $n \geq 2$ of players. Let $Eu_i(p_\sigma) = Eu_i(\sigma)$ and $\sigma = (\sigma_i, \sigma_{-i})$ for notational convenience, where $\sigma_{-i} = (\sigma_1, ..., \sigma_{i-1}, \sigma_{i+1}, ..., \sigma_n) \in \times_{k \in N \backslash i} \Delta_{\mathbb{S}_k}$.

For $i \in N$, mixed strategy $\sigma_i \in \Delta_{\mathbb{S}_i}$ is a *best response* to the $n-1$-tuple $\sigma_{-i} \in \underset{k \in N \backslash i}{\times} \Delta_{\mathbb{S}_k}$ of mixed strategies of other players if

$$Eu_i(\sigma_i, \sigma_{-i}) \geq Eu_i(\sigma_i', \sigma_{-i}) \text{ for all } \sigma_i' \in \Delta_{\mathbb{S}_i}.$$

For every player $i \in N$, let $BR_i : \underset{j \in N \backslash i}{\times} \Delta_{\mathbb{S}_j} \twoheadrightarrow \Delta_{\mathbb{S}_i}$ denote the associated best correspondence by

$$BR_i(\sigma_{-i}) = \{\sigma_i \in \Delta_{\mathbb{S}_i} : Eu_i(\sigma_i, \sigma_{-i}) \geq Eu_i(\sigma_i', \sigma_{-i}) \text{ for all } \sigma_i' \in \Delta_{\mathbb{S}_i}\}.$$

It can be shown [39, Chapter 8, pp. 250-1], that for every σ_{-i}, there is a subset $\mathbb{S}'_i \subseteq \mathbb{S}_i$ such that $BR_i(\sigma_{-i}) = \Delta_{\mathbb{S}'_i}$, where this latter contains all mixed strategies σ_i placing non-zero probability only on those strategies $s_i \in \mathbb{S}'_i$. That is,

$$\Delta_{\mathbb{S}'_i} := \{\sigma_i \in \Delta_{\mathbb{S}_i} : s'_i \notin \mathbb{S}'_i \Rightarrow \sigma_i(s'_i) = 0\}.$$

Then, the whole correspondence

$$(BR_1, ..., BR_n) = BR : \times_{i \in N} \Delta_{\mathbb{S}_i} \twoheadrightarrow \times_{i \in N} \Delta_{\mathbb{S}_i}$$

is upper hemicontinuous [39, p. 950], and thus fulfils the conditions required by Kakutani's theorem.

Theorem 4.3.1. *(Theorem 1, p. 457 [33]) Let $F : C \twoheadrightarrow C$ be a upper hemicontinuos correspondence, then it exists $x \in C$ such that $x \in F(x)$.*

Since a mixed strategy Nash equilibrium is a fixed point of correspondence BR, that is, a n-tuple $\sigma = (\sigma_1, ..., \sigma_n) \in \underset{i \in N}{\times} \Delta_{\mathbb{S}_i}$ satisfying

$$\sigma_i \in BR_i(\sigma_1, ..., \sigma_{i-1}, \sigma_{i+1}, ..., \sigma_n) \text{ for all } i \in N,$$

the set of equilibria is not empty. Nash equilibrium, whether in pure or mixed strategies, does not seem to prevail as the outcome of real-life interactions. The reason is that individuals commonly try to cooperate, while Nash equilibrium assumes that players choose their actions independently. In addition, the issue of Nash equilibrium with non-additive randomized strategies, outlined in the remaining of this section, is difficult to be interpreted. Specifically, randomizing accordingly to an additive probability may be viewed in terms of frequency of pure strategies played in repeated games. However, if randomization occurs according to a non-additive probability, then it is not immediate how this can apply to any real-life interaction. Apart from this conceptual concern, it is clear that how to condition non-additive probabilities as analysed in the previous Section is crucial for defining Nash equilibrium with non-additive random strategies, with players maximizing their Choquet expected utility.

Formally, suppose that player i randomly chooses an action in $\mathbb{S}_i$ with respect to a non-additive probability γ_i. Concerning the interpretation, this probability need not be the actual distribution according to which she randomly selects her action. Indeed, γ_i might be the distribution that guides her choice as perceived by other players or by an outside observer. But most importantly, the notion of Nash equilibrium assumes that players choose their action independently. This means that the knowledge of each player, beyond the description of the game, consists solely of her action. Independence of γ_i would therefore mean that the knowledge of her own action does not change her belief regarding the probability over other players' actions. In terms of conditioning, it entails that the probability over other players actions, conditional on any subset of player i's actions, coincides with the unconditional distribution. Let $\mathcal{F}_i$ be the partition of $\mathbb{S}$ whose atoms are $\{s_i\}\times\mathbb{S}_{-i}$, $s_i \in \mathbb{S}_{-i}$. The partition $\mathcal{F}_i$ represents the knowledge available to player i.

Definition 4.4. A non-additive probability γ over $\mathbb{S}$ realizes γ_i, $i = 1, ..., n$ as independent probabilities if

(a) for every i and every $A \subseteq \mathbb{S}_i$, $\gamma(A \times \mathbb{S}_{-i}) = \gamma_i(A)$ and

(b) for every $B \subseteq \mathbb{S}_{-i}$, $\gamma(B \times \mathbb{S}|\mathcal{F}_i) = \gamma(B \times \mathbb{S}_i)$.

In order for γ_i, $i = 1, ..., n$ to be realized as independent probabilities, there must be a probability γ over the product space $\mathbb{S}$ satisfying these two conditions. Condition (a) states that the marginal of γ over $\mathbb{S}_i$ coincides with γ_i. Condition (b) states that knowing $\mathcal{F}_i$, player i does not change her belief about others' actions. In other words, the conditional probability knowing $\mathcal{F}_i$, $\gamma(B \times \mathbb{S}|\mathcal{F}_i)$, coincides with $\gamma(B \times \mathbb{S}_i)$.

Note that in the additive case, there is a unique probability that realizes $p_i, i = 1, ..., n$ as independent probabilities. This is the product probability shown above. Conversely, in the non-additive case the product probability (however defined) shall not generically realize $\gamma_i, i = 1, ...n$ as independent. There is no proof thus far of the conjecture that for any probabilities $\gamma_i, i = 1, ...n$, there is γ over $\mathbb{S}$ that realizes them as independent. Moreover, there is no guarantee that there is a unique probability that does it. The definition of independence of non-additive probabilities paves the way to the definition of Nash equilibrium.

Coming to the geometric approach to conditioning described in the previous section, Nash equilibrium requires, on top of incentive compatibility conditions, that players would choose their actions independently of each other. When playing the mixed action γ_i, player i's payoff is $E(u_i|\mathcal{F}_i)$, where the expectation is taken with respect to a probability γ that realized the (non-additive) mixed actions γ_i as independent. Note that in case there are multiple probabilities that realize γ_i as independent, there may be multiple expected payoffs with the same set of mixed actions. In equilibrium, $E(u_i|\mathcal{F}_i)$ should be greater than or equal to the expected payoff guaranteed by any specific action $s_i \in \mathbb{S}_i$. However, given the action $s_i \in \mathbb{S}_i$, all other players still select their actions independently of each other. Thus, the payoff associated with action $s_i \in \mathbb{S}_i$ is the expectation of player i's payoff taken with respect to a probability that realizes $(\gamma_j)_{j \neq i}$ as independent.

Bibliography

[1] M. Aigner. *Combinatorial Theory.* Springer, 1997. Reprint of the 1979 Edition.

[2] J. Albizuri, J. Arin, and J. Rubio. An axiom system of a value of games in partition function form. *International Game Theory Review*, 7:63 – 72, 2005.

[3] R. Aumann, S. Hart, and M. Perry. *Conditioning and the Sure Thing Principle.* Center for Rationality DP-393. The Hebrew University of Jerusalem, 2005.

[4] M. Ben-Or and N. Linial. Collective coin flipping, robust voting schemes and minima of Banzhaf values. In *Proceedings of the 26th Conference on Foundations of Computer Science*, pages 408–4016, 1985.

[5] E. M. Bolger. A set of axiom of a value for partition function form games. *The Annals of Statistics*, 18:37 – 44, 1989.

[6] A. Chateauneuf and J. Y. Jaffray. Some characterizations of lower probabilities and other monotone capacities through the use of Mobius inversion. *Mathematical Social Sciences*, 17(3):263 – 283, 1989.

[7] G. Choquet. Theory of capacities. *Annales de l'institut Fourier*, 5:131 – 295, 1954.

[8] T. Chunquiao and C. Xiaohong. Intuitionistic fuzzy choquet integral operator for multicriteria decision making. *Expert Systems with Applications*, 37(1):149–157, 2010.

[9] B. A. Davey and H. A. Priestley. *Introduction to Lattices and Order (Second edition)*. Cambridge University Press, 2002.

[10] A. P. Dempster. Upper and lower probabilities induced by a multivlued mapping. *Ann. Math. Statist.*, 38:325 – 339, 1967.

[11] A. P. Dempster. A generalization of Bayesian inference. *J. Roy. Statist. Soc., Series B*, 30:205 – 247, 1968.

[12] J. Dow and S. R. d. C. Werlang. Nash equilibrium under Knightian uncertainty: breaking down backward induction. *Journal of Economic Theory*, 64:305 – 324, 1993.

[13] J. Eichberger and D. Kelsey. Non-additive beliefs and strategic equilibria. *Games and Economic Behavior*, 30:183 – 215, 1995.

[14] R. Fagin and J. Y. Halpern. A new approach to updating beliefs. *Proceedings of the 6th conference on uncertainty in AI*, pages 317 – 325, 1990.

[15] P. Ghirardato. On the indipendence for non-additive measures, with a Fubini theorem. *Journal of Economic Theory*, 73:261 – 291, 1996.

[16] I. Gilboa. Additivivations of non-additive measures. *Econometrica*, 57:1153 – 1169, 1989.

[17] I. Gilboa and E. Lehrer. Global games. *International Journal of Game Theory*, (20):120–147, 1990.

[18] I. Gilboa and E. Lehrer. The value of information - an axiomatic approach. *Journal of Mathematics Economics*, 20:443 – 459, 1991.

[19] I. Gilboa and D. Schmeidler. Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18:141 – 153, 1989.

[20] I. Gilboa and D. Schmeidler. Updating ambiguous beliefs. *Journal of Economic Theory*, 59:33 – 49, 1993.

[21] I. Gilboa and D. Schmeidler. Additive representations of non-additive measures and the choquet integral. *Annals of Operation Research*, 52:43 – 65, 1994.

[22] M. Grabisch. The interaction and Möbius representations od fuzzy measures on finite spaces. In M. Grabisch, T. Murofushi, and M. Sugeno, editors, *Fuzzy Measures and Integals - Theory and applications*, pages 70 – 93. Physica - Verlag, Springer, 2000.

[23] M. Grabisch. The lattice of embedded subsets. *Discrete Applied Mathematics*, 158(5):479–488, 2010.

[24] M. Grabish. The application of fuzzy integrals in multicriteria decision making. *European Journal of Operation Research*, 89:445 – 456, 1995.

[25] M. Grabish and C. Labreuche. To be symmetric or asymmetric? a dilemma in decision making. *Studies in Fuzziness and Soft Computing*, 51:179 – 194, 2000.

[26] M. Grabish and F. Yukiniko. A coalition formation value for games in fartition function form. *European Journal of Operational Research*, 221(1):175 – 185, 2012.

[27] R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete Mathematics: A Foundation for Computer Science (2nd Edition)*. Addison-Wesley Professional, 1994.

[28] H. Haller. Non-additive beliefs in solvable games. *Theory and Decisions*, 49:313 – 338, 2000.

[29] J. Y. Halpern and M. N. Tuttle. Knowledge probability and adversaries. *Journal of the ACM*, 40(4):917 – 962, 1993.

[30] J. C. Harsányi. A simplified bargaining model for the n-person cooperative game. *International Economic Review*, 4(2):190–220, 1963.

[31] S. Hart and D. Schmeidler. Existence of correlated equilibria. *Mathematical of Operations Research*, 14(1):18 – 25, 1989.

[32] J. Ju. The consensus value: a new solution concept for cooperative games. *Social Chioce and Welfare*, 28(4):685 – 703, 2007.

[33] S. Kakutani. A generalization of brouwer's fixed point theorem. *Duke Mathematical Journal*, 8(3):457 – 459, 1941.

[34] L. Koczy. A recursive core for partition function form games. *Theory and Decision*, 63:41 – 51, 2007.

[35] D. Kreps. *Note on the theory of Choice*. Westview Press, Boulder, London, 1988.

[36] E. Lehrer. Updating non-additive probabilities - a geometric approach. *Games and Economic Behaviour*, 50:42 – 57, 2004.

[37] I. Macho-Stadler, D. Perez-Castillo, and D. Wettstein. Sharing the surplus: an extension of the shapley value for enviroments with externalities. *Journal of Economic Theory*, 135:339 – 356, 2007.

[38] G. Markowsky. Primes, irreducibles and extremal lattices. *Order*, 9(3):265–290, 1992.

[39] A. Mass-Colell, M. D. Whinston, and J. R. Green. *Microeconomic Theory*. Oxford University Press, 1995.

[40] R. B. Myerson. Values of games in partition function form. *International Journal of Game Theory*, 6(1):23–31, 1977.

[41] H. T. Nguyen and W. A. On decision making using belief functions. *Advances in the Dempster-Shafer Theory of evidence*, 1:311 – 330, 1994.

[42] G. Owen. *Game Theory (Second edition)*. Academic Press, Inc., 1982.

[43] B. Peleg and P. Sudhölter. *Introduction to the Theory of Cooperative Games*. Springer, 2007.

[44] K. Pham Do and N. H. The shapley value for partition function form games. *International Game theory review*, 9:353 – 360, 2007.

[45] Y. Rébillé. Decision making over necessity measure through the choquet integral criterion. *Fuzzy Sets and Sysyems*, 157(1):3025–3039, 2006.

[46] J. Rosenmuller. On core and value. *Mathematical Methods of Operations Research*, 9:84 – 104, 1971.

[47] J. Rosenmuller. Some properties of convex set functions, part ii. *Mathematical Methods of Operations Research*, 17:287 – 307, 1972.

[48] G. Rossi. Worth-sharing throught Mobius inversion. *Homo economicus*, 24(3/4):411 – 433, 2007.

[49] G. C. Rota. On the foundations of combinatorial theory I: theory of Möbius functions. *Z. Wahrscheinlichkeitsrechnung u. verw. Geb.*, (2):340–368, 1964.

[50] L. S. Savage. *The Foundations of Statistics (second edition)*. Dover, New York, 1972.

[51] D. Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 57(3):571 – 587, 1989.

[52] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.

[53] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27:379 – 423, 623 – 656, 1948.

[54] L. S. Shapley. Cores of convex games. *International Journal of Game Theory*, 1(1):11–26, 1971.

[55] L. S. Shapley. A value for n-person games. In A. E. Roth, editor, *The Shapley value – Essays in honor of Lloyd S. Shapley*, pages 31–40. Cambridge University Press, 1988. Reprinted from *Contributions to the Theory of Games*, vol. 2, eds. H. Kuhn and A. W. Tucker, Princeton University Press 1953, pp. 307–317.

[56] L. S. Shapley and P. Dubey. Mathematical properties of the Banzhaf power index. *Mathematics of Operational Reaserch*, 4(2):99–131, 1979.

[57] M. Stern. *Semimodular Lattices*. Cambridge University Press, 1999.

[58] R. M. Thrall and W. F. Lucas. N-person games in partition function form. *Naval Research Logistics Quarterly*, 10(1):281–298, 1963.

[59] J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 1944.

[60] R. J. Weber. Probabilistic values for games. In A. E. Roth, editor, *The Shapley value – Essays in honor of Lloyd S. Shapley*, pages 101–119. Cambridge University Press, 1988.

[61] H. J. Witta-Jacobsen, T. Tranes, and B. Sloth. Nash equilibrium with lower probabilities. *Theory and Decision*, 12:37 – 66, 1997.

Ringraziamenti

Desidero ringraziare il Prof. Rossi per il tempo, la passione e la competenza che ha messo a mia disposizione senza risparmiarsi. Ringrazio inoltre il Prof. Ferri per la disponibilità ad accettare il ruolo di correlatore e i preziosi consigli e il Prof. Parmeggiani per alcuni importanti commenti sul primo capitolo della Tesi. Ringrazio i miei genitori per avermi sempre spronato a dare il meglio. Un ringraziamento particolare va alla mia ragazza Beatrice che mi è sempre stata vicino sia nei momenti di sconforto che in quelli di gioia, per non aver mai smesso di credere in me e soprattutto per tutto l'amore che mi hai dato e continui a darmi ogni giorno. Questa tesi è dedicata a te.