Université Toulouse III - Paul Sabatier

THESE
Pour l'obtention du titre de DOCTEUR EN INFORMATIQUE

Possibilistic Decision Theory : From Theoretical Foundations to Influence Diagrams Methodology

Candidat : Wided GUEZGUEZ

JURY

Directeurs : Nahla BEN AMOR (ISG Tunis)
Hélène Fargier (IRIT Toulouse)

Rapporteurs : Zied ELOUEDI (ISG Tunis)
Salem BENFERHAT (Université d'Artois)

Examinateurs : Boutheina BEN YAGHLANE (IHEC Tunis)
Marie Christine LAGASQUIE (IRIT Toulouse)
Régis SABBADIN (INRA Toulouse)
Olivier SPANJAARD (LIP 6 Paris)

Mai 2012

Abstract

The field of decision making is a multidisciplinary field in relation with several disciplines such as economics, operations research, etc.. Theory of expected utility has been proposed to model and solve decision problems. These theories have been questioned by several paradoxes (Allais, Ellsberg) who have shown the limits of its applicability. Moreover, the probabilistic framework used in these theories is not appropriate in particular situations (total ignorance, qualitative uncertainty). To overcome these limitations, several studies have been developed basing on the use of Choquet and Sugeno integrals as decision criteria and a non classical theory to model uncertainty. Our main idea is to use these two lines of research to develop, within the framework of sequential decision making, decision models that are based on Choquet integrals as decision criteria and the possibility theory to represent uncertainty. Our goal is to develop graphical decision models that represent compact models for decision making when uncertainty is represented using possibility theory. We are particularly interested by possibilistic decision trees and influence diagrams and their evaluation algorithms.

Keywords

Sequential decision making, uncertainty, Choquet integrals, possibility theory, decision trees, influence diagrams.

Résumé

Le domaine de prise de décision est un domaine multidisciplinaire en relation avec plusieurs disciplines telles que l'économie, la recherche opérationnelle, etc. La théorie de l'utilité espérée a été proposée pour modéliser et résoudre les problèmes de décision. Ces théories ont été mises en cause par plusieurs paradoxes (Allais, Ellsberg) qui ont montré les limites de son applicabilité. Par ailleurs, le cadre probabiliste utilisé dans ces théories s'avère non approprié dans certaines situations particulières (ignorance totale, incertitude qualitative). Pour pallier ces limites, plusieurs travaux ont été élaborés concernant l'utilisation des intégrales de Choquet et de Sugeno comme critères de décision d'une part et l'utilisation d'une théorie d'incertitude autre que la théorie des probabilités pour la modélisation de l'incertitude d'une autre part. Notre idée principale est de profiter de ces deux directions de recherche afin de développer, dans le cadre de la décision séquentielle, des modèles de décision qui se basent sur les intégrales de Choquet comme critères de décision et sur la théorie des possibilités pour la représentation de l'incertitude. Notre objectif est de développer des modèles graphiques décisionnels, qui représentent des modèles compacts et simples pour la prise de décision dans un contexte possibiliste. Nous nous intéressons en particulier aux arbres de décision et aux diagrammes d'influence possibilistes et à leurs algorithmes d'évaluation.

Mots-clés

Prise de décision séquentielle, incertitude, intégrales de Choquet, théorie de possibilité, arbres de décision, diagrammes d'influence.

Acknowledgments

I would like to thank my advisor Mrs Nahla Ben Amor for her patience in supervising this work in a concerned way. I would like to address special thanks for you, for your support, your advise, your productive comments and for the many hours spent helping this work. Thank you for your endless patience in improving my writing, for your comments on chapter drafts and for your great efforts to explain things clearly and simply.

I would like to express my sincere gratitude to my co-advisor Mrs *Hélène* Fargier for her support in the fulfillment of my PhD. Her competence and her generosity have allowed me to work in ideal conditions. I am deeply indebted to you for your encouragements and for all the interested discussion that have always been of great interest during the elaboration of this work.

I would like to express my consideration to Mr Didier Dubois and Mr Régis Sabbadin, who received me several times in their laboratories. Thank you for your interest to this work and you valuable discussions and comments.

Thanks to all the members of my laboratory LARODEC in Institut Supérieur de Gestion de Tunis and to my friends for all good and bad times we get together.

Special thanks go to the members of IRIT laboratory for their friendly support.

It is difficult to overstate my gratitude to my extended family for providing a loving environment for me which help me during the elaboration of this work.
Lastly, and most importantly, I wish to thank my parents and my husband for their steadfast support, their love and patience. I dedicate this work to them and to my daughter Sarra.

Table des matières

-

Table des figures

Liste des tableaux

Introduction

Decision making process is an important topic in Artificial Intelligence. Modeling decision problems is a tedious task since in real world problems several types of uncertainty related to decision maker's behavior and states of nature should be considered.

The most famous decision criterion proposed by decision theory is the expected utility. Despite the success of this decision model, it has some limits since it is unable to represent all decision makers behaviors as it has been highlighted by Allais [1] and Ellsberg [30]. To respond to these paradoxes, alternative decision models like those based on Choquet and Sugueno integrals have been proposed [13, 78].

Besides, most of available decision models refer to probability theory for the representation of uncertainty. However, this framework is appropriate only when all numerical data are available, which is not always possible. Indeed, there are some situations, like the case of total ignorance, which are not well handled and which can make the probabilistic reasoning unsound. Several non classical theories of uncertainty have been proposed in order to deal with uncertain and imprecise data such as evidence theory [72], Spohn's ordinal conditional functions [76] and possibility theory [89] issued from fuzzy sets theory [88].

The aim of this thesis is to study different facets of the *possibilistic decision theory* from its theoretical foundations to sequential decisions problems with possibilistic graphical decision models.

Our choice of the possibilistic framework is motivated by the fact that the possibility theory offers a natural and simple model to handle uncertain information. In fact, it is an appropriate framework for experts to express their opinions about uncertainty numerically using possibility degrees or qualitatively using total pre-order on the universe of discourse.

In the first part of this thesis, we provide in Chapter 1 a study of existing classical decision models namely Maximax, Maximin, Minimax regret, Laplace and Hurwicz decision criteria under total uncertainty. Then, we present expected decision theories (i.e. expected utility and subjective expected utility) and we detail non expected decision models.

We develop in Chapter 2 possibilistic decision theory (i.e *optimistic and pessimistic utility (U_{pes} and U_{opt}), binary utility (PU), possibilistic likely dominance (LN and $L\Pi$) and order of magnitude expected utility (OMEU)*) by detailing the axiomatic system of each criterion. We then give special attention to Choquet based criteria by developing particular properties of possibilistic Choquet integrals. More precisely, we propose necessity-based Choquet integrals for cautious decision makers and possibility-based Choquet integrals for adventurous decision makers.

The second part of our work is dedicated to sequential decision making where a decision maker should choose a sequence of decisions that are executed successively. Several graphical decision models can be used to model in a compact manner sequential decision making. We can in particular mention decision trees [65], influence diagrams [43] and valuation based systems [73], etc. Most of these graphical models refer to probability theory as uncertainty framework and to expected utility as decision criterion. This motivates us to study possibilistic graphical decision models using results of the first part on possibilistic decision criteria.

To this end, we first give in Chapter 3 an overview on standard decision trees and influence diagrams and on their evaluation algorithms. Then, we formally define *possibilistic decision trees* by studying in Chapter 4 the complexity of finding the optimal strategy in the case of different possibilistic decision criteria. We show that except for the Choquet based decision criteria, the application of *dynamic programming* is possible for most possibilistic decision criteria since pessimistic and optimistic utilities, binary utility, likely dominance and order of magnitude expected utility satisfy the crucial *monotonicity* property needed for the application of this algorithm.

For the particular case of Choquet based criterion we show that the problem is NP-hard and we develop a *Branch and Bound* algorithm. We also characterize some particular cases where dynamic programming can be applied.

In order to show the efficiency of the studied algorithms in the case of possibilistic decision trees with Choquet integrals, we propose in Chapter 5 an experimental study aiming to compare results provided by dynamic programming w.r.t those of Branch and Bound.

Possibistic decision trees inherit the same limits than those of standard decision trees, namely the fact that they are no longer appropriate to model huge decision problems since they will grow exponentially. Hence, our interest of proposing in Chapter 6 another possibilistic graphical decision models i.e. *possibilistic influence diagrams*.

We, especially distinguish two classes of possibilistic influence diagrams (homogeneous and heterogeneous) depending on the quantification of chance and utility nodes. Indeed, homogeneous possibilistic influence diagrams concern the case where chance and value nodes are quantified in the same setting contrarily to the case of heterogeneous ones. Then, we propose indirect evaluation algorithms for different kinds of possibilistic influence diagrams : the first algorithm is based on the transformation of influence diagrams into possibilistic decision trees and the second one into possibilistic networks [6] which are a possibilistic counterpart of Bayesian networks [46]. These indirect approaches allow us to benefit from already developed evaluation algorithms for possibilistic decision trees and also for possibilistic networks.

Première partie

Possibilistic decision criteria based on Choquet integrals under uncertainty

Chapitre 1

An Introduction to Decision Theories : Classical Models

1.1 Introduction

Decision making is a multidisciplinary domain that concerns several disciplines such that economy, psychology, operational research and artificial intelligence. So, decision making is a complex domain characterized by uncertainty since it is related to several other domains. In addition, decision makers cannot express their uncertainty easily because necessary informations are not always available.

In general, we can distinguish three situations of decision making according to the uncertainty relative to the states of nature :

- A situation of total uncertainty when no information is available.
- A situation of probabilistic uncertainty when it exists a probability function (objective or subjective) that quantifies uncertainty about states of nature.
- A situation of non probabilistic uncertainty when probability theory cannot be used to quantify uncertainty. Then, non classical theories of uncertainty such as fuzzy sets theory [88], possibility theory [89] and evidence theory [72] can be used.

According to the situation, several decision criteria have been proposed in the literature. These criteria can be classified into two classes :

1. Quantitative decision approaches and
2. Qualitative decision approaches.

The most famous quantitative decision models are based on expected utility. These classical decision models are well developed and axiomatized by Von Neumann and Morgenstern [57] in context of objective probabilistic uncertainty and by Savage [68] when the probability is subjective. However, these classical decision models have some limits since they cannot represent all the decision maker behavior and all kinds of uncertainty.

In order to overcome these limitations, new models have been developed : Choquet "non expected utility", possibilistic decision theory [26, 27], Sugeno integrals [26, 69, 78], etc.

This chapter gives a survey of decision theories. It is organized as follows : Section 1.2 presents basic definitions and notations. Uncertainty in decision problems is discussed in Section 1.3. Section 1.4 details the framework of decision making under total uncertainty. Expected decision models will be developed in Section 1.5. Non expected decision models will be introduced in Section 1.6.

1.2 Definitions and notations

A decision problem consists on a choice between a list of possible alternatives considering expert knowledge's about states of nature and his preferences about possible results of his decision expressed via a set of utilities. The choice of the optimal strategy is also influenced by the nature of the decision maker that can be :

- Optimistic : Decision maker chooses the decision that has the maximal payoff even if it is a risky choice.
- Pessimistic : Decision maker chooses the least risky decision.
- Neutral : Decision maker is neutral w.r.t loss and gain.

Decision making is the identification and the choice of some alternatives based on preferences of the decision maker and the decision environment. This process aims to select a strategy that is optimal w.r.t the decision maker's satisfaction.

Let us give some useful notations :

- The set of *states of nature* is denoted by S. $|S|$ denotes the cardinality of the set S.
- The set of subsets of S represents *events*. This set is denoted by E (i.e. $E = 2^S$).
- The set of *consequences* is denoted by $\mathcal{C}$.
- An *act*, called also an *action* (a decision), is denoted by f and it assigns a consequence to each state of nature it is a mapping $f : S \longmapsto \mathcal{C}$ from S to $\mathcal{C}$. We have : $\forall s \in S, f(s) \in \mathcal{C}$.
- The set of *acts* is denoted by F.
- A *constant act* is an act that gives the same consequence whatever the state of nature. F^{const} is the set of constant acts.
- An *utility function* (denoted by u) is a mapping from $\mathcal{C}$ to U, i.e. $u : \mathcal{C} \to U$ where $U = \{u_1, \ldots, u_n\}$ is a totally ordered subset of $\mathbb{R}$ such that $u_1 \leq \cdots \leq u_n$.
- The *worst utility* (resp. the *best utility*) is denoted by $u_\perp$ (resp. $u_\top$).

1.3 Uncertainty in decision problems

There are several situations of information availability about states of nature. These situations are characterized by different forms of uncertainty, namely :

- **Total uncertainty** where no information is available about the states of nature.
- **Probabilistic uncertainty** where uncertainty can be modeled via a probability distribution. We can distinguish two types of probability :

1. *Objective probability :* also called by frequentest probability, it indicates the relative frequency of the realization of events. A situation of uncertainty characterized by objective probability is called a situation of *risk*. In this report, as in literature, we will use the term risk to indicate probabilistic uncertainty and the term uncertainty to indicate all types of uncertainty.
2. *Subjective probability :* models the personal degree of belief that the events occur.

– **Non probabilistic uncertainty** where probability theory cannot be used to model uncertainty. Several non probabilistic uncertainty theories (named also non classical theories of uncertainty) have been developed such as *fuzzy sets theory* [88], *imprecise probabilities* [84], *possibility theory* [18, 20, 90] *evidence theory* [72] and *rough set theory* [59].
  In some situations, decision makers are unable to give exact numerical values to quantify decision problems but they can only provide an order relation between different values. This order relation can be represented by numerical values which have no sense but which express only the order. These situations are characterized by *qualitative uncertainty*, which can be represented using possibility theory.

1.4 Decision criteria under total uncertainty

Several decision criteria have been developed for decision under total uncertainty, regarding the decision maker behavior (optimistic, pessimistic and neutral). Among the most used ones, we will detail the Maximin, Maximax, Minimax regret, Laplace and Hurwicz decision criteria.

1.4.1 Maximin and Maximax criteria

Maximin and *Maximax* are a non probabilistic decision criteria defined by Wald [85, 86]. Maximin represents the *pessimistic* behavior of the decision maker. Decisions are ranked according to their worst outcomes. Indeed, the optimal decision is the one whose worst outcome is at least as good as the worst outcome of any other decisions.

Symmetrically, Maximax represents the *optimistic* behavior of the decision maker since decisions are ranked according to their best outcomes. Formally, these decision criteria can be defined as follows :

Definition 1.1 *The Maximin criterion (denoted by a_*) is expressed by :*

$$a_* = \max_{f \in F} a_{min}(f) \tag{1.1}$$

where $a_{min}(f) = \min_{s \in S} u(f(s))$.

The Maximax criterion (denoted by a^*) is expressed by :

$$a^* = \max_{f \in F} a_{max}(f) \tag{1.2}$$

where $a_{max}(f) = \max_{s \in S} u(f(s))$.

Example 1.1 *Let us consider a decision maker who should choose what he will buy between ice cream, cold drinks and newspapers. His satisfaction depends on the climate (nice, rain and snow). Table 1.1 represents utilities of each choice.*

Choice	Climate		
	Nice	Rain	Snow
Ice cream	500	300	50
Drinks	200	400	150
Newspapers	100	250	450

Table 1.1 – Utilities of drink choice problem

If we use the Maximin decision criterion, we have $a_* = 150$ and the optimal decision is to buy drinks. While if we use the Maximax decision criterion, we have $a^* = 500$ and the optimal decision is to buy ice cream.

Maximin and Maximax are very simple to compute, but they are not discriminant since we can have the same values (a_* or a^*) for different decisions (e.g. in the example 1.1 if the maximal utility for the three choices is equal to 500 then we will have $a^* = 500$ for each decision and we cannot choose between them).

1.4.2 Minimax regret criterion

In 1951, Savage [68, 80] proposed a decision model based on the *regret* (also called opportunity loss). Indeed, the regret is obtained by computing the difference between the utility of the current consequence and the maximal one for the same state. The Minimax regret approach is to minimize the worst case regret.

Definition 1.2 *The Minimax regret criterion (denoted by r_*) is computed as follows :*

$$r_* = \min_{f \in F} r(f) \tag{1.3}$$

where $r(f) = \max_{s \in S} r(f, s)$ and $r(f, s) = [\max_{f' \in F} u(f'(s))] - u(f(s))$.

Example 1.2 *Let us continue with the example 1.1. Table 1.2 represents the matrix of regrets which computes $r(f(s))$ for each state and each decision. In this example the decision maker will choose to buy drinks since $r_* = 300$.*

d_i	Nice	Rain	Snow	$r(f)$
Ice cream	0	100	400	400
Drinks	300	0	300	150
Newspapers	400	150	0	400

TABLE 1.2 – The matrix of regrets

Note that like the Maximin criterion, the Minimax regret criterion models pessimism but it is more sophisticated since it compares choices based on their regrets considering other choices. Nevertheless, the Minimax regret can lead the same minimal regrets for different decisions.

1.4.3 Laplace criterion

Laplace criterion (also called *Laplace insufficient reason* criterion) is justified by the fact that if no probabilities have been defined then there is no reason to not consider that any state $s \in S$ is more or less likely to occur than any other state [75]. Laplace criterion is the first model that used probability theory to represent uncertainty about states of nature, we have $\forall s \in S, pr(s) = 1/|S|$ (principle of equiprobability).

Definition 1.3 *The Laplace decision criterion (denoted by Lap^*) is computed as follows :*

$$Lap^* = \max_{f \in F} Lap(f) \tag{1.4}$$

where $Lap(f) = \frac{\sum_{s \in S} u(f(s))}{|S|}$.

Example 1.3 *Using the same example 1.1, we have :*
$Lap(buy\ ice\ cream) = 283.333$, $Lap(buy\ drinks) = 250$ *and*
$Lap(buy\ newspapers) = 266.666$.
So, the optimal decision according to Laplace criterion is to buy ice cream.

Laplace criterion uses the sum and the division operators, so it assumes that we have numerical utilities, which is not always the case. This point will be detailed in Chapter 2 where we will present qualitative decision theories. Note that Laplace criterion may give the same value for two different situations as it is presented in the following example

Example 1.4 *Let* $S = \{s_1, s_2\}$, $U = \{-100, 0, 100\}$ *and* $F = \{f, g\}$. *Table* 1.3 *represents the utility of each act in* F *for each state in* S *:*

	f	g
s_1	100	0
s_2	-100	0

TABLE 1.3 – Utilities for act f and g

$Lap(f) = \frac{(100+(-100))}{2} = 0$ *and* $Lap(g) = \frac{(0+0)}{2} = 0$.

Therefore, the assumption that all the states are equiprobable is not reasonable since we can have the same average for different decisions.

1.4.4 Hurwicz decision criterion

It is also called the criterion of realism or weighted average decision criterion [44]. In fact, it is a compromise between optimistic and pessimistic decision criteria. The computation of the Hurwicz is based on the coefficient α which is a value in the interval $[0, 1]$ that expresses the behavior of the decision maker such that if α is close to 1 (resp. 0) then the decision maker is optimistic (resp. pessimistic).

Definition 1.4 *The Hurwicz decision criterion (denoted by* H*) is computed as follows :*

$$H(f) = \alpha \ \min_{s \in S} u(f(s)) \ + \ (1-\alpha) \ \max_{s \in S} u(f(s)) \qquad (1.5)$$

Example 1.5 *Using the same example 1.1, we have for* $\alpha = 0.8$ *:*
$H(buy\ ice\ cream) = (0.8 * 50) + (0.2 * 500) = 140$*,*
$H(buy\ drinks) = (0.8 * 150) + (0.2 * 400) = 200$ *and*
$H(buy\ newspapers) = (0.8 * 100) + (0.2 * 450) = 170$*.*
So, the optimal decision according to the Hurwicz criterion is to buy drinks.

1.5 Expected decision theories

As we have seen in the Section 1.3, there exist several ways to represent uncertainty according to available information. This variety of uncertainty modeling's leads to different decision making criteria. Mainly, we distinguish quantitative and qualitative decision criteria. In this section, we focus on historical, quantitative decision criteria (see chapter 2 for more details about qualitative decision criteria).

Quantitative decision criteria can be used when uncertainty is represented by numerical values. The principal quantitative decision criterion is the *expected decision model* introduced by Bernoulli and developed by von Neumann and Morgenstern [57]. Despite its success, the expected decision model has some limits which were the subject of several works that proposed extensions of expected models for instance rank dependent utility and more generally *non expected decision models* [69]. These quantitative decision criteria are detailed in the sequel of this section.

1.5.1 Von Neumann and Morgenstern's decision model

In 1738, Bernoulli published a work entitled "St. Petersburg proceedings" in which he presented the paradox (known by "St. Petersburg paradox") which shows via a game that the expected value (i.e. the expected payoffs) of a small value may be infinite [4]. The principle of this game is as follows : *It tosses a coin, if face appears then we win* 2 *dollars and we stop the game, else we stay in the game and a new toss is made. If face appears then we win* 4 *dollars and we stop the game, else we throw the coin again. If face appears, we receive* 8 *dollars and so on.*
So in this game if a player wants to win 2^n dollars then he must made $n-1$ times tails

before face. In this case, the probability to win 2^n dollars is $1/2^n$. The expected monetary value of this game is $\sum_{i=1}^{+\infty} \frac{1}{2^i} * 2^i = \sum_{i=1}^{+\infty} 1 = +\infty$. Since the expected monetary value of this game is infinite then a player can pay any amount to play this game which is not reasonable.

In this example Bernoulli shows that the decision criterion based on the expectation value should be refined via the notion of *utility*. He argued that the utility function be a logarithmic function given its property of decreasing marginal utilities. In fact, this utility function allows a non linear processing of consequences which avoids the paradoxal nature of the game presented by Saint-Petersbourg.

Expected Utility (EU)

In 1944 [57], Von Neumann and Morgenstern (VNM) have developed the proposition of Bernoulli and they proposed *Expected Utility* theory (denoted by EU) and defined necessary conditions that guarantee the existence of a utility function. The EU model concerns decision making under risk, i.e. it is assumed that an objective probability distribution on the state of nature S is known. As we have seen, an act assigns a consequence to each state of nature and a utility is affected to each consequence by a utility function. So, $\forall s \in S, \exists u_i \in U \ s.t \ u(f(s)) = u_i$.

Formally, we have for each utility $u_i \in U$, $pr(u_i) = \sum_{s \in S} pr(s)$ such that $u(f(s)) = u_i$. Uncertainty, about the set of utilities $U = \{u_1, \ldots, u_n\}$ can be represented via a probabilistic lottery L denoted by $L = \langle \lambda_1/u_1, \ldots, \lambda_n/u_n \rangle$ where λ_i is *the uncertainty degree that the decision leads to an outcome of utility* u_i ($\lambda_i = L(u_i)$).
The size of a simple lottery is simply the number of its outcomes.

In a decision problem, each possible strategy can be represented by a lottery. Especially, in VNM approach λ_i is an objective probability and the expected utility of a lottery L is defined as follows :

Definition 1.5 *The expected utility of a lottery L (denoted by EU) is computed as follows :*

$$EU(L) = \sum_{u_i \in \mathcal{U}} u_i * L(u_i) \tag{1.6}$$

Example 1.6 *Let $L = \langle 0.1/10, 0.6/20, 0.3/30 \rangle$ and $L' = \langle 0.6/10, 0.4/30 \rangle$ be two probabilistic lotteries. Using Equation (1.6), we have $EU(L) = (0.1 * 10) + (0.6 * 20) + (0.3 * 30) = 22$ and $EU(L') = (0.6 * 10) + (0.4 * 30) = 18$ so L is preferred to L'.*

Von Neumann and Morgenstern axiomatization

Von Neumann and Morgenstern have proposed an axiomatic system (denoted by S_{EU}) to characterize a preference relation $\succeq$ between probabilistic lotteries.

Let L, L' and L'' be three probabilistic lotteries, the axiomatic system S_{EU} is defined as follows :

1. Axiom $1^{S_{EU}}$. **Completeness (Orderability)** : It is always possible to state either that $L \succeq L'$ or $L \succeq L'$.
2. Axiom $2^{S_{EU}}$. **Reflexivity** : Any lottery L is always at least as preferred as itself : $L \succeq L$.
3. Axiom $3^{S_{EU}}$. **Transitivity** : If $L \succeq L'$ and $L' \succeq L"$ then $L \succeq L"$.
4. Axiom $4^{S_{EU}}$. **Continuity** : If L' is between L and L'' in preference then there is a probability p for which the rational agent (DM) will be indifferent between the lottery L' and the lottery in which L comes with probability p, L'' with probability $(1-p)$.
$$L \succeq L' \succeq L'' \Rightarrow \exists p,\ s.t\ \langle p/L, (1-p)/L''\rangle \sim L'.$$
5. Axiom $5^{S_{EU}}$. **Substitutability** : If a DM is indifferent between two lotteries L and L', then there is a more complex lottery in which L can be substituted with L'.
$$(L \sim L') \Rightarrow \exists p,\ s.t\ \langle p/L, (1-p)/L''\rangle \sim \langle p/L', (1-p)/L''\rangle.$$
6. Axiom $6^{S_{EU}}$. **Decomposability** : Compound lotteries can be reduced to simpler lotteries using the laws of probability.
$$\langle p/L, (1-p)/\langle q/L', (1-q)/L''\rangle\rangle \Rightarrow \langle p/L, (1-p)q/L', (1-p)(1-q)/L''\rangle.$$
7. Axiom $7^{S_{EU}}$. **Independence** : If a DM prefers L to L', then he must prefer the lottery in which L occurs with a higher probability.
$$L \succeq L' \Rightarrow \forall p \in [0,1] \Longleftrightarrow \langle p/L, (1-p)/L"\rangle > \langle p/L', (1-p)/L"\rangle.$$

This independence axiom is the central axiom of the expected utility model. It can be interpreted as follows : The DM who prefers L to L' and who should make a choice between two mixtures $pL + (1-p)L''$ and $pL' + (1-p)L''$ will operate as follows : If an event with a probability $(1-p)$ happens, he will obtain L'' apart of his choice. However, if the complementary event happens the decision maker has to choose between L and L'. If the agent prefers L to L' then he will prefer the mixture $pL + (1-p)L''$ to $pL' + (1-p)L''$ according to the independence axiom.

The existence of a utility function according to VNM axioms is stated by the following theorem :

Theorem 1.1 *If the preference relation $\succeq$ satisfies completeness, reflexivity, transitivity, continuity and independence axioms then it exists a utility function $u : \mathcal{C} \to \mathbb{R}$ such that :*

$$L \succeq L' \quad \Leftrightarrow \quad \sum_{c \in C} u(c)L(c) \geq \sum_{c \in C} u(c)L'(c) \tag{1.7}$$

where, $\forall c \in C, L(c)$ (resp. $L'(c)$) is the probability degree to have the utility $u(c)$ from L (resp. L').

1.5.2 Savage decision model

In the VNM model, the hypothesis of the existence of objective probabilities is a strong assumption which is not guaranteed in all situations. So, an extension of expected decision theory based on subjective probability has been proposed in the literature [68].

Subjective Expected Utility (SEU)

Subjective expected utility is indeed based on the use of subjective probabilities to represent uncertainty. This theory was developed by Savage in 1954 [68]. In this decision theory, subjective probability is used to model uncertainty.

Definition 1.6 *The subjective expected utility of an act f (denoted by SEU) is computed as follows :*

$$SEU(f) = \sum_{s \in S} pr(s) * u(f(s)). \tag{1.8}$$

Example 1.7 *Let the set of states of nature $S = \{s_1, s_2, s_3\}$ such that $pr(s_1) = 0.5$, $pr(s_2) = 0.3$ and $pr(s_3) = 0.2$. The decision maker should choose between the act f and g that assign an utility to each state in S as it is represented in Table* 1.4.

Acts/States	s_1	s_2	s_3
f	20	10	30
g	10	20	30

TABLE 1.4 – The set of utilities

Using Equation (1.8), we have $SEU(f) = (0.5 * 20) + (0.3 * 10) + (0.2 * 30) = 19$ *and* $SEU(g) = (0.5 * 10) + (0.3 * 20) + (0.2 * 30) = 17$*, so* f *is preferred to* g*.*

Savage axiomatization

The second axiomatic system is the one proposed by Savage (denoted by S_{SEU}) [68]. This system gives necessary conditions that should be verified by a preference relation $\succeq$ between acts.

Before the development of the set of axioms, let us define the following notation concerning acts : $fAh(s)$: the act f is applied if a state s pertains to the event A while the act h is applied if $s \in A^c$. An event A is null iff $\forall f, \forall g, fAg \succeq g$ and $gAf \succeq g$.

1. Axiom $1^{S_{SEU}}$. $\succeq$ is complete and transitive.
2. Axiom $2^{S_{SEU}}$. (Sure Thing Principle) For any $f, g, h, h' \in F$ and not null event $A \subseteq S$:
$$fAh \succeq gAh \ \ iff \ \ fAh' \succeq gAh'.$$
3. Axiom $3^{S_{SEU}}$. For not null event $A \subseteq S$, and for any $f, g \in F^{const}$, $\forall h \in E$ we have :
$$fAh \succeq gAh \ \ iff \ \ f \succeq g.$$
4. Axiom $4^{S_{SEU}}$. For any $A, B \in S$ and for $f, g, f', g' \in F^{const}$ such that $f \succ g$ and $f' \succ g'$, we have :
$$fAg \succeq fBg \ \ iff \ \ f'Ag' \succeq f'Bg'.$$
5. Axiom $5^{S_{SEU}}$. There exist $f, g \in F^{const}$ such that $f \succ g$.
6. Axiom $6^{S_{SEU}}$. For any $f, g \in F$ such that $f \succ g$ and for any $h \in F^{const}$ there exists a finite partition P of the set S such that for any $H \in P$:
   (a) $[hHf] \succ g$ and
   (b) $f \succ [hHg]$.

The key axiom of Savage is the *Sure Thing Principle (**STP**)* (Axiom $2^{S_{SEU}}$). This axiom is interpreted by the fact that if an act is preferred when an event E is occurred then it will still preferred whatever the act in the case of complementary event. The sure thing principle (STP) axiom is considered as a strong condition and a weak version, named by *weak sure thing principle (WSTP)* has been proposed (see e.g. [80]) :

$$fEj \succ gEj \quad \Rightarrow \quad fEj' \succeq gEj'. \tag{1.9}$$

The existence of a utility function according to Savage axioms is stated by the following theorem :

Theorem 1.2 *If the preference relation $\succeq$ satisfies Axiom $1^{S_{SEU}}$ to Axiom $6^{S_{SEU}}$ then it exists a utility function $u : \mathcal{C} \to \mathbb{R}$ and a probability distribution Pr deduced from the preference relation between acts such that :*

$$f \succeq g \quad \Leftrightarrow \quad \sum_{s \in S} u(f(s))pr(s) \geq \sum_{s \in S} u(g(s))pr(s), \forall f, g \in F \tag{1.10}$$

1.6 Beyond expected utility decision models

Expected utility decision models cannot represent all decision makers behaviors because of their linear processing of probabilities. In fact, Allais [1] and Ellsberg [30] have presented experiences where EU and SEU cannot be used. In addition, probability theory cannot represent all kinds of uncertainty.

To overcome these limits, some extensions of expected utility have been developed like Rank Dependent Utility (RDU), Choquet and Sugeno integrals that we will detail below.

1.6.1 Rank Dependent Utility (RDU)

In 1953, Allais has shown the contradiction of the independence axiom of the VNM's system (Axiom $7^{S_{EU}}$) with the following counter example [1].

Counter Example 1.1 *Suppose that an agent will choose between the following lotteries :*

1. ***L1** : win 1 M with certainty ($L1 = \langle 1/1 \rangle$).*

2. ***L2*** *: win* 1 *M with a probability* 0.89*,* 5 *M with a probability* 0.1 *and* 0 *with a probability* 0.01 *(*$L2 = \langle 0.01/0, 0.89/1, 0.1/5\rangle$*).*

Then, he should choose between :

1. ***L1'*** *: win* 1 *M with a probability* 0.11 *and* 0 *with a probability* 0.89 *(*$L1' = \langle 0.89/0, 0.11/1\rangle$*).*
2. ***L2'*** *: win* 5 *M with a probability* 0.1 *and* 0 *with a probability* 0.9 *(*$L2' = \langle 0.9/0, 0.1/5\rangle$*).*

Clearly, an agent will choose $L1$ *instead of* $L2$ *and* $L2'$ *instead of* $L1'$*.*

Consider now the following game with four lotteries.

1. ***P*** *: win* 1 *M with a probability* 1 *(*$P = \langle 1/1\rangle$*).*
2. ***Q*** *: win* 0 *M with a probability* $1/11$ *and* 5 *M with a probability* $10/11$ *(*$Q = \langle 0.09/0, 0.9/5\rangle$*).*
3. ***R*** *: win* 0 *M with a probability* 1 *(*$R = \langle 1/0\rangle$*).*
4. ***S*** *: win* 1 *M with a probability* 1 *(*$S = \langle 1/1\rangle$*).*

We can represent lotteries $L1, L2, L1'$ *and* $L2'$ *as follows :*

$$L1 = 0.11P + 0.89S$$

$$L2 = 0.11Q + 0.89S$$

$$L1' = 0.11P + 0.89R$$

$$L2' = 0.11Q + 0.89R$$

As we have said the lottery $L1$ *is preferred to the lottery* $L2$*, thus* $0.11P + 0.89S \succeq 0.11Q + 0.89S$.

According to the independence axiom (Axiom $7^{S_{EU}}$*), this preference relation is equivalent to* $P \succeq Q$ *with* $p = 0.11,\ (1-p) = 0.89$ *and* $L" = S$*. Normally, we should have :* $L1' = 0.11P + 0.89R \succeq L2' = 0.11Q + 0.89R$ *i.e.* $L1' \succeq L2'$*. But in this experience, the agent has chosen* $L2'$*, hence contradiction : The independence axiom is not respected.*

As a solution to this paradox, Quiggin has developed *Rank Dependent Utility* [64] which is based on the use of a non linear processing of probabilities via a transformation function of probabilities (φ) which transforms cumulative probability. Generally, small probability degrees are neglected by decision makers and they have not an important impact on their choices. Based on this hypothesis, Buffon [11] has proposed to deal with a non linear probabilities which leads to a non linear treatment of consequences in decision making.
The *Rank Dependent Utility* criterion (denoted by *RDU*) is defined as follows :

Definition 1.7 *The Rank Dependent Utility of a lottery $L = \langle pr_1/u_1, \ldots, pr_n/u_n\rangle$ is computed as follows :*

$$RDU(L) = u_1 + \sum_{i=2}^{n} \varphi(\sum_{k=i}^{n} pr_k)\,[u_i - u_{i-1}] \tag{1.11}$$

where $\varphi(pr_k)$ is a transformation function of the probability pr_k.

Example 1.8 *Let $L = \langle 0.1/10, 0.6/20, 0.3/30\rangle$ and $L' = \langle 0.6/10, 0.4/30\rangle$ be two probabilistic lotteries. φ is a transformation function of probability such that :*

- *If $0 \leq pr_k < 0.35$ then $\varphi(pr_k) = 0$.*
- *If $0.35 \leq pr_k < 0.7$ then $\varphi(pr_k) = 0.3$.*
- *If $0.7 \leq pr_k \leq 1$ then $\varphi(pr_k) = 1$.*

Using Equation (1.11), we have $RDU(L) = 10+\varphi(0.6+0.3)*(20-10)+\varphi(0.3)*(30-10) = 20$ and $RDU(L') = 10 + \varphi(0.4) * (30 - 10) = 16$ so L is preferred to L'.

1.6.2 Choquet expected utility (CEU)

In 1961, Ellsberg has shown the contradiction of the Sure Thing Principle (Axiom $2^{S_{SEU}}$) via the following counter example [30] :

Counter Example 1.2 *Suppose that we have a box containing 90 balls (30 red (R), 60 blue (B) or yellow (Y)). The agent should choose between 4 decisions :*

- *x_1 : Bet on the fact that the drawn ball is red.*
- *x_2 : Bet on the fact that the drawn ball is blue.*
- *x_3 : Bet on the fact that the drawn ball is red or yellow.*
- *x_4 : Bet on the fact that the drawn ball is blue or yellow.*

Profits of each decision are presented in Table 1.5.

	R	B	Y
x_1	100	0	0
x_2	0	100	0
x_3	100	0	100
x_4	0	100	100

TABLE 1.5 – Profits

The majority prefer x_1 to x_2 or x_4 to x_3.

If we suppose that we can have 100$ as a profit if the ball is yellow then the decision x_1 will be similar to x_3 and x_2 similar to x_4. Normally, the preference relation will still unchanged since we have modified a constant profit in the two decisions, we will have $x_3 \succeq x_4$.
In fact, there is no a couple p and u such that $SEU(x_1) > SEU(x_2)$ and $SEU(x_4) > SEU(x_3)$. Hence Contradiction and the STP axiom is not satisfied.

The paradox of Allais and Ellsberg can be captured by the use of a Choquet integrals as a decision criterion based on a weakening of Savage's sure thing principle. In fact, Choquet expected utility allow the representation of the behaviors unlighted by Allais and Ellsberg.

Following [38] and [69], Choquet integrals appear as a right way to extend expected utility to non Bayesian models. Like the EU model, this model is a numerical, compensatory, way of aggregating uncertain utilities. But it does not necessarily resort on a Bayesian modeling of uncertain knowledge. Indeed, this approach allows the use of any monotonic set function [1] as a way of representing the DM's knowledge.

More precisely, Choquet integrals are defined according to a *capacity* (denoted by μ) which is a fuzzy measure $\mu : A \to [0,1]$ where A is a subset of the state of nature S.

Let X be a measurable function from some set T to $\mathbb{R}$, Choquet integrals are defined as follows :

$$\int_{Ch} X d_\mu = \int_{-\infty}^{0} [\mu(X > t) - 1]dt + \int_{0}^{\infty} \mu(X > t)dt. \tag{1.12}$$

If X is a finite set of values such that $x_1 \leq x_2 \leq \cdots \leq x_n$, Equation (1.12) may be written as follows :

$$\int_{Ch} X d_\mu = x_1 + \sum_{i=2}^{n} (x_i - x_{i-1})\mu(X \geq x_i). \tag{1.13}$$

When the measurable function X is a utility function u, the Choquet expected utility of Equation (1.12) (denoted by CEU) writes :

$$\int_{Ch} u d_\mu = \int_{-\infty}^{0} [\mu(u > t) - 1]dt + \int_{0}^{\infty} \mu(u > t)dt. \tag{1.14}$$

Given a lottery L, CEU may be also expressed by :

$$Ch_\mu(L) = \sum_{i=1}^{n} (u_i - u_{i-1})\mu(L \geq u_i) = u_1 + \sum_{i=2}^{n} (u_i - u_{i-1})\mu(u \geq u_i). \tag{1.15}$$

1. This kind of set function is often called capacity or fuzzy measure.

Given a decision f, CEU may be also expressed by :

$$Ch_{\mu}(f) = u_1 + \sum_{i=2}^{n}(u_i - u_{i-1})\mu(F_i) \quad (1.16)$$

where $F_i = \{s, u(f(s)) \geq u_i)$.
The fuzzy measure μ used in the definition of the Choquet expected utility may be any fuzzy measure.

Note that in the probabilistic case the capacity μ is the probability measure and the CEU corresponds to the particular case of *the Expected Utility (EU)* (see Equation 1.6) whereas if the capacity μ is a transformed probability (via a transformation function of probability φ) then CEU is simply collapse to *the Rank Dependent Utility (RDU)* (see Equation 1.11).

Example 1.9 *Let us consider the two lotteries of example 1.6 i.e.* $L = \langle 0.1/10, 0.6/20, 0.3/30 \rangle$ *and* $L' = \langle 0.6/10, 0.4/30 \rangle$*. Using Equation (1.15), we have* $Ch_{pr}(L) = 10 + (20 - 10) * 0.9 + (30 - 20) * 0.3 = 22$ *and* $Ch_{pr}(L') = 10 + (30 - 10) * 0.4 = 18$. *The value of* $Ch_{pr}(L)$ *(respectively* $Ch_{pr}(L')$*) is equal to* $EU(L)$ *(respectively* $EU(L')$*) since as we have mentioned* Ch_{pr} *is similar to* EU*.*

1.6.3 Sugeno integrals

In several cases, the decision maker cannot express his uncertainty by numerical values but he can only give an order between events. So, uncertainty is qualitative and quantitative decision models cannot be applied anymore.

Sugeno integrals [78, 79] are the qualitative counterpart of Choquet integrals requiring a totally ordered scale of uncertainty.

These integrals are used for qualitative decision theory based on any qualitative fuzzy measure μ. Sugeno integrals can be defined as follows :

$$S_{\mu}(f) = \max_{c \in C} \min(\mu(F_c), u(c)) \quad \forall f \in F \quad (1.17)$$

with $F_c = \{s \in S, u(f(s)) \geq u(c)\}$.

Note that if μ is a possibility measure Π (or a necessity measure N) [18] then S_{μ} is a possibilistic decision criterion that we will develop in Chapter 2.

1.7 Conclusion

As we have seen in this chapter, there exist several classical decision theories. In fact, the use of the appropriate decision criterion depends on the context of the decision problem namely on the nature of uncertainty (total uncertainty, numerical and ordinal) and on the behavior of decision makers (pessimistic, optimistic and neutral). EU-based decision models are well developed and axiomatized but they present some limits. We have presented some extensions of EU decision models, especially rank dependent utility, Choquet and Sugeno integrals. In next chapters, we will develop possibilitic decision theories that aim to avoid limits of EU decision models by using possibility theory for the representation of uncertainty and non expected decision models.

Chapitre 2

Possibilistic Decision Theory

2.1 Introduction

Probability theory is the fundamental uncertainty theory used in classical decision theory. Despite its fame, probability theory presents some limits since it cannot model qualitative uncertainty and total ignorance is represented by equiprobability which formalizes randomness rather than ignorance. In order to avoid limits of probability theory, non classical uncertainty theories have been developed. Possibility theory offers a suitable framework to handle uncertainty since it allows the representation of qualitative uncertainty. Decision criteria based on possibility theory have been developed such as fuzzy sets theory [88], possibility theory [89] and evidence theory [72].

In this chapter, we will first give some basic elements of possibility theory in Section 2.2. Pessimistic and optimistic utilities will be detailed in Section 2.3 and binary utilities will be developed in Section 2.4. Section 2.5 and 2.6 are respectively devoted to possibilistic likely dominance and to Order of Magnitude expected utility. Finally, Section 2.7 presents a deep study of possibilistic Choquet integrals with necessity and possibility measures.

Principal results of this chapter are published in [7, 9].

2.2 Basics of possibility theory

Possibility theory is a non classical theory of uncertainty devoted to handle imperfect informations. This section gives some basic elements of this theory, for more details see [18, 89, 90].

2.2.1 Possibility distribution

The basic building block of possibility theory is the notion of possibility distribution. It is denoted by π and it is a mapping from the universe of discourse Ω to a bounded linearly ordered scale L exemplified by the unit interval $[0,1]$, i.e. $\pi : \Omega \to [0,1]$.
The particularity of the possibilistic scale L is that it can be interpreted in twofold : in an *ordinal* manner, i.e. when the possibility degree reflect only an ordering between the possible values and in a *numerical* interpretation, i.e. when possibility distributions are related to upper bounds of imprecise probability distributions.

The function π represents a flexible restriction of the values $\omega \in \Omega$ with the conventions represented in Table 2.1.

$\pi(\omega) = 0$	ω is impossible
$\pi(\omega) = 1$	ω is totally possible
$\pi(\omega) > \pi(\omega')$	ω is more possible than ω' (or is more plausible)

TABLE 2.1 – Conventions for possibility distribution π

An important property relative to possibility distribution is the **normalization** stating that at least one element of Ω should be fully possible i.e. :

$$\exists \omega \in \Omega \ s.t \ \pi(\omega) = 1 \tag{2.1}$$

Possibility theory is driven by the principle of minimal specificity. A possibility distribution π' is more specific than π iff $\pi' \leq \pi$, which means that any possible value for π' should be at least as possible for π. Then, π' is more informative than π.
In the possibilistic framework, extreme forms of partial knowledge can be represented as follows :

– *Complete knowledge :*

$$\exists \omega \in \Omega, \pi(\omega) = 1 \qquad and \qquad \forall \omega' \neq \omega, \pi(\omega') = 0. \tag{2.2}$$

– *Complete ignorance :*

$$\forall \omega \in \Omega, \pi(\omega) = 1. \tag{2.3}$$

Example 2.1 *Let us consider the problem of dating the fossil. The universe of discourse related to this problem is the set of geological era defined by*
$\Omega = \{Cenozoic(Ceno), Mesozoic(Meso), Paleozoic(Paleo)\}$.
Suppose that a geologist gave his opinion on the geological era (E) of a fossil in the form of a possibility distribution π_1 i.e. :

$$\pi_1(Ceno) = 0.4, \quad \pi_1(Meso) = 1, \quad \pi_1(Paleo) = 0.3.$$

For instance, the degree 0.3 *represents the degree of possibility that the geological era of F is Paleozoic. π_1 is normalized since $max(0.4, 1, 0.3) = 1$.*
$\pi_1(E = Meso) = 1$ means that it is fully possible that the fossil dates from the Mesozoic era.

2.2.2 Possibility and necessity measures

In probability theory, for any event $\psi \subset \Omega$, $P(\neg\psi) = 1 - P(\psi)$. The expression *It is not probable that "not ψ"* means that *It is probable that ψ*. On the contrary, *it is possible that* ψ does not entail anything about the possibility of $\neg\psi$.

Probability is self dual, which is not the case of possibility theory where the description of uncertainty about ψ needs two dual measures : The *possibility measure* $\Pi(\psi)$ and the *necessity measure* $N(\psi)$ detailed below.
These two dual measures are defined as follows :

Possibility measure

Given a possibility distribution π, the possibility measure is defined by :

$$\Pi(\psi) = max_{\omega \in \psi} \pi(\omega) \quad \forall \psi \subseteq \Omega. \tag{2.4}$$

$\Pi(\psi)$ is called the possibility degree of ψ, it corresponds to the possibility to have one of the models of ψ as the real world [20]. Table 2.2 gives main properties of possibility measure.

$\Pi(\psi) = 1$ and $\Pi(\neg\psi) = 0$	ψ is certainly true
$\Pi(\psi) = 1$ and $\Pi(\neg\psi) \in]0, 1[$	ψ is somewhat certain
$\Pi(\psi) = 1$ and $\Pi(\neg\psi) = 1$	total ignorance
$\Pi(\psi) > \Pi(\varphi)$	ψ is more plausible than φ
$max(\Pi(\psi), \Pi(\neg\psi)) = 1$	ψ and $\neg\psi$ cannot be both impossible
$\Pi(\psi \vee \varphi) = \max(\Pi(\psi), \Pi(\varphi))$	disjunction axiom
$\Pi(\psi \wedge \varphi) \leq \min(\Pi(\psi), \Pi(\varphi))$	conjunction axiom

TABLE 2.2 – Possibility measure Π

Necessity measure

The necessity measure represents the dual of the possibility measure. Formally, $\forall \psi \subseteq \Omega$:

$$N(\psi) = 1 - \Pi(\neg\psi) = min_{\omega \notin \psi}(1 - \pi(\omega)). \tag{2.5}$$

Necessity measure corresponds to the certainty degree to have one of the models of ψ as the real world. This measure evaluates at which level ψ is certainly implied by our knowledge represented by π. Table 2.3 represents a summary of main properties of this measure.

$N(\psi) = 1$ and $N(\neg\psi) = 0$	ψ is certainly true
$N(\psi) \in]0,1[$ and $N(\neg\psi) = 0$	ψ is somewhat certain
$N(\psi) = 0$ and $N(\neg\psi) = 0$	total ignorance
$min(N(\psi), N(\neg\psi)) = 0$	unique link existing between $N(\psi) and N(\neg\psi)$
$N(\psi \wedge \varphi) = \min(N(\psi), N(\varphi))$	conjunction axiom
$N(\psi \vee \varphi) \geq \max(N(\psi), N(\varphi))$	disjunction axiom

TABLE 2.3 – Necessity measure N

Possibilistic conditioning

The conditioning consists in revising our initial knowledge, represented by a possibility distribution π, which will be changed into another possibility distribution $\pi' = \pi(.|\psi)$ with $\psi \neq \emptyset$ and $\Pi(\psi) > 0$. The two interpretations of the possibilistic scale induce two definitions of the conditioning :

– min-based conditioning relative to the ordinal setting :

$$\pi(\omega|_m\psi) = \begin{cases} 1 & if\ \pi(\omega) = \Pi(\psi)\ \ and\ \ \omega \in \psi \\ \pi(\omega) & if\ \pi(\omega) < \Pi(\psi)\ \ and\ \ \omega \in \psi \\ 0 & otherwise. \end{cases} \tag{2.6}$$

– product-based conditioning relative to the numerical setting :

$$\pi(\omega|_p\psi) = \begin{cases} \frac{\pi(\omega)}{\Pi(\omega)} & if\ \omega \in \psi \\ 0 & otherwise. \end{cases} \tag{2.7}$$

Example 2.2 *Let us consider the problem of fossil's dating. Suppose that the geologist makes a radioactivity's test on the fossil which help them to know the fossil's breed represented by the following set :*
$Breed = \{Mammal, Fish, Bird\}$.
Suppose that we have a fully certain piece of information indicating that the breed of the fossil is mammal.
Then, $\psi' = \{Ceno \wedge Mammal, Meso \wedge Mammal, Paleo \wedge Mammal\}$ *and*
$\Pi(\psi) = max(0.2, 0.3, 0.5) = 0.5$. *Using Equations* 2.6 *and* 2.7*, new distributions are presented in Table* 2.4 *where* $\psi = (Era \wedge Breed)$.

Era	Breed	$\pi(\psi)$	$\pi(\psi \mid_p \psi')$	$\pi(\psi \mid_m \psi')$
Ceno	Mammal	0.2	0.2	0.4
Ceno	Fish	1	0	0
Ceno	Bird	0	0	0
Meso	Mammal	0.3	0.3	0.6
Meso	Fish	0.7	0	0
Meso	Bird	0.7	0	0
Paleo	Mammal	0.5	1	1
Paleo	Fish	0.2	0	0
Paleo	Bird	1	0	0

TABLE 2.4 – Possibilistic conditioning

2.2.3 Possibilistic lotteries

Dubois et al. [24, 25] have proposed a possibilistic counterpart of VNM's notion of lottery (Chapter 1 Section 1.5). In the possibilistic framework, an act can be represented by a possibility distribution on U, also called a *possibilistic lottery*, and denoted by $\langle \lambda_1/u_1, \ldots, \lambda_n/u_n \rangle$ where $\lambda_i = \pi(u_i)$ is the possibility that the decision leads to an outcome of utility u_i. When there is no ambiguity, we shall forget about the utility degrees that receive a possibility degree equal to 0 in a lottery, i.e. we write $\langle \lambda_1/u_1, \ldots, \lambda_n/u_n \rangle$ instead of $\langle \lambda_1/u_1, \ldots, 0/u_i, \ldots \lambda_n/u_n \rangle$. The set of possibilistic lotteries is denoted by $\mathcal{L}$.

A *possibilistic compound lottery* $\langle \lambda_1/L_1, \ldots, \lambda_m/L_m \rangle$ (also denoted by $\lambda_1 \wedge L_1 \vee \cdots \vee \lambda_m \wedge L_m$) is a possibility distribution over a subset of $\mathcal{L}$. The possibility $\pi_{i,j}$ of getting a utility degree $u_j \in U$ from one of its sub−lotteries L_i depends on the possibility λ_i of getting L_i and on the conditional possibility $\lambda^i_j = \pi(u_j \mid L_i)$ of getting u_j from L_i i.e. $\pi_{i,j} = \lambda_j \otimes \lambda^i_j$, where $\otimes$ is equal to min in the case of qualitative scale and it is equal to $*$ if the scale is numerical. Hence, the possibility of getting u_j from a compound lottery $\langle \lambda_1/L_1, \ldots, \lambda_m/L_m \rangle$ is the *max*, over all L_i, of $\pi_{i,j}$. Thus, [24, 25] have proposed to reduce $\langle \lambda_1/L_1, \ldots, \lambda_m/L_m \rangle$ into a simple lottery denoted by, *Reduction*$(\langle \lambda_1/L_1, \ldots, \lambda_m/L_m \rangle)$. Formally, we have :

$$Reduction(\langle \lambda_1/L_1, \ldots, \lambda_m/L_m \rangle) = \langle \underset{j=1..m}{max} (\lambda_j \otimes \lambda^j_1)/u_1, \ldots, \underset{j=1..m}{max} (\lambda_j \otimes \lambda^j_n)/u_n \rangle. \tag{2.8}$$

where $\otimes$ is the product operator in the case of quantitative possibility theory and the min operator in the case of its qualitative counterpart.

Example 2.3 *Let* $L_1 = \langle 0.2/10, 0.9/20, 1/30 \rangle$*,* $L_2 = \langle 1/10, 0.1/20, 0.1/30 \rangle$ *and* $L_3 = \langle 1/10, 0.15/20, 0.5/30 \rangle$ *be three possibilistic lotteries.*

$L_4 = \langle 0.2/L_1, 1/L_2, 0.5/L_3 \rangle$ *is a compound possibilistic lottery that will be reduced into a simple possibilistic lottery* L'_4*. We have :*

- *In qualitative setting :*
  - $L'_4(10) = \max[\min(0.2, 0.2), \min(1, 1), \min(0.5, 1)] = 1$,
  - $L'_4(20) = \max[\min(0.2, 0.9), \min(1, 0.1), \min(0.5, 0.15)] = 0.2$ *and*
  - $L'_4(30) = \max[\min(0.2, 1), \min(1, 0.1), \min(0.5, 0.5)] = 0.5$.

  So, $L'_4 = \langle 1/10, 0.2/20, 0.5/30 \rangle$.
- *In numerical setting :*
  - $L'_4(10) = \max[(0.2 * 0.2), (1 * 1), (1 * 0.5)] = 1$,
  - $L'_4(20) = \max[(0.2 * 0.9), (1 * 0.1), (0.5 * 0.15)] = 0.18$ *and*
  - $L'_4(30) = \max[(0.2 * 1), (1 * 0.1), (0.5 * 0.5)] = 0.25$.

  So, $L'_4 = \langle 1/10, 0.18/20, 0.25/30 \rangle$.

Like the probabilistic case, the reduction of a compound possibilistic lottery is polynomial in the size of the compound lottery since min (or $*$) and max are polynomial operations.

2.3 Pessimistic and optimistic utilities

Under the assumption that the utility scale and the possibility scale are commensurate and purely ordinal, [25] have proposed qualitative pessimistic and optimistic utility degrees for evaluating any simple lottery $L = \langle \lambda_1/u_1, \ldots, \lambda_n/u_n \rangle$ (possibly issued from the reduction of a compound lottery).

2.3.1 Pessimistic utility (U_{pes})

The pessimistic criterion was originally proposed by Whalen [21], it supposes that the decision maker is happy when bad consequences are hardly plausible.

Definition 2.1 *The pessimistic utility of a possibilistic lottery* $L = \langle \lambda_1/u_1, \ldots, \lambda_n/u_n \rangle$ *(denoted by* U_{pes}*) is computed as follows :*

$$U_{pes}(L) = \min_{i=1..n} \max(u_i, N(L \geq u_i)) \tag{2.9}$$

where $N(L \geq u_i) = 1 - \Pi(L < u_i) = 1 - \underset{j=1,i-1}{max} \lambda_j$.

Example 2.4 *Let a possibilistic lottery* $L = \langle 0.5/0.4, 1/0.6, 0.2/0.8 \rangle$*, using Equation (2.9) we have* $U_{pes}(L) = max(0.5, min(0.6, 0.5), min(0.8, 0)) = 0.5$.

Particular values for U_{pes} are as follows :
- If L assigns the possibility 1 to $u_\perp$ (the worst utility) then $U_{pes}(L) = 0$.
- If L assigns the possibility 1 to $u_\top$ (the best utility) and 0 to all other utilities then $U_{pes}(L) = 1$.

2.3.2 Optimistic utility (U_{opt})

The optimistic criterion was originally defined by Yager [21], it captures the optimistic behavior of the decision maker. It estimates to what extend it is possible that a possibilistic lottery reaches a good utility.

Definition 2.2 *The optimistic utility of a possibilistic lottery $L = \langle \lambda_1/u_1, \ldots, \lambda_n/u_n \rangle$ (denoted by U_{opt}) is computed as follows :*

$$U_{opt}(L) = \max_{i=1..n} \min(u_i, \Pi(L \geq u_i)) \tag{2.10}$$

where $\Pi(L \geq u_i) = \underset{j=i..n}{max} \lambda_j$.

Example 2.5 *Let a lottery $L = \langle 0.5/0.4, 1/0.6, 0.2/0.8 \rangle$, using Equation (2.10) we have* $U_{opt}(L) = max(min(0.5, 1), min(0.6, 1), min(0.8, 0.2)) = 0.6$.

Particular values for U_{opt} are as follows :
- If L assigns 1 to $u_\top$ then $U_{opt}(L) = 1$.
- If L assigns 1 to $u_\perp$ and 0 to all other utilities then $U_{opt}(L) = 0$.

U_{pes} and U_{opt} are qualitative decision criteria that represent particular cases of Sugeno integrals in the context of possibility theory. More precisely, if the fuzzy measure in the Sugeno formula (Chapter 1 Section 1.6) is a necessity measure N then the Sugeno integral is pessimistic utility. If this capacity is a possibility measure Π then the Sugeno integral is optimistic utility.

2.3.3 Axiomatization of pessimistic and optimistic utilities

As we have seen in Chapter 1, there exist two basic axiomatic systems for expected utilities (S_{EU} and S_{SEU}). Pessimistic and optimistic utilities were axiomatized in the style of VNM [24, 25] to characterize preference relations between possibilistic lotteries. They have been axiomatized in the style of Savage by [26, 27, 28].

These relations between Sugeno integrals and qualitative utilities have lead to an axiomatic systems of Sugeno integral that generalizes the ones of U_{pes} and U_{opt} (see [26] for more details).

Axiomatization of pessimistic utility (U_{pes}) in the style of VNM

Let $\succeq$ be a preference relation on the set of possibilistic lotteries $\mathcal{L}$. The axiomatic system of U_{pes} (denoted by S_P) was proposed by [27], it is defined as follows :

- Axiom 1^{S_P}. **Total pre-order** : $\succeq$ is reflexive, transitive and complete.
- Axiom 2^{S_P}. **Certainty equivalence** : if the belief state is a crisp set $A \subseteq U$, then there is $u \in U$ such that $\{u\} \sim A$.
- Axiom 3^{S_P}. **Risk aversion** : if π is more specific than π' then $\pi \succeq \pi'$.
- Axiom 4^{S_P}. **Independence** : if $\pi_1 \sim \pi_2$ then $\langle \lambda/\pi_1, \mu/\pi \rangle \sim \langle \lambda/\pi_2, \mu/\pi \rangle$.
- Axiom 5^{S_P}. **Reduction of lotteries** :
  $\langle \lambda/u, \mu/(\alpha/u, \beta/u') \rangle \sim \langle max(\lambda, min(\mu, \alpha))/u, min(\mu, \beta)/u' \rangle$.
- Axiom 6^{S_P}. **Continuity** : $\pi' \leq \pi \Rightarrow \exists \lambda \in [0, 1]$ s.t. $\pi' \sim \langle 1/\pi, \lambda/u_\perp \rangle$

In [24], Dubois et al. have done a deeper study of pessimistic utilities and have shown that only four axioms are needed for this decision criterion. The authors have shown that qualitative utilities require preference and uncertainty scales equipped with the maximum, the minimum and an order reversing operations. An improved set of axioms have been proposed for pessimistic utilities that does not include the axiom concerning the reduction of lotteries (Axiom 5^{S_P}) since this axiom is implicitly obtained by the definition of possibilistic lotteries. Another result in [24] is that the axiom of certainty equivalence (Axiom 2^{S_P}) is redundant and it is a direct consequence of (Axiom 1^{S_P}, Axiom 4^{S_P} and Axiom 6^{S_P}).

The work of Dubois et al. [24] has led to a new set of axioms (denoted by S'_P) which contains (Axiom 1^{S_P}, Axiom 3^{S_P} and Axiom 4^{S_P}) and a new form of the axiom of continuity (Axiom $6^{S'_P}$) :

- Axiom $6^{S'_P}$. **Continuity** : $\forall \pi, \exists \lambda \in [0, 1]\ \pi \sim \langle 1/u_\top, \lambda/u_\perp \rangle$.

Theorem 2.1 *A preference relation $\succeq$ on $\mathcal{L}$ satisfies the axiomatic system S'_P iff there exists a utility function $u : L \rightarrow [0, 1]$ such that :*

$$L \succeq L' \quad iff \quad U_{pes}(L) \geq U_{pes}(L'). \tag{2.11}$$

Axiomatization of pessimistic utility (U_{pes}) in the style of Savage

The axiomatic system of U_{pes} in the context of Savage is denoted by S_{PS}, it contains the following axioms :

- Axiom $1^{S_{PS}}$. is the Axiom $1^{S_{SEU}}$ from S_{SEU} concerning ranking of acts.
- Axiom $2^{S_{PS}}$. **(Weak compatibility with constant acts)** : Let x and y be two constant acts ($x = z$ and $y = w$) $\forall E \subseteq S$ and $\forall h$ $z \leq w \Rightarrow xEh \leq yEh$
- Axiom $3^{S_{PS}}$. is the Axiom $5^{S_{SEU}}$ from S_{SEU} concerning the non triavility.
- Axiom $4^{S_{PS}}$. **(Restricted max dominance)** : Let f and g be any two acts and y be a constant act of value y : $f \succ g$ and $f \succ y \Rightarrow f \succ g \vee y$.
- Axiom $5^{S_{PS}}$. **(Conjunctive dominance)** : $\forall f, g, h$ $g \succ f$ and $h \succ f \Rightarrow g \wedge h \succ f$.

The restricted max dominance axiom (Axiom $4^{S_{PS}}$) means that if an act f is preffered to an act g and also to the constant act y then, even if the worst consequences of g are improved to the value y, the act f is still preferred to g. Indeed, a strengthened form of the conjunctive dominance is expressed by the axiom Axiom $5^{S_{PS}}$. Notice that pessimistic utility does not satisfy the STP axiom but its weaker version namely the axiom $WSTP$ (Chapter 1 Section 1.5).

Theorem 2.2 *A preference relation $\succeq$ on acts satisfies the axiomatic system S_{PS} iff there exists a utility function $u : C \rightarrow [0, 1]$ and a possibility distribution $\pi : S \rightarrow [0, 1]$ such that $\forall f, g \in F$: $f \succeq g$ iff $U_{pes}(f) \geq U_{pes}(g)$.*

After presenting the axioms of pessimistic utility in the style of VNM and Savage, we will proceed to represent those concerning optimistic utility.

Axiomatization of optimistic utility (U_{opt}) in the style of VNM

In [25], Dubois et al. presented an axiomatic system (denoted by S_O) that characterizes U_{opt}. This system is obtained from S_P by substituting Axiom 2^{S_P} and Axiom 4^{S_P} by their diametrical counterparts i.e. Axiom 2^{S_O} and Axiom 4^{S_O} :

- Axiom 2^{S_O}. **Uncertainty attraction** : if $\pi' \geq \pi$ then $\pi' \succeq \pi$.
- Axiom 4^{S_O}. $\forall \pi, \exists \lambda \in [0, 1]$ s.t. $\pi \sim \langle \lambda/u_\top, 1/u_\perp \rangle$.

In an analogous way to the pessimistic case, the Axiom $6^{S'_O}$ is defined for optimistic utility to improve its axiomatic system [24] :

- Axiom $6^{S'_O}$. **Continuity** : $\pi' \leq \pi \Rightarrow \exists \lambda \in [0, 1]$ s.t. $\pi \sim \langle 1/\pi', \lambda/u_\perp \rangle$.

Theorem 2.3 *A preference relation $\succeq$ on $\mathcal{L}$ satisfies the axiomatic system S_O and Axiom $6^{S'_O}$ iff there exists an optimistic utility function $u : L \to [0,1]$ such that :*

$$L \succeq L' \;\; iff \;\; U_{opt}(L) \geq U_{opt}(L'). \tag{2.12}$$

Axiomatization of optimistic utility (U_{opt}) in the style of Savage

The axiomatic system of U_{opt} in the context of Savage is denoted by S_{OS}, it shares some similar axioms to S_{PS} (Axiom $1^{S_{PS}}$, Axiom $2^{S_{PS}}$ and Axiom $3^{S_{PS}}$). S_{OS} is as follows :

- Axiom $1^{S_{OS}}$ is Axiom $1^{S_{PS}}$.
- Axiom $2^{S_{OS}}$ is Axiom $2^{S_{PS}}$.
- Axiom $3^{S_{OS}}$ is Axiom $3^{S_{PS}}$.
- Axiom $4^{S_{OS}}$. **(Restricted conjunctive dominance)** : Let f and g be any two acts and y be a constant act of value y : $g \succ f$ and $y \succ f \Rightarrow g \wedge y \succ f$.
- Axiom $5^{S_{OS}}$. **(disjunctive dominance)** : $\forall f, g, h \; f \succ g$ and $f \succ h \Rightarrow f \succ g \vee h$.

Axiom $4^{S_{OS}}$ is the dual property of the restricted max dominance which holds for the conjunction of two acts and a constant one. It allows a partial decomposability of qualitative utility with respect to the conjunction of acts in the case where one of them is constant.

The second particular axiom in S_{OS} is the axiom of disjunctive dominance which express that the decision maker focuses on the "best" plausible states.

Theorem 2.4 *A preference relation $\succeq$ on acts satisfies the axiomatic system S_{OS} iff there exists a utility function $u : C \to [0,1]$ and a possibility distribution $\pi : S \to [0,1]$ such that $\forall f, g \in F$:*

$$f \succeq g \;\; iff \;\; U_{opt}(f) \geq U_{opt}(g). \tag{2.13}$$

2.4 Binary utilities (PU)

Giang and Shenoy [36] criticized pessimistic and optimistic utilities presented by Dubois et al. in [21]. Their argument is based on the fact that proposed frameworks for possibilistic utilities are based on axioms (i.e Axiom 2^{S_P} and Axiom 2^{S_O}) relative to uncertainty attitude contrary to the VNM axiomatic system based on risk attitude which does not make a sense in the possibilistic framework since it represents uncertainty rather than risk.

Moreover, to use pessimistic and optimistic utilities, the decision maker should classify himself as either pessimistic or optimistic which is not always obvious and even this classification is done it can lead to unreasonable decision.

That is why Giang and Shenoy [37] have proposed a bipolar criterion which encompasses both the pessimistic and optimistic utilities. Claiming that the lotteries that realize in the best prize or in the worst prize play an important role in decision making, these authors have proposed a bipolar model in which the utility of an outcome is a pair $u = \langle \overline{u}, \underline{u} \rangle$ where $\max(\overline{u}, \underline{u}) = 1$: the utility is binary i.e. $\overline{u}$ is interpreted as the possibility of getting the ideal, good reward (denoted by $\top$) and $\underline{u}$ is interpreted as the possibility of getting the anti ideal, bad reward (denoted by $\perp$).

Because of the normalization constraint $\max(\overline{u}, \underline{u}) = 1$, the set
$U = \{\langle \overline{u}, \underline{u} \rangle \in [0,1]^2, \max(\lambda, \mu) = 1\}$ is totally ordered :

$$\langle \overline{u}, \underline{u} \rangle \succeq_b \langle \overline{v}, \underline{v} \rangle \quad iff \quad \begin{cases} \overline{u} = \overline{v} = 1 \ and \ \underline{u} \leq \underline{v} \\ \quad\quad or \\ \overline{u} \geq \overline{v} \ and \ \underline{u} = \underline{v} = 1 \\ \quad\quad or \\ \overline{u} = 1, \underline{v} = 1 \ and \ \overline{v} < 1 \end{cases} \tag{2.14}$$

Each $u_i = \langle \overline{u_i}, \underline{u_i} \rangle$ in the utility scale is thus understood as a small lottery $\langle \overline{u_i}/\top, \underline{u_i}/\perp \rangle$. A lottery $\langle \lambda_1/u_1, \ldots, \lambda_n/u_n \rangle$ can be view as a compound lottery, and its utility is computed by reduction using Equation2.15.

Definition 2.3 *The binary utility of a lottery $L = \langle \lambda_1/u_1, \ldots, \lambda_n/u_n \rangle$ (denoted by PU) is computed as follows :*

$$\begin{aligned} &PU(\langle \lambda_1/u_1, \ldots, \lambda_n/u_n \rangle) \\ &= Reduction(\lambda_1/\langle \overline{u_1}/\top, \underline{u_1}/\perp \rangle, \ldots, \lambda_n/\langle \overline{u_n}/\top, \underline{u_n}/\perp \rangle) \\ &= \langle \max_{j=1..n}(\min(\lambda_j, \overline{u_j}))/\top, \max_{j=1..n}(\min(\lambda_j, \underline{u_j}))/\perp \rangle \end{aligned} \tag{2.15}$$

We thus get, for any lottery L a binary utility $PU(L) = \langle \overline{u}, \underline{u} \rangle$ in U. Lotteries can then be compared according to Equation (2.14) :

$$L \succeq L' \quad iff \quad Reduction(L) \succeq Reduction(L'). \tag{2.16}$$

Example 2.6 *Let $u_1 =< 1, 0 >$, $u_2 =< 1, 0.5 >$, $u_3 =< 1, 0.7 >$ and $u_4 =< 1, 1 >$. Let L and L' be two corresponding lotteries such that :*

$L = \langle 0.7/u_1, 1/u_2, 0.5/u_3, 0.5/u_4\rangle$ and $L' = \langle 0.5/u_1, 0.7/u_2, 0/u_3, 1/u_4\rangle$.
Using equation 2.15, we have $PU(L) = \langle 1, 0.5\rangle$ and $PU(L') = \langle 1, 1\rangle$. So, $L' \succ L$.

Axiomatization of binary utilities (PU)

The preference relation $\succeq_{PU}$ satisfies the following axiomatic system (denoted by S_{PU}) in the style of Von Neumann and Morgenstern decision theory :

- Axiom $1^{S_{PU}}$. **Total pre-order** : $\succeq_{PU}$ is reflexive, transitive and complete.
- Axiom $2^{S_{PU}}$. **Qualitative monotonicity** $\succeq_{PU}$ satisfies the following condition :

$$\langle \lambda/u_\top, \mu/u_\bot\rangle \succeq \langle \lambda'/u_\top, \mu'/u_\bot\rangle \ \ if \ \ \begin{cases} (1 \geq \lambda \geq \lambda' \ \ and \ \ \mu = \mu' = 1) \ \ or \\ (\lambda = 1 \ \ and \ \ \lambda' < 1) \ \ or \\ (\lambda = \lambda' = 1 \ \ and \ \ \mu' \geq \mu) \end{cases} \tag{2.17}$$

- Axiom $3^{S_{PU}}$. **Substitutability** : if $L \sim L'$ then $\langle \lambda/L, \mu/L"\rangle \sim \langle \lambda/L', \mu/L"\rangle$.
- Axiom $4^{S_{PU}}$. **Continuity** : $\forall c \in C$, $\exists L \in \mathcal{L}$ s.t $c \sim L$.

Theorem 2.5 *A preference relation $\succeq$ on $\mathcal{L}$ satisfies the axiomatic system S_{PU} iff there exists a binary utility such that :*

$$L \succeq L' \ \ iff \ \ PU(L) \geq PU(L'). \tag{2.18}$$

2.5 Possibilistic likely dominance (LN, $L\Pi$)

When the scales evaluating the utility and the possibility of the outcomes are not commensurate, [29, 31] propose to prefer, among two possibilistic decisions, the one that is more likely to overtake the other. Such a rule does not assign a global utility degree to the decisions, but draws a pairwise comparison. Although designed on a Savage-like framework rather than on lotteries, it can be translated on lotteries. This rule states that given two lotteries $L_1 = \langle \lambda_1^1/u_1^1, \ldots, \lambda_n^1/u_n^1\rangle$ and $L_2 = \langle \lambda_1^2/u_1^2, \ldots, \lambda_n^2/u_n^2\rangle$, L_1 is as least as good as L_2 as soon as the likelihood (here, the necessity or the possibility) of *the utility of L_1 is as least as good as the utility of L_2* is greater or equal to the likelihood of *the utility of L_2 is as least as good as the utility of L_1*. Formally :

Definition 2.4 *$\geq_{LN}$ and $\geq_{L\Pi}$ are defined as follows :*

$$L_1 \geq_{LN} L_2 \;\; iff \;\; N(L_1 \geq L_2) \geq N(L_2 \geq L_1). \tag{2.19}$$

$$L_1 \geq_{L\Pi} L_2 \;\; iff \;\; \Pi(L_1 \geq L_2) \geq \Pi(L_2 \geq L_1) \tag{2.20}$$

where $\Pi(L_1 \geq L_2) = \sup_{u_i^1, u_i^2 \text{ s.t. } u_i^1 \geq u_i^2} min(\lambda_i^1, \lambda_i^2)$ and
$N(L_1 \geq L_2) = 1 - \sup_{u_i^1, u_i^2 \text{ s.t. } u_i^1 < u_i^2} min(\lambda_i^1, \lambda_i^2)$.

The preference order induced on the lotteries is not transitive, but only quasitransitive : obviously $L_1 >_N L_2$ and $L_2 >_{LN} L_3$ implies $L_1 >_{LN} L_3$ (resp. $L_1 >_{L\Pi} L_2$ and $L_2 >_{L\Pi} L_3$ implies $L_1 >_{L\Pi} L_3$) but it may happen that $L_1 \sim_{LN} L_2$, $L_2 \sim_{LN} L_3$ (resp. $L_1 \sim_{L\Pi} L_2$, $L_2 \sim_{L\Pi} L_3$) and $L_1 >_{LN} L_3$ (resp. $L_1 >_{L\Pi} L_3$).

Example 2.7 *Let the set of states of nature $S = \{s_1, s_2, s_3\}$ such that $\Pi(s_1) = 0.3$, $\Pi(s_2) = 0.7$ and $\Pi(s_3) = 1$ and the set of utilities $U = \{2, 3, 5\}$. The lotteries L and L' are as follows $L = \langle 1/2, 0.7/3, 0.3/5 \rangle$ and $L' = \langle 0.7/2, 0.3/3, 1/5 \rangle$.*

We have $[L \succ L'] = \{s_1, s_2\}$ and $[L' \succ L] = \{s_3\}$.
$\Pi(\{s_1, s_2\}) = 0.7 < \Pi(\{s_3\}) = 1$, so $L' >_{L\Pi} L$.
$N(\{s_1, s_2\}) = 0 < N(\{s_3\}) = 0.3$, so $L' >_{LN} L$.

Axiomatization of possibilistic likely dominance

In 2003, Dubois et al. [29] have developed the axiomatic system (denoted by S_L) of likely dominance rule in the context of Savage decision theory [29].
In fact, this axiomatic system is a relaxed Savage framework augmented by the ordinal invariance axiom. A preference relation $\succeq_{LN}$ or $\succeq_{L\Pi}$ satisfies the following axioms :

- Axiom 1^{S_L}. **Weak pre-order** : $\succeq$ is irreflexive, quasitransitive and complete.
- Axiom 2^{S_L}. **Weak Sure Thing Principle** : $fAh \succ gAh \;\; iff \;\; fAh' \succeq gAh'$.
- Axiom 3^{S_L}. (The third axiom of Savage (Axiom $3^{S_{SEU}}$)).
- Axiom 4^{S_L}. There exist three acts f, g and $h \in F^{const}$ such that $f \succ g \succ h$.
- Axiom 5^{S_l}. **Ordinal Invariance** $\forall f, f', g, g'$ four acts : if (f, g) and (f', g') are state wise order equivalent [1] iff $f' \succeq g'$.

1. Two pairs of acts (f, g) and (f', g') are called state wise order equivalent iff $\forall s \in S, f(s) \geq_p g(s)$ iff $f'(s) \geq_p g'(s)$ s.t $\geq_p$ is a preference relation among constant acts then $f \geq g$.

The Axiom 2^{S_L} is the weak version of the Sure Thing Principle (i.e Axiom $2^{S_{SEU}}$).

Theorem 2.6 *A preference relation $\succeq$ on $\mathcal{L}$ satisfies the axiomatic system S_L iff there exists a utility function such that :*

$$L \succeq L' \ \ iff \ \ LN(L) \geq LN(L'). \tag{2.21}$$

$$L \succeq L' \ \ iff \ \ L\Pi(L) \geq L\Pi(L'). \tag{2.22}$$

2.6 Order of Magnitude Expected Utility ($OMEU$)

Order of Magnitude Expected Utility theory relies on a qualitative representation of beliefs, initially proposed by Spohn [77], via *Ordinal Conditional Functions*, and later popularized under the term *kappa-rankings*. $\kappa : 2^{\Omega} \rightarrow Z^{+} \cup \{+\infty\}$ is a kappa-ranking if and only if :

S1 $\min_{\omega \in \Omega} \kappa(\{\omega\}) = 0$

S2 $\kappa(A) = \min_{\omega \in A} \kappa(\{\omega\})$ if $\emptyset \neq A \subseteq A$, $\kappa(\emptyset) = +\infty$

Note that an event A is more likely than an event B if and only if $\kappa(A) < \kappa(B)$: kappa-rankings have been termed as "disbelief functions". They receive an interpretation in terms of order of magnitude of "small" probabilities. "$\kappa(A) = i$" is equivalent to $P(A)$ is of the same order of ε^i, for a given fixed infinitesimal ε. As pointed out by [22], there exists a close link between kappa-rankings and possibility measures, insofar as any kappa-ranking can be represented by a possibility measure, and vice versa.

Order of magnitude utilities have been defined in the same way [62, 87]. Namely, an order of magnitude function $\mu : X \rightarrow Z^{+} \cup \{+\infty\}$ can be defined in order to rank outcomes $x \in X$ in terms of degrees of "dissatisfaction". Once again, $\mu(x) < \mu(x')$ if and only if x is more desirable than x', $\mu(x) = 0$ for the most desirable consequences, and $\mu(x) = +\infty$ for the least desirable consequences. μ is interpreted as : $\mu(x) = i$ is equivalent to say that the utility of x is of the same order of ε^i, for a given fixed infinitesimal ε. An *order of magnitude expected utility* (OMEU) model can then be defined (see [62, 87] among others). Considering an order of magnitude lottery $L = \langle \kappa_1/\mu_1, \ldots, \kappa_n/\mu_n \rangle$ as representing a some probabilistic lottery, it is possible to compute the order of magnitude of the expected utility of this probabilistic lottery : it is equal to $\min_{i=1,n}\{\kappa_i + \mu_i\}$. Hence the definition of the OMEU value of a κ lottery $L = \langle \kappa_1/\mu_1, \ldots, \kappa_n/\mu_n \rangle$:

Definition 2.5 *The order of magnitude of the expected utility of a lottery L is computed as follows :*

$$OMEU(L) = \min_{i=1,n}\{\kappa_i + \mu_i\}. \tag{2.23}$$

Example 2.8 *Let us consider a two lotteries $L = \langle 1.2/2, 0/4, 3/5, 5/7\rangle$ and $L' = \langle 0/2, 1/4, 3.6/5, 0.5/6\rangle$. Using Equation (2.23) we have :*
$OMEU(L) = min(3.2, 4, 8, 11) = 3.2$ and $OMEU(L') = min(2, 5, 8.6, 6.5) = 2$
so $L' \succ_{OMEU} L$.

According to the interpretation of kappa ranking in terms of order of magnitude of probabilities, the product of infinitesimal the conditional probabilities along the paths lead to a sum of the kappa levels. Hence the following principle of reduction of the kappa lotteries :

$$\begin{aligned}&Reduction(\kappa_1 \wedge L_1 \vee \cdots \vee \kappa_m \wedge L_m)\\ &= \langle \underset{j=1..m}{min}(\kappa_1^j + \kappa_j)/u_1, \ldots, \underset{j=1..m}{min}(\kappa_n^j + \kappa_j)/u_n\rangle\end{aligned} \tag{2.24}$$

Axiomatization of order of magnitude of the expected utility

In [35], Giang and Shenoy have proposed axioms relative to the preference relation w.r.t the OMEU criterion. These axioms are analogous to the ones proposed by von Neumann and Morgenstern and similar to those presented in [50].

The preference relation $\succeq_{OMEU}$ satisfies the following system of axioms denoted by S_{OMEU} :

- Axiom $1^{S_{OMEU}}$. The preference relation between lotteries is complete and transitive.
- Axiom $2^{S_{OMEU}}$. (**Reduction of compound lotteries**) Any compound lottery is indifferent to a simple lottery whose disbelief degrees are calculated according to Spohn's calculus :
  A compound lottery denoted by $L_c = \langle \lambda_1/L_1, \ldots, \lambda_m/L_m\rangle$ where
  $L_i = \langle k_{i1}/\mu_1, \ldots, k_{in}/\mu_n\rangle$ for $1 \leq i \leq k$ is indifferent to the simple lottery $L_s = \langle k_1/\mu_1, \ldots, k_n/\mu_n\rangle$ where : $k_j = min_{1\leq i\leq k}\{\lambda_i + k_{ij}\}$.
- Axiom $3^{S_{OMEU}}$. (**Substitutability**) If $L_i \sim L_{i'}$ then
  $\langle \lambda_1/L_1, \ldots, \lambda_i/L_i, \ldots \lambda_m/L_m\rangle \sim \langle \lambda_1/L_1, \ldots, \lambda_i/L'_i, \ldots \lambda_m/L_m\rangle$.
- Axiom $4^{S_{OMEU}}$. (**Quasi-continuity**) For each utility $u_i \in U$ there exists a qualitative lottery that is indifferent to it.
- Axiom $5^{S_{OMEU}}$. (**Transitivity**) $\forall L_i, L_j, L_k \in \mathcal{L}$ if $L_i \succ L_j$ and $L_j \succ L_k$ then $L_i \succ L_k$.

– Axiom $6^{S_{OMEU}}$. (**Qualitative monotonicity**) Let two standard lotteries $L = \langle k_1/\mu_1, k_2/\mu_2\rangle$ and $L' = \langle k'_1/\mu_1, k'_2/\mu_2\rangle$:

$$L \succ L' \;\; iff \;\; \begin{cases} k_1 = k'_1 = 0 \; and \; k_r > k'_r \; \forall r \neq 1 \\ k_1 = 0 \; and \; k'_1 > 0 \\ k_1 < k'_1 \; and \; k_2 = k'_2 \end{cases} \tag{2.25}$$

Theorem 2.7 *A preference relation $\succeq$ on $\mathcal{L}$ satisfies the system of axioms S_{OMEU} iff :*

$$L \succeq L' \;\; iff \;\; OMEU(L) \geq OMEU(L'). \tag{2.26}$$

2.7 Possibilistic Choquet integrals

Possibilistic Choquet integrals allow the representation of different behaviors of decision makers according to the nature of the fuzzy measure μ in Equation (1.15) defined in Chapter 1.

Indeed Possibility-based Choquet integrals allow to represent behaviors of adventurous possibilistic decision makers by considering the fuzzy measure μ as a possibility measure Π as stated by the following definition :

Definition 2.6 *The possibility-based Choquet integrals of a lottery L (denoted by $Ch_\Pi(L)$) is computed as follows :*

$$Ch_\Pi(L) = \Sigma_{i=n,1}(u_i - u_{i-1}) \; . \; \Pi(L \geq u_i). \tag{2.27}$$

Example 2.9 *Let $L = \langle 1/10, 0.2/20, 0.7/30\rangle$ and $L' = \langle 1/10, 0.1/30\rangle$ be two possibilistic lotteries, we have $Ch_\Pi(L) = 10 + (20-10) * 0.7 + (30-20) * 0.7 = 24$ and $Ch_\Pi(L') = 10 + (30-10) * 0.1 = 12$. So, $L \succ L'$.*

Necessity-based Choquet integrals allow to represent behaviors of cautious possibilistic decision makers by considering the fuzzy measure μ as a necessity measure.

Definition 2.7 *The necessity based Choquet integrals of a lottery L (denoted by $Ch_N(L)$) is computed as follows :*

$$Ch_N(L) = \Sigma_{i=n,1}(u_i - u_{i-1}) \; . \; N(L \geq u_i). \tag{2.28}$$

Example 2.10 *Let $L = \langle 0.3/10, 0.5/20, 1/30\rangle$ and $L' = \langle 1/10, 0.5/20, 0.2/30\rangle$ be two possibilistic lotteries, using Equation (2.28), we have $Ch_N(L) = 10 + (20-10)*(1-0.3) + (30-20)*(1-0.5) = 22$ and $Ch_N(L') = 10 + (20-10)*(1-1) = 10$. So, $L \succ L'$.*

2.7.1 Axiomatization of possibilistic Choquet integrals

The key axiom of Choquet expected utility is based on the notion of *comonotony* (the terminology of comonotony comes from common monotony). Formally, we say that two acts f and g of F are *comonotonic* if there exists no pair s and s' of S such that :

$$f(s) \succ f(s') \;\; and \;\; g(s) \prec g(s').$$

Note that any positive linear combination of two comonotonic acts preserves the initial order between these acts. Basing on this property of comonotonic acts, [38] and [69] have proposed the *comonotonic sure thing principle*. The axiomatic system of Choquet Expected Utility denoted by S_{Ch} in the style of Savage contains the following axioms :

- Axiom $1^{S_{Ch}}$. **(Weak order)** : the preference relation $\succeq$ is a weak order.
- Axiom $2^{S_{Ch}}$. **(Continuity)** : $\forall f, g, h \in F$, If $f \succeq g$ and $g \succeq h$ then there exist α and $\beta \in [0,1]$ such that $\alpha f + (1-\alpha)h \succeq g$ and $g \succeq \beta f + (1-\beta)h$.
- Axiom $3^{S_{Ch}}$. **(Comonotonic sure thing principle)** : Let f and g be two acts of F and A_i with $(i = 1 \dots n)$ a partition on S. We have $f = (x_1, A_1; \dots; x_k, A_k; \dots; x_n, A_n)$ and $g = (y_1, A_1; \dots; y_k, A_k; \dots; y_n, A_n)$ such that $(x_1 \leq \dots \leq x_k \leq \dots \leq x_n)$ and $(y_1 \leq \dots \leq y_k \leq \dots \leq y_n)$ and $\exists i (i = 1 \dots n)$ such that $x_i = y_i = u_c$ then :

  $$f \succ g \Longleftrightarrow f' \succ g'.$$

  f' and g' are two acts obtained from f and g by replacing the common utility u_c by a new value that guarantee the ascending order of x_i and y_i.

Theorem 2.8 *A preference relation $\succeq$ on $\mathcal{L}$ satisfies the axiomatic system S_{Ch} iff there exists a utility function such that :*

$$L \succeq L' \;\; iff \;\; Ch_N(L) \geq Ch_N(L'). \tag{2.29}$$

and

$$L \succeq L' \;\; iff \;\; Ch_\Pi(L) \geq Ch_\Pi(L'). \tag{2.30}$$

In 2006, *Rébillé* has provided an axiomatization of a preference relation of a decision maker that ranks necessity measures according to their Choquet's expected utilities [66]. This axiomatic system has been developed under risk in a similar way than the one of Von Neumann and Morgenstern's approach [57]. Nevertheless, in its axiomatic system *Rébillé* proposed a linear mixture of possibilistic lotteries by probability degrees which is not allowed in our work since we use only possibility degrees to model uncertainty.

2.7.2 Properties of possibilistic Choquet integrals

We propose now some additional properties of possibilistic Choquet integrals that are particularly useful to study the behavior of this decision criteria in sequential decision making.

Proposition 2.1 *Given a lottery* $L = \langle \lambda_1/u_1, \ldots, \lambda_n/u_n \rangle$, *an utility* u_i *s.t.* $u_i \leq \max_{u_j \in L, \lambda_j > 0} u_j$ *and a lottery* $L' = \langle \lambda'_1/u_1, \ldots, \lambda'_n/u_n \rangle$ *s.t.* $\lambda'_i \geq \lambda_i$ *and* $\forall j \neq i, \lambda'_j = \lambda_j$, *it holds that* $Ch_N(L') \leq Ch_N(L)$.

We provide the proof of Proposition 2.1 in the ordinal setting.

Proof. [Proof of Proposition 2.1]
We suppose without loss of generality that the u_j are ranked by increasing order, i.e. that $j < l \;\; iff \;\; u_j < u_l$.
Let n be the index of the greater u_k such that $\lambda_k > 0$. Hence there is a λ_j, $j \leq k$ such that $\lambda_j = 1$ and thus $1 - max(\lambda_1, \ldots, \lambda_j) = 0$ for any $j > k$.
Since $u_i \leq max_{u_j \in L, \lambda_j > 0} u_j$, $i \leq k$, L and L' are written as follows :
$L = \langle \lambda_1/u_1, \ldots, \lambda_{i-1}/u_{i-1}, \lambda_i/u_i, \lambda_{i+1}/u_{i+1}, \ldots, \lambda_n/u_n \rangle$
$L' = \langle \lambda_1/u_1, \ldots, \lambda_{i-1}/u_{i-1}, \lambda' i/u_i, \lambda_{i+1}/u_{i+1}, \ldots, \lambda_n/u_n \rangle$

$Ch_N(L)$ can be decomposed in 3 terms V_1, V_2, V_3, i.e. $Ch_N(L) = V_1 + V_2 + V_3$ where :
$V_1 = u_1 + (u_2 - u_1)(1 - \lambda_1) + \cdots + (u_i - u_{i-1})(1 - max(\lambda_1, \ldots, \lambda_{i-1}))$
$V_2 = (u_{i+1} - u_i)(1 - max(\lambda_1, \ldots, \lambda_i)) + (u_{i+2} - u_{i+1})(1 - max(\lambda_1, \ldots, \lambda_i, \lambda_{i+1})) + \cdots + (u_k - u_{k-1})(1 - max(\lambda_1, \ldots, \lambda_i, \lambda_{i+1}, \ldots, \lambda_{k-1}))$
$V_3 = (u_{k+1} - u_k)(1 - max(\lambda_1, \ldots, \lambda_k)) + \cdots + (u_n - u_{n-1})(1 - max(\lambda_1, \ldots, \lambda_n))$
Since $(1 - max(\lambda_1, \ldots, \lambda_j) = 0$ for any $j > k$, $V_3 = 0$: $Ch_N(L) = V_1 + V_2$.
$Ch_N(L')$ can also be decomposed into 3 terms V'_1, V'_2, V'_3, i.e.
$Ch_N(L') = V'_1 + V'_2 + V'_3$ where :

$V'_1 = u_1 + (u_2 - u_1)(1 - \lambda_1) + \cdots + (u_i - u_{i-1})(1 - max(\lambda_1, \ldots, \lambda_{i-1})) = V_1$
$V'_2 = (u_{i+1} - u_i)(1 - max(\lambda_1, \ldots, \lambda_{i-1}, \lambda'_i)) + (u_{i+2} - u_{i+1})(1 - max(\lambda_1, \ldots, \lambda_{i-1}, \lambda'_i, \lambda_{i+1})) + \cdots + (u_k - u_{k-1})(1 - max(\lambda_1, \ldots, \lambda_{i-1}, \lambda'_i, \lambda_{i+1}, \ldots, \lambda_{k-1}))$
$V'_3 = (u_{k+1} - u_k)(1 - max(\lambda_1, \ldots, \lambda_k)) + \cdots + (u_n - u_{n-1})(1 - max(\lambda_1, \ldots, \lambda_n)) = V_3 = 0$
As a consequence, it holds that : $Ch_N(L) - Ch_N(L') = V_2 - V'_2$.
Since $\lambda'_i \geq \lambda_i$, $1 - max(\lambda_1, \ldots, \lambda_{i-1}, \lambda'_i, \ldots, \lambda_j)$ is lower than $1 - max(\lambda_1, \ldots, \lambda_{i-1}, \lambda_i, \ldots, \lambda_j)$, for any j.
Thus $V_2 \geq V'_2$ and $Ch_N(L) \geq Ch_N(L')$. ■

Example 2.11 *Let $L = \langle 1/10, 0.2/20, 0.5/30 \rangle$ and $L' = \langle 0.2/5, 1/10, 0.4/20, 0.1/35 \rangle$ be two possibilistic lotteries such that $\max_{u_i \in L} = 30$, $L(10) = L'(10) = 1$ and $L(20) = 0.2 < L'(20) = 0.4$. We have $Ch_N(L) = 10$ and $Ch_N(L') = 9$.*

This emphasizes the pessimistic character of Ch_N : adding to a lottery any consequence that is not better that its best one decreases its evaluation.

Note that Proposition 2.1 is invalid for possibility-based Choquet integrals as it is shown in the following counter example :

Counter Example 2.1 *Let $U = \{10, 20, 30\}$, $L = \langle 1/10, 0.5/20, 0.2/30 \rangle$ and $L' = \langle 1/10, 0.8/20, 0.2/30 \rangle$. Using Equation (2.27), we have $Ch_\Pi(L) = 10 + (20-10)*0.5 + (30-20)*0.2 = 17$ and $Ch_\Pi(L') = 10 + (20-10)*0.8 + (30-20)*0.2 = 20$ so even necessary conditions in the proposition 2.1 are verified in L and L' we have $Ch_\Pi(L') > Ch_\Pi(L)$.*

As a consequence of the Proposition 2.1, we get the following result :

Proposition 2.2 *Let L_1, L_2 be two lotteries such that $\max_{u_i \in L_2, \lambda_i > 0} u_i \leq \max_{u_i \in L_1, \lambda_i > 0} u_i$. It holds that :*

$$Ch_N(Reduction(\langle 1/L_1, 1/L_2 \rangle)) \leq Ch_N(L_1).$$

Proof. [Proof of Proposition 2.2]
We provide the proof of Proposition 2.2 in the ordinal setting.
Let L_1, L_2 be two lotteries such that $max_{u_i \in L_2, \lambda_i > 0} u_i \leq max_{u_i \in L_1, \lambda_i > 0} u_i$.

Let $L = Reduction(\langle 1/L_1, 1/L_2\rangle)$. From the definition of the reduction (Equation 4.2), it hold that $\lambda_i = \max(\min(1, \lambda_1^j), \min(1, \lambda_2^j)) = \max(\lambda_1^j, \lambda_2^j), \forall i$.

Since $max_{u_j \in L_2, \lambda_j^2 > 0} u_j \leq max_{u_j \in L_1, \lambda_j^1 > 0} u_j$, we can get L from L_1 by increasing each λ_j^1 to the value $\max(\lambda_j^1, \lambda_j^2)$, for any j such that $\lambda_j^1 < \lambda_j^2$. According to proposition 2.1, this is done without increasing the value of the Choquet integral of L, hence $Ch_N(L) \leq Ch_N(L_1)$.
Formally, $L^0 = L_1$, then for $j = 1 \dots n$,
$L^j = \langle \lambda_1^j/u_1, \dots, \lambda_n^j/u_n\rangle$ such that for any $k \neq j$, $\lambda_k^j = \lambda_k^{j-1}$ and $\lambda_j^j = \max(\lambda_j^{j-1}, \lambda_2^{j-1})$.
By construction, $L^n = L$. Thanks to Proposition 2.1, $Ch_N(L^j) \leq Ch_N(L^{j-1}), j = 1, n$.
Then $Ch_N(L) \leq Ch_N(L_1)$. ■

Example 2.12 *Let* $L_1 = \langle 0.2/5, 1/10, 0.4/20, 0.1/35\rangle$ *and*
$L_2 = \langle 1/10, 0.2/20, 0.5/30\rangle$ *be two possibilistic lotteries.*
We can check that $\underset{u_i \in L_2, \lambda_i > 0}{max} u_i = 30 \leq \underset{u_i \in L_1, \lambda_i > 0}{max} u_i = 35$.
We have $Reduction(\langle 1/L_1, 1/L_2\rangle) = \langle 0.2/5, 1/10, 0.4/20, 0.5/30, 0.1/35\rangle$
$\Rightarrow Ch_N(\langle 0.2/5, 1/10, 0.4/20, 0.5/30, 0.1/35\rangle) = 9 = Ch_N(L_1)$.

No such property holds for Ch_Π, as shown by the following counter example :

Counter Example 2.2 *Let* $L_1 = \langle 0.2/0, 1/2, 0.5/9\rangle$ *and* $L_2 = \langle 0.4/4, 1/7\rangle$ *be two possibilistic lotteries, we can check that* $\underset{u_i \in L_2, \lambda_i > 0}{max} u_i = 7 \leq \underset{u_i \in L_1, \lambda_i > 0}{max} u_i = 9$.
We have $Reduction(\langle 1/L_1, 1/L_2\rangle) = \langle 0.2/0, 1/2, 0.4/4, 1/7, 0.5/9\rangle$
$\Rightarrow Ch_\Pi(\langle 0.2/0, 1/2, 0.4/4, 1/7, 0.5/9\rangle) = 8$. *Moreover,* $Ch_\Pi(L_1) = 5.5$ *which contradicts the Proposition 2.2.*

It is simple to verify that the Proposition 2.1 and 2.2 are valid for ordinal and numerical settings of possibility theory. This validity is due to the fact that $\forall \lambda_i, (\lambda_i * 1) = \lambda_i$ and $\min(1, \lambda_i) = \lambda_i$.

2.8 Software for possibilistic decision making

In this section, we propose a software implementing possibilistic decision criteria studied in this chapter. This software, implemented with Matlab 7.10, allows the construction of possibilistic lotteries and their reduction (in the case of qualitative and numerical possibilitic setting). Using this software, possibilistic lotteries can be compared w.r.t any possibilistic decision criteria i.e. U_{pes}, U_{opt}, PU, LN, $L\Pi$, $OMEU$, Ch_N and Ch_Π.

Figure 2.1 is relative to the main menu allowing the construction of possibilistic lotteries and the qualitative and numerical reduction of a possibilistic compound lottery.

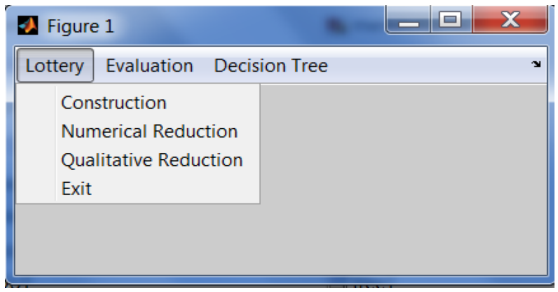

FIGURE 2.1 – Possibilistic lottery

For instance, Figure 2.2 is relative to the qualitative reduction of a compound lottery.

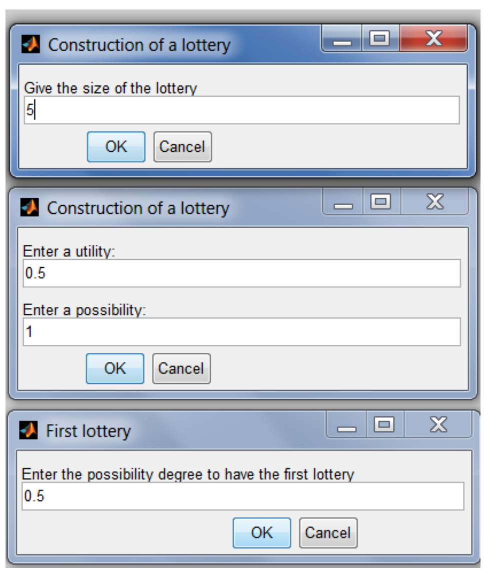

FIGURE 2.2 – Reduction of compound lottery

The reduced lottery is displayed in a table with two columns : the first one for utilities and the second one for possibilities as shown in Figure 2.3.

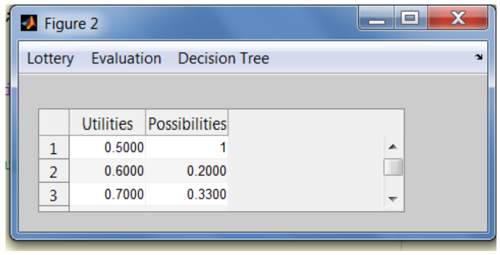

FIGURE 2.3 – The reduced lottery

Once a possibilistic lottery is constructed, its value is computed according to any possibilistic decision criterion as it is shown in Figure 2.4. Then, any two possibilistic lotteries can be compared w.r.t any decision criterion studied in this chapter. For instance, Figure 2.5 presents the result of comparison of two possibilistic lotteries w.r.t possibility-based Choquet integrals.

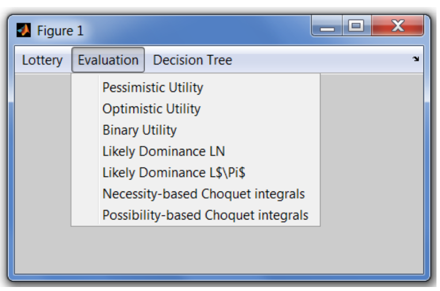

FIGURE 2.4 – Possibilistic decision criteria

2.9 Conclusion

In this chapter, we have presented a survey on possibilistic decision theory which overcomes some weakness lied to the use of probability theory to model uncertainty in classical

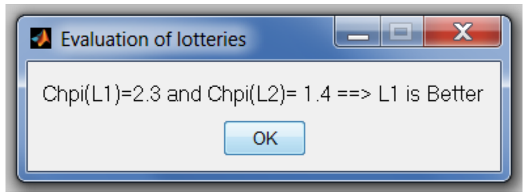

FIGURE 2.5 – Comparison of two lotteries

decision theories. We especially focus on main possibilistic decision criteria with their axiomatization in the style of VNM and Savage. We also detailed some properties of possibilistic Choquet integrals.

Deuxième partie

Graphical decision models under uncertainty

Chapitre 3

Graphical Decision Models

3.1 Introduction

In the first part of this thesis, we have been interested by one-stage decision making. In multi-stage decision making (also called sequential decision making), several actions (decisions) should be executed successively. The consequence of an action executed at step t will be the state of nature in step $t+1$. A *strategy* (a policy) is a function that assigns a decision to each state.

Several graphical models can be used for sequential decision making, such as *decision trees* [65], *influence diagrams* [43], *valuation based systems* [73], etc. These tools offer a direct or a compact representation of sequential decision problems and they represent intuitive and simple tools to deal with decision problems of greater size.

Given a sequential decision problem, the question is how to find a strategy that is optimal w.r.t a decision criterion. Depending on the graphical models, different algorithms have been proposed :

- *Dynamic programming* was initially introduced by Richard Bellman in 1940. The main contribution of Bellman is that he sets the optimization problem in a recursive form [2]. The method proposed by Bellman is the *backward induction method* that consists in handling the problem from the end (in time), so the last decisions are first considered then the process follows backwards in time until the first decision step.
- *Resolute choice* was introduced by McClennen in 1990. The resolute choice behavior must be adopted by decision makers using non expected utility criteria [54]. According to McClennen [54] : *"The theory of resolute choice is predicated on the notion that the single agent who is faced with making decisions over time can achieve a cooperative arrangement between his present self and his relevant future selves that satisfies the principle of intra personal optimality."*

This chapter is organized as follows : in Section 3.2, probabilistic decision trees which are the oldest decision graphical models will be developed. Then, influence diagrams will be presented in Section 3.3. Both of these decision formalisms will be presented with their evaluation algorithms.

3.2 Decision trees

Decision trees proposed by Raiffa in 1968 [65] are the pioneer of graphical decision models. They allow a direct modeling of sequential decision problems by representing in a simple graphical way all possible scenarios.

Decision trees were used in several real world applications, we can for illustration mention :

- Health : in the department of physical medicine and rehabilitation of Wayne state university school of medicine (USA), decision trees have been used to identify potential mental health problems and to guide decision making for referrals [52].
- Environment : Gerber products, the well known baby products company, have used decision trees to decide whether to continue using one kind of plastic or not according to the opinion of several organizations such as the environmental group, the consumer products safety commission [12].
- Energy : Energy star which is a joint program of the U.S. environmental protection agency and the U.S. department of energy. Decision trees have been used to improve the quality, the reliability and speed decisions in the domain of energy.

3.2.1 Definition of decision trees

A decision tree is composed of a graphical component and a numerical one as detailed below.

> Graphical component

A decision tree is a tree $\mathcal{T} = (\mathcal{N}, \mathcal{E})$ which has a numerical part. The set of nodes $\mathcal{N}$ contains three kinds of nodes :

- $\mathcal{D} = \{D_0, \dots, D_m\}$ is the set of decision nodes (represented by rectangles). The labeling of the nodes is supposed to be in accordance with the temporal order i.e. if D_i is a descendant of D_j, then $i > j$. Generally, the root node of the tree is a decision node, denoted by D_0.
- $\mathcal{LN} = \{LN_1, \dots, LN_k\}$ is the set of leaves, also called utility leaves : $\forall LN_i \in \mathcal{LN}$, $u(LN_i)$ is the utility of being eventually in node LN_i. For the sake of simplicity we assume that only leave nodes lead to utilities.
- $\mathcal{C} = \{C_1, \dots, C_n\}$ is the set of chance nodes represented by circles.

For any $N_i \in \mathcal{N}$, $Succ(N_i) \subseteq \mathcal{N}$ denotes the set of its children. Moreover, for any $D_i \in \mathcal{D}, Succ(D_i) \subseteq \mathcal{C} : Succ(D_i)$ is the set of actions that can be decided when D_i is observed. For any $C_i \in \mathcal{C}, Succ(C_i) \subseteq \mathcal{LN} \cup \mathcal{D} : Succ(C_i)$ is indeed the set of outcomes of the action C_i - either a leaf node is observed, or a decision node is observed (and then a new action should be executed).

The size $|\mathcal{T}|$ of a decision tree is its number of edges which is equal to the number of its nodes minus 1.

> **Numerical component**

The numerical component of decision trees valuates the edges outgoing from chance nodes and assigns utilities to leaves nodes.

In classical probabilistic decision trees the uncertainty pertaining to the possible outcomes of each $C_i \in \mathcal{C}$ is represented by a conditional probability distribution p_i on $Succ(C_i)$, such that $\forall N_i \in Succ(C_i), p_i(N_i) = P(N_i|path(C_i))$ where $path(C_i)$ denotes all the value assignments to chance and decision nodes on the path from the root node to C_i. To each chance node $C_i \in \mathcal{C}$ we can associate a probabilistic lottery L_{C_i} relative to its outcomes.

Example 3.1 *The decision tree of Figure 3.1 is defined by* $D = \{D_0, D_1, D_2\}$, $\mathcal{C} = \{C_1, C_2, C_3, C_4, C_5, C_6\}$ *and* $\mathcal{LN} = U = \{0, 1, 2, 3, 4, 5\}$. *Corresponding lotteries to chance nodes are* $L_{C_1} = \langle 0.6/L_{D_1}, 0.4/L_{D_2}\rangle$, $L_{C_2} = \langle 0.3/1, 0.7/2\rangle$, $L_{C_3} = \langle 1/1, 0/5\rangle$, $L_{C_4} = \langle 0.2/0, 0.8/4\rangle$, $L_{C_5} = \langle 0.4/1, 0.6/4\rangle$ *and* $L_{C_6} = \langle 0.5/2, 0.5/5\rangle$.

3.2.2 Evaluation of decision trees

A decision tree is considered as a finite set of strategies. Formally, we define a strategy as a function δ from $\mathcal{D}$ to $\mathcal{C} \cup \{\perp\}$. $\delta(D_i)$ is the action to be executed when a decision node D_i is observed. $\delta(D_i) = \perp$ means that no action has been selected for D_i (because either D_i cannot be reached or the strategy is partially defined). Admissible strategies must be :
- *sound* : $\forall D_i \in \mathcal{D}, \delta(D_i) \in Succ(D_i) \cup \{\perp\}$.
- *complete* : (i) $\delta(D_0) \neq \perp$ and (ii) $\forall D_i$ s.t. $\delta(D_i) \neq \perp, \forall N \in Succ(\delta(D_i))$, either $\delta(N) \neq \perp$ or $N \in \mathcal{LN}$.

Let Δ be the set of sound and complete strategies that can be built from the decision tree, then any strategy δ in Δ can be view as a connected subtree of the decision tree whose arcs are of the form $(D_i, \delta(D_i))$.

Evaluating a decision tree consists in finding the optimal strategy δ^* within Δ w.r.t a decision criterion O. Formally, $\forall \delta_i \in \Delta$ we have $\delta^* \succeq_O \delta_i$ (i.e. δ^* is preffered to any strategy

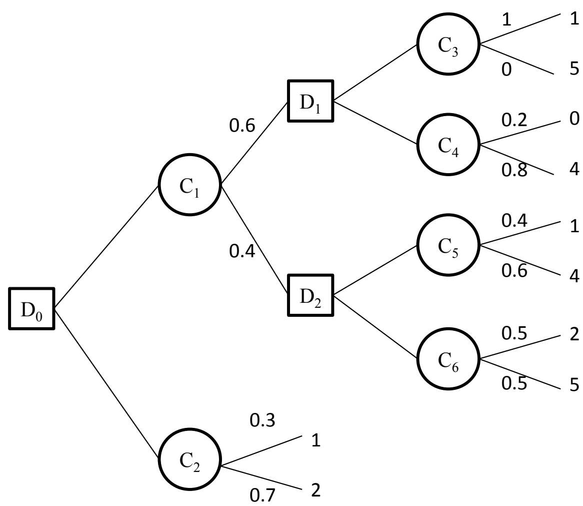

FIGURE 3.1 – Example of a probabilistic decision tree

$\delta_i \in \Delta$ w.r.t a decision criterion O). In probabilistic decision trees, the decision criterion O corresponds to the expected utility EU (see Chapter 1).

The size $|\delta|$ of a strategy δ is the sum of its number of nodes and edges, it is obviously lower than the size of the decision tree.

Strategies can be evaluated and compared thanks to the notion of lottery reduction. Recall indeed that leaf nodes in $\mathcal{LN}$ are labeled with utility degrees. Then a chance node can be seen as a simple probabilistic lottery (for the most right chance nodes) or as a compound lottery (for the inner chance nodes). Each strategy δ_i is a compound lottery L_i and can be reduced to an equivalent simple one. Formally, the composition of lotteries will be applied from the leafs of the strategy to its root, according to the following recursive definition for any node $N_i \in \mathcal{N}$:

$$L(N_i, \delta) = \begin{cases} L(\delta(N_i), \delta) \ \ if \ \ N_i \in \mathcal{D} \\ Reduction(< pr_i(X_j)/L(X_j, \delta)_{X_j \in Succ(N_i)} >) \ if \ N_i \in \mathcal{C} \\ < 1/u(N_i) > \ \ \ if \ \ N_i \in \mathcal{LN} \end{cases} \tag{3.1}$$

Equation (3.1) is simply the adaptation of lottery reduction to strategies, we can then compute $Reduction(\delta) = L(D_0, \delta)$: $Reduction(\delta)(u_i)$ that is simply the probability of getting utility u_i when δ is applied from D_0.

Example 3.2 *Let us evaluate the decision tree in Figure 3.1 using the expected utility criterion EU. As shown in Table* 3.1*, we can distinguish* 5 *possible strategies* *($\Delta = \{\delta_1, \delta_2, \delta_3, \delta_4, \delta_5\}$) where each strategy δ_i is characterized by a lottery L_i :*

δ_i	L_i	$EU(L_i)$
$\delta_1 = \{(D_0, C_1), (D_1, C_3), (D_2, C_5)\}$	$\langle 0.76/1, 0.24/4, 0/5 \rangle$	1.72
$\delta_2 = \{(D_0, C_1), (D_1, C_3), (D_2, C_6)\}$	$\langle 0.6/1, 0.2/2, 0.2/5 \rangle$	2
$\delta_3 = \{(D_0, C_1), (D_1, C_4), (D_2, C_5)\}$	$\langle 0.12/0, 0.16/1, 0.72/4 \rangle$	3.04
$\delta_4 = \{(D_0, C_1), (D_1, C_4), (D_2, C_6)\}$	$\langle 0.12/0, 0.2/2, 0.48/4, 0.2/5 \rangle$	**3.32**
$\delta_5 = \{(D_0, C_2)\}$	$\langle 0.3/1, 0.7/2 \rangle$	1.7

TABLE 3.1 – Exhaustive enumeration of possible strategies in Figure 3.1

From Table 3.1*, we can see that the optimal strategy in this decision tree is $\delta^* = \delta_4$ with $EU(\delta^*) = 3.32$ corresponding to bold lines in Figure 3.2.*

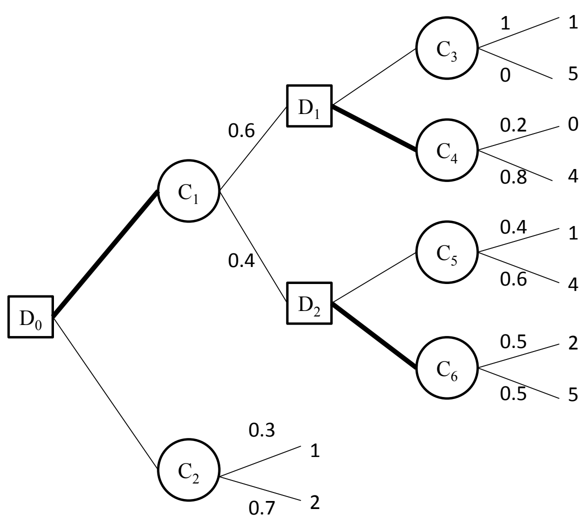

FIGURE 3.2 – The optimal strategy $\delta^* = \{(D_0, C_1), (D_1, C_4), (D_2, C_6)\}$

The number of potential strategies in a probabilistic decision tree is in $O(2^{\sqrt{n}})$ as we will prove in the next chapter (Proof of Proposition 5). Given the large number of strategies in the decision tree, an exhaustive enumeration of all possible strategies to find the best one is intractable. As an alternative method, Bellman proposed a recursive method of dynamic programming called *backward search method* or *backward induction method* [2].

It is important to note that dynamic programming can be applied only when the crucial property of *monotonicity* or *weak monotonicity* is satisfied by the decision criterion which is the *EU* criterion. This property states that if a probabilistic lottery L is preferred to the lottery L' w.r.t a decision criterion O then the compound lottery $\langle \alpha/L, (1-\alpha)/L" \rangle$ is preferred to $\langle \alpha/L', (1-\alpha)/L" \rangle$ w.r.t O ($\alpha \in [0,1]$ and $L"$ is a probabilistic lottery). This property will be deeply studied in the next chapter.

The principle of backwards reasoning procedure (called ProgDyn) is depicted in a recursive manner by Algorithm 3.1. When each chance node is reached, an optimal sub-strategy is build for each of its children - these sub-strategies are combined w.r.t. their probability degrees, and the resulting compound lottery (corresponding to the compound strategy) is reduced : we get an equivalent simple lottery, representing the current optimal sub-strategy. When a decision node X is reached, a decision Y^* leading to a sub-strategy optimal w.r.t EU is selected among all the possible decisions $Y \in Succ(X)$, by comparing the simple lotteries equivalent to each sub strategy.

Note that $L[u_i]$ is the probability degree to have the utility u_i in the lottery L and $Succ(N).first$ is the first node in the set of successors $Succ(N)$.

Clearly, Algorithm 3.1 crosses each edge in the tree only once. When the comparison of simple lotteries (Line (2)) and the reduction operation on a 2-level lottery (Line (1)) can be performed in polytime, its complexity is polynomial w.r.t the size of the tree.

Example 3.3 *Let us reconsider the decision tree in the example 3.1. Principal steps for the evaluation of this decision tree using the dynamic programming function (Algorithm 3.1) are detailed in what follows :*

- *Initially, we have $\delta = \emptyset$ and $N = D_0$ with $succ(D_0) = \{C_1, C_2\}$.*
- *For $Y = C_1$, $L_{C_1} = ProgDyn(C_1, \delta)$ since $succ(C_1) = \{D_1, D_2\}$ we have $Y = D_1$ and $Y = D_2$.*
- *For $Y = D_1$, we have $L_{D_1} = ProgDyn(D_1, \delta)$ and $succ(D_1) = \{C_3, C_4\}$:*
  1. *If $Y = C_3$ then $L_{C_3} = \langle 0/0, 1/1, 0/2, 0/3, 0/4, 0/5 \rangle$ and $EU(L_{C_3}) = 1$.*

Algorithm 3.1: ProgDyn

```
Data: In : a node X, In/Out : a strategy δ
Result: A lottery L
begin
    for i ∈ {1,...,n} do L[u_i] ← 0
    if N ∈ LN then L[u(N)] ← 1
    if N ∈ C then
        % Reduce the compound lottery
        foreach Y ∈ Succ(N) do
            L_Y ← ProgDyn(Y, δ)
            for i ∈ {1,...,n} do
                L[u_i] ← max(L[u_i], (λ_N(Y) * L_Y[u_i])) (Line (1))
        if N ∈ D then
            % Choose the best decision
            Y* ← Succ(N).first
            foreach Y ∈ Succ(N) do
                L_Y ← ProgDyn(Y, δ)
                if EU(L_Y) > EU(L_Y*) then Y* ← Y (Line (2))
            δ(N) ← Y*
            L ← L_Y*
        return L
end
```

2. *If* $Y = C_4$ *then* $L_{C_4} = \langle 0.2/0, 1/1, 0/2, 0/3, 0.8/4, 0/5\rangle$ *and* $EU(L_{C_4}) = 3.2$*. Since* $EU(L_{C_4}) > EU(L_{C_3})$*, so* $Y^* = C_4$*,* $\delta(D_1) = C_4$ *and* $L_{D_1} = \langle 0.2/0, 1/1, 0/2, 0/3, 0.8/4, 0/5\rangle$.

– *For* $Y = D_2$*, we have* $L_{D_2} = ProgDyn(D_2, \delta)$ *and* $succ(D_2) = \{C_5, C_6\}$ *:*

1. *If* $Y = C_5$ *then* $L_{C_5} = \langle 0/0, 0.4/1, 0/2, 0/3, 0.6/4, 0/5\rangle$ *and* $EU(L_{C_5}) = 2.8$.
2. *If* $Y = C_6$ *then* $L_{C_6} = \langle 0/0, 0/1, 0.5/2, 0/3, 0/4, 0.5/5\rangle$ *and* $EU(L_{C_6}) = 3.5$*. Since* $EU(L_{C_6}) > EU(L_{C_5})$*, so* $Y^* = C_6$*,* $\delta(D_2) = C_6$ *and* $L_{D_2} = \langle 0/0, 1/1, 0.5/2, 0/3, 0/4, 0.5/5\rangle$.

$\Rightarrow L_{C_1} = \langle 0.6/L_{D_1}, 0.4/L_{D_2}\rangle = \langle 0.12/0, 0/1, 0.2/2, 0/3, 0.48/4, 0.2/5\rangle$ *and* $EU(L_{C_1}) = 3.32$.

– *For* $Y = C_2$*,* $L_{C_2} = ProgDyn(C_2, \delta)$ *we have :*
$L_{C_2} = \langle 0/0, 0.3/1, 0.7/2, 0/3, 0/4, 0/5\rangle$ *and* $EU(L_{C_2}) = 1.7$.
$\Rightarrow EU(L_{C_1}) > EU(L_{C_2})$*, so* $Y^* = C_1$*,* $\delta(D_0) = C_1$ *and*
$\delta^* = \{(D_0, C_1), (D_1, C_4), (D_2, C_6)\}$ *with* $EU(\delta^*) = 3.32$ *(see this optimal strategy in*

Figure 3.3).
Obviously, the value of $EU(\delta^*)$ obtained by dynamic programming is equal to the one obtained by exhaustive enumeration in example 3.2.

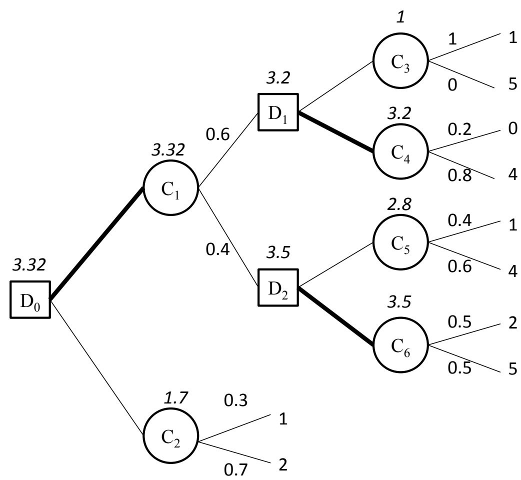

FIGURE 3.3 – The optimal strategy $\delta^* = \{(D_0, C_1), (D_1, C_4), (D_2, C_6)\}$ using dynamic programing

3.3 Influence diagrams

Despite its popularity, decision trees have some limits since they are not appropriate in the case of huge decision problems.

Influence diagrams (IDs) were proposed by Howard and Matheson in 1981 [43] as an alternative to decision trees since they represent a compact graphical model to represent decision maker's belief and preferences about a sequence of decisions to be made under probabilistic uncertainty without a real restriction on its forms.

Influence diagrams are used in several real applications in an efficient manner :
- Automated extraction : in [55], a new methodology for automated extraction of the optimal pathways from IDs has been developed in order to help specialists to relate all available pieces of evidence and consequences of choices.
- Medical diagnosis sector : A new technique for improving medical diagnosis for cancer patients has been proposed in [3, 40].

– Financial sector : IDs were applied in the investment domain in order to allows to the investors to construct optimal investment portfolios [82].
– Web semantic : [56] developed a personalized retrieval model based on influence diagrams that aims to integrate the user profile in the retrieval process. [83] suggested a framework for assessing interoperability on the systems communicating over the semantic web using influence diagrams.

3.3.1 Definition of influence diagrams

As decision trees, influence diagrams are composed of a graphical component and a numerical one.

$>$ **Graphical component**

The graphical component (or qualitative component) is a directed acyclic graph (DAG) denoted by $G = (N, A)$ where A is a set of arcs in the graph and N is a set of nodes partitioned into three subsets C, D and V such that :
– $D = \{D_1, \ldots, D_m\}$ is a set of decision nodes which depict decision and have a temporal order, namely the first decision to make must precede all other decision nodes and the last decision should not be followed by any other decision. Decision nodes are represented by rectangles.
– $C = \{C_1, \ldots, C_n\}$ is a set of chance nodes which represent relevant uncertain factors for decision problem. Chance nodes are represented by circles. The set of chance nodes C is partitioned into three subsets [47] :

1. SC_0 is the set of chance nodes observed prior to any decision.
2. SC_i is the set of chance nodes observed after D_i that is taken and before that the decision D_{i+1} is taken.
3. SC_m is the set of chance nodes never observed or observed too late to have an impact on any decision (i.e. observed after the decision D_m). We have :

$$SC_0 \prec D_1 \prec SC_1 \prec \cdots \prec SC_{m-1} \prec D_m \prec SC_m$$

– $V = \{V_1, \ldots, V_k\}$ is a set of value nodes which represent utilities to be maximized, they are represented by lozenges.
In what follows, we use the same notation for nodes of the influence diagram and variables of the decision problem represented by this influence diagram e.g. the variable represented

by the node C_i is also denoted by C_i. Moreover, c_{ij} (resp. d_{ij}, v_{ij}) denotes the j^{th} value of the variable C_i (resp. D_i, V_i).

The set of arcs A contains two kinds of arcs (see Figure 3.4).

– *conditional arcs* have as target chance or value nodes (first, second and fourth type of arc in Figure 3.4). Only conditional arcs that have as target chance nodes represent probabilistic dependencies.

– *informational arcs* have as target decision nodes and they imply time precedence (third and fifth type of arc in Figure 3.4).

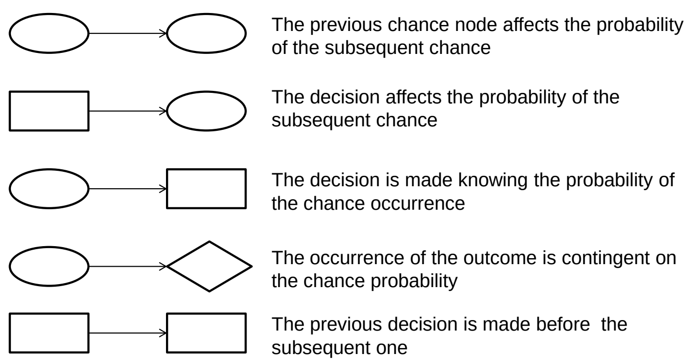

FIGURE 3.4 – Types of arcs in an influence diagram

Example 3.4 *Let us state a simple decision problem of Medical Diagnosis [74] : A physician is trying to decide on a policy for treating patients suspected of suffering from a disease D. D causes a pathological state P that in turn causes a symptom S to be exhibited. The physician first observes whether or not a patient is exhibiting symptom S. Based on this observation, he either treats the patient (for D and P) or not. Physician's utility function depends on his decision to treat (Tr) or not, the presence or absence of the disease D and of the pathological state P. Figure 3.5 presents an influence diagram for the medical diagnosis problem, it contains three chance nodes (S, P and D), one decision node (Tr) and one value node (V). Only the arc that has as target the decision node Tr is an informational arc.*

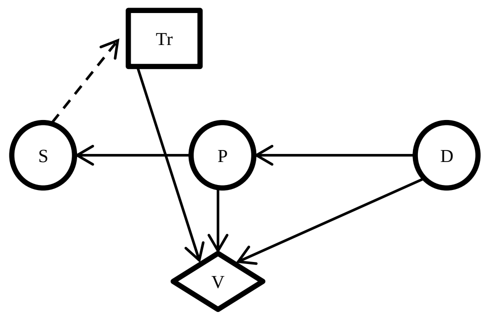

FIGURE 3.5 – The graphical component of the influence diagram for the medical diagnosis problem

The graphical component of an ID encodes different conditional independences between chance nodes [47]. More precisely, a chance node C_i depends only on chance nodes belonging to their parents (the set of parent of C_i is denoted by $Pa(C_i)$).

> **Numerical component**

The numerical component (or quantitative component) of IDs evaluates the different links in the graph. Namely, each conditional arc which has as target a chance node C_i is quantified by a conditional probability distribution of C_i in the context of its parents (denoted by $Pa(C_i)$). Such conditional probabilities should respect the following normalization constraints :

– If $Pa(C_i) = \emptyset$ (C_i is a root) then the a priori probability relative to C_i should satisfy :

$$\sum_{c_{ij} \in \Omega_{C_i}} P(c_{ij}) = 1 \tag{3.2}$$

where : Ω_{C_i} is the domain of C_i.

– If $Pa(C_i) \neq \emptyset$, then the relative conditional probability relative to C_i in the context of any instance $pa(C_i)$ of its parents $Pa(C_i)$ should satisfy :

$$\sum_{c_{ij} \in \Omega_{C_i}} P(c_{ij} \mid pa(C_i)) = 1. \tag{3.3}$$

Chance nodes represent uncertain variables characterizing decision problem. Each decision's alternative may have several consequences according to random variables. The set of consequences is characterized by a utility function. In IDs, consequences are represented by different combinations of value node's parents. So, each value node $V_i \in V$ is characterized by a utility function in the context of its parents that assigns a numerical utility to each instantiation $pa(V_i)$ of its parents $Pa(V_i)$.

Jensen [46, 49] gives the following proposition characterizing the d-separation criterion for influence diagrams.

Proposition 3.1 *Let $C_l \in SC_i$ and D_j be a decision variable s.t $i < j$. Then*
(i) C_l and D_j are d-separated i.e :

$$P(C_l \mid D_j) = P(C_l). \tag{3.4}$$

(ii) Let W be any set of variables prior to D_j in the temporal ordering. Then, C_l and D_j are d-separated given W i.e :

$$P(C_l \mid D_j, W) = P(C_l \mid W). \tag{3.5}$$

Note that the d-separation property for influence diagrams is slightly different from the one defined for Bayesian network [46, 49] since, utility nodes and links into decision nodes are ignored.

Example 3.5 *Let us present the numerical component of the influence diagram introduced in example 3.4. Table* 3.2 *represents a priori and conditional probabilities for chance nodes S, P and D. Table* 3.3 *represents the set of utilities for the value node V, in the context of its parents (Tr, P and D).*

D	$P(D)$	P	D	$P(P \mid D)$	S	P	$P(S \mid P)$
d	0.1	p	d	0.8	s	p	0.7
d̃	0.9	p	d̃	0.15	s	p̃	0.2
		p̃	d	0.2	s̃	p	0.3
		p̃	d̃	0.85	s̃	p̃	0.8

Table 3.2 – A priori and conditional probabilities

Physician's Utilities	States			
	pathological state (p)		no pathological state ($\tilde{\text{p}}$)	
	disease (d)	no disease ($\tilde{\text{d}}$)	disease (d)	no disease ($\tilde{\text{d}}$)
Treat (tr)	10	6	8	4
Not treat ($\tilde{\text{tr}}$)	0	2	1	10

TABLE 3.3 – Physician's utilities

As mentioned above, decision nodes act differently from chance nodes, thus it is meaningless to specify prior probability distribution on them. Moreover, it has no meaning to attach a probability distribution to children nodes of a decision node D_i unless a decision d_{ij} has been taken.

Therefore what is meaningful is $P(c_{ij} \mid do(d_{ij}))$, where $do(d_{ij})$ is the particular operator defined by Pearl [63], and not $P(c_{ij}, d_{ij})$. When iterating this reasoning we can bunch the whole decision nodes together and express the joint probability distribution of different chance nodes conditioned by decision nodes. This means that if we fix a particular configuration of decision nodes, say d, we get a Bayesian network representing $P(C \mid do(d))$ i.e the joint probability relative to C, in the context of decision's configuration d. In other words, the joint distribution relative to C remains the same when varying d. Thus, using the chain rule relative to Bayesian network [46, 49], we can infer the following chain rule relative to influence diagrams [47] :

$$P(C \mid D) = \Pi_{C_i \in C} P(C_i \mid Pa(C_i)). \tag{3.6}$$

Example 3.6 *Let us present the chain rule of the influence diagram in Figure 3.5 using the equation 3.6.*

3.3.2 Evaluation of influence diagrams

Given an influence diagram, the identification of its optimal policy can be ensured via evaluation algorithms which allow to generate the best strategy yielding to the highest expected utility. In 1990, Cooper has shown that the problem of evaluation of ID is NP-hard [15]. Within influence diagrams evaluation algorithms, we can distinguish :

– *Direct methods* [70, 81] operate directly on influence diagrams. These methods are based on two main operations i.e. arc reversal using Bayes theorem and node removal through some value preserving reduction.

P	D	S	$P(D)$	$P(P \mid D)$	$P(S \mid P)$	$P(P, D, S \mid Tr)$
p	d	s	0.1	0.8	0.7	0.056
p	d	$\tilde{s}$	0.1	0.8	0.3	0.024
p	$\tilde{d}$	s	0.9	0.15	0.7	0.0945
p	$\tilde{d}$	$\tilde{s}$	0.9	0.15	0.3	0.0405
$\tilde{p}$	d	s	0.1	0.2	0.2	0.004
$\tilde{p}$	d	$\tilde{s}$	0.1	0.2	0.8	0.016
$\tilde{p}$	$\tilde{d}$	s	0.9	0.85	0.2	0.153
$\tilde{p}$	$\tilde{d}$	$\tilde{s}$	0.9	0.85	0.8	0.612

TABLE 3.4 – The chain rule of the influence diagram in Figure 3.5

– *Indirect methods* transform influence diagrams into a secondary structure used to ensure computations. We can in particular mention the transformation into Bayesian networks [14] and into decision trees [71] that we will detail in what follows :

Evaluation of influence diagrams using Bayesian networks

This method, proposed by Cooper [14], is based on transforming influence diagrams into Bayesian networks [46] as secondary structure following these three steps :

1. Transform each decision node into a chance node characterized by an equi-probability, as follows :
$$P(D_i \mid Pa(D_i)) = \frac{1}{\mid dom(D_i) \mid} \tag{3.7}$$
where $dom(D_i)$ is the set of possible instance of D_i.
2. Transform the value node V into a binary chance node with two values False (F) and True (T).
3. Convert the utility function associated to V into a probability function as follows, $\forall pa(V) \in Pa(V)$:
$$P(v = T \mid pa(V)) = \frac{U(pa(V)) + K_2}{k_1} \tag{3.8}$$
where
$$K_1 = U_{max} - U_{min} \tag{3.9}$$
and
$$K_2 = -U_{min}. \tag{3.10}$$

U_{max} and U_{min} are the maximal utility and the minimal utility levels, respectively. Obviously, $P(v = F \mid pa(V)) = 1 - P(v = T \mid pa(V))$.

Once the BN is constructed, optimal strategy will be found through inference in BNs. In general, inference in Bayesian networks is an NP-hard problem. Several propagation algorithms have been proposed, the fundamental one was developed by Pearl for singly connected networks [60, 61]. Jensen was developed the propagation algorithm for multiply connected networks known as junction tree propagation algorithm [46, 47].

Let us start with solving a single decision problem i.e. the influence diagram contains one decision node D_m. Let E be the set of evidences, it contains the set of chance nodes in Bayesian network with known values. Solving this decision problem amounts to determine the instantiation of D_m that maximizes the expected utility computed as follows :

$$MEU(D_m, E) = \max_{D_m} \left[\sum_{Pa'(V)} u(Pa(V))P(Pa'(V)|D_m, E) \right]. \tag{3.11}$$

Where $Pa'(V)$ is the set of chance nodes in the set of parents of the node V $(Pa(V))$. Using the equation 3.8, we obtain by replacing $u(Pa(V))$:

$MEU(D_m, E) = \max_{D_m} \left[\sum_{Pa'(V)}(K_1 P(V = T \mid Pa(V) - K_2)P(Pa'(V) \mid D_m, E)\right]$. We have :

$$MEU(D_m, E) = K_1 * max_{D_m}[P(V = T \mid D_m, E)] - K_2. \tag{3.12}$$

So, the maximization of expected utility requires the calculation of $P(V = T|D_m, E)$ for a given instantiation of D_m. This conditional probability is computed using the appropriate Bayesian network inference algorithm according to the structure of the BN.

In the case of multiple decision problem, i.e. the influence diagram contains several decision nodes $D_1, \ldots, D_m$, for each decision node D_i in D, uninstantiated chance nodes are removed from $Pa(D_i)$ and $Pa'(D_i)$ because the selection of the optimal decision for D_i must be made in light of available information. The set of evidence E should be updated in the light of the previous step including decisions $D_1, \ldots, D_{i-1}$ that have been made.

Formally, the maximal expected utility of a set of decisions node D in light of evidence E is computed using a recursive version of equation 3.12 :

$$MEU(D, E) = K_1 * f(D, E) - K_2. \tag{3.13}$$

where :

$$f(D,E) = \max_{D.first} \left[\sum_{Pa'(Dr.first)} f(Dr, e \cup Pa(Dr.first)) * P(Pa'(Dr.first) \mid D.first, E) \right] \tag{3.14}$$

where : $D.first$ is the first decision to be made in D and Dr is the remaining decisions in D when the first one is removed. We have $Pa'(\emptyset) = Pa(V)$, $f(\emptyset, E) = P(V = T \mid Pa(V))$ and $P(\emptyset \mid D.first, E) = 1$.

Example 3.7 *Let us continue with the Medical Diagnosis's example.*
After the transformation of the ID, the decision node Tr will become a chance node Tr, its a priori probability distribution is presented in Table 3.5*.*
The new chance node V is characterized by a conditional probability distribution detailed in Table 3.6*.*
Figure 3.6 presents the obtained Bayesian network.

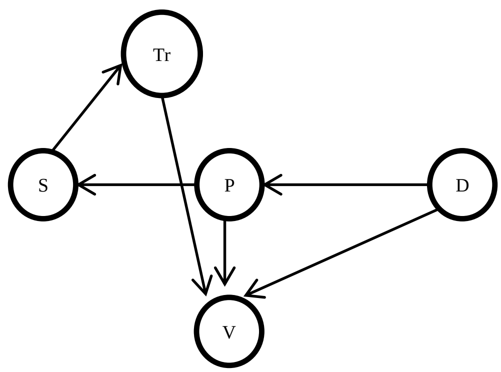

FIGURE 3.6 – The Bayesian network corresponding to the influence diagram in Figure 3.5

Tr	$P(Tr)$
tr	1/2
$\tilde{tr}$	1/2

TABLE 3.5 – A priori probability distribution for Tr

D	Tr	P	$P(v = T \mid D, Tr, P)$	$P(v = F \mid D, Tr, P)$
d	tr	p	1	0
$\tilde{d}$	tr	p	0.6	0.4
d	tr	$\tilde{p}$	0.8	0.2
$\tilde{d}$	tr	$\tilde{p}$	0.4	0.6
d	$\tilde{tr}$	p	0	1
$\tilde{d}$	$\tilde{tr}$	p	0.2	0.8
d	$\tilde{tr}$	$\tilde{p}$	0.1	0.9
$\tilde{d}$	$\tilde{tr}$	$\tilde{p}$	1	0

TABLE 3.6 – Conditional probability distribution for V

We have $k_1 = U_{max} - U_{min} = 10 - 0 = 10$ *and* $k_2 = 0$.
For instance, if the evidence is that $S = s$, *then* $MEU(Tr, S = s) = 10 * max_{Tr}[P(v = T \mid Tr, S = s)] = 7.988$ *meaning that the best decision is* $Tr = tr$.

Evaluation of influence diagram using decision trees

The transformation of an influence diagram into a decision tree requires a reordering of chance nodes in the diagram based on the concept of *decision window* [71].

A decision window of a decision node D_i is the set of chance nodes observed between the decision node D_i and D_{i+1} (is the set SC_i as it is detailed in section 3.3.1).

The principal of the transformation of an influence diagram into a decision tree can be summarized as follows [71] :

- Find each arc from a chance node in one decision window to a chance node in an earlier decision window. These arcs are called reversible arcs.
- Reverse these reversible arcs.
- Develop the decision tree according to the reordering of chance and decision nodes. A priori and conditional probabilities relative to chance nodes in the decision tree are computed from those of the influence diagram. Similarly, utilities are the same as those in the numerical component of the influence diagram.

Once the decision tree is constructed, the optimal strategy will be found through dynamic programming (see Algorithm 3.1).

Example 3.8 *Let us continue with the Medical Diagnosis's example. The influence diagram in Figure 3.5 is transformed into the decision tree presented in Figure 3.7. Table* 3.9, 3.10 *and* 3.11 *represent probability tables relative to chance nodes in this tree.*

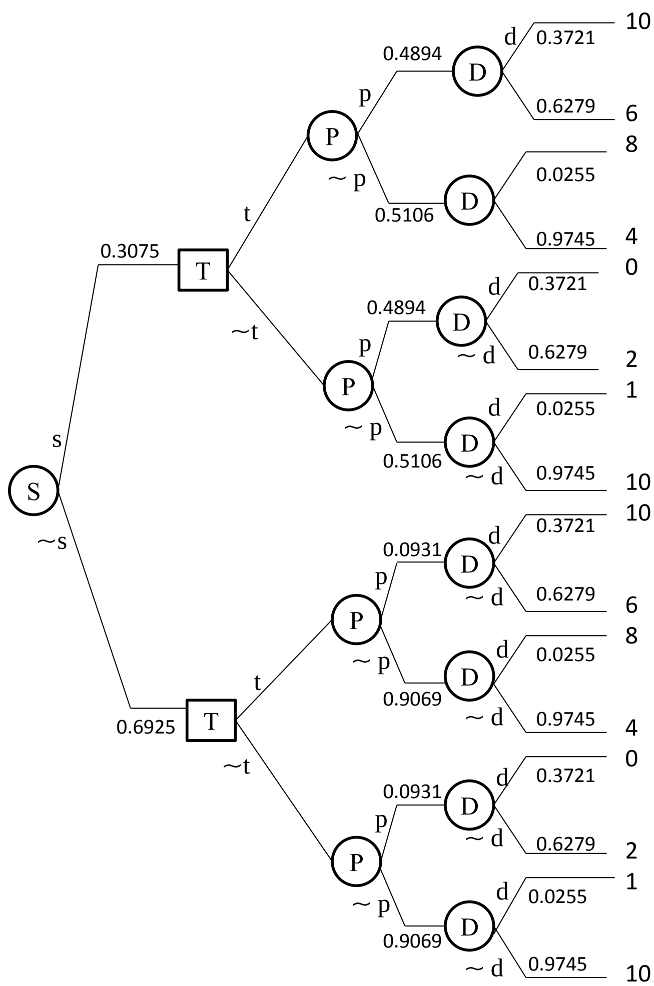

FIGURE 3.7 – The corresponding decision tree to the influence diagram in Figure 3.5

To compute these probabilities, we have used the probabilities $P(P \cap D)$ represented in Table 3.7 *and $P(P \cap S)$ represented in Table* 3.8.

In fact, probabilities in Table 3.7 *are used to compute* $P(P = p) = 0.08 + 0.135 = 0.215$

D	P	$P(D \cap P)$
d	p	0.08
$\tilde{d}$	p	0.135
d	$\tilde{p}$	0.02
$\tilde{d}$	$\tilde{p}$	0.765

TABLE 3.7 – Probabilities for $P(D \cap P)$

and $P(P = \tilde{p}) = 0.02 + 0.765 = 0.785$. *These probabilities are used to compute* $P(P \cap S)$ *represented in Table 3.8.*

P	S	$P(P \cap S)$
p	s	0.1505
p	$\tilde{s}$	0.0645
$\tilde{p}$	s	0.157
$\tilde{p}$	$\tilde{s}$	0.628

TABLE 3.8 – Probabilities for $P(P \cap S)$

We can conclude from Table 3.8 that $P(S = s) = 0.1505 + 0.157 = 0.3075$ *and* $P(S = \tilde{s}) = 0.0645 + 0.628 = 0.6925$ *as it is represented in Table* 3.9*.*

S	$P(S)$
s	0.3075
$\tilde{s}$	0.6925

TABLE 3.9 – A priori probabilities for S

D	P	$P(D\|P)$
d	p	0.3721
$\tilde{d}$	p	0.6279
d	$\tilde{p}$	0.0255
$\tilde{d}$	$\tilde{p}$	0.9745

TABLE 3.10 – Conditional probabilities $P(D|P)$

If the patient exhibits the symptom S *then the optimal strategy in the tree represented*

P	S	$P(P\|S)$
p	s	0.4894
p	$\tilde{s}$	0.0931
$\tilde{p}$	s	0.5106
$\tilde{p}$	$\tilde{s}$	0.9069

TABLE 3.11 – Conditional probabilities $P(P|S)$

in Figure 3.7 is to treat the patient as it is shown in Figure 3.8. The expected utility of this strategy is 7.988 *(*$((5.7593 * 0.3075) + (8.9776 * 0.6925))$*) which is equal to the maximal expected utility found in example* 3.7 *where the influence diagram was evaluated using a Bayesian network.*

3.4 Conclusion

In this chapter, we have developed two probabilistic decision models which are decision trees and influence diagrams where the decision criterion is the expected utility (EU). These models allow the representation of sequential decision problems.

As we have seen in Chapter 2, possibilistic decision theory presents an interesting alternative to the classical decision theory used as a framework in standard decision trees and influence diagrams.

In the Chapter 4 and 5, we propose a deep study of possibilistic decision trees and in Chapter 6 we will study the possibilistic counterpart of influence diagrams.

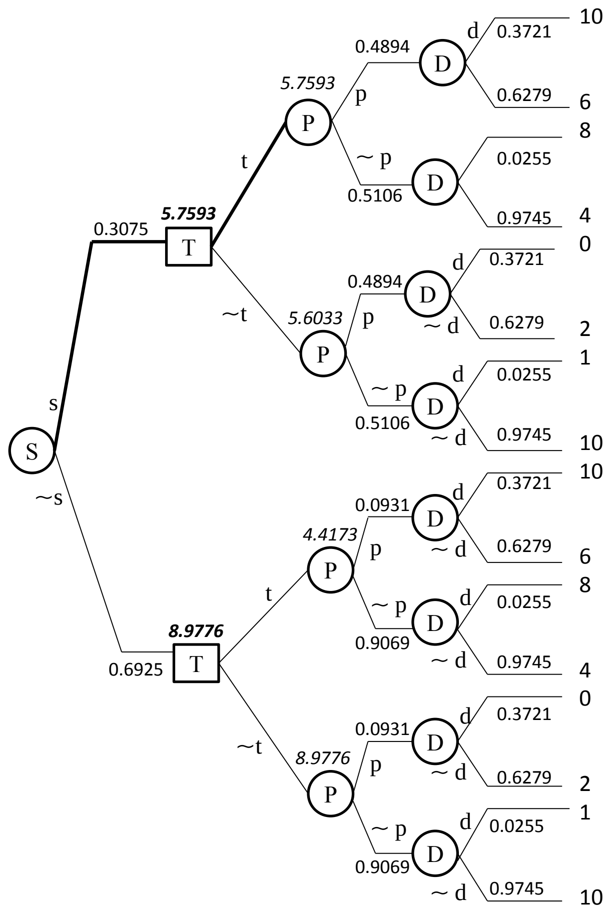

Figure 3.8 – Optimal strategy if the patient exhibits the symptom S

Chapitre 4

Possibilistic Decision Trees

4.1 Introduction

As we have seen in the previous chapter, graphical decision models provide intuitive representations of decision problems under uncertainty. Most of these models were developed in the probabilistic framework.

In this chapter, we develop possibilistic decision trees by studying the complexity of decision making in possibilistic decision trees for each possibilistic decision criteria presented in Chapter 2.

This chapter is organized as follows : in Section 4.2, possibilistic decision trees will be developed. Section 4.3 will detail these graphical models with qualitative possibilistic utilities. Decision trees with possibilistic likely dominance will be detailed in Section 4.4 and those with order of magnitude expected utility in Section 4.5. Possibilistic decision trees with Choquet integrals will be developed in Section 4.6 and polynomial cases of possibilistic Choquet integrals will be presented in Section 4.7.

Principle results of this chapter are published in [9].

4.2 Possibilistic decision trees

Possibilistic decision trees have the same graphical component as probabilistic ones (see Section 3.2 in Chapter 3) i.e. it is composed of a set of nodes $\mathcal{N}$ and a set of edges $\mathcal{E}$. Like probabilistic decision trees, the set of nodes $\mathcal{N}$ contains three kinds of nodes i.e. $\mathcal{N} = \mathcal{D} \cup \mathcal{C} \cup \mathcal{LN}$ where $\mathcal{D}$ is the set of decision nodes, $\mathcal{C}$ is the set of chance nodes and $\mathcal{LN}$ is the set of leaves. This is not the case of the numerical component which relies in the possibilistic framework :

– Arcs issuing from chance nodes are quantified by possibility degrees in the context of their parents. Formally, for any $C_i \in \mathcal{C}$, the uncertainty pertaining to the more or less possible outcomes of each C_i is represented by a *conditional possibility distribution* π_i on $Succ(C_i)$, such that $\forall N \in Succ(C_i), \pi_i(N) = \Pi(N|path(C_i))$. To each node $C_i \in C$, a possibilistic lottery L_{C_i} is associated relative to its outcomes.

– Then, a utility is assigned to each leaf nodes which can be numerical (e.g. currency gain) or ordinal (e.g. satisfaction) according to the decision criterion.

Example 4.1 *The decision tree of Figure 4.1 is defined by*
$D = \{D_0, D_1, D_2\}$, $\mathcal{C} = \{C_1, C_2, C_3, C_4, C_5, C_6\}$ *and* $\mathcal{LN} = U = \{0, 1, 2, 3, 4, 5\}$. *Correspon-*

ding possibilistic lotteries to chance nodes are $L_{C_1} = \langle 1/L_{D_1}, 0.5/L_{D_2}\rangle$*,* $L_{C_2} = \langle 1/1, 0.7/2\rangle$*,* $L_{C_3} = \langle 1/1, 0/5\rangle$*,* $L_{C_4} = \langle 0.2/0, 1/4\rangle$*,* $L_{C_5} = \langle 1/1, 0.3/4\rangle$ *and* $L_{C_6} = \langle 1/2, 0.5/5\rangle$*.*

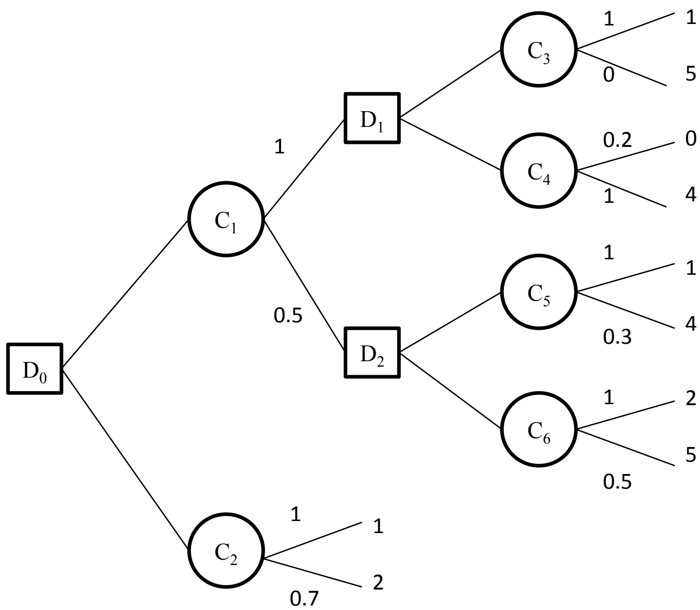

FIGURE 4.1 – Example of possibilistic decision tree

As we have seen in the previous chapter, solving decision trees amounts at building an optimal strategy δ^* in Δ (the set of sound and complete strategies).

Like in probabilistic decision trees, strategies can be evaluated and compared thanks to the notion of possibilistic lottery reduction : each chance node can be seen as a simple lottery (for the most right chance nodes) or as a compound lottery (for the inner chance nodes). Each strategy is thus a compound lottery and can be reduced to an equivalent simple one. Formally, the composition of possibilistic lotteries will be applied from the leafs of the strategy to its root, according to the following recursive definition for any N_i in $\mathcal{N}$:

$$L(N_i, \delta) = \begin{cases} L(\delta(N_i), \delta) \quad if \quad N_i \in \mathcal{D} \\ Reduction(\langle \pi_i(X_j)/L(X_j, \delta)_{X_j \in Succ(N_i)}\rangle) \ if \ N_i \in \mathcal{C} \\ < 1/u(N_i) > \quad if \quad N_i \in \mathcal{LN} \end{cases} \tag{4.1}$$

where $Reduction(\langle \pi_i(X_j)/L(X_j,\delta)_{X_j \in Succ(N_i)} \rangle)$ is defined by the following equation as we have seen in Chapter 2 :

$$Reduction(\langle \lambda_1/L_1, \dots, \lambda_m/L_m \rangle) = \langle \max_{j=1..m} (\lambda_j \otimes \lambda_1^j)/u_1, \dots, \max_{j=1..m} (\lambda_j \otimes \lambda_n^j)/u_n \rangle \qquad (4.2)$$

$\otimes$ is the product operator in the case of numerical possibility theory and the min operator in the case of its qualitative counterpart.

Since, the operators max, $*$ and min used in the reduction operation are polytime, Equation (4.2) defines a polytime computation of the reduced lottery.

Proposition 4.1 *For any strategy δ in Δ, a simple possibilistic lottery reduction equivalent to δ can be computed in polytime.*

Proof. [Proof of Proposition 4.1]
Let $\delta \in \Delta = \{(D_0, \delta(D_0)), \dots, (D_i, \delta(D_i)), \dots, (D_n, \delta(D_l))\}$ be a complete and sound strategy.
We first compute the compound lottery corresponding to δ, merging each decision node D_i in δ with the chance node in $\delta(D_i)$, say C_i^δ. We get a compound lottery $L = \{C_0^\delta, \dots, C_i^\delta, \dots, C_l^\delta\}$; the merging is performed linearly in the number of decision nodes in the strategy.

Then we can suppose without loss of generality that the nodes are numbered in such a way that $i < j$ implies that C_i^δ does not belong to the subtree rooted C_j^δ (we label the nodes from the root to the leaves).

Then, for $i = m$ to 1, we replace each compound lottery $C_i^\delta = \langle pr_i(X_{i1})/X_{i1}, \dots, pr_i(X_{ik_i})/X_{ik_i} \rangle$ by its reduction, where $Succ(C_i^\delta) = \{X_{i1}, \dots, X_{ik_i}\}$ is the set of successors of $C_i^\delta$and $k_i = |Succ(C_i^\delta)|$. Because we proceed from the leaves to the root, the X_{i1} are simple lotteries. Since the min and max operation are linear, the reduction of this 2 level compound lottery is linear in the size of the compound lottery. The size of the resulting compound lottery is bounded by the sum of the size of the elementary lotteries before reduction, and thus linear. In any case, it is bounded by the number of levels in the scale, which is itself bounded by the number of edges and leaves in the tree (for the case where all the possibility degrees and all the utility degrees are different). Hence a complexity of the reduction is bounded by $O(|E + LN|)$, where E is the number of edges and LN is the number of leave nodes in the strategy.

Thanks to the backward recursion, each node in the strategy is visited only once. Thus a global complexity is bounded by $O(l.(E + LN))$, where l the number of chance nodes in the strategy. ■

We are now in position to compare strategies, and thus to define the notion of optimality. Let O be one of the possibilistic decision criteria defined in Chapter 2 (i.e. depending on the application, $\geq_O$ is either $\geq_{L\Pi}$, or $\geq_{LN}$, or the order induced by U_{pes}, or by U_{opt}, etc.). A strategy $\delta \in \Delta$, is said to be optimal w.r.t. $\geq_O$ iff :

$$\forall \delta' \in \Delta, Reduction(\delta) \geq_O Reduction(\delta'). \tag{4.3}$$

Notice that this definition does not require the full transitivity (nor the completeness) of $\geq_O$ and is meaningful as soon as the strict part of $\geq_O$ or $>_O$, is transitive. This means that it is applicable to the preference relations that rely on the comparison of global utilities (qualitative utilities, binary utility and Choquet integrals) but also to $\geq_{LN}$ and $\geq_{L\Pi}$. We show in the following that the complexity of the problem of optimization depends on the criterion at work.

Like probabilistic decision trees, the simplest solving method of possibilistic decision trees consists on an exhaustive enumeration of all possible strategies in the decision tree which will be compared w.r.t decision criterion. The following example illustrates this process using Ch_N.

Example 4.2 *Let us evaluate the decision tree in Figure 4.1 using necessity-based Choquet integrals as a decision criterion in the context of qualitative possibility theory.*
We can distinguish, in Table 4.1*,* 5 *possible strategies (*$\Delta = \{\delta_1, \delta_2, \delta_3, \delta_4, \delta_5\}$*) where* L_i *is the lottery of the strategy* δ_i *:*

δ_i	L_i	$Ch_N(L_i)$
$\delta_1 = \{(D_0, C_1), (D_1, C_3), (D_2, C_5)\}$	$\langle 1/1, 0.3/4, 0/5 \rangle$	1
$\delta_2 = \{(D_0, C_1), (D_1, C_3), (D_2, C_6)\}$	$\langle 1/1, 0.5/2, 0.5/5 \rangle$	1
$\delta_3 = \{(D_0, C_1), (D_1, C_4), (D_2, C_5)\}$	$\langle 0.2/0, 0.5/1, 1/4 \rangle$	2.3
$\delta_4 = \{(D_0, C_1), (D_1, C_4), (D_2, C_6)\}$	$\langle 0.2/0, 0.5/2, 1/4, 0.5/5 \rangle$	2.6
$\delta_5 = \{(D_0, C_2)\}$	$\langle 1/1, 0.7/2 \rangle$	1.7

TABLE 4.1 – Exhaustive enumeration of possible strategies in Figure 6.2

So, the optimal strategy in this decision tree is δ_4 *with* $Ch_N(\delta_4) = 2.6$ *as it is shown in Figure 4.2.*

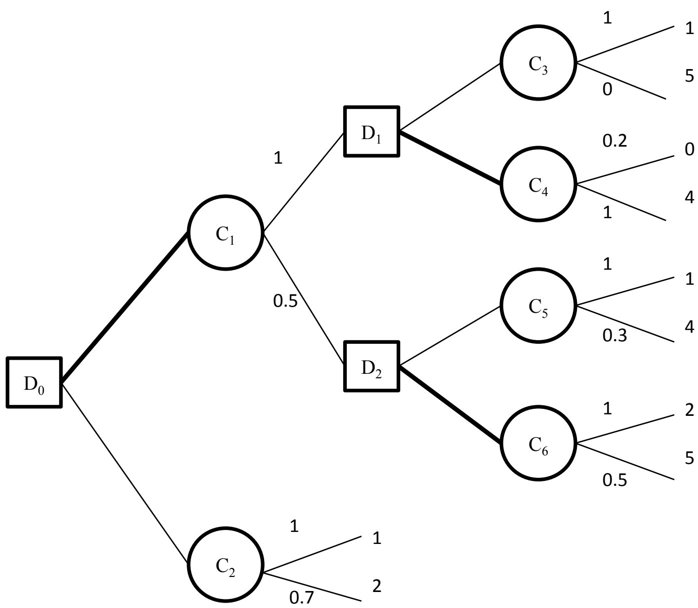

FIGURE 4.2 – The optimal strategy $\delta^* = \{(D_0, C_1), (D_1, C_4), (D_2, C_6)\}$

Finding optimal strategies in possibilistic decision trees via an exhaustive enumeration of Δ is a highly computational task. For instance, in a possibilistic decision tree with n decision nodes and a branching factor equal to 2, the number of potential strategies is in $O(2^{\sqrt{n}})$ (exactly like probabilistic decision trees since the two kinds of decision trees have the same graphical component). Based on the work of [39], we can propose the following result :

Proposition 4.2 *In a possibilistic decision tree with* n *nodes and a branching factor equal to* 2*, the number of potential strategies is in* $O(2^{\sqrt{n}})$.

Proof. [Proof of Proposition 4.2]
Suppose that we have a binary decision tree such that we have 4^i decision nodes in depth $2i$ (1 decision node in depth 0,. . ., 16 decision nodes in depth 4). We will proceed by backward induction to compute the number of strategies according to the depth in the

decision tree.
For decision nodes which have no decision nodes in its successors we distinguish 2 strategies. Then we proceed by recurrence and the number of strategies starting by a chance node is equal to the product of the numbers of strategies beginning from its children. For decision nodes, the number of strategies is equal to the sum of the number of strategies of its children.

The total number of strategies is equal to a sequence $2u_{k-1}^2$ when k is the number of decision nodes in a path from a decision node to a utility node. The general term of this sequence is equal to $2^{(2^{k+1}-1)}$. So, the number of strategies in the decision tree pertains to $O(2^{\sqrt{n}})$. ■

For standard probabilistic decision trees, where the goal is to maximize expected utility (EU), an optimal strategy can be computed in polytime (with respect to the size of the tree) via the dynamic programming which builds the best strategy backwards, optimizing the decisions from the leaves of the tree to its root (see Algorithm 4.1).

Regarding possibilistic decision trees, Garcia and Sabbadin [33] have shown that such a method can also be used to get a strategy maximizing U_{pes} and U_{opt}. The reason is that like EU, these possibilistic decision criteria satisfy the key property of *weak monotonicity* stating that the combination of L (resp. L') with L", does not change the initial order induced by O between L and L' - this allows dynamic programming to decide in favor of L or L' before considering the compound decision.

Formally for any decision criterion O over possibilistic lotteries, $\geq_O$ is said to be weakly monotonic iff whatever L, L' and L", whatever (α,β) such that $\max(\alpha, \beta) = 1$:

$$L \succeq_O L' \Rightarrow \langle \alpha/L, \beta/L" \rangle \succeq_O \langle \alpha/L', \beta/L" \rangle. \tag{4.4}$$

Given any preference order $\succeq_O$ (satisfying the weak monotonicity property) among possibilistic lotteries, the possibilistic counterpart of dynamic programming algorithm (Algorithm 3.1) is depicted by Algorithm 4.1. When each chance node is reached, an optimal sub-strategy is built for each of its children - these sub-strategies are combined w.r.t. their possibility degrees, and the resulting compound strategy is reduced : we get an equivalent simple lottery, representing the current optimal sub-strategy. When a decision node X is reached, a decision Y^* leading to a sub-strategy optimal w.r.t $\succeq_O$ is selected among all the possible decisions $Y \in Succ(X)$, by comparing the simple lotteries equivalent to each sub strategies.

This procedure crosses each edge in the tree only once. When the comparison of simple lotteries by $\succeq_O$ (Line (2)) and the reduction operation on a 2-level lottery (Line (1)) can be performed in polytime, its complexity is polynomial w.r.t the size of the tree as stated by the following Proposition :

Proposition 4.3 *If $\succeq_O$ satisfies the monotonicity property, then dynamic programing computes a strategy optimal w.r.t O in polynomial time with respect to the size of the decision tree.*

Proof. [Proof of Proposition 4.3]
The principle of the Backward induction method at work in dynamic programming is to eliminate sub-strategies that are not better than the optimal sub-strategies. The principle of monotonicity writes :
$L \succeq_O L' \Rightarrow \langle \alpha/L, \beta/L" \rangle \succeq_O \langle \alpha/L', \beta/L" \rangle$.
It guarantees that the elimination of sub-strategies that are not strictly better than their concurrents is sound and complete for the decision trees of size 2. Notice that $L \succ_O L'$ does not imply that L' does not belong to an optimal strategy but it implies that if L' belongs to an optimal strategy, so does L. When, a unique strategy among the optimal one is searched for, the algorithm can forget about L'.

The sequel on the proof is direct, by recursion on the depth on the decision tree. Let us denote $\langle \alpha/L, \beta/L" \rangle$ by L_1 and $\langle \alpha/L', \beta/L" \rangle$ by L_2. Indeed, from $L \geq_O L' \Rightarrow L_1 \succeq_O L_2$, we get that $L \geq_O L' \Rightarrow \langle \gamma L_1, \delta L_3 \rangle \succeq_O \langle \gamma L_2, \delta L_3 \rangle$ and so on. ■

Algorithm 4.1: Dynamic programming

```
Data: In : a node X, In/Out : a strategy δ
Result: A lottery L
begin
    for i ∈ {1, ..., n} do L[u_i] ← 0
    if N ∈ LN then L[u(N)] ← 1
    if N ∈ C then
        % Reduce the compound lottery
        foreach Y ∈ Succ(N) do
            L_Y ← ProgDyn(Y, δ)
            for i ∈ {1, ..., n} do
                L[u_i] ← max(L[u_i], (π_N(Y) ⊗ L_Y[u_i])) (Line (1))
        if N ∈ D then
            % Choose the best decision
            Y* ← Succ(N).first
            foreach Y ∈ Succ(N) do
                L^Y ← ProgDyn(Y, δ)
                if L_Y >_O L_{Y*} then Y* ← Y (Line (2))
            δ(N) ← Y*
            L ← L_{Y*}
        return L
end
```

In Line 1 of Algorithm 4.1, $\otimes$ is the min operator in the case of qualitative possibility theory and the product operator in the case of numerical possibility theory. We will see in the following that, beyond U_{pes} and U_{opt} criteria, several other criteria satisfy the monotonicity property and that their optimization can be managed in polytime by dynamic programming. The possibilistic Choquet integrals, on the contrary, do not satisfy weak monotonicity ; we will show that they lead to NP-Complete decision problems.

Formally, for any of the possibilistic optimization criteria, the corresponding decision problem can be defined as follows :

Definition 4.1 ***[DT-OPT-**O**]***(Strategy optimization w.r.t. an optimization criterion O in possibilistic decision trees)*
INSTANCE : A possibilistic Decision Tree $\mathcal{T}$, a level α.
QUESTION : Does there exist a strategy $\delta \in \Delta$ such as Reduction(δ) $\geq_O \alpha$?

For instance DT-OPT-Ch_N (resp. DT-OPT-Ch_Π, DT-OPT-U_{pes}, DT-OPT-U_{opt}, DT-OPT-PU, DT-OPT-LN, DT-OPT-$L\Pi$ and DT-OPT-$OMEU$) corresponds to the optimization of the possibilistic qualitative utility Ch_N (resp. Ch_Π, U_{pes} and U_{opt}, PU, LN, $L\Pi$ and $OMEU$). Each one of these decision problems will be studied in what follows.

4.3 Qualitative possibilistic utilities (U_{pes}, U_{opt}, PU)

Possibilistic qualitative utilities U_{pes} and U_{opt} satisfy the weak monotonicity principle. Although not referring to a classical, real-valued utility scale, but to a 2 dimensional scale, this is also true in the case of PU.

Proposition 4.4 *$\succeq_{PU}$, $\succeq_{U_{pes}}$ and $\succeq_{U_{opt}}$ satisfy the weak monotonicity property.*

This proposition is not explicitly proved in the literature although it is a common knowledge in qualitative possibilistic decision theory (see [25, 37]).

Proof. [Proof of Proposition 4.4]
Weak monotonicity of PU
Consider any three lotteries L, L' and L''. We can suppose without loss of generality that they are in a reduced form, i.e. :
$L = \langle \overline{u}/\top, \underline{u}/\bot \rangle$,
$L' = \langle \overline{v}/\top, \underline{v}/\bot \rangle$,
$L'' = \langle \overline{w}/\top, \underline{w}/\bot \rangle$.
Let $L_1 = Reduction(\langle \alpha/L, \beta/L'' \rangle)$ and $L_2 = Reduction(\langle \alpha/L', \beta/L'' \rangle)$.
According to the reduction operation, we get :

- $L_1 = \langle \overline{u_1}/\top, \underline{u_1}/\bot \rangle$, where $\overline{u_1} = max(min(\alpha, \overline{u}), min(\beta, \overline{w}))$ and $\underline{u_1} = max(min(\alpha, \underline{u}), min(\beta, \underline{w}))$.
- $L_2 = \langle \overline{u_2}/\top, \underline{u_2}/\bot \rangle$, where $\overline{u_2} = max(min(\alpha, \overline{v}), min(\beta, \overline{w}))$ and $\underline{u_2} = max(min(\alpha, \underline{v}), min(\beta, \underline{w}))$.

Suppose that $L \geq_{PU} L'$. Recall that $max(\alpha, \beta) = 1$ and that $L \geq_{PU} L'$ arises in 3 cases (i.e. (i) $\overline{u} = \overline{v} = 1$ and $\underline{u} \leq \underline{v}$, (ii) $\overline{u} \geq \overline{v}$ and $\underline{u} = \underline{v} = 1$, (iii) $\overline{u} = 1$, $\overline{v} < 1$ and $\underline{v} = 1$). Hence 6 different cases. For each of them, we show that $L_1 \geq_{PU} L_2$ can be deduced :

- Case 1 : $\overline{u} = \overline{v} = 1$ and $\underline{u} \leq \underline{v}$, $\alpha = 1$. $\overline{u} = \overline{v} = 1$ and $\alpha = 1$ implies that $\overline{u_1} = max(min(\alpha, \overline{u}), min(\beta, \overline{w})) = 1$ and $\overline{u_2} = max(min(\alpha, \overline{v}), min(\beta, \overline{w})) = 1$ $\underline{u} \leq \underline{v}$ and $\alpha = 1$ implies that $\underline{u_1} = min(\beta, \underline{w}) = \underline{u_2}$. Hence $L_1 =_{PU} L_2$.
- Case 2 : $\overline{u} = \overline{v} = 1$ and $\underline{u} \leq \underline{v}$, $\beta = 1$, $\overline{u} = \overline{v} = 1$ and $\beta = 1$ implies that $\overline{u_1} = max(min(\alpha, \overline{u})), min(\beta, \overline{w}) = max(\alpha, \overline{w})$ and $\overline{u_2} = max(min(\alpha, \overline{v}), min(\beta, \overline{w})) =$

$max(\alpha, \overline{w}) = max(\alpha, \overline{w}) = \overline{u_1}$. $\underline{u} \leq \underline{v}$ and $\beta = 1$ implies that $\underline{u_1} = max(min(\alpha, \underline{u}), \underline{w})$ and $\underline{u_1} = max(min(\alpha, \underline{v}), \underline{w})$; since $\underline{u} \leq \underline{v}$, we get $\underline{u_1} \leq \underline{u_2}$.

Recall that $max(\overline{w}, \underline{w}) = 1$. When $\overline{w} = 1$, we get $\overline{u_1} = \overline{u_2} = 1$ and $\underline{u_1} \leq \underline{u_2}$, and thus $L_1 \geq_{PU} L_2$. When $\underline{w} = 1$, $\underline{u_1} = \underline{u_2} = 1$ and $\overline{u_1} = \overline{u_2}$. Hence $L_1 =_{PU} L_2$.

– Case 3 : $\overline{u} \geq \overline{v}$ and $\underline{u} = \underline{v} = 1$, $\alpha = 1$. This case is similar to case 1 (exchanging the roles of the positive utilities and of the negative utilities.

– Case 4 : $\overline{u} \geq \overline{v}$ and $\underline{u} = \underline{v} = 1$, $\beta = 1$. This case is similar to case 2 (exchanging the roles of the positive utilities and of the negative utilities.

– Case 5 : $\overline{u} = 1$, $\overline{v} < 1$, $\underline{v} = 1$, $\alpha = 1$. Then : $\overline{u_1} = 1$, $\underline{u_1} = max(\underline{u}, min(\beta, \underline{w}))$, $\overline{u_2} = max(\overline{v}, min(\beta, \overline{w}))$ and $\underline{u_2} = 1$. That is to say $\overline{u_1} = 1 \geq \overline{u_2}$ and $\underline{u_1} \leq \underline{u_2} = 1$. Thus $L_1 \geq_{PU} L_2$.

– Case 6 : $\overline{u} = 1$, $\overline{v} < 1$, $\underline{v} = 1$, $\beta = 1$. Then : $\overline{u_1} = max(\alpha, \overline{w})$ $\underline{u_1} = max(min(\alpha, \underline{u}), \underline{w})$ and $\overline{u_2} = max(min(\alpha, \overline{v}), \overline{w})$

$\underline{u_2} = max(\alpha, \underline{w})$. When $\overline{w} = 1$ (resp. $\underline{w} = 1$) we get $\overline{u_1} = \overline{u_2} = 1$ and $\underline{u_1} \leq \underline{u_2}$ ($\underline{u_1} = \underline{u_2} = 1$ $\overline{u_1} \geq \overline{u_2}$). Hence $L_1 \leq_{PU} L_2$.

So, in any case, $L \geq_{PU} L'$ implies that $L_1 \leq_{PU} L_2$, i.e.
$L_1 = Reduction(\langle \alpha/L, \beta/L" \rangle) \geq L_2 = Reduction(\langle \alpha/L', \beta/L" \rangle)$.
As a consequence $L \geq_{PU} L'$ implies that $\langle \alpha/L, \beta/L" \rangle \geq \langle \alpha/L', \beta/L" \rangle$.

Weak monotonicity of U_{pes}

Consider any three lotteries L, L' and $L"$. We can, without loss of generality, suppose that L, L' and $L"$ are constant lotteries (thanks to certainty equivalence axiom [25]) i.e. $L =< 1/u >$, $L' =< 1/u' >$ and $L" =< 1/u" >$: any utility degree different from u (resp. u', resp. $u"$) receives a possibility degree equal to 0.
If $L \sim_{U_{pes}} L'$ then from the independence axiom [25] we have $\langle \alpha/L, \beta/L" \rangle \sim_{U_{pes}} \langle \alpha/L', \beta/L" \rangle$ (under the assumption that
$max(\alpha, \beta) = 1$).
We thus only have to consider the case $L >_{U_{pes}} L'$. Since $U_{pes}(L) = u$ and $U_{pes}(L') = u'$, this implies that $u > u'$. Let :
$L_1 = Reduction((\alpha \wedge L) \vee (\beta \wedge L")) =< \alpha/u, \beta/u" >$ and
$L_2 = Reduction((\alpha \wedge L') \vee (\beta \wedge L")) =< \alpha/u', \beta/u" >$.
Three cases are to be considered : $u" \geq u > u'$, $u > u' \geq u"$ and $u > u" > u'$

– Case 1 : $u" \geq u > u'$. Then :
$U_{pes}(L_1) = max(min(u", 1 - \alpha), min(u, 1)) = max(min(u", 1 - \alpha), u)$ and $U_{pes}(L_2) = max(min(u", 1 - \alpha), u')$.
Obviously, $u > u'$ implies $max(min(u", 1 - \alpha), u) \geq max(min(u", 1 - \alpha), u')$, i.e. $U_{pes}(L_1) \geq U_{pes}(L_2)$.

– Case 2 : $u > u' \geq u$". Then :
$U_{pes}(L_1) = max(min(u, 1-\beta), u")$ and $U_{pes}(L_2) = max(min(u', 1-\beta), u")$. Obviously, $u > u'$ implies $max(min(u, 1-\beta), u") \geq max(min(u', 1-\beta), u")$,
i.e. $U_{pes}(L_1) \geq U_{pes}(L_2)$.

– Case 3 : $u > u" > u'$. Hence :
$U_{pes}(L_1) = max(min(u, 1-\beta), u")$ and $U_{pes}(L_2) = max(min(u", 1-\alpha), u')$.
Recall that $max(\alpha, \beta) = 1$. If $\alpha = 1$, $U_{pes}(L_2) = u'$; from $u" > u'$ we then get : $max(min(u, 1-\beta), u") \geq u'$, i.e. : $U_{pes}(L_1) \geq U_{pes}(L_2)$. If $\beta = 1$, $U_{pes}(L_1) = u"$. From $u" > u'$, we get
$u" \geq max(min(u", 1-\alpha), u')$ i.e. $U_{pes}(L_1) \geq U_{pes}(L_2)$.

So, $U_{pes}(L_1) \geq U_{pes}(L_2)$: in any case, $\langle \alpha/L, \beta/L" \rangle \geq_{U_{Pes}} \langle \alpha/L', \beta/L" \rangle$.

Weak monotonicity of U_{opt}

The proof is similar to the previous one. Consider any three lotteries L, L' and $L"$. We can without loss of generality suppose that L, L' and $L"$ are constant lotteries (thanks to certainty equivalence axiom [25]) i.e. $L =< 1/u >$, $L' =< 1/u' >$ and $L" =< 1/u" >$: any utility degree different from u (resp. u', resp. $u"$) receives a possibility degree equal to 0.

If $L \sim_{U_{opt}} L'$ then from the independence axiom [25] we have $\langle \alpha/L, \beta/L" \rangle \sim_{U_{opt}} \langle \alpha/L', \beta L" \rangle$ (under the assumption that $max(\alpha, \beta) = 1$).

We thus only have to consider the case $L >_{U_{opt}} L'$. Because $U_{opt}(L) = u$ and $U_{opt}(L') = u'$, this implies that $u > u'$. Let :

$L_1 = Reduction((\alpha \wedge L) \vee (\beta \wedge L")) =< \alpha/u, \beta/u" >$.

$L_2 = Reduction((\alpha \wedge L') \vee (\beta \wedge L")) =< \alpha/u', \beta/u" >$.

Three cases are to be considered : $u" \geq u > u'$, $u > u' \geq u"$ and $u > u" > u'$.

– Case 1 : $u" \geq u > u'$.
Then $U_{opt}(L_1) = \max(\min(u, \max(\alpha, \beta)), \min(u", \beta))$ and $U_{opt}(L_2) = \max(\min(u', \max(\alpha, \beta)), \min(u", \beta))$.
Recall that $\max(\alpha, \beta) = 1$.
Thus : $U_{opt}(L_1) = \max(\min(u, 1), \min(u", \beta)) = \max(u, \min(u", \beta))$ and $U_{opt}(L_2) = \max(\min(u', 1), \min(u", \beta)) =$
$\max(u', \min(u", \beta))$.
$u > u'$ implies that $\max(u, \min(u", \beta)) \geq \max(u', \min(u", \beta))$, i.e. $U_{opt}(L_1) > U_{opt}(L_2)$.

– Case 2 : $u > u' \geq u"$.
Then $U_{opt}(L_1) = \max(\min(u, \alpha), \min(u", 1)) = \max(\min(u, \alpha), u")$ and
$U_{opt}(L_2) = \max(\min(u', \alpha), \min(u", 1)) = \max(\min(u', \alpha), u")$. $u > u'$ implies $\max(\min(u, \alpha), u") \geq \max(\min(u', \alpha), u")$, i.e. $U_{opt}(L_1) \geq U_{opt}(L_2)$.

– Case 3 : $u > u" > u'$.

Hence $U_{opt}(L_1) = max(min(u, \alpha), min(u", 1)) = max(min(u, \alpha), u")$ and $U_{opt}(L_2) = max(min(u', 1), min(u", \beta)) =$
$max(u', min(u", \beta))$. If $\alpha = 1$, then : $U_{opt}(L_1) = max(u, u") = u$ and $U_{opt}(L_2) = max(u', min(u", \beta))$. $u > u"$, so $u > min(u", \beta)$; moreover $u > u'$, so $u > max(min(u", \beta), u')$, i.e. $U_{opt}(L_1) > U_{opt}(L_2)$.

If $\beta = 1$, $U_{opt}(L_1) = max(min(u, \alpha), u")$ and $U_{opt}(L_2) = max(u', u") = u"$. $u" \geq u"$, so $max(min(u, \alpha), u") \geq u"$, i.e. $U_{opt}(L_1) \geq U_{opt}(L_2)$.

So, $U_{opt}(L_1) \geq U_{opt}(L_2)$: in any case, $\langle \alpha/L, \beta/L" \rangle \geq_{U_{opt}} \langle \alpha/L', \beta/L" \rangle$. ■

As a consequence, dynamic programming (i.e. Algorithm 4.1) applies to the optimization of these criteria in possibilistic decision trees. It is also known that dynamic programming applies to the optimization of U_{pes}, U_{opt} and PU in possibilistic Markov decision processes [67] and thus to decision trees.We can then derive the following corollary :

Corollary 4.1 *DT-OPT-U_{pes}, DT-OPT- U_{opt} and DT-OPT-PU belong to P.*

4.4 Possibilistic likely dominance $(LN, L\Pi)$

We show now that possibilistic likely dominance satisfies the weak monotonicity principle.

Proposition 4.5 *$\succeq_{L\Pi}$ and $\succeq_{LN}$ satisfy the weak monotonicity principle.*

In [23], the authors have defined the likely dominance decision rule and have presented its axiomatic system in the context of Savage decision theory. In what follows we develop a formal proof for Proposition 4.5 which is a direct consequence of the basic axiom of Weak Sure Thing Principle (Axiom 2^{S_L}).

Proof. [Proof of Proposition 4.5]

Ordinal setting

Consider any three lotteries L, L' and $L"$. We can suppose without loss of generality that they are in a reduced form, i.e. :

$L = \langle \lambda_1/u_1, \ldots, \lambda_n/u_n \rangle$,
$L' = \langle \lambda_1^{'}/u_1^{'}, \ldots, \lambda_n^{'}/u_n^{'} \rangle$,
$L" = \langle \lambda_1^{"}/u_1^{"}, \ldots, \lambda_n^{"}/u_n^{"} \rangle$.

Let $L_1 = Reduction(\alpha \wedge L \vee \beta \wedge L")$. According to the definition of the reduction (see section 2.2), the possibility of getting a utility degree $u_k \in U$ from L_1 is equal to $\lambda_k^1 = max(min(\alpha, \lambda_k), min(\beta, \lambda_k^"))$.
Let $L_2 = Reduction(\alpha \wedge L' \vee \beta \wedge L")$. According to the definition of the reduction, the possibility of getting a utility degree $u_k \in U$ from L_2 is equal to $\lambda_k^2 = max(min(\alpha, \lambda_k^{'}), min(\beta, \lambda_k^"))$.

Weak monotonicity of $\geq_{L\Pi}$

Suppose that $L \geq_{L\Pi} L'$, i.e. $\Pi(L \geq L') \geq \Pi(L' \geq L)$. Consider the set of utility degrees receiving a possibility equal to 1 in $L : U = \{u_i, \lambda_i = 1\}$ and the set of utility degrees receiving a possibility equal to 1 in $L' : U' = \{u_i, \lambda_i^{'} = 1\}$. These sets are not empty since the distributions are normalized. $\Pi(L \geq L') \geq \Pi(L' \geq L)$ if and only if $max_{u \in U} \geq min_{u \in U'}$.
Let $U_1 = \{u_k, \lambda_k = 1\}$ and $U_2 = \{u_k, \lambda_k' = 1\}$.

– If $\alpha = 1$: we have $U \subseteq U_1$ and $U' \subseteq U_2$. Hence $max_{u \in U}$ belongs to U_1 and $min_{u \in U'}$ belongs to U_2, $max_{u \in U_1} \geq max_{u \in U}$ and $min_{u \in U_2} \leq min_{u \in U'}$. Thus $max_{u \in U_1} \geq min_{u \in U_2}$, i.e. $L_1 \geq_{L\Pi} L_2$.

– If $\alpha < 1$ and $\beta = 1$: let u_i be any of the degrees that receive a degree 1 in $L"$. Since $\beta = 1$, u_i belongs to both U_1 and U_2. Thus $\Pi(L_1 \geq L_2) = \Pi(L_2 \geq L_1) = 1$.

So, $L \geq_{L\Pi} L'$ implies that $L_1 \geq_{L\Pi} L_2$, i.e. that $Reduction(\alpha \wedge L \vee \beta \wedge L") \geq_{L\Pi} Reduction(\alpha \wedge L' \vee \beta \wedge L")$. Which means that $L \geq_{L\Pi} L'$ implies that $(\alpha \wedge L \vee \beta \wedge L") \geq_{L\Pi} (\alpha \wedge L' \vee \beta \wedge L")$.

$\Rightarrow$ Weak monotonicity is satisfied by $\geq_{L\Pi}$.

Weak monotonicity of $\geq_{LN}$

Suppose that $L \geq_{LN} L'$, i.e. $N(L \geq L') \geq N(L' \geq L)$. Consider the set of utility degrees receiving a possibility equal to 1 in $L : U = \{u_i, \lambda_i = 1\}$ and the set of utility degrees receiving a possibility equal to 1 in $L' : U' = \{u_i, \lambda_i^{'} = 1\}$. These sets are not empty since the distributions are normalized. $N(L \geq L')$ is equal to zero as soon as $max_{u \in U'} \geq min_{u \in U}$. $N(L \geq L')$ is positive (and $N(L' \geq L)$ is null) iff $min_{u \in U} > max_{u \in U'}$.
Thus $N(L \geq L') \geq N(L' \geq L)$ when either $min_{u \in U} > max_{u \in U'}$ or
$min_{u \in U} \leq max_{u \in U'}$ and $min_{u \in U} \leq max_{u \in U}$.

– If $\beta = 1$: let u_i be any of the degrees that receive a degree 1 in $L"$. Since $\beta = 1$, u_i belongs to both U_1 and U_2. Thus $N(L_1 \geq L_2) = N(L_2 \geq L_1) = 0$.

– If $\beta < 1$ then $\alpha = 1$ and thus $U \subseteq U_1$ and $U_2 \subseteq U_L'$. In particular $min_{u \in U}$ belongs to U_1 and $max_{u \in U'}$ belongs U_2. Thus $\Pi(L_1 \geq L_2) = 1 : N(L_2 \geq L_1) = 0$. This implies

that $N(L_1 \geq L_2) \geq N(L_2 \geq L_1)$.
So, $L \geq_{LN} L'$ implies that $L_1 \geq_{LN} L_2$, i.e. that $Reduction(\alpha \wedge L \vee \beta \wedge L") \geq_{LN} Reduction(\alpha \wedge L' \vee \beta \wedge L")$. Which means that $L \geq_{LN} L'$ implies that $(\alpha \wedge L \vee \beta \wedge L") \geq_{LN} (\alpha \wedge L' \vee \beta \wedge L")$.
$\Rightarrow$ Weak monotonicity is satisfied by $\geq_{LN}$.

Cardinal setting
According to the definition of reduction and in the case of $\otimes = *$, the possibility of getting a utility degree $u_k \in U$ from the lottery L_1 is equal to $\lambda_k^1 = max((\alpha * \lambda_k), (\beta * \lambda_k^{"}))$. Concerning the lottery L_2 we have $\lambda_k^2 = max((\alpha * \lambda'_k), (\beta * \lambda_k^{"}))$.

Note that the reasoning of the proof in the ordinal setting is also valid for the cardinal setting concerning the weak monotonicity of $\geq_{L\pi}$ and $\geq_{LN}$. ■

Algorithm 4.1 is thus sound and complete for $\geq_{L\Pi}$ and $\geq_{LN}$, and provides in polytime any possibilistic decision tree with a strategy optimal w.r.t these criteria ($\geq_{L\pi}$ *and* $\geq_{LN}$).

Proposition 4.5 allows the definition of the following corollary :

Corollary 4.2 *DT-OPT-LN and DT-OPT-L*Π *belong to P.*

It should be noticed that, contrarily to what can be done with the three previous rules, the likely dominance comparison of two lotteries will be reduced to a simple comparison of aggregated values (Line (2)). Anyway, since only one best strategy is looked for, the transitivity of $>_{LN}$ (resp. $>_{L\Pi}$) guarantees the correctness of the procedure - the non transitivity on the indifference is not a handicap when only one among the best strategies is looked for. The difficulty would be raised if we were looking for all the best strategies.

4.5 Order of magnitude expected utility (OMEU)

We shall now define kappa decision trees : for any $C_i \in \mathcal{C}$ the uncertainty pertaining to the more or less possible outcomes of each C_i is represented by a kappa degree $\kappa_i(N) = Magnitude(P(N|past(C_i))), \forall N \in Succ(C_i)$ (with the normalization condition that the degree $\kappa = 0$ is given to at least one N in $Succ(C_i)$). According to the interpretation of kappa ranking in terms of order of magnitude of probabilities, the product of infinitesimal the conditional probabilities along the paths lead to a sum of the kappa levels. Hence the following principle of reduction of the kappa lotteries :

$$Reduction(\langle \kappa_1/L_1, \ldots, \kappa_m/L_m \rangle =$$
$$\langle \min_{j=1..m} (\kappa_1^j + \kappa_j)/u_1, \ldots, \min_{j=1..m} (\kappa_n^j + \kappa_j)/u_n \rangle \quad (4.5)$$

Like qualitative utilities and possibilistic likely dominance rule, OMEU satisfies the weak monotonicity principle :

Proposition 4.6 *$\succeq_{OMEU}$ satisfies the weak monotonicity property.*

Proof. [Proof of Proposition 4.6]
Consider any tree L, L' and $L"$ be 3 kappa lotteries. We can suppose without loss of generality that they are in reduced form, i.e. that :
$L = \langle \kappa_1/\mu_1, \ldots, \kappa_n/\mu_n \rangle$,
$L' = \langle \kappa'_1/\mu_1, \ldots, \kappa'_n/\mu_n \rangle$ and
$L" = \langle \kappa_1"/\mu_1, \ldots, \kappa_n"/\mu_n \rangle$. It holds that : $OMEU(L) = \min_{i=1,n}\{\kappa_i + u_i\}$ and $OMEU(L') = \min_{i=1,n}\{\kappa'_i + u_i\}$.
Let $L_1 = Reduction(\langle \alpha/L, \beta/L" \rangle$. According to the reduction definition, the kappa ranking of utility degree $u_k \in U$ from L' is equal to : $\kappa_k = \min((\alpha + \kappa_k), (\beta + \kappa_k"))$.
Thus : $OMEU(L_1) = \min_{i=1..n} min[(\kappa_i + \alpha), (\kappa_i" + \beta)] + u_i$.

Similarly, let $L_2 = Reduction(\langle \alpha/L', \beta/L" \rangle$.
It holds that $OMEU(L_2) = min_{i=1..n} min[(\kappa'_i + \alpha), (\kappa_i" + \beta)] + u_i$.
Suppose that $L \geq_{OMEU} L'$, i.e. that $\min\limits_{i=1..n}\{\kappa_i + u_i\} \leq \min\limits_{i=1..n}\{\kappa'_i + u_i\}$.
Then $\min\limits_{i=1..n}\{\kappa_i + u_i\} + \alpha \leq \min\limits_{i=1..n}\{\kappa'_i + u_i\} + \alpha$.
Then $\min\limits_{i=1..n}\{\kappa_i + u_i + \alpha\} \leq \min\limits_{i=1..n}\{\kappa'_i + u_i + \alpha\}$.
As a consequence, we get :
$\min(\min\limits_{i=1..n}\{\kappa_i" + u_i + \beta\}, \min\limits_{i=1..n}\{\kappa_i + u_i + \alpha\}) \leq$
$\min(\min\limits_{i=1..n}\{\kappa_i" + u_i + \beta\}, \min\limits_{i=1..n}\{\kappa'_i + u_i + \alpha\})$.
By associativity of the min operation, we get :
$\min\limits_{i=1..n} \min(\{\kappa_i" + u_i + \beta\}, \{\kappa_i + u_i + \alpha\}) \leq$
$\min\limits_{i=1..n} \min(\{\kappa_i" + u_i + \beta\}, \{\kappa'_i + u_i + \alpha\})$.
Hence :
$\min\limits_{i=1..n} min[(\kappa_i + \alpha), (\kappa_i" + \beta)] + u_i \leq \min\limits_{i=1..n} min[(\kappa'_i + \alpha), (\kappa_i" + \beta)] + u_i$.

That is to say that $OMEU(L_1) \leq OMEU(L_2)$.

We have shown that $L \geq_{OMEU} L'$ implies that $Reduction(\langle \alpha/L, \beta/L" \rangle) \geq_{OMEU} Reduction(\langle \alpha/L', \beta/L" \rangle)$.
Thus $L \geq_{OMEU} L'$ implies that $\langle \alpha/L, \beta/L" \rangle \geq_{OMEU} \langle \alpha/L', \beta/L" \rangle$. ■

As a consequence of Proposition 4.6, dynamic programming (Algorithm 4.1) is appropriate for the optimization of Order of Magnitude Expected Utility. We can then give the following corollary :

Corollary 4.3 *DT-OPT-OMEU belongs to P.*

4.6 Possibilitic Choquet integrals (Ch_Π and Ch_N)

Contrary to qualitative utilities, binary possibilistic utility and likely dominance, the situation is much lesser comfortable when the aim is to optimize a possibilistic Choquet integral (either Ch_N or Ch_Π). The point is that the possibilistic Choquet integrals (as many other Choquet integrals) do not satisfy the monotonicity principle in both ordinal and numerical settings as illustrated by counter example 4.1 and 4.2, respectively.

Counter Example 4.1 (Ordinal setting)

– *Necessity-based Choquet integrals :*

Let us consider these three possibilistic lotteries
$L = \langle 0.2/0, 0.5/0.51, 1/1 \rangle$, $L' = \langle 0.1/0, 0.6/0.5, 1/1 \rangle$ *and* $L" = \langle 0.01/0, 1/1 \rangle$.
$L_1 = \langle \alpha/L, \beta/L" \rangle$ *and* $L_2 = \langle \alpha/L', \beta/L" \rangle$, *with* $\alpha = 0.55$ *and* $\beta = 1$. *Using Equation 4.2 we have :* $Reduction(L_1) = \langle 0.2/0, 0.5/0.51, 1/1 \rangle$ *and* $Reduction(L_2) = \langle 0.1/0, 0.55/0.5, 1/1 \rangle$.
Computing $Ch_N(L) = 0.653$ *and* $Ch_N(L') = 0.650$ *we get* $L \geq_{Ch_N} L'$.
But $Ch_N(Reduction(L_1)) = 0.653 < Ch_N(Reduction(L_2)) = 0.675$, *i.e.* $\langle \alpha/L, \beta/L" \rangle <_{Ch_N} \langle \alpha/L', \beta/L" \rangle$ *:*
$\rightarrow$ *This contradicts the monotonicity property.*

– *Possibility-based Choquet integrals :*

Let us consider these three lotteries
$L = \langle 1/0, 0.5/0.51, 0.2/1 \rangle$, $L' = \langle 1/0, 0.6/0.5, 0.1/1 \rangle$ *and* $L" = \langle 1/0, 0.49/0.51 \rangle$.
$L_1 = \langle \alpha/L, \beta/L" \rangle$ *and* $L_2 = \langle \alpha/L', \beta/L" \rangle$, *with* $\alpha = 1$ *and* $\beta = 0.55$. *Using Equation 4.2 we have :* $Reduction(L_1) = \langle 1/0, 0.5/0.51, 0.2/1 \rangle$ *and*
$Reduction(L_2) = \langle 1/0, 0.6/0.5, 0.49/0.51, 0.1/1 \rangle$.
Computing $Ch_\Pi(L) = 0.353$ *and* $Ch_\Pi(L') = 0.350$ *we get* $L >_{Ch_\Pi} L'$.
But $Ch_\Pi(Reduction(L_1)) = 0.3530 < Ch_\Pi(Reduction(L_2)) = 0.3539$, *i.e.* $\langle \alpha/L, \beta/L" \rangle <_{Ch_\Pi}$

$\langle \alpha/L', \beta/L" \rangle$:
$\rightarrow$ *This contradicts the monotonicity property.*

Counter Example 4.2 (Numerical setting)

– *Necessity-based Choquet integrals :*

Let us consider these three lotteries $L = \langle 0.2/0, 0.5/0.51, 1/1 \rangle$, $L' = \langle 0.1/0, 0.6/0.5, 1/1 \rangle$ *and* $L" = \langle 0.01/0, 1/1 \rangle$.
$L_1 = \langle \alpha/L, \beta/L" \rangle$ *and* $L_2 = \langle \alpha/L', \beta/L" \rangle$, *with* $\alpha = 0.55$ *and* $\beta = 1$. *Using equation 4.2 we have :* $Reduction(L_1) = \langle 0.11/0, 0.275/0.51, 1/1 \rangle$ *and* $Reduction(L_2) = \langle 0.055/0, 0.33/0.5, 1/1 \rangle$.
Computing $Ch_N(L) = 0.653$ *and* $Ch_N(L') = 0.650$ *we get* $L \geq_{Ch_N} L'$.
But $Ch_N(Reduction(L_1)) = 0.809 < Ch_N(Reduction(L_2) = 0.45$,
i.e. $\langle \alpha/L, \beta/L" \rangle <_{Ch_N} \langle \alpha/L', \beta/L" \rangle$:
$\rightarrow$ *This contradicts the monotonicity property.*

– *Possibility based Choquet integrals :*

Let us consider these three lotteries $L = \langle 1/0, 0.5/0.51, 0.2/1 \rangle$, $L' = \langle 1/0, 0.6/0.5, 0.1/1 \rangle$ *and* $L" = \langle 1/0, 0.49/0.51 \rangle$.
$L_1 = \langle \alpha/L, \beta/L" \rangle$ *and* $L_2 = (\langle \alpha/L', \beta/L" \rangle$, *with* $\alpha = 1$ *and* $\beta = 0.55$. *Using Equation 4.2 we have :* $Reduction(L_1) = \langle 1/0, 0.5/0.51, 0.2/1 \rangle$ *and* $Reduction(L_2) = \langle 1/0, 0.6/0.5, 0.26/0.51, 0.1/1 \rangle$.
Computing $Ch_\Pi(L) = 0.353$ *and* $Ch_\Pi(L') = 0.350$ *we get* $L >_{Ch_\Pi} L'$.
But $Ch_\Pi(Reduction(L_1)) = 0.3530 < Ch_\Pi(Reduction(L_2)) = 0.3516$, *i.e.* $\langle \alpha/L, \beta/L" \rangle <_{Ch_\Pi} \langle \alpha/L', \beta/L" \rangle$:
$\rightarrow$ *This contradicts the monotonicity property.*

Based on these counterexamples, it is clear that dynamic programming cannot be applied for DT-OPT-Ch_N and DT-OPT-Ch_Π problems. In Proposition 4.7, we show that these problems are in fact NP-hard.

Proposition 4.7 *DT-OPT-*Ch_N *and DT-OPT-*Ch_Π *are NP-hard.*

To present the proof of this proposition, we will use the properties of possibilistic Choquet integrals developed in Chapter 2 (Propositions 1 and 2) for the case of Ch_N and a part of the work of [39] for the case of Ch_Π.

The hardness of the problem is obtained by a polynomial reduction from a $3SAT$ problem to DT-OPT-Ch_N (resp. DT-OPT-Ch_Π) as shown in the following proof.

Algorithm 4.2: Necessity-based transformation

Data: A CNF $Cl = \{Cl_1, \ldots, Cl_m\}$ on $X = \{X_1, \ldots, X_n\}$
Result: Possibilistic decision tree $\Pi\mathcal{T}$
begin
  Fix $\epsilon \in [0, 1]$ such that $\epsilon^n < 1/2$
  foreach $X_i \in X$ **do**
    $u_{x_i} \leftarrow 2(n-i)+1$
    $\lambda_{x_i} \leftarrow \epsilon^{i+1}$
    $u_{\neg x_i} \leftarrow 2(n-i)+2$
    $\lambda_{\neg x_i} \leftarrow \epsilon^i$
  Let $u_\top \leftarrow (2 * n) + 1$
  Create a decision node D_0 as the root of $\Pi\mathcal{T}$
  Create a chance node H as the unique child of D_0
  foreach $X_i \in X$ **do**
    Create a decision node D_{X_i} with two children C_{x_i} and $C_{\neg x_i}$:
    C_{x_i} is the simple lottery $\langle 1/u_\top, \lambda_{x_i}/u_{x_i} \rangle$
    $C_{\neg x_i}$ is the simple lottery $\langle 1/u_\top, \lambda_{\neg x_i}/u_{\neg x_i} \rangle$
    Add D_{X_i} to the children on H, with a possibility degree equal to 1
  foreach $Cl_i = \{l_1, l_2, l_3\} \in Cl$ **do**
    Create a decision node D_{Cl_i} with as 3 children $C^i_{l_1}$, $C^i_{l_2}$, $C^i_{l_3}$
    $C^i_{l_j}$ is the simple lottery $\langle 1/u_\top, \lambda_{l_1}/u_{l_1} \rangle$
    $C^i_{l_j}$ is the simple lottery $\langle 1/u_\top, \lambda_{l_2}/u_{l_2} \rangle$
    $C^i_{l_j}$ is the simple lottery $\langle 1/u_\top, \lambda_{l_3}/u_{l_3} \rangle$
  Add D_{Cl_i} to the children on H, with a possibility degree equal to 1
  return $\Pi\mathcal{T}$
end

Algorithm 4.3: Possibility-based transformation

Data: 3_SAT problem
Result: Possibilistic decision tree $\Pi\mathcal{T}$
begin
  Fix ϵ such that $0 < \epsilon < 1$
  Create a decision tree D_0 as the root of $\Pi\mathcal{T}$
  Create a chance node H
  Add H as a child of D_0
  $T_0 \leftarrow CreateTN(1, 0)$
  foreach $X_i \in X$ **do**
    Create a decision node D_{X_i}
    Add a chance node C_{x_i} as a child of D_{X_i}
    Associate to C_{x_i} the lottery $\langle 1/0 \rangle$
    Add a chance node $C_{\neg x_i}$ as a child of D_{X_i}
    Associate to $C_{\neg x_i}$ the lottery $\langle 1/0 \rangle$
  foreach $Cl_i \in Cl$ **do**
    foreach *literal* $l_j \in Cl_i$ **do**
      Associate to C_{l_j} the lottery $< \epsilon^i / \Sigma_{k=0}^{i-1} 10^k >$
  return $\Pi\mathcal{T}$
end

Proof. [Proof of Proposition 4.7]
We first prove that DT-OPT-Ch_N (resp DT-OPT-Ch_Π) belongs to NP class.

Membership to NP

The membership of DT-OPT-Ch_N (resp. DT-OPT-Ch_Π) to NP is straightforward. In fact, there is a polynomial algorithm for the determination of an optimal strategy w.r.t Ch_N (resp. Ch_Π) in a possibilistic decision tree by *an oracle machine*. This algorithm will guess a strategy δ for the decision tree that will be reduced into a lottery L. This lottery will be evaluated w.r.t the decision criterion i.e. $Ch_N(L)$ (resp. $Ch_\Pi(L)$) will be computed. According to the Definition 4.1, the final step of the algorithm is to check that $Ch_N(L) \geq \alpha$ (resp. $Ch_\Pi(L) \geq \alpha$).

Since the reduction operation is linear in the size of the compound lottery and the computation of the Necessity-based Choquet value (resp. the Possibility-based Choquet value) is linear in the number of utility levels in the utility scale, the full procedure is polynomial. Hence $DT - OPT - Ch_N$ (resp. $DT - OPT - Ch_\Pi$) belongs to NP.

NP Hardness of DT-OPT-Ch_N

The Hardness of the problem is obtained by the following polynomial reduction from a $3SAT$ to DT-OPT-Ch_N. A $3SAT$ problem is a set of a 3 CNF on $X = \{x_1, \ldots, x_n\}$ which represents also the set of literals $L = \{l_1, \ldots, l_n\}$. The set of clauses is denoted by $Cl = \{Cl_1, \ldots, Cl_m\}$ where each clause Cl_i is defined by $Cl_i = \{l_i^1, l_i^2, l_i^3\}$.

The principle of the transformation is as follows :

– For each literal $l \in L$ we define a utility u_l and a possibility degree λ_{x_i}. We also define a utility degree $u_\top$ that will be greater than all u_l.

– For each $x_i \in X$ we associate a decision node D_{x_i} with two chance nodes $C_{x_i} =< 1/u_\top, \lambda_{x_i}/u_{x_i} >$ and $C_{\neg x_i} =< 1/u_\top, \lambda_{\neg x_i}/u_{\neg x_i} >$ to $\neg x_i$ as children. The first one represents the choice x_i and the second one the choice $\neg x_i$.

– For each $Cl^i = \{l_1^i, l_2^i, l_3^i\} \in Cl$ we define a decision node D_{Cl_i} with three chance nodes as children :

$C_{l_1}^i =< 1/u_\top, \lambda_{l_1}/u_{l_1} >$ (meaning that the satisfaction of the clause is ensured by the choice l_1),

$C_{l_2}^i =< 1/u_\top, \lambda_{l_2}/u_{l_2} >$ (meaning that the satisfaction of the clause is ensured by the choice l_2)

$C_{l_3}^i =< 1/u_\top, \lambda_{l_3}/u_{l_3} >$ (meaning that the satisfaction of the clause is ensured by the choice l_3).

When selecting a chance node for D_{Cl_i}, a strategy specifies how it intends to satisfy clause Cl_i.

This reduction, outlined in Algorithm 4.2, is performed in $O(m + n)$. In fact, the decision tree contains $m+n+1$ decision nodes, $3m+2n+1$ chance nodes and $(3m+2n)\times 2$ leaves.

A strategy δ can select the literals in a consistent manner (in this case, if l is chosen for X_i, $\neg l$ is never chosen for a D_{Cl_i}) or in a contradictory manner (i.e. δ selects l in some decision node in the tree and $\neg l$ for some others). By construction, there is a bijection between the non contradictory strategies, if any, and the models of the formula.

The simple lottery equivalent to a strategy δ is the following : $\pi(\top) = 1$, $\pi(u_l) = \lambda_l$ if literal l is chosen for some decision node, $\pi(u_l) = 0$ otherwise.

– The set of simple lotteries equivalent to contradictory strategies is included in L_{NC} s.t :

$\mathcal{L}_{NC} = \{L : \pi_L(u_\top) = 1, \forall l \in L, \pi_L(u_l) \in \{0, \lambda_l\} \text{ and } \min(\pi_L(u_l), \pi_L(u_{\neg l})) = 0\}$.

– The set of simple lotteries equivalent to contradictory strategies is included in L_C

s.t :
$\mathcal{L}_C = \{L : \pi_L(u_\top) = 1, \forall l \in L, \pi_L(u_l) \in \{0, \lambda_l\}, \exists l \in L \text{ s.t. } \min(\pi_L(u_l), \pi_L(u_{\neg l})) \neq 0\}$.
The principle of the proof is to set the values of the λ_l's and the u_l's in such a way that the Choquet value of the worst of the non contradictory lotteries is greater than the Choquet value of the best contradictory lottery. To this extend, we choose an $\epsilon \in [0, 1]$ such that $\epsilon^n < 1/2$. Then we set $\lambda_{x_i} = \epsilon^{i+1}$, $u_{\neg x_i} = 2(n-i)+2$, $\lambda_{\neg x_i} = \epsilon^i$, $u_\top = (2*n)+1$.

It holds that :

– The worst non contradictory lottery in $\mathcal{L}_{NC}$, denoted by $L^{\downarrow}_{NC}$, is such as all the positive literals are possible and the possibility of any negative literal is equal to 0 i.e.

$$L^{\downarrow}_{NC} = \langle \lambda_{x_n}/u_{x_n}, \ldots, \lambda_{x_1}/u_{x_1}, 1/u_\top \rangle$$

(for the sake of simplicity we omitted terms where possibility degrees are equal to 0). This holds since according to the proposed codification, positive literal have always a utility lower than their negative version.

– The best contradictory lottery in $\mathcal{L}_C$, denoted by $L^{\uparrow}_C$, is such as all negative literals are possible and the possibility of any positive literal is equal to 0 except for x_1 (the less valuable positive literal) i.e.

$$L^{\uparrow}_C = \langle \lambda_{\neg x_n}/u_{\neg x_n}, \ldots, \lambda_{x_1}/u_{x_1}, \lambda_{\neg x_1}/u_{\neg x_1}, 1/u_\top \rangle$$

(terms with 0 degrees are omitted). This holds since (i) according to the proposed codification negative literals always have a utility greater than their positive version and (ii) the less the number of utilities in the lottery receiving a non negative possibility degree, the greater the Choquet value (Proposition 2.2).

– Considering $L^{\downarrow}_{NC}$, the utilities that receive a positive degree of possibility are, by increasing order :
$u_{x_n} < u_{x_{n-1}} < \cdots < u_{x_1} < u_\top$ (all $\neg x_i$, receive a possibility degree equal to 0). Hence :

$$Ch_N(L^{\downarrow}_{NC}) = u_{x_n} + (u_{x_{n-1}} - u_{x_n})(1 - \lambda_{x_n}) + (u_{x_{n-2}} - u_{x_{n-1}})(1 - max(\lambda_{x_n}, \lambda_{x_{n-1}}))$$
$$+ \cdots + (u_{x_1} - u_{x_2})(1 - max(\lambda_{x_n}, \ldots, \lambda_{x_2})) + (u_\top - u_{x_1})(1 - max(\lambda_{x_n}, \ldots, , \lambda_{x_2}, \lambda_{x_1}))$$
$$= 1 + 2(1 - \lambda_{x_n}) + 2(1 - \lambda_{x_{n-1}}) + \cdots + 2(1 - \lambda_{x_1})$$
$$= 2n + 1 - 2(\lambda_{x_n} + \cdots + \lambda_{x_1})$$ and :

– Considering $L^{\uparrow}_C$, the utilities that receive a positive degree of possibility are, by increasing order : $u_{\neg x_n} < u_{\neg x_{n-1}} < \cdots < u_{\neg x_2} < u_{x_1} < u_{\neg x_1} < u_\top$ (all x_i, $i > 1$ receive a possibility degree equal to 0). Hence :

$Ch_N(L_C^{\uparrow}) = u_{\neg x_n} + (u_{\neg x_{n-1}} - u_{\neg x_n})(1 - \lambda_{\neg x_n})$
$+(u_{\neg x_{n-2}} - u_{\neg x_{n-1}})(1 - max(\lambda_{\neg x_n}, \lambda_{\neg x_{n-1}})) + \dots$
$+(u_{\neg x_2} - u_{\neg x_3})(1 - max(\lambda_{\neg x_n}, \dots, \lambda_{\neg x_3})) + (u_{x_1} - u_{\neg x_2})(1 - max(\lambda_{\neg x_n}, \dots, \lambda_{\neg x_2}))$
$+(u_{\neg x_1} - u_{x_1})(1 - max(\lambda_{\neg x_n}, \dots, \lambda_{\neg x_2}, \lambda_{x_1}))$
$+(u_{\top} - u_{\neg x_1})(1 - max(\lambda_{\neg x_n}, \dots, \lambda_{\neg x_2}, \lambda_{x_1}, \lambda_{\neg x_1}))$
$= 2 + 2(1 - \lambda_{\neg x_n}) + 2(1 - \lambda_{\neg x_{n-1}}) + \dots + 2(1 - \lambda_{\neg x_3})$
$+(1 - \lambda_{\neg x_2}) + (1 - \lambda_{x_1}) + (1 - \lambda_{\neg x_1})$
$= 2 + 2(1 - \lambda_{\neg x_n}) + 2(1 - \lambda_{\neg x_{n-1}}) + \dots$
$+(1 - \lambda_{\neg x_2}) + (1 - \lambda_{\neg x_2}) = 2(1 - \lambda_{\neg x_n}) + 2(1 - \lambda_{\neg x_{n-1}}) + \dots + 2(1 - \lambda_{\neg x_2}) + (1 - \lambda_{\neg x_1})$
$= 2 + 2(n-1) - 2(\lambda_{\neg x_n} + \dots + \lambda_{\neg x_2}) + 1 - \lambda_{\neg x_1}$
$= 2.n + 1 - 2(\lambda_{\neg x_n} + \dots + \lambda_{\neg x_2}) - \lambda_{\neg x_1}$

It follows that $Ch_N(L_{NC}^{\downarrow}) - Ch_N(L_C^{\uparrow})$
$= 2n + 1 - 2(\lambda_{x_n}, \dots, \lambda_{x_1})$
$= -2n - 1 + 2(\lambda_{\neg x_n}, \dots, \lambda_{\neg x_2}) + \lambda_{\neg x_1}$
$= 2(\lambda_{\neg x_n}, \dots, \lambda_{\neg x_2}) + \lambda_{\neg x_1} - 2(\lambda_{x_n}, \dots, \lambda_{x_1})$
$= \lambda_{\neg x_1} - 2\lambda_{x_n}$(since by definition $\lambda_{\neg x_i} = \lambda_{x_{i-1}}$).
Recall that $\lambda_{\neg x_1} = \epsilon$ and $\lambda_{x_n} = \epsilon^{n+1}$: $Ch_N(L_{NC}^{\downarrow}) - Ch_N(L_C^{\uparrow})$ is equal to $\epsilon - 2.\epsilon^{n+1}$. Since we have chosen ϵ in $[0, 1]$ is such a way that $\epsilon^n < 1/2$, we get $Ch_N(L_{NC}^{\downarrow}) - Ch_N(L_C^{\uparrow}) > 0$. This shows that $Ch_N(L_{NC}^{\downarrow}) > Ch_N(L_C^{\uparrow})$.
Hence the Choquet value of any non contradictory strategy, if such a strategy exists, is greater than $Ch_N(L_C^{\uparrow})$. Moreover, the CNF is consistent iff there exists a non contradictory strategy. Hence, it is consistent iff there exist a strategy with a Choquet value greater than $\alpha = Ch_N(L_C^{\uparrow})$.

NP Hardness of DT-OPT-Ch_Π

The hardness of the problem is proved by a polynomial reduction from $3SAT$ to DT-OPT-Ch_Π. In the following, we will use a constant $0 < \epsilon < 1$. Obviously, $\epsilon^i < 1, i = 1, m$ and $i < j$ implies that $\epsilon^i > \epsilon^j$.
A possibilistic decision tree is built with a root node D_0 having as unique child a chance node that branches on n decision nodes $D_i, i = 1, n$ (with a possibility degree equal to 1 for each). Each D_i must makes a decision on the value on X_i : it has two children, C_{x_i} and $C_{\neg x_i}$, which are chance node. Consider any literal $l \in \{x_1, \dots, x_n, \neg x_1, \dots, \neg x_n\}$ and the corresponding chance node C_l For the purpose of normalization of the possibility distribution, C_l is linked, with a possibility degree equal to 1, to a leave labeled with utility 0. In addition, for any Cl_i in Cl satisfied by l, a leave node labeled by Cl_i is added as a

child of C_l, with a possibility degree equal to ϵ^i and a utility degree equal $\Sigma_{k=0}^{i-1} 10^k$. For Cl_1 (resp. Cl_2, Cl_3,. . .,Cl_m) the associated utility is 1 (resp. 11, 111,. . ., $\underbrace{1..1}_{m\ terms}$)).

This reduction, outlined by Algorithm 4.3, is performed in $O(n + m)$.
One can check that :

- There is a bijection between the interpretation of the CNF and the admissible strategies
- Ch_Π value of a strategy δ is equal to $\sum_{i=1,m,\delta \text{ satisfies } Cl_i} 10^{i-1} * \epsilon^i$

The CNF is consistent iff there exists a strategy that satisfies all clauses. Indeed, its Choquet value will be equal to $\sum_{i=1,m,} 10^{i-1} * \epsilon^i$ which is the greater possible Choquet value. This means that the proposed reduction approach from a $3SAT$ problem to a decision tree ensures that the optimal strategy has the maximal possibility-based Choquet value. ■

Example 4.3 (resp. 4.4) illustrates the polynomial transformation of a $3SAT$ problem to DT-OPT-Ch_N (resp. DT-OPT-Ch_Π) described in the previous proof using the algorithm 4.2 (resp algorithm 4.3).

Example 4.3 *To illustrate the transformation algorithm in the case of necessity-based Choquet integrals, we will consider the case of* $3SAT = ((x_1 \vee x_2 \vee x_3) \wedge (\neg x_1 \vee \neg x_2 \vee \neg x_3))$ *and* $\epsilon = 0.7$*. Using algorithm 4.2, we obtain the decision tree represented in Figure 4.3. Details of possibility distributions and utilities are as follows :*

- *For* x_1 *we have* $u_{x_1} = 2(3-1)+1 = 5$ *and* $\lambda_{x_1} = 0.7^2 = 0.49$.
- *For* $\neg x_1$ *we have* $u_{\neg x_1} = 2(3-1)+2 = 6$ *and* $\lambda_{\neg x_1} = 0.7^1 = 0.7$.
- *For* x_2 *we have* $u_{x_2} = 2(3-2)+1 = 3$ *and* $\lambda_{x_2} = 0.7^3 = 0.343$.
- *For* $\neg x_2$ *we have* $u_{\neg x_2} = 2(3-2)+2 = 4$ *and* $\lambda_{\neg x_2} = 0.7^2 = 0.49$.
- *For* x_3 *we have* $u_{x_3} = 2(3-3)+1 = 1$ *and* $\lambda_{x_3} = 0.7^4 = 0.2401$.
- *For* $\neg x_3$ *we have* $u_{\neg x_3} = 2(3-3)+2 = 2$ *and* $\lambda_{\neg x_3} = 0.7^3 = 0.343$.

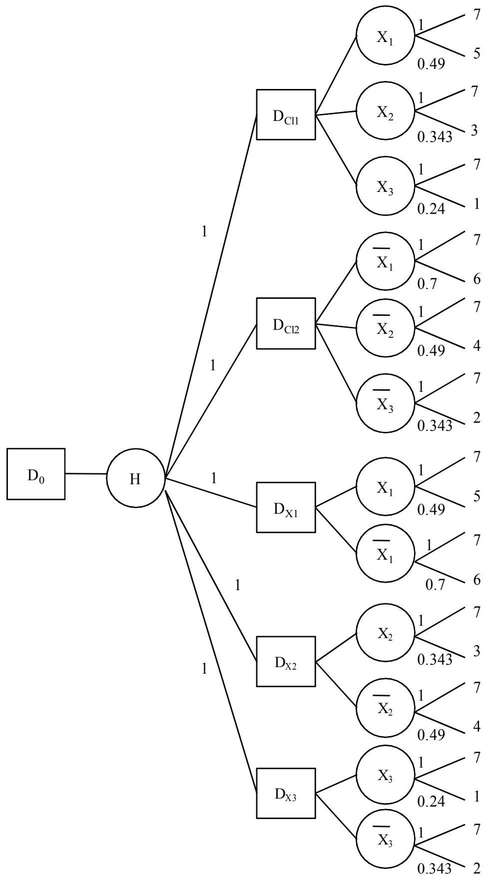

FIGURE 4.3 – Transformation of the CNF $((x_1 \vee x_2 \vee x_3) \wedge (\neg x_1 \vee \neg x_2 \vee \neg x_3))$ to a decision tree with $\epsilon = 0.7$.

Example 4.4 *Let us consider the* $3SAT = ((x_1 \vee x_2 \vee x_3) \wedge (\neg x_1 \vee \neg x_2 \vee \neg x_3))$ *with* $\epsilon = 0.2$. *Using Algorithm 4.3, we obtain the decision tree represented in Figure 4.4 such that :*

- $u(Cl_1) = \sum_{k=0}^{0} 10^k = 1$ *and* $\pi(Cl_1) = (0.2)^1 = 0.2$.
- $u(Cl_2) = \sum_{k=0}^{1} 10^k = 11$ *and* $\pi(Cl_2) = (0.2)^2 = 0.04$.

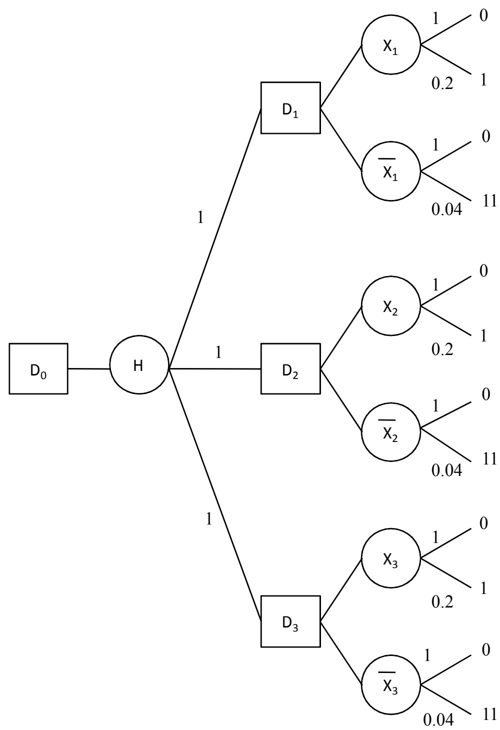

FIGURE 4.4 – Transformation of the CNF $((x_1 \vee x_2 \vee x_3) \wedge (\neg x_1 \vee \neg x_2 \vee \neg x_3))$ with $\epsilon = 0.2$.

4.7 Polynomial cases of possibilistic Choquet integrals

It is important to note that Proposition 4.7 is not true for all possibility distributions. We can in particular distinguish three classes of decision problems (denoted by Binary-Class, Max-Class and Min-Class) where DT-OPT-Ch_Π and DT-OPT-Ch_N are polynomial and dynamic programming can be applied to find the optimal strategy.

4.7.1 Binary possibilistic lotteries

The first polynomial case of possibilistic Choquet integrals (denoted by *Binary-Class*) concerns binary lotteries defined as follows :

Definition 4.2 *Let $U = \{u_1, u_2\}$ be the set of possible utilities composed of only two utilities*

u_1 and u_2 such that $u_1 < u_2$. In the case of Binary-Class, each lottery $L \in \mathcal{L}$ is as follows : $L = \langle \lambda_1/u_1, \lambda_2/u_2 \rangle$.

Proposition 4.8 *DT-OPT-Ch_N (resp. DT-OPT-Ch_Π) is polynomial in the case of Binary-Class.*

Proof. [Proof of Proposition 4.8]
In what follows, we present the proof in numerical setting (the same principle is valid for the ordinal setting).

Necessity-based Choquet integrals

$U = \{u_1, u_2\}$ and $u_1 < u_2$
We will consider the following three possibilistic lotteries :
$L = \langle \lambda_1/u_1, \lambda_2/u_2 \rangle$, $L' = \langle \lambda'_1/u_1, \lambda'_2/u_2 \rangle$ and $L" = \langle \lambda"_1/u_1, \lambda"_2/u_2 >$.
$Ch_N(L) = u_1 + (u_2 - u_1)(1 - \lambda_1)$, $Ch_N(L') = u_1 + (u_2 - u_1)(1 - \lambda'_1)$
$Ch_N(L) \geq Ch_N(L') \Rightarrow \lambda'_1 \geq \lambda_1$.

Note that $L^1 = \alpha L + \beta L"$ and $L^2 = \alpha L' + \beta L"$
– If $\alpha = 1$: $L^1 = \langle max(\lambda_1, \beta\lambda"_1)/u_1, max(\lambda_2, \beta\lambda"_2)/u_2 >$ and
$L^2 = \langle max(\lambda'_1, \beta\lambda"_1)/u_1, max(\lambda'_2, \beta\lambda"_2)/u_2 >$.
$Ch_N(L^1) = u_1 + (u_2 - u_1)(1 - max(\lambda_1, \beta\lambda"_1))$ and
$Ch_N(L^2) = u_1 + (u_2 - u_1)(1 - max(\lambda'_1, \beta\lambda"_1))$. Two cases are possible :
– If $\lambda_1 > \beta\lambda"_1$ and $\lambda'_1 > \beta\lambda"_1 \Rightarrow Ch_N(L^1) > Ch_N(L^2)$
since $\lambda'_1 \geq \lambda_1$
– If $\lambda_1 < \beta\lambda"_1$ and $\lambda'_1 < \beta\lambda"_1 \Rightarrow Ch_N(L^1) = Ch_N(L^2)$
– If $\beta = 1$: $L^1 = \langle \max(\alpha\lambda_1, \lambda"_1)/u_1, \max(\alpha\lambda_2, \lambda"_2)/u_2 >$ and
$L^2 = \langle \max(\alpha\lambda'_1, \lambda"_1)/u_1, \max(\alpha\lambda'_2, \lambda"_2)/u_2 >$.
$Ch_N(L^1) = u_1 + (u_2 - u_1)(1 - \max(\alpha * \lambda_1, \lambda"_1))$ and
$Ch_N(L^2) = u_1 + (u_2 - u_1)(1 - \max(\alpha * \lambda'_1, \lambda"_1))$. Since $\lambda'_1 \geq \lambda_1$ then $Ch_N(L^1) \geq Ch_N(L^2)$.

Possibility-based Choquet integrals

Let three possibilistic lotteries : $L = \langle \lambda_1/u_1, \lambda_2/u_2 >$, $L' = \langle \lambda'_1/u_1, \lambda'_2/u_2 >$ and $L" = \langle \lambda"_1/u_1, \lambda"_2/u_2 >$.
Let $L_1 = \alpha L + \beta L"$ and $L_2 = \alpha L' + \beta L"$ with $max(\alpha, \beta) = 1$.
We have $Ch_\Pi(L) = u_1 + (u_2 - u_1) * \lambda_2$ and $Ch_\Pi(L') = u_1 + (u_2 - u_1) * \lambda'_2$, so if we suppose that $Ch_\Pi(L) > Ch_\Pi(L')$ then $\lambda_2 > \lambda'_2$.

There are two possible cases :

– Case 1 : If $\alpha = 1$

We have $L_1 = L + \beta L"$ and $L_2 = L' + \beta L"$.

$L_1 = \langle \max(\lambda_1, \beta\lambda"_1)/u_1, \max(\lambda_2, \beta\lambda"_2)/u_2 >$, and

$L_2 = \langle \max(\lambda'_1, \beta\lambda"_1)/u_1, \max(\lambda'_2, \beta\lambda"_2)/u_2 >$ so $Ch_\Pi(L_1) = u_1 + (u_2 - u_1) * max(\lambda_2, \beta\lambda"_2)$ and

$Ch_\Pi(L_2) = u_1 + (u_2 - u_1) * max(\lambda'_2, \beta\lambda"_2)$.

Since we have $\lambda_2 > \lambda'_2 \Rightarrow Ch_\Pi(L_1) \geq Ch_\Pi(L_2)$.

– Case 2 : If $\beta = 1$

We have $L_1 = \alpha L + L"$ and $L_2 = \alpha L' + L"$. $L_1 = \langle \max(\alpha\lambda_1, \lambda"_1)/u_1, \max(\alpha\lambda_2, \lambda"_2)/u_2 >$, and $L_2 = \langle \max(\alpha\lambda'_1, \lambda"_1)/u_1, \max(\alpha\lambda'_2, \lambda"_2)/u_2 >$ so $Ch_\Pi(L_1) = u_1 + (u_2 - u_1) * \max(\alpha\lambda_2, \lambda"_2)$ and $Ch_\Pi(L_2) = u_1 + (u_2 - u_1) * \max(\alpha\lambda'_2, \lambda"_2)$.

Since we have $\lambda_2 > \lambda'_2 \Rightarrow Ch_\Pi(L_1) \geq Ch_\Pi(L_2)$.

■

4.7.2 The maximal possibility degree is affected to the maximal utility

The second polynomial case of possibilistic Choquet integrals (denoted by *Max-Class*) concerns possibilistic lotteries where the maximal possibility degree namely 1 is affected to the maximal utility in the lottery. This class is defined as follows :

Definition 4.3 *Let $U = \{u_1, \ldots, u_n\}$ be the set of possible utilities where u_{max} is the maximal utility in a possibilistic lottery L such that $u_{max} \leq u_n$. In the case of Max-Class, each lottery $L \in \mathcal{L}$ is as follows : $L = \langle \lambda_1/u_1, \ldots, 1/u_{max} \rangle$.*

Proposition 4.9 *DT-OPT-Ch_Π (resp. DT-OPT-Ch_N with $\alpha = 1$) is polynomial in the case of Max-Class.*

Proof. [Proof of Proposition 4.9]

In what follows, we present the proof in numerical setting (the same principle is valid for the ordinal setting).

Necessity-based Choquet integrals

We will consider the case where $\alpha = 1$

– Case 1 : Let us consider three lotteries L, L' and L" having the same maximal utility u_n :
$L = \langle\lambda_1/u_1, \ldots, 1/u_n\rangle, L' = \langle\lambda'_1/u_1, \ldots, 1/u_n\rangle$ and
$L" = \langle\lambda"_1/u_1, \ldots, 1/u_n\rangle$. Then,
$Ch_N(L) = u_1 + \cdots + (u_n - u_{n-1})(1 - \max(\lambda_1, \ldots, \lambda_{n-1}))$
$Ch_N(L') = u_1 + \cdots + (u_n - u_{n-1})(1 - \max(\lambda'_1], \ldots, \lambda'_{n-1}))$

$L_1 = \langle\alpha/L, \beta/L"\rangle$ and $L_2 = \langle\alpha/L', \beta/L"$ are two compound lotteries, we have :
$L_1 = \langle\max(\lambda_1, \beta\lambda"_1)/u_1, \ldots, 1/u_n\rangle$
$L_2 = \langle\max(\lambda'_1, \beta\lambda"_1)/u_1, \ldots, 1/u_n\rangle$
$Ch_N(L_1) = u_1 + \ldots$
$+ (u_n - u_{n-1})(1 - \max(\max(\lambda_1, \beta\lambda"_1), \ldots, \max(\lambda_{n-1}, \beta\lambda"_{n-1})))$
$Ch_N(L_2) = u_1 + \ldots$
$+ (u_n - u_{n-1})(1 - \max(\max(\lambda'_1, \beta\lambda"_1), \ldots, \max(\lambda'_{n-1}, \beta\lambda"_{n-1})))$
$\Rightarrow If \;\; Ch_N(L) > Ch_N(L') \;\; then \;\; Ch_N(L_1) \geq Ch_N(L_2)$

– Case 2 : L and L' have the same maximal utility denoted by u_{max} and L'' has as maximal utility u_i.
– If $u_{max} > u_i$
$L = \langle\lambda_1/u_1, \ldots, \lambda_i/u_i, \ldots, 1/u_{max}\rangle$
$L' = \langle\lambda'_1/u_1, \ldots, \lambda'_i/u_i, \ldots, 1/u_{max}\rangle$ and
$L" = \langle\lambda"_1/u_1, \ldots, \lambda"_i/u_i, \ldots, 1/u_{max}\rangle$. Then,

$Ch_N(L) = u_1 + \cdots + (u_i - u_{i-1})(1 - \max(\lambda_1, \ldots, \lambda_{i-1})) + \cdots + (u_{max} - u_{max-1})(1 - \max(\lambda_1, \ldots, \lambda_{max-1}))$
$Ch_N(L') = u_1 + \cdots + (u_i - u_{i-1})(1 - \max(\lambda'_1, \ldots, \lambda'_{i-1})) + \cdots + (u_{max} - u_{max-1})(1 - \max(\lambda'_1, \ldots, \lambda'_{max-1}))$
$L" = \langle\lambda"_1/u_1, \ldots, 1/u_i\rangle$.

$L_1 = \langle\alpha/L, \beta/L"\rangle$ and $L_2 = \langle\alpha/L', \beta/L"$ are two compound lotteries, we have :
$L_1 = \langle\max(\lambda_1, \beta\lambda"_1)/u_1, \ldots, 1/u_i, \ldots, 1/u_{max}\rangle$
$L_2 = \langle\max(\lambda'_1, \beta\lambda"_1)/u_1, \ldots, 1/u_i, \ldots, 1/u_{max}\rangle$
$Ch_N(L_1) = u_1 + \ldots$
$+ (u_i - u_{i-1})(1 - \max(\max(\lambda_1, \beta\lambda"_1), \ldots,$
$\max(L[u_{i-1}], \beta L''[u_{i-1}])))$
$Ch_N(L_2) = u_1 + \ldots$
$+ (u_n - u_{n-1})(1 - \max(\max(\lambda'_1, \beta\lambda"_1), \ldots, \max(\lambda'_{i-1}, \beta\lambda"_{i-1})))$

$\Rightarrow If \;\; Ch_N(L) > Ch_N(L') \;\; then \;\; Ch_N(L_1) \geq Ch_N(L_2)$.

– If $u_{max} < u_i$

$L = \langle \lambda_1/u_1, \dots, 1/u_{max}\rangle$,
$L' = \langle \lambda'_1/u_1, \dots, 1/u_{max}\rangle$ and
$L'' = \langle \lambda"_1/u_1, \dots, \lambda"_{max}/u_{max}, \dots, 1/u_i\rangle$. Then,

$Ch_N(L) = u_1 + \dots + (u_{max} - u_{max-1})(1 - \max(\lambda_1, \dots, \lambda_{max-1}))$
$Ch_N(L') = u_1 + \dots + (u_{max} - u_{max-1})(1 - \max(\lambda'_1, \dots, \lambda'_{max-1}))$

$L_1 = \langle \alpha/L, \beta/L"\rangle$ and $L_2 = \langle \alpha/L', \beta/L"$ are two compound lotteries, we have :
$L_1 = \langle \max(\lambda_1, \beta\lambda"_1)/u_1, \dots, 1/u_{max}, \dots, 1/u_i\rangle$
$L_2 = \langle \max(\lambda'_1, \beta\lambda"_1)/u_1, \dots, 1/u_{max}, \dots, 1/u_i\rangle$
$Ch_N(L_1) = u_1 + \dots + (u_{max} - u_{max-1})(1 - \max(\max(\lambda_1, \beta\lambda"_1),$
$\dots, \max(\lambda_{max-1}, \beta\lambda"_{max-1})))$
$Ch_N(L_2) = u_1 + \dots + (u_{max} - u_{max-1})(1 - \max(\max(\lambda'_1, \beta\lambda"_1),$
$\dots, \max(\lambda'_{max-1}], \beta\lambda"_{max-1})))$
$\Rightarrow If \;\; Ch_N(L) > Ch_N(L') \;\; then \;\; Ch_N(L_1) \geq Ch_N(L_2)$

– Case 3 : L and $L"$ have the same maximal utility denoted by u_{max} and L' has as maximal utility u_i.

– If $u_{max} > u_i$

$L = \langle \lambda_1/u_1, \dots \lambda_i/u_i, \dots, 1/u_{max}\rangle$,
$L' = \langle \lambda'_1/u_1, \dots, 1/u_i >$ and
$L" = \langle \lambda"_1/u_1, \dots, \lambda"_i/u_i, \dots, 1/u_{max}\rangle$. Then,

$Ch_N(L) = u_1 + \dots + (u_i - u_{i-1})(1 - \max(\lambda_1, \dots, \lambda_{i-1})) + \dots + (u_{max} - u_{max-1})(1 -$
$\max(\lambda_1, \dots, \lambda_{max-1}))$
$Ch_N(L') = u_1 + \dots + (u_i - u_{i-1})(1 - \max(\lambda'_1, \dots, \lambda'_{i-1}))$.

$L_1 = \langle \alpha/L, \beta/L"\rangle$ and $L_2 = \langle \alpha/L', \beta/L"$ are two compound lotteries, we have :
$L_1 = \langle \max(\lambda_1, \beta\lambda"_1)/u_1, \dots, \max(\lambda_i, \beta\lambda"_i)/u_i, \dots, 1/u_{max}\rangle$
$L_2 = \langle \max(\lambda'_1, \beta\lambda"_1)/u_1, \dots, 1/u_i\rangle$
$Ch_N(L_1) = u_1 + \dots$
$+ (u_i - u_{i-1})(1 - \max(\max(\lambda_1, \beta\lambda"_1), \dots, \max(\lambda_{i-1}, \beta\lambda_{i-1}))) + \dots + (u_{max} -$
$u_{max-1})(1 - \max(\max(\lambda_1, \beta\lambda"_1),$
$\dots, \max(\lambda_{max-1}, \beta\lambda"_{max-1})))$
$Ch_N(L_2) = u_1 + \dots$

$+ (u_i - u_{i-1})(1 - \max(\max(\lambda'_1, \beta\lambda"_1), \dots, \max(\lambda"_{i-1}, \beta\lambda"_{i-1})))$
$\Rightarrow If \;\; Ch_N(L) > Ch_N(L') \;\; then \;\; Ch_N(L_1) \geq Ch_N(L_2).$

– If $u_{max} < u_i$
$L = \langle \lambda_1/u_1, \dots, 1/u_{max} \rangle$,
$L' = \langle \lambda'_1/u_1, \dots, \lambda'_{max}/u_{max}, \dots, 1/u_i \rangle$ and
$L" = \langle \lambda"_1/u_1, \dots, 1/u_{max} \rangle$. Then,

$Ch_N(L) = u_1 + \dots + (u_{max} - u_{max-1})(1 - \max(\lambda_1, \dots, \lambda_{max-1}))$
$Ch_N(L') = u_1 + \dots + (u_{max} - u_{max-1})(1 - \max(\lambda'_1, \dots, \lambda'_{max-1})) + \dots + (u_i - u_{i-1})(1 - \max(\lambda'_1, \dots, \lambda'_{i-1}))$

$L_1 = \langle \alpha/L, \beta/L" \rangle$ and $L_2 = \langle \alpha/L', \beta/L"$ are two compound lotteries, we have :
$L_1 = \langle \max(\lambda_1, \beta\lambda"_1)/u_1, \dots, 1/u_{max} \rangle$
$L_2 = \langle \max(\lambda'_1, \beta\lambda"_1)/u_1, \dots, 1/u_{max}, \dots, 1/u_i \rangle$
$Ch_N(L_1) = u_1 + \dots + (u_{max} - u_{max-1})(1 - \max(\max(\lambda_1, \beta\lambda"_1),$
$\dots, \max(\lambda_{max-1}, \beta\lambda"_{max-1})))$
$Ch_N(L_2) = u_1 + \dots + (u_{max} - u_{max-1})(1 - \max(\max(\lambda'_1, \beta\lambda"_1),$
$\dots, \max(\lambda'_{max_1}, \beta\lambda"_{max-1})))$
$\Rightarrow If \;\; Ch_N(L) > Ch_N(L') \;\; then \;\; Ch_N(L_1) \geq Ch_N(L_2).$

– Case 4 : L has a maximal utility denoted by u_{max} and L' and L'' have the same maximal utility u_i.
– If $u_{max} > u_i$
$L = \langle \lambda_1/u_1, \dots, \lambda_i/u_i, \dots, 1/u_{max} \rangle$,
$L' = \langle \lambda'_1/u_1, \dots, 1/u_i \rangle and$
L"= $\langle \lambda"_1/u_1, \dots, 1/u_i \rangle$. Then,

Ch$_N(L) = u_1 + \dots + (u_i - u_{i-1})(1 - \max(\lambda_1, \dots, \lambda_{i-1})) + \dots + (u_{max} - u_{max-1})(1 - \max(\lambda_1, \dots, \lambda_{max-1}))$
$Ch_N(L') = u_1 + \dots + (u_i - u_{i-1})(1 - \max(\lambda'_1, \dots, \lambda'_{i-1})) \Rightarrow Ch_N(L) > Ch_N(L')$

$L_1 = \langle \alpha/L, \beta/L" \rangle$ and $L_2 = \langle \alpha/L', \beta/L"$ are two compound lotteries, we have :
$L_1 = \langle \max(\lambda_1, \beta\lambda"_1)/u_1, \dots, \max(\lambda_i, \beta)/u_i, \dots, 1/u_{max} \rangle$
$L_2 = \langle \max(\lambda'_1, \beta\lambda"_1)/u_1, \dots, 1/u_i >$
$Ch_N(L_1) = u_1 + \dots$
$+ (u_i - u_{i-1})(1 - \max(\max(\lambda_1, \beta\lambda"_1), \dots, \max(\lambda_{i-1}, \beta) + (u_{max} - u_{max-1})(1 -$

$\max(\max(\lambda_1, \beta\lambda"_1), \ldots, \max(\lambda_{max-1}, \beta\lambda"_{max-1})))$
$Ch_N(L_2) = u_1 + \ldots$
$+ (u_i - u_{i-1})(1 - \max(\max(\lambda'_1, \beta\lambda"_1), \ldots, \max(\lambda'_{i-1}, \beta\lambda"_{i-1})))$
$\Rightarrow Ch_N(L_1) > Ch_N(L_2)$.

– If $u_{max} < u_i$
$L = \langle \lambda_1/u_1, \ldots, 1/u_{max} \rangle$,
$L' = \langle \lambda'_1/u_1, \ldots, \lambda'_{max}/u_{max}, \ldots, 1/u_i \rangle$ and
$L'' = \langle \lambda"_1/u_1, \ldots, \lambda"_{max}/u_{max}, \ldots, 1/u_i \rangle$. Then,

$Ch_N(L) = u_1 + \cdots + (u_{max} - u_{max-1})(1 - \max(\lambda_1, \ldots, \lambda_{max-1}))$
$Ch_N(L') = u_1 + \cdots + (u_{max} - u_{max-1})(1 - \max(\lambda'_1, \ldots, \lambda'_{max-1})) + \cdots + (u_i - u_{i-1})(1 - \max(\lambda'_1, \ldots, \lambda'_{i-1}))$
$Ch_N(L') > Ch_N(L)$

$L_1 = \langle \alpha/L, \beta/L" \rangle$ and $L_2 = \langle \alpha/L', \beta/L"$ are two compound lotteries, we have :
$L_1 = \langle \max(\lambda_1, \beta\lambda"_1)/u_1, \ldots, 1/u_{max} \rangle$
$L_2 = \langle \max(\lambda'_1, \beta\lambda"_1)/u_1, \ldots, \max(\lambda'_{max}, \beta\lambda"_{max}), \ldots, 1/u_i \rangle$
$Ch_N(L_1) = u_1 + \cdots + (u_{max} - u_{max-1})(1 - \max(\max(\lambda_1, \beta\lambda"_1),$
$\ldots, \max(\lambda_{max-1}, \beta\lambda"_{max-1})))$
$Ch_N(L_2) = u_1 + \cdots + (u_{max} - u_{max-1})(1 - \max(\max(\lambda'_1, \beta\lambda"_1),$
$\ldots, \max(\lambda'_{max-1}, \beta\lambda"_{max-1}) + \cdots + (u_i - u_{i-1})(1 - \max(\max(\lambda'_1, \beta\lambda"_1), \ldots, \max(\lambda'_{i-1}, \beta\lambda"_{i-1})))$
$\Rightarrow Ch_N(L_2) > Ch_N(L_1)$.

Possibility-based Choquet integrals

Let three possibilistic lotteries $L = \langle \lambda_1/u_1, \ldots, 1/u_n \rangle$,
$L' = \langle \lambda'_1/u_1, \ldots, 1/u_n \rangle$ and $L" = \langle \lambda"_1/u_1, \ldots, 1/u_n \rangle$. Then,

$Ch_\Pi(L) = u_1 + (u_2 - u_1) * 1 + \cdots + (u_n - u_{n-1}) * 1$
and $Ch_\Pi(L') = u_1 + (u_2 - u_1) * 1 + \cdots + (u_n - u_{n-1}) * 1$.
$\Rightarrow Ch_\Pi(L) = Ch_\Pi(L')$. There are two cases :

– Case 1 : If $\alpha = 1$
$L_1 = \langle 1/L, \beta/L" \rangle$ and $L_2 = \langle 1/L', \beta/L"$ are two compound lotteries, we have :
$L_1 = \langle \max(\lambda_1, \beta\lambda"_1)/u_1, \ldots, \max(1, \beta)/u_n \rangle$, and
$L_2 = \langle \max(\lambda'_1, \beta\lambda"_1)/u_1, \ldots, \max(1, \beta)/u_n \rangle$.

$\Rightarrow Ch_\Pi(L_1) = Ch_\Pi(L_2)$.

– Case 2 : If $\beta = 1$: similar to the first case.

■

4.7.3 The maximal possibility degree is affected to the minimal utility

The third polynomial case of possibilistic Choquet integrals (denoted by *Min-Class*) concerns possibilistic lotteries where the maximal possibility degree namely 1 is affected to the minimal utility in the lottery. This class is defined as follows :

Definition 4.4 *Let $U = \{u_1, \dots u_n\}$ be the set of possible utilities where u_{min} is the minimal utility in a possibilistic lottery such that $u_{min} \leq u_n$. In the case of Min-Class, each lottery $L \in \mathcal{L}$ is as follows : $L = \langle 1/u_{min}, \dots, \lambda_n/u_n \rangle$.*

Proposition 4.10 *DT-OPT-Ch_N is polynomial in the case of Min-Class.*

Proof. [Proof of Proposition 4.10]

In what follows, we present necessary proof in numerical setting (the same principle is valid for the ordinal setting).

We will consider the case where $\alpha = 1$

– Case 1 : L, L' and $L"$ have the same minimal utility u_i.

$L = \langle 1/u_i, \dots, \lambda_n/u_n \rangle \Rightarrow Ch_N(L) = u_i$

$L' = \langle 1/u_i, \dots, \lambda'_n/u_n \rangle \Rightarrow Ch_N(L') = u_j$

$L" = \langle 1/u_i, \dots, \lambda"_n/u_n \rangle$,

$L_1 = \langle 1/u_i, \dots, \max(\lambda_n, \beta\lambda"_n)/u_n \rangle$,

$L_2 = \langle 1/u_i, \dots, \max(\lambda'_n, \beta\lambda"_n)/u_n \rangle$

$\Rightarrow Ch_N(L_1) = Ch_N(L_2) = u_i$.

– Case 2 : L and L' have the same minimal utility u_i but not $L"$.

$L = \langle 1/u_i, \dots, \lambda_n/u_n \rangle$ and $L' = \langle 1/u_i, \dots, \lambda'_n/u_n \rangle$

$\Rightarrow Ch_N(L) = Ch_N(L') = u_i$ and $L" = \langle 1/u_j, \dots, \lambda"_n/u_n \rangle$.

– If $L" = \langle 1/u_j, \dots, \lambda"_n/u_n \rangle$ (i.e. $u_i < u_j$)

$L_1 = \langle 1/u_i, \dots, \max(\alpha\lambda'_n, \beta\lambda"_n)/u_n \rangle$,

$L_2 = \langle 1/u_i, \dots, \max(\alpha\lambda'_n, \beta\lambda"_n)/u_n \rangle$

$\Rightarrow Ch_N(L_1) = Ch_N(L_2) = u_i$.

– If $L" = \langle 1/u_j, \dots, \lambda"_i/u_i, \dots, \lambda"_n/u_n \rangle$ (i.e. $u_i > u_j$)

$L_1 = \langle \beta/u_j, \dots, \max(\alpha\lambda_n, \beta\lambda"_n)/u_n >$,

$L_2 = \langle \beta/u_j, \ldots, \max(\alpha\lambda'_n, \beta\lambda"_n)/u_n >$
$\Rightarrow Ch_N(L_1) = Ch_N(L_2)$.

– Case 3 : L and L' don't have the same minimal utility.
$L = \langle 1/u_i, \ldots, \lambda_n/u_n\rangle$ and $L' = \langle 1/u_j, \ldots, \lambda'_n/u_n\rangle$
$\Rightarrow Ch_N(L) = u_i$ and $Ch_N(L') = u_j$.

– If $L" = \langle 1/u_i, \ldots, \lambda"_n/u_n\rangle$ (i.e. $u_i > u_j$)
$L_1 = \langle 1/u_i, \ldots, \max(\alpha\lambda'_n, \beta\lambda"_n)/u_n\rangle$,
$L_2 = \langle 1/u_j, \ldots, \max(\alpha\lambda'_i, \beta\lambda"_i)/u_i, \ldots, \max(\alpha\lambda'_n, \beta\lambda"_n)/u_n\rangle$
$= \langle 1/u_j, \ldots, \max(\lambda'_i, \beta)/u_i, \ldots, \max(\alpha\lambda'_n, \beta\lambda"_n)/u_n\rangle$
$\Rightarrow Ch_N(L_1) = u_i > Ch_N(L_2) = u_j$.

– If $L" = \langle 1/u_i, \ldots, \lambda"_j/u_j, \ldots, \lambda"_n/u_n\rangle$ (i.e. $u_i < u_j$)
$L^1 = \langle 1/u_i, \ldots, \max(\alpha\lambda'_n, \beta\lambda"_n)/u_n >$,
$L^2 = \langle \beta/u_i, \beta\lambda"_{i+1}/u_{i+1}, \ldots, \beta\lambda"_{j-1}/u_{j-1}, 1/u_j,$
$\ldots, \max(\alpha\lambda'_n, \beta\lambda"_n)/u_n >$
$\Rightarrow Ch_N(L_1) = u_i$
$Ch_N(L_2) = u_i + (u_{i+1} - u_i)(1-\beta) + (u_{i+2} - u_{i+1})(1 - max(\beta, \beta\lambda"_{i+1})) + \cdots + (u_j - u_{j-1})(1 - \max(\beta, \ldots, \beta\lambda"_{j-1})) + (u_{j+1} - u_j)(1 - \max(\beta, \ldots, 1)) + \cdots + (u_n - u_{n-1})(1 - max(\beta, \ldots, 1))$
$= u_i + (u_{i+1} - u_i)(1-\beta) + (u_{i+2} - u_{i+1})(1 - max(\beta, \beta\lambda"_{i+1})) + \cdots + (u_j - u_{j-1})(1 - \max(\beta, \ldots, \beta\lambda"_{j-1})) + (u_{j+1} - u_j)(1-1) + \cdots + (u_n - u_{n-1})(1-1)$
$= u_i + (u_{i+1} - u_i)(1-\beta) + (u_{i+2} - u_{i+1})(1 - \max(\beta, \beta\lambda"_{i+1})) + \cdots + (u_j - u_{j-1})(1 - \max(\beta, \ldots, \beta\lambda"_{j-1}))$
Let $X = (u_{i+1} - u_i)(1-\beta) + (u_{i+2} - u_{i+1})(1 - max(\beta, \beta L''[u_{i+1}])) + \cdots + (u_j - u_{j-1})(1 - \max(\beta, \ldots, \beta\lambda"_{j-1}))$
$X \geq 0$ since $(u_{i+1} - u_i)(1-\beta) \geq 0$
$(u_{i+2} - u_{i+1})(1 - \max(\beta, \beta\lambda"_{i+1}) \geq 0$
$\ldots$
$(u_j - u_{j-1})(1 - \max(\beta, \ldots, \beta\lambda"_{j-1})) \geq 0$
$\Rightarrow Ch_N(L_2) = u_i + X$ (s.t. $X \geq 0$)
$\Rightarrow Ch_N(L_1) \leq Ch_N(L_2)$

– Case 4 : L and L' have not the same minimal utility i.e.
$L = \langle 1/u_j, \ldots, \lambda_n/u_n\rangle$ and
$L' = \langle 1/u_i, \ldots, \lambda'_n/u_n\rangle$
$\Rightarrow Ch_N(L) = u_j$ and $Ch_N(L') = u_i$

– If $L" = \langle 1/u_i, \ldots, \lambda"_n/u_n >$ (i.e. $u_i > u_j$) : similar to the first item in the previous case

– If $L" = \langle 1/u_i, \ldots, \lambda"_j/u_j, \ldots, \lambda"_n/u_n\rangle$ (i.e. $u_i < u_j$) : similar to the second item in the previous case

– Case 5 : The minimal utility in L (resp. $L', L"$) is u_i (resp. u_j, u_k).

– If $u_i > u_j > u_k$,

$L = \langle 1/u_i, \ldots, \lambda_n/u_n\rangle$,

$L' = \langle 1/u_j, \ldots, \lambda'_i/u_i, \ldots, \lambda'_n/u_n\rangle$

and $L" = \langle 1/u_k, \ldots, \lambda"_j/u_j, \ldots, \lambda"_i/u_i, \ldots, \lambda"_n/u_n\rangle$

$\Rightarrow Ch_N(L) = u_i$ and $Ch_N(L') = u_j, Ch_N(L) > Ch_N(L')$

$L^1 = \langle \beta/u_k, \ldots, \beta\lambda_j/u_j, \ldots, 1/u_i, \ldots, \max(\lambda_n, \beta\lambda"_n)/u_n\rangle$

$L^2 = \langle \beta/u_k, \ldots, 1/u_j, \ldots, 1/u_i, \ldots, \max(\lambda'_n, \beta\lambda"_n)/u_n\rangle$

$Ch_N(L^1) = u_k + (u_{k+1} - u_k)(1-\beta) + \cdots + (u_{i+1} - u_i)(1-1)$

$Ch_N(L^2) = u_k + \cdots + (u_{j+1} - u_j)(1-1) < Ch_N(L^1)$.

– If $u_i > u_k > u_j$

$L = \langle 1/u_i, \ldots, \lambda_n/u_n\rangle$,

$L' = \langle 1/u_j, \ldots, \lambda'_k/u_k, \ldots, \lambda'_i/u_i, \ldots, \lambda'_n/u_n\rangle$

and $L" = \langle 1/u_k, \ldots, \lambda"_i/u_i, \ldots, \lambda"_n/u_n\rangle$

$\Rightarrow Ch_N(L) = u_i$ and $Ch_N(L') = u_j, Ch_N(L) > Ch_N(L')$

$L_1 = \langle \beta/u_k, \ldots, 1/u_i, \ldots, \max(\lambda_n, \beta\lambda"_n)/u_n\rangle$

$L_2 = \langle 1/u_j, \ldots, \max(\lambda'_k, \beta\lambda"_k/u_k, \ldots, \max(\lambda'_n, \beta\lambda"_n)/u_n >$

$Ch_N(L_1) = u_k + \cdots + (u_i - u_{i-1})(1 - \max(\beta, \ldots, \max(\lambda_{i-1}, \beta\lambda"_{i-1})))$

$Ch_N(L_2) = u_j$.

$\Rightarrow Ch_N(L_1) > Ch_N(L_2)$.

– If $u_j > u_i > u_k$

$L = \langle 1/u_i, \ldots, \lambda_j/u_j, \ldots, \lambda_n/u_n\rangle$,

$L' = \langle 1/u_j, \ldots, \lambda'_n/u_n\rangle$

and $L" = \langle 1/u_k, \ldots, \lambda"_i/u_i, \ldots, \lambda"_j/u_j, \ldots, \lambda"_n/u_n\rangle$

$\Rightarrow Ch_N(L) = u_i$ and $Ch_N(L') = u_j, Ch_N(L') > Ch_N(L)$

$L_1 = \langle \beta/u_k, \ldots, 1/u_i, \ldots, \max(\lambda_n, \beta\lambda"_n)/u_n >$

$L_2 = \langle \beta/u_k, \ldots, \beta\lambda"_i/u_i, \ldots, 1/u_j, \ldots, \max(\lambda'_n, \beta\lambda"_n)/u_n >$

$Ch_N(L_1) = u_k + \cdots + (u_{i+1} - u_i)(1-1)$

$Ch_N(L_2) = u_k + \cdots + (u_{i+1} - u_i)(1 - \max(\beta, \ldots, \max(\lambda'_i, \beta\lambda"_i))) + \cdots + (u_{j+1} - u_j)(1-1)$

$Ch_N(L_2) = Ch_N(L^1) + (u_{i+1} - u_i)(1 - \max(\beta, \ldots, \max(\lambda'_i, \beta\lambda"_i))) + \cdots + (u_{j+1} - u_j)(1-1)$

$\Rightarrow Ch_N(L^2) > Ch_N(L^1)$.

– If $u_k > u_i > u_j$

$L = \langle 1/u_i, \ldots, \lambda_k/u_k, \ldots, \lambda_n/u_n\rangle$,

$L' = \langle 1/u_j, \ldots, \lambda'_i/u_i, \ldots, \lambda'_k/u_k, \ldots, \lambda'_n/u_n\rangle$

and $L" = \langle 1/u_k, \ldots, \lambda"_n/u_n\rangle$

$\Rightarrow Ch_N(L) = u_i$ and $Ch_N(L') = u_j, Ch_N(L) > Ch_N(L')$
$L^1 = \langle 1/u_i, \ldots, \max(\lambda_k, \beta)/u_k, \ldots, \max(\lambda_n, \beta\lambda"_n)/u_n \rangle$
$L^2 = \langle 1/u_j, \ldots, \max(\lambda'_n, \beta\lambda"_n)/u_n \rangle$
$Ch_N(L^1) = u_i$
$Ch_N(L^2) = u_j$
$\Rightarrow Ch_N(L^1) > Ch_N(L^2)$.

– If $u_k > u_j > u_i$
$L = \langle 1/u_i, \ldots, \lambda_j/u_j, \ldots, \lambda_k/u_k, \ldots, \lambda_n/u_n \rangle$,
$L' = \langle 1/u_j, \ldots,'_k /u_k, \ldots, \lambda'_n/u_n \rangle$
and $L" = \langle 1/u_k, \ldots, \lambda"_n/u_n \rangle$
$\Rightarrow Ch_N(L) = u_i$ and $Ch_N(L') = u_j, Ch_N(L') > Ch_N(L)$
$L^1 = \langle 1/u_i, \ldots, \lambda_j/u_j, \ldots, \max(\lambda_n, \beta\lambda"_n)/u_n \rangle$
$L^2 = \langle 1/u_j, \ldots, \max(\lambda'_n, \beta\lambda"_n)/u_n \rangle$
$Ch_N(L^1) = u_i$
$Ch_N(L^2) = u_j$
$\Rightarrow Ch_N(L^1) > Ch_N(L^2)$.

■

4.8 Conclusion

In this chapter, we have developed possibilistic decision trees where possibilistic decision criteria presented in Chapter 2 are used. We have proposed a full theoretical study of the complexity of the problem of finding an optimal strategy in possibilistic decision trees. Table 4.8 summarizes the results of this study.

U_{pes}	U_{opt}	PU	$L\Pi$	LN	$OMEU$	Ch_N	Ch_Π
P	P	P	P	P	P	NP-hard	NP-hard

TABLE 4.2 – Results about the Complexity of $\Pi Tree - OPT$ for the different possibilistic criteria

In fact, we have developed necessary proofs for each decision criterion in order to show if the monotonicity property is verified (to apply dynamic programming in order to find the optimal strategy) or not. Then, we have shown that strategy optimization in possibilistic decision trees is a polynomial problem for most of possibilistic decision criteria except for possibilistic Choquet integrals. Indeed, we have shown that the problem of finding a strategy optimal w.r.t possibility-based or necessity-based Choquet integrals is NP-hard via a reduction from a $3SAT$ problem. Nevertheless, we have identified three particular cases when these criteria satisfy the monotonicity property.

In next chapter, we develop an alternative solving approach for possibilistic decision tree with Choquet integrals since dynamic programming cannot be applied. More precisely, we propose an implicit enumeration approach via a Branch and Bound algorithm.

Chapitre 5

Solving algorithm for Choquet-based possibilistic decision trees

5.1 Introduction

In the previous chapter, we have shown that the problem of finding an optimal strategy w.r.t possibilistic Choquet integrals is NP-hard. As a consequence, the application of dynamic programming may lead to suboptimal strategies.

As an alternative, we propose to proceed by implicit enumeration via a *Branch and Bound* algorithm based on an optimistic evaluation of the Choquet value of possibilistic decision trees.

In order to study the feasibility of our proposed solutions for finding the optimal strategy w.r.t possibilistic decision criteria in decision trees, we propose also an experimental study.

In what follows, Section 5.1 presents the Branch and Bound algorithm and Section 5.2 gives our experimental results.

The main results of this chapter are published in [7, 8].

5.2 Solving algorithm for non polynomial possibilistic Choquet integrals

As stated by Proposition 4.8, 4.9 and 4.10, dynamic programming can be applied for only some particular classes of Choquet-based possibilistic decision trees i.e. Binary-Class, Max-Class and Min-Class. As an alternative, we propose to proceed by implicit enumeration via a *Branch and Bound algorithm*. Our choice was motivated by the success of this approach with the Rank Dependent Utility (RDU) criterion [39] where the implicit enumeration outperforms the resolute choice [58].

The Branch and Bound algorithm (denoted by BB and outlined by Algorithm 5.1) takes as argument a partial strategy δ and an upper bound of the best Choquet value it can reach. It returns the value Ch_N^{opt} (respectively Ch_Π^{opt}) of the best strategy denoted by δ^{opt}. The initial parameters of this algorithm are :

– The empty strategy ($\delta(D_i) = \perp, \forall D_i$) for δ.

– The value of the strategy provided by dynamic programming algorithm for δ^{opt}. Indeed, even not necessarily providing an optimal strategy, this algorithm may provide a good one, at least from a consequentialist point of view.

At each step, the current partial strategy, δ, is developed by the choice of an action for

some unassigned decision node. When several decision nodes need to be developed, the one with the minimal rank (i.e. the former one according to the temporal order) is developed first. The recursive procedure stops when either the current strategy is complete (then δ^{opt} and Ch_N^{opt} may be updated (resp. Ch_Π^{opt})) or proves to be worst than δ^{opt} in any case.
To this extent, we call a function that computes a lottery (denoted by $Lottery(\delta)$) that overcomes all those associated with the complete strategies compatible with δ and use $Ch_N(Lottery(\delta))$ (resp. $Ch_\Pi(Lottery(\delta)))$ as an upper bound of the Choquet value of the best strategy compatible with δ the evaluation is sound, because whatever L, L', if L overcomes L', then $Ch_N(L) \geq Ch_N(L')$ (resp. $Ch_\Pi(L) \geq Ch_\Pi(L')$).
Whenever $Ch_N(Lottery(\delta)) \leq Ch_N^{opt}$ (resp. $Ch_\Pi(Lottery(\delta)) \leq Ch_\Pi^{opt}$), the algorithm backtracks, yielding the choice of another action for the last decision nodes considered. Moreover when δ is complete, $Lottery(\delta)$ returns $L(D_0, \delta)$; the upper bound is equal to the Choquet value when computed for a complete strategy.

Function *Lottery* (Algorithm 5.2) inputs a partial strategy. It proceeds backwards, assigning a simple lottery $\langle 1/u(NL_i)\rangle$ to each leaf in the decision tree.

In the Branch and Bound algorithm (Algorithm 5.1), the fuzzy measure μ may be the possibility measure Π or the necessity measure N according to the problem at hand.

Algorithm 5.1: BB

Data: A (possibly partial) strategy δ, its Choquet value Ch_μ^δ

Result: Ch_μ^{opt} *% also memorizes the best strategy found so far,* δ^{opt}

begin
    if $\delta = \emptyset$ **then** $\mathcal{D}_{pend} = \{D_1\}$ **else**
        $D_{pend} = \{D_i \in \mathcal{D}$ s.t. $\delta(D_i) = \perp$ and $\exists D_j, \delta(D_j) \neq \perp$ and $D_i \in Succ(\delta(D_j))\ \}$
    if $\mathcal{D}_{pend} = \emptyset$ *(% δ is a complete strategy)* **then**
        if $Ch_\mu^\delta > Ch_\mu^{opt}$ **then**
        $\delta^{opt} \leftarrow \delta$
        return Ch_μ^δ
    else
        $D_{next} \leftarrow arg\ \min_{D_i \in D_{pend}}\ i$
        foreach $C_i \in Succ(D_{next})$ **do**
            $\delta(D_{next}) \leftarrow C_i$
            $Eval \leftarrow Ch_\mu(Lottery(D_0, \delta))$
            if $Eval > Ch_\mu^{opt}$ **then**
                $Ch_\mu^{opt} \leftarrow \max(Ch_\mu^\delta(\delta), Eval)$
    return Ch_μ^{opt}
end

Algorithm 5.2: Lottery

Data: a node X, a (possibly partial) strategy δ

Result: L^X % *$L^X[u_i]$ is the possibility degree to have the utility u_i*

begin

- **for** $i \in \{1,..,n\}$ **do** $L^X[u_i] \leftarrow 0$
- **if** $X \in \mathcal{LN}$ **then** $L^X[u(X)] \leftarrow 1$
- **if** $X \in \mathcal{C}$ **then**
  - **foreach** $Y \in Succ(X)$ **do**
    - $L^Y \leftarrow Lottery(Y, \delta)$
    - **for** $i \in \{1,..,n\}$ **do** $L^X[u_i] \leftarrow \max(L^X[u_i], \pi_X(Y) \otimes L^Y[u_i])$
    - % $\otimes = \min$ *in the ordinal setting ;*
    - % $\otimes = *$ *in the numerical setting*
- **if** $X \in \mathcal{D}$ **then**
  - **if** $\delta(X) \neq \perp$ **then** $L^X = Lottery(\delta(X), \delta)$ **else**
    - **if** $|Succ(X)| = 1$ **then**
      - $L^X = Lottery(\delta(Succ(X)), \delta)$
    - **else**
      - **foreach** $Y \in Succ(X) \cap N_\delta$ **do**
        - $L^Y \leftarrow Lottery(Y, \delta)$
        - **for** $i \in \{1,..,n\}$ **do** $G^c_Y[u_i] \leftarrow 1 - max_{u_j < u_i} L^Y[u_j]$
      - % *Compute the upper envelop of the cumulative functions)*
      - **for** $i \in \{1,..,n\}$ **do** $G^c[u_i] \leftarrow \max_{Y \in Succ(X) \cap N_\delta} G^c_Y[u_i]$
      - % Compute $Rev(G^c)$
      - $L^X[u_n] \leftarrow 1$
      - **for** $i \in \{n-1,..,1\}$ **do** $L^X[u_i] \leftarrow 1 - G^c[u_{i+1}]$
- **return** L^X

end

At each chance node C_i, we perform a composition of the lotteries in $Succ(C_i)$ according to the principle of reduction of possibilistic lotteries presented in Chapter 2 (Section 2.2). At each decision node D_i we choose a lottery that overcomes all those in $Succ(D_i)$. To this end, let us use the following notations and definitions :

– Given a simple lottery $L \in \mathcal{L}$, G^c_L is the *possibilistic decumulative* function of a lottery

L such that $\forall u \in U$:

$$G^c_L(u) = \begin{cases} N(L \geq u) & if \ \mu = N \\ \Pi(L \geq u) & if \ \mu = \Pi. \end{cases} \quad (5.1)$$

– Given a set $\mathcal{G} = \{G^c_{L_1}, \ldots, G^c_{L_k}\}$ of decumulative functions, the *upper envelop* of $\mathcal{G}$ is the decumulative function $G^c_{\mathcal{G}}$ defined by :

$$\forall u \in U, G^c_{\mathcal{G}}(u) = max_{G^c_{L_i} \in \mathcal{G}} G^c_{L_i}(u). \quad (5.2)$$

– Given a decumulative function $G^c_{\mathcal{G}}$ on U, $Rev(G^c)$ gives a lottery defined by :

$$Rev(G^c)(u_i) = \begin{cases} 1 & if \quad i = n \\ 1 - G^c_{\mathcal{G}}(u_{i+1}) & if \quad i \in \{1, \ldots, n-1\}. \end{cases} \quad (5.3)$$

Proposition 5.1 *The possibilistic decumulative function associated to a lottery $Rev(G^c)$ is equal to G^c.*

Proof. [Proof of Proposition 5.1]
We have by definition, $G^c_{Rev(G^c)}(u_1) = 1 - G^c(u_1)$. Moreover, $\forall i = 2, n, Rev(G^c)(u_i) \geq Rev(G^c)(u_{i-1})$.
Hence $G^c_{Rev(G^c)}(u_i) = 1 - max_{j=1,i-1} Rev(G^c)(u_j) = 1 - Rev(G^c)(u_{i-1})$.
Since $Rev(G^c)u_{i-1} = 1 - G^c(u_i)$, we get $G^c_{Rev(G^c)}(u_i) = G^c(u_i)$.
Thus $G^c_{Rev(G^c)} = G^c$. ■

As a consequence : Given a set $\{L_1, \ldots, L_k\} \subseteq \mathcal{L}$ of simple lotteries over U, $\mathcal{G} = \{G^c_{L_1}, \ldots, G^c_{L_k}\}$ the set of their decumulative functions, we can check that : $Rev(G^c_{\mathcal{G}})$ overcomes any lottery $L_i \in \{L_1, \ldots, L_k\}$.

Hence, the Choquet value of $Lottery(D_0, \delta)$ is an upper bound of the Choquet value of the best complete strategy compatible with δ, which proofs the correctness of our algorithm.

Example 5.1 *Let us consider the possibilistic decision tree in Figure 5.1.*

The exhaustive enumeration of this tree (Figure 5.1) leads to 5 *possible strategies* $\Delta = \{\delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \}$ *represented in Table 5.1. The optimal strategy is* δ_3 *with* $Ch_N(\delta_3) = 0.675$.

Let us start by the dynamic programming algorithm (Algorithm 4.1) :
– Initially, we have $\delta = \emptyset$ *and* $N = D_0$ *with* $succ(D_0) = \{C_1, C_2\}$.

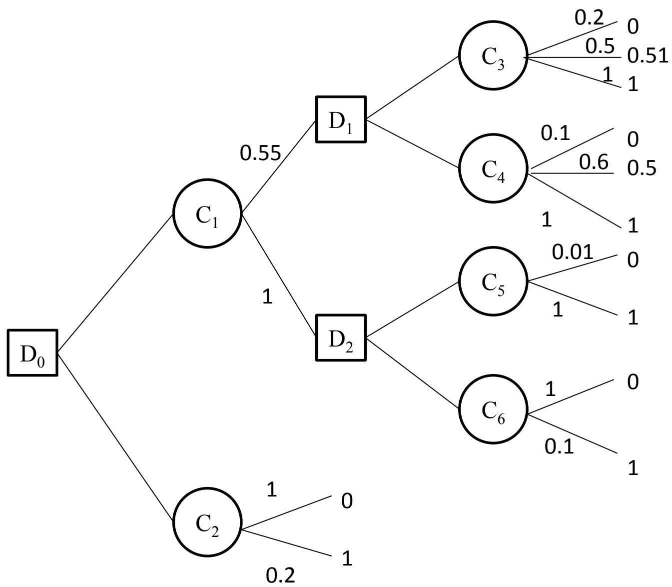

FIGURE 5.1 – Possibilistic decision tree

δ_i	L_i	$Ch_N(L_i)$
$\delta_1 = \{(D_0, C_1), (D_1, C_3), (D_2, C_5)\}$	$\langle 0.2/0, 0.5/0.51, 1/1\rangle$	0.653
$\delta_2 = \{(D_0, C_1), (D_1, C_3), (D_2, C_6)\}$	$\langle 1/0, 0.5/0.51, 0.55/1\rangle$	0
$\delta_3 = \{(D_0, C_1), (D_1, C_4), (D_2, C_5)\}$	$\langle 0.1/0, 0.55/0.5, 1/1\rangle$	0.675
$\delta_4 = \{(D_0, C_1), (D_1, C_4), (D_2, C_6)\}$	$\langle 1/0, 0.55/0.5, 1/1\rangle$	0
$\delta_5 = \{(D_0, C_2)\}$	$\langle 1/0, 0.2/1\rangle$	0

TABLE 5.1 – Exhaustive enumeration of possible strategies in Figure 5.1

– *For* $Y = C_1$, $L_{C_1} = ProgDyn(C_1, \delta)$ *since* $succ(C_1) = \{D_1, D_2\}$ *we have* $Y = D_1$ *and* $Y = D_2$.

– *For* $Y = D_1$, *we have* $L_{D_1} = ProgDyn(D_1, \delta)$ *and* $succ(D_1) = \{C_3, C_4\}$ *:*

1. *If* $Y = C_3$ *then* $L_{C_3} = \langle 0.2/0, 0.5/0.51, 1/1\rangle$ *and* $Ch_N(L_{C_3}) = 0.653$.
2. *If* $Y = C_4$ *then* $L_{C_4} = \langle 0.1/0, 0.6/0.5, 1/5\rangle$ *and* $Ch_N(L_{C_4}) = 0.650$. *Since* $Ch_N(L_{C_3}) > Ch_N(L_{C_4})$, *so* $Y^* = C_3$, $\delta(D_1) = C_3$ *and*

$L_{D_1} = \langle 0.2/0, 0.5/0.51, 1/1 \rangle$.

– *For* $Y = D_2$, *we have* $L_{D_2} = ProgDyn(D_2, \delta)$ *and* $succ(D_2) = \{C_5, C_6\}$ *:*

1. *If* $Y = C_5$ *then* $L_{C_5} = \langle 0.01/0, 1/1 \rangle$ *and*
   $Ch_N(L_{C_5}) = 0.99$.
2. *If* $Y = C_6$ *then* $L_{C_6} = \langle 1/0 \rangle$ *and*
   $Ch_N(L_{C_6}) = 0$. *Since* $Ch_N(L_{C_5}) > Ch_N(L_{C_6})$, *so* $Y^* = C_5$, $\delta(D_2) = C_5$ *and* $L_{D_2} = \langle 0.01/0, 1/1 \rangle$.

$\Rightarrow L_{C_1} = \langle 0.55/L_{D_1}, 1/L_{D_2} \rangle$ *and* $Ch_N(L_{C_1}) = 0.653$.

– *For* $Y = C_2$, $L_{C_2} = ProgDyn(C_2, \delta)$ *we have :*

$L_{C_2} = \langle 1/0, 0.2/1 \rangle$ *and* $Ch_N(L_{C_2}) = 0$.

$\Rightarrow Ch_N(L_{C_1}) > Ch_N(L_{C_2})$, *so* $Y^* = C_1$, $\delta(D_0) = C_1$ *and*

$\delta^* = \{(D_0, C_1), (D_1, C_3), (D_2, C_5)\}$ *with* $Ch_N(\delta^*) = 0.653$.

Note that the value of $Ch_N(\delta^*)$ *obtained by dynamic programming is different from the one obtained by exhaustive enumeration (i.e.* 0.675*) since as we have seen in Chapter 4 dynamic programming does not guarantee optimal solutions since possibilistic Choquet integrals does not satisfy the monotonicity property.*

We propose now to apply the Branch and Bound algorithm (Algorithm 5.1) for the evaluation. The major steps of this algorithm can be summarized as follows (we start with the solution provided by dynamic programming i.e. $Ch_N(\delta^{opt}) = 0.653$*) :*

– $\delta = \emptyset$ *and* $Ch_N^{opt} = 0.653$ *(lower bound given by dynamic programming).*

BB calls $Ch_N(Lottery(D_0, (D_0, C_1)))$

We have $G^c_{C_3} = \langle 1/0, 0.8/0.51, 0.5/1 \rangle$ *and* $G^c_{C_4} = \langle 1/0, 0.9/0.5, 0.4/1 \rangle$.

So $G^c = \langle 1/0, 0.9/0.5, 0.8/0.51, 0.5/1 \rangle$ *and* $L^{D_1} = \langle 0.1/0, 0.2/0.5, 0.5/0.51, 1/1 \rangle$.

We have $G^c_{C_5} = \langle 1/0, 0.99/1 \rangle$ *and* $G^c_{C_6} = \langle 1/0, 0/1 \rangle$.

So $G^c = \langle 1/0, 0.99/1 \rangle$ *and* $L^{D_2} = \langle 0.01/0, 1/1 \rangle$.

So, $Lottery(D_0, (D_0, C_1)) = (0.1/0, 0.2/0.5, 0.5/0.51, 1/1)$

and $Eval = Ch_N(Lottery(D_0, (D_0, C_1))) = 0.703 > 0.653$.

– $\delta = (D_0, C_1)$ *and* $Ch_N^{opt} = 0.653$.

BB calls $Ch_N(Lottery(D_0, ((D_0, C_1), (D_1, C_3))))$.

$Lottery(D_0, ((D_0, C_1), (D_1, C_3))) = \langle 0.2/0, 0.5/51, 1/1 \rangle$

and $Eval = Ch_N(Lottery(D_0, ((D_0, C_1), (D_1, C_3)))) = 0.653 = 0.653$.

$\delta = (D_0, C_1)$ *and* $Ch_N^{opt} = 0.653$.

BB calls $Ch_N(Lottery(D_0, ((D_0, C_1), (D_1, C_4))))$

$Lottery(D_0, ((D_0, C_1), (D_1, C_4))) = \langle 0.1/0, 0.55/0.5, 1/1 \rangle$

and $Eval = Ch_N(Lottery(D_0, ((D_0, C_1), (D_1, C_4)))) = 0.675 > 0.653$.

– $\delta = ((D_0, C_1), (D_1, C_4))$ *and* $Ch_N^{opt} = 0.1$.
BB calls $Ch_N(Lottery(D_0, ((D_0, C_1), (D_1, C_4), (D_2, C_5))))$,
$Lottery(D_0, ((D_0, C_1), (D_1, C_4), (D_2, C_5))) = \langle 0.1/0, 0.55/0.5, 1/1 \rangle$
and $Eval = Ch_N(Lottery(D_0, ((D_0, C_1), (D_1, C_4), (D_2, C_5)))) = 0.675 > 0.653$.
– $\delta = ((D_0, C_1), (D_1, C_4), (D_2, C_5))$ *and* $Ch_N^{opt} = 0.653$.
There is no more pending decision node. $\delta^{opt} \leftarrow ((D_0, C_1), (D_1, C_4), (D_2, C_5))$ *and* $Ch_N^{opt} = 0.675$.
– $\delta = ((D_0, C_1), (D_1, C_4))$ *and* $Ch_N^{opt} = 0.675$.
BB calls $Ch_N(Lottery(D_0, ((D_0, C_1), (D_1, C_4), (D_2, C_6))))$,
$Lottery(D_0, ((D_0, C_1), (D_1, C_4), (D_2, C_6))) = \langle 1/0, 0.55/0.5, 0.55/1 \rangle$
and $Eval = Ch_N(Lottery(D_0, ((D_0, C_1), (D_1, C_4), (D_2, C_6)))) = 0 < 0.675$.
– $\delta = ((D_0, C_1), (D_1, C_4), (D_1, C_5))$ *and* $Ch_N^{opt} = 0.675$.
There is no more pending decision node, $\delta^{opt} \leftarrow ((D_0, C_1), (D_1, C_4), (D_2, C_5))$ *and* $Ch_N^{opt} = 0.675$.

The algorithm eventually terminates with $\delta^{opt} = ((D_0, C_1), (D_1, C_4), (D_2, C_5))$ *and* $Ch_N^{opt} = 0.675$ *corresponds to the optimal strategy obtained by exhaustive enumeration (see Table 5.1).*

5.3 Experimental results

In order to show the feasibility of the studied algorithms in the case of possibilistic decision trees using Choquet integrals, we propose an experimental study aiming at :

– Compare results provided by dynamic programming w.r.t those of Branch and Bound by computing the regret of applying the first algorithm even if it does not guarantee the optimal values (as it is the case with Branch and Bound).

– Compare the execution CPU time of dynamic programming and Branch and Bound for polynomial cases of possibilistic Choquet integrals (i.e. Binary-Class and Max-Class for Ch_Π, Binary-Class, Min-Class and Max-Class for Ch_N).

To this end, we have implemented both dynamic programming and Branch and Bound algorithms in Matlab 7.10.0. The experimental study was carried out on a PC with Duo CPU 210 GHz and 4.00 GO (RAM).

The first step of our experimental study concerns the generation of binary possibilistic decision trees with ND decision nodes, NC chance nodes ($NC = ND * 2$) and NV utilities ($NV = ND + NC + 1$). The depth of generated decision trees is $ND - 1$ (see Figure 5.2) such that at each level i ($0 \leq i \leq ND - 1$) we have 2^i nodes.

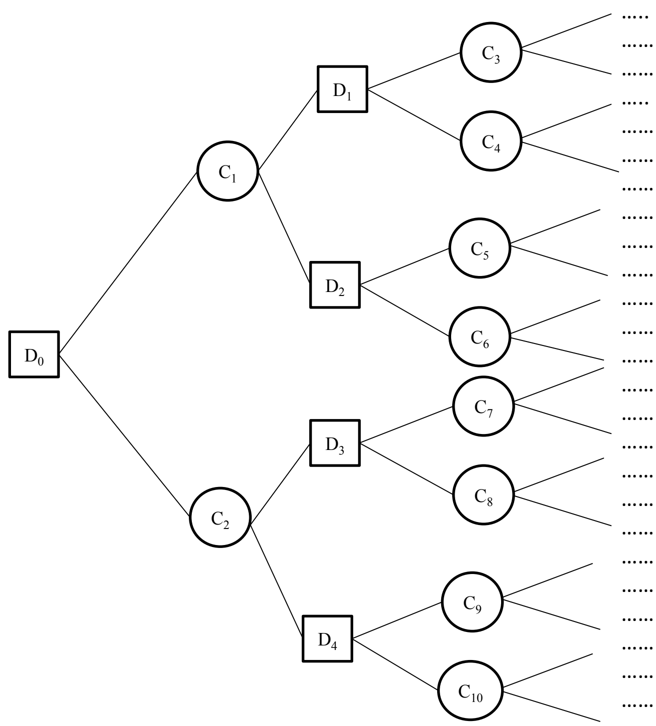

FIGURE 5.2 – Structure of constructed decision trees

Following this reasoning, we propose to consider 4 cases namely $ND = 5$, $ND = 21$, $ND = 85$ and $ND = 341$. This means that the size of generated trees will be 31, 127, 511, 2047 respectively. For utilities, we have randomly chosen values in the set $U = \{0, 1, \ldots, 20\}$ (with a numerical interpretation). Conditional possibilities relative to chance nodes are also chosen randomly in $[0, 1]$ ensuring the possibilistic normalization.

Using these parameters, we have generated randomly a sample of 50 possibilistic decision trees for each tree size (i.e. 31, 127, 511 and 2047).

Quality of solutions provided by dynamic programming

Since the application of dynamic programming in the case of possibilistic Choquet integrals can lead to a suboptimal strategy, we propose to estimate their quality by comparing them to exact values generated by Branch and Bound. More precisely, we compute for different trees the closeness value (denoted by $Closeness$) equal to $\frac{V_{DP}}{V_{BB}}$ such that V_{DP} is the possibilistic Choquet integrals relative to the optimal strategy provided by dynamic programming and V_{BB} by Branch and Bound. Clearly within the randomly generated trees some of them correspond to particular cases where the two approaches are equivalent i.e. Binary-Class, Min-Class and Max-Class.

		Tree Size			
	Setting	31	127	511	2047
Ch_N	Qualitative	0.998	0.843	0.632	0.190
	Numerical	0.987	0.765	0.473	0.25
Ch_Π	Qualitative	1	0.85	0.693	0.32
	Numerical	0.946	0.727	0.487	0.21

Table 5.2 – The closeness value with Ch_N and Ch_Π

	Tree Size			
Decision criterion	31	127	511	2047
Ch_N	99%	78%	75%	71%
Ch_Π	98%	80%	73%	71%

Table 5.3 – The percentage of polynomial cases

The experimental results, summarized in Table 5.2, confirm that the closeness value is close to 1 for smallest decision trees (31 nodes) for the case of Ch_N and Ch_Π in qualitative and numerical settings. This means that for small trees, dynamic programming gives a very good approximation of optimal strategies (about 99% for tree size equal to 31 and 80% for tree size equal to 127). This good approximation can be explained by the large number of polynomial cases for smallest decision trees (about 99% for 31 nodes and 80% for 127 nodes) as it is presented in Table 5.3. For large trees, the closeness decreases approaching to 0 for trees having 2047 nodes. Clearly the number of polynomial cases also decreases in this case (about 70%).

Execution CPU time

Table 5.4 (resp. Table 5.5) gives different average execution CPU time for each size of possibilistic decision trees with Ch_N (resp. Ch_Π) in both qualitative and numerical settings.

First, we note that we have the same trend regarding the execution CPU time for Ch_N and also for Ch_Π in qualitative and numerical setting i.e. it increases according to the size

		Tree Size			
	Algorithm	31	127	511	2047
Qualitative setting	Dynamic Programming	0.119	3.160	69.605	1.6295e+003
	Branch and Bound	0.276	7.144	121.751	2.6413e+006
Numerical setting	Dynamic Programming	0.106	2.859	68.383	1.3541e+003
	Branch and Bound	0.409	5.976	120.5095	2.3658e +006

Table 5.4 – Execution CPU time for Ch_N (in seconds)

		Tree Size			
	Algorithm	31	127	511	2047
Qualitative setting	Dynamic Programming	0.1216	2.905	66.1384	3.9629e+003
	Branch and Bound	0.284	5.967	118.178	4.0624e+003
Numerical setting	Dynamic Programming	0.1226	2.5654	65.7173	13043e+003
	Branch and Bound	0.314	5.484	118.559	2.323e+006

Table 5.5 – Execution CPU time for Ch_Π (in seconds)

of the tree.

These results also show that dynamic programming is faster than Branch and Bound algorithm since initially it computes the lower bound using dynamic programming.

5.4 Conclusion

In this chapter, we have proposed a Branch and Bound algorithm to find optimal strategies in possibilistic decision trees when the decision criteria are possibilistic Choquet integrals. In fact in such a case we have shown that the application of dynamic programming can lead to sub-optimal solutions. Then, we have performed experiments on different decision trees built randomly in order to study the quality of solutions provided by dynamic programming by comparing them to those of the Branch and Bound algorithm. We have also compared the two algorithms w.r.t their execution CPU time. In the next chapter we will study another graphical possibilistic model, namely *possibilistic influence diagrams*.

Chapitre 6

Possibilistic Influence Diagrams : Definition and Evaluation Algorithms

6.1 Introduction

After developing possibilistic decision trees in Chapter 4 and 5, we are now interested by the study of the possibilistic counterpart of influence diagrams in order to benefit from the simplicity of these graphical decision models.

Depending on the quantification of chance and decision nodes, we distinguish two kinds of possibilistic influence diagrams namely homogeneous and heterogeneous ones. For these two classes, we propose indirect evaluation algorithms transforming them into possibilistic decision trees (developed in the previous chapter) or into possibilistic networks [5].

This chapter is organized as follows : in Section 6.2, possibilistic influence diagrams will be developed. In Section 6.3, we propose two evaluation algorithms of these graphical decision models via a transformation into possibilistic decision trees or into possibilistic networks.

6.2 Possibilistic influence diagrams

Roughly speaking, possibilitic influence diagrams, denoted by $\Pi ID_{\otimes}^{u}$, have the same graphical component as standard ones (seen in Chapter 3) i.e. they are composed of a set of nodes $\mathcal{N} = D \cup C \cup V$ where D is the set of decision node, C is the set of chance nodes and V is the set of value nodes and a set of arcs A (informational and conditional arcs). This is not the case of the numerical component which relies on the possibilistic framework such that :

- For each chance node $C_i \in C$, we should provide conditional possibility degrees $\Pi(c_{ij} \mid pa(C_i))$ of each instance c_{ij} of C_i in the context of each instance of its parents. In order to satisfy the normalization constraint, these conditional distributions should satisfy :
$$max_{c_{ij} \in D_{c_i}} \Pi(c_{ij} \mid pa(C_i)) = 1. \tag{6.1}$$
In what follows ΠID_{*}^{u} (resp. ΠID_{min}^{u}) denotes possibilistic influence diagrams where conditional possibility distributions are modeled in the numerical (resp. qualitative) setting.
- For each value node $V_i \in V$, a set of utilities U is defined in the context of each instantiation $pa(V_i)$ of its parents $Pa(V_i)$. In what follows $\Pi ID_{\otimes}^{*}$ (resp. $\Pi ID_{\otimes}^{min}$) denotes possibilistic influence diagrams where utilities are numerical (resp. qualitative).

– Likewise standard influence diagrams, decision nodes in possibilistic IDs are not quantified.

Since decision nodes are not quantified, they act differently from chance nodes, thus for a given chance node C_i and a decision node D_i, it is meaningless to consider $\Pi(c_{ij}, d_{ij})$. In fact, what is meaningful is $\Pi(c_{ij} \mid do(d_{ij}))$ where $do(d_{ij})$ is the particular operator defined by Pearl [63]. Using chain rules relative to possibilistic networks [5] and to standard influence diagrams [47], the following chain rule for possibilistic IDs can be inferred :

$$\pi(C \mid D) = \bigotimes_{C_i \in C} \Pi(C_i \mid Pa(C_i)) \tag{6.2}$$

where $\bigotimes$ is the min operator in the case of qualitative possibility theory and the product operator in the case of numerical possibility theory.

Like standard influence diagrams, a general proof of Equation 6.2 concerning the chain rule of possibilistic influence diagrams can be done by considering a particular configuration d of decisions. If this configuration is inserted in the possibilistic influence diagram then we will get a possibilistic network representing $\Pi(C|d)$. Using the chain rule of possibilistic networks [6], we obtain two cases w.r.t the interpretation of the uncertainty scale :

1. For numerical setting, we have $\Pi(C|d)$ is the product of all possibility potentials attached to the decision variables instantiated to d.
2. For qualitative setting, we have $\Pi(C|d)$ is the minimum of all possibility potentials attached to the decision variables instantiated to d.

In other words, possibilistic influence diagrams are a compact representation of the joint distribution relative to chance nodes conditioned by a configuration of decision nodes.

Different combinations between the quantification of chance and utility nodes in influence diagrams offer several kinds of possibilistic influence diagrams which can be grouped into two principal classes [42] :

1. **Homogeneous possibilistic influence diagrams** where chance and value nodes are quantified in the same setting. Within this class, we can distinguish two variants :
   - *Product-based possibilistic influence diagrams*, denoted by ΠID_*^*, where both dependencies between chance nodes and value nodes are quantified in a genuine numerical setting.
   - *Min-based possibilistic influence diagrams*, denoted by ΠID_{min}^{min}, where both dependencies between chance nodes and value nodes are quantified in a qualitative setting used for encoding an ordering between different states of the world [41, 42].
2. **Heterogeneous possibilistic influence diagrams** where chance and value nodes are not quantified in the same setting. Depending on this quantification, there are

two possible quantifications and heterogeneous possibilistic influence diagrams will be denoted by ΠID_{*}^{min} and ΠID_{min}^{*}.

Different kinds of possibilistic influence diagrams are summarized in Table 6.1.

U/ Π	Qualitative	Numerical
Qualitative	ΠID_{min}^{min}	ΠID_{*}^{min}
Numerical	ΠID_{min}^{*}	ΠID_{*}^{*}

TABLE 6.1 – Classification of possibilistic influence diagrams

The following example presents a min-based possibilistic influence diagram.

Example 6.1 *The influence diagram of Figure 6.1 is defined by* $D = \{D1, D2\}$, $\mathcal{C} = \{A1, A2\}$ *and* $\mathcal{V} = \{U\}$.

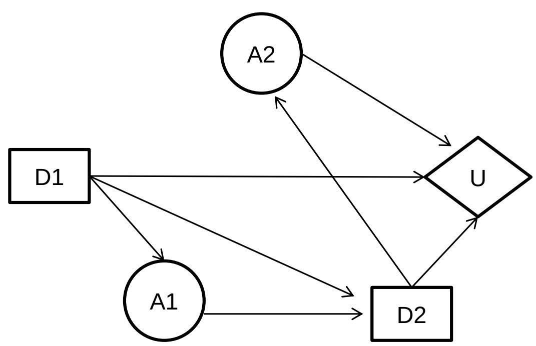

FIGURE 6.1 – The graphical component of the influence diagram

Conditional possibilities are represented in Tables 6.2 *and* 6.3*. Table* 6.4 *represents the set of utilities for the value node* U*.*

$A1$	$D1$	$\pi(A1 \mid D1)$
T	T	1
T	F	0.4
F	T	0.2
F	F	1

TABLE 6.2 – Conditional possibilities for $A1$

$A2$	$D2$	$\pi(A2 \mid D2)$
T	T	0.3
T	F	1
F	T	1
F	F	0.4

TABLE 6.3 – Conditional possibilities for $A2$

$D1$	$D2$	$A2$	$u(D1, D2, A2)$
T	T	T	0.2
T	T	F	0.3
T	F	T	0.4
T	F	F	0.6
F	T	T	1
F	T	F	0
F	F	T	0.1
F	F	F	0.7

TABLE 6.4 – The utility function $u(D1, D2, A2)$

Let us represent in Table 6.5 *the chain rule of the possibilistic influence diagram in Figure 6.1 using the Equation 6.2 in the qualitative setting of possibility theory.*

6.3 Evaluation of possibilistic influence diagrams

Given a possibilistic influence diagram, it should be evaluated to determine optimal decisions δ^*. Contrarily to standard influence diagrams where the decision criterion is the maximal expected utility MEU, we can here use the panoply of possibilitic decision criterion (already presented in Chapter 2) under the constraint to respect the semantic underlying the influence diagram (i.e. qualitative or quantitative).

$A1$	$A2$	$D1$	D_2	$\pi(A_1, A_2 \mid D_1, D_2)$
T	T	T	T	0.3
T	T	T	F	1
T	T	F	T	0.3
T	T	F	F	0.4
T	F	T	T	1
T	F	T	F	0.4
T	F	F	T	0.4
T	F	F	F	0.4
F	T	T	T	0.2
F	T	T	F	0.2
F	T	F	T	0.3
F	T	F	F	1
F	F	T	T	0.2
F	F	T	F	0.2
F	F	F	T	1
F	F	F	F	0.4

TABLE 6.5 – The chain rule of the possibilistic influence diagram in Figure 6.1

More precisely, *possibilistic likely dominance* (LN and $L\Pi$) and *possibilistic Choquet integrals* (Ch_N and Ch_Π) can be used with product-based possibilistic influence diagrams since they can be defined with numerical possibility theory.

Besides, *pessimistic and optimistic utilities* (U_{pes}, U_{opt}) and *binary utilities* (PU) can be used in min-based influence diagrams since they are a purely ordinal possibilistic decision criteria.

It is important to note that only possibilistic Choquet integrals can be used as decision criteria in heterogeneous possibilistic influence diagrams, since they are appropriate to handle heterogeneous informations.

Table 6.6 indicates for each kind of possibilistic influence diagrams, the possibilistic decision criteria that can be used.

Few works were interested to this problem. Garcia et al. [33, 34] have proposed two methods for the evaluation of possibilistic IDs using pessimistic and optimistic utilities. Their first work consists on an indirect method based on the transformation of possibilistic

	U_{pes}	U_{opt}	PU	LN	$L\Pi$	$OMEU$	Ch_N	Ch_Π
ΠID_{min}^{min}	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$
ΠID_{*}^{*}				$\surd$	$\surd$		$\surd$	$\surd$
ΠID_{*}^{min}							$\surd$	$\surd$
ΠID_{min}^{*}							$\surd$	$\surd$

TABLE 6.6 – Adaptation of possibilistic decision criteria to different kinds of possibilistic influence diagrams

influence diagrams into possibilistic decision trees and the in the second one, they proposed a variable elimination algorithm. Note that influence diagrams are developed using order of magnitude expected utility ($OMEU$) as a decision criterion [53]. Therefore, a variable elimination algorithm was used to compute the optimal strategy in an order of magnitude influence diagram.

We propose now to consider all possible decision criteria via two indirect evaluation methods : the first one is based on the transformation of possibilistic influence diagrams into possibilistic decision trees and the second one is based on their transformation into possibilistic networks [5, 6].

6.3.1 Evaluation of influence diagrams using possibilistic decision trees

A possibilistic influence diagram can be unfold into a possibilistic decision tree using transformation method similar to the one proposed for the case of standard influence diagrams (detailed in Chapter 3).

This transformation may lead to new dependencies between chance nodes which will be quantified using initial possibility distributions. Then, utility values are the same as those in the influence diagram and they will be affected to each leaf in the decision tree.

Once the possibilistic decision tree is constructed, it should be evaluated to find the optimal strategy. This obviously depends on the decision criterion, more precisely :

- *Dynamic programming* algorithm (Algorithm 4.1 in Chapter 4) should be applied for other possibilistic decision criteria (i.e. U_{opt}, U_{pes}, PU, L_N, L_Π, $OMEU$) and for polynomial cases of possibilistic Choquet integrals (i.e. Binary-Class, Max-Class and Min-Class).
- *Branch and Bound* algorithm (Algorithm 5.1 in Chapter 5) should be applied in the

case of possibilistic Choquet integrals (i.e. Ch_N and Ch_Π).

Example 6.2 *The possibilistic decision tree in Figure 6.2 corresponds to the transformation of the influence diagrams of Figure 6.1.*

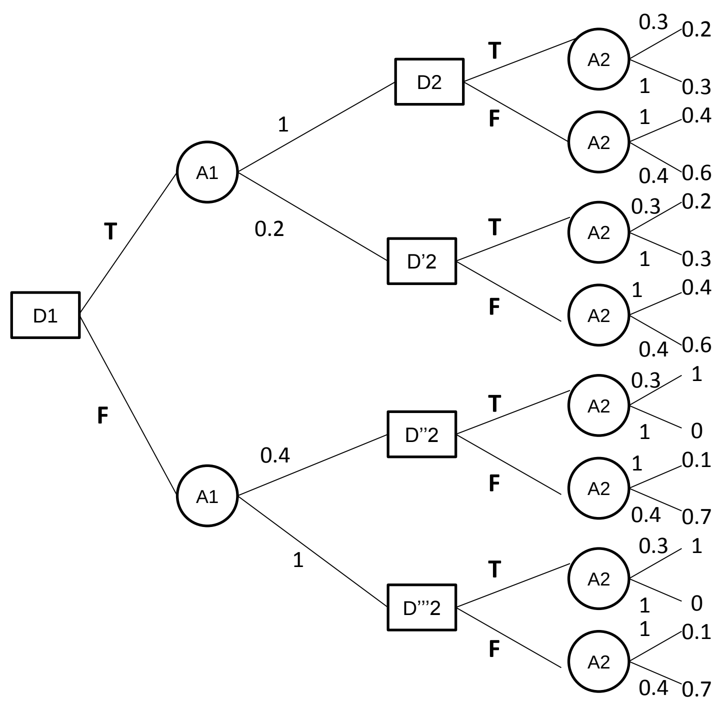

FIGURE 6.2 – The possibilistic decision tree corresponding to the transformation of the influence diagrams of Figure 6.1

Suppose that we will use the optimistic utility criterion U_{opt} *(Equation 2.10) as decision criterion, then the application of the dynamic programing algorithm generates two optimal strategies* $\delta_1^* = \{(D1 = T), (D2 = F)\}$ *and* $\delta_2^* = \{(D1 = F), (D2 = F)\}$ *with* $U_{opt}(\delta_1^*) = 0.4$ *and* $U_{opt}(\delta_2^*) = 0.4$ *as it is presented in Figure 6.3.*

Clearly, the size of the decision tree can grow exponentially w.r.t the size of the influence diagrams since we should duplicate several parts of the decision tree in order to represent all possible scenarios. This drawback of decision trees encourages us to explore another track by transforming influence diagrams into compact structures which are possibilistic networks.

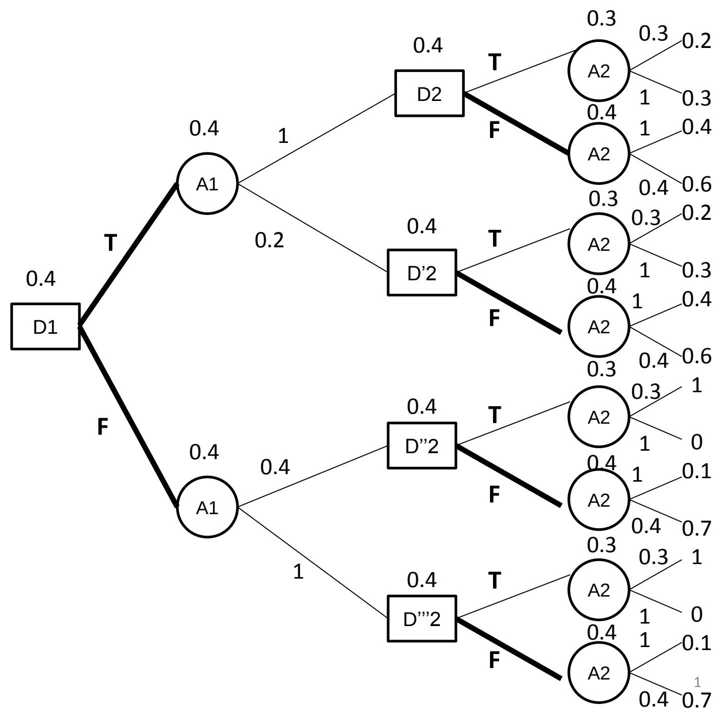

FIGURE 6.3 – Optimal strategies in possibilistic decision tree in Figure 6.2

6.3.2 Evaluation of influence diagrams using possibilistic networks

The idea of this evaluation method is to adapt the Cooper's method [14] to our context by morphing the initial influence diagram into a possibilistic network, then to use it in order to perform computations in a local manner via propagation algorithms. In fact possibilistic networks [32] are possibilitic counterparts of Bayesian networks and can be defined in a min-based or product-based version.

Moreover, several propagation algorithms are available with such a networks ; some of them are adaptations of Pearl and Jensen algorithms [5, 6, 10, 32] and others are dedicated to min-based possibilistic networks like the anytime algorithm [5, 6].

It is important to note that, a propagation algorithm in a possibilistic network is a form of dynamic programming. So, this indirect method of evaluation can be used only in the case of decision criteria that satisfy the monotonicity property (i.e. U_{pes}, U_{opt}, PU, LN, $L\Pi$, $OMEU$ and polynomial classes of possibilistic Choquet integrals).

In our work, we offer the possibility of evaluating possibilistic influence diagrams with several value nodes using a pretreatment on the influence diagram before its transformation into a possibilistic network. The pretreatment step consists on the reduction of the number of value nodes to one (denoted by V_r) that will inherit the parents of all value nodes. The value node V_r will have the minimum of utilities, formally :

$$u(V_r \mid pa(V_r)) = \min_{i=1...k} u(V_i \mid pa(V_r)) \tag{6.3}$$

The key idea of the proposed algorithm is to transform decision and the value node into chance nodes in order to obtain possibilistic networks and then, perform propagation in this secondary structure.

New chance nodes obtained from the transformation of decision nodes should be characterized by total ignorance namely :

$$\Pi(d_{ij}|pa(D_i)) = 1, \quad \forall D_i \in D \tag{6.4}$$

Value nodes will be transformed into a new binary chance nodes which will be quantified according to the nature of utilities, we can distinguish two cases :

1. Assuming that utilities and possibilities are commensurable and uncertainty scale is $[0, 1]$, binary chance nodes issued from the transformation of value nodes should be quantified as follows :
$$\Pi(V_r = T|pa(V_r)) = u(pa(V_r)). \tag{6.5}$$
and
$$\Pi(V_r = F|pa(V_r)) = 1. \tag{6.6}$$
2. If utilities and possibilities are not commensurable then each utility should be transformed into the scale $[0, 1]$:
$$\Pi(V_r = T|pa(V_r)) = \frac{u(pa(V_r)) - U_{min}}{U_{max} - U_{min}} \tag{6.7}$$
and
$$\Pi(V_r = F|pa(V_r)) = 1. \tag{6.8}$$
where U_{max} (resp. U_{min}) is the maximal utility in $U(pa(V_r))$ (resp. is the minimal utility in $U(pa(V_r))$).

Once the possibilistic network is constructed then the optimal strategy will computed iteravely via appropriate propagation algorithms as we will detail after.

In order to illustrate the computation phase, we consider the case of qualitative possibilistic utilities (i.e. U_{opt}, U_{pes} and PU) when the uncertainty scale is purely ordinal which is beneficial in the case of possibilistic decision making since the qualitative aspect of these decision criteria is the particularity of this theory.

The evaluation of possibilistic influence diagrams starts by the instantiation of the last decision D_m that maximizes the qualitative utility taking into account a set of evidence E that contains the set of nodes with known values. Then for each decision D_i, iterating backwards with $i = m-1, \ldots, 1$ (w.r.t the temporal order of decisions) and considering a set of evidence updated with selected instantiations of decisions in previous steps.

Given a min-based possibilistic network, we propose the following result to compute the optimal instantiation of the decision D_i maximizing the optimistic utility (U_{opt}) :

Proposition 6.1 *The optimal instantiation of the decision D_i maximizing the optimistic utility in a possibilistic network is determined as follows :*

$$U^*_{opt}(D_i, E) = \max_{D_i} \Pi(V_r = T | D_i, E). \tag{6.9}$$

Proof. [Proof of Proposition 6.1]
Using the definition of the optimistic utility in Chapter 2, $U_{opt}(\delta)$ can be expressed by : $U_{opt}(\delta) = \max_{c \in C} min(\Pi(c \mid \delta(c)), u(c, \delta(c)))$.

In a possibilistic network and at a stage i, $U^*_{opt}(D_i, E)$ can be computer as follows : $U^*_{opt}(D_i, E) = \max_{D_i} \left[\max_{Pa'(V_r)} \min(U(Pa(V_r)), \Pi(Pa'(V_r)|D_i, E))\right]$
where $Pa'(V_r)$ is the set of chance nodes in the parents of V_r ($Pa'(V) \subset Pa(V)$). Since $U(Pa(V_r)) = \Pi(V_r = T|Pa(V_r))$, we obtain :
$U^*_{opt}(D_i, E) = \max_{D_i} \left[\max_{Pa'(V_r)} \min(\Pi(V_r = T|Pa(V_r)), \Pi(Pa'(V_r)|D_i, E))\right]$
$U^*_{opt}(D_i, E) = \max_{D_i} \Pi(V_r = T|D_i, E)$. ■

The computation of $\Pi(V_r = T|Dm, E)$ is ensured via propagation algorithms depending on the DAG structure. In fact in the case of singly connected DAGs (DAG which contain no loops) the possibilistic adaptation of Pearl's algorithm is used and the possibilistic adaptation of junction trees propagation are appropriate for multiply connected DAGs (DAG which can contain loops). If these two algorithms are blocked in min-based possibilistic networks, the anytime algorithm can be used [5, 6]. The following proposition is available for the case of pessimistic utilities :

Proposition 6.2 *The optimal instantiation of the decision D_i maximizing the pessimistic* *utility in a possibilistic network is determined as follows :*

$$U^*_{pes}(D_i, E) = \max_{D_i} \min_{Pa'(V_r)} \Pi(V_r = T|Pa'(V_r), D_i, E). \tag{6.10}$$

where $Pa'(V_r)$ is the set of chance node in the parents of the value node V_r.

Proof. [Proof of Proposition 6.2]
Pessimistic utility of a strategy δ is expressed as follows :
$U_{pes}(\delta) = \min_{c \in C} \max(n\Pi(c \mid \delta(c)), u(c, \delta(c)))$ where n is transformation function such that $n\Pi(\delta \geq u_i) = \Pi(\delta < u_i)$. The pessimistic utility of a decision D_i is computed in a possibilistic network as follows :
$U^*_{pes}(D_i, E) = \max_{D_i} \left[\min_{Pa'(V_r)} \max(U(Pa(V_r)), \Pi(Pa'(V_r)|D_i, E))\right]$
where $Pa'(V_r)$ is the set of chance nodes in the parents of V_r ($Pa'(V_r) \in Pa(V_r)$). Since $U(Pa(V_r)) = \Pi(V_r = T|Pa(V_r))$, we obtain :
$U^*_{pes}(D_i, E) = \max_{D_i} \left[\min_{Pa'(V_r)} \max(\Pi(V_r = T|Pa(V_r)), \Pi(Pa'(V_r)|D_i, E))\right]$
$U^*_{pes}(D_i, E) = \max_{D_i} \min_{Pa'(V_r)} \Pi(V_r = T|Pa'(V_r), D_i, E)$. ■

In a previous work [42], we have shown that in the case of binary utilities the transformation of the value node V_r into a chance node includes also the transformation of binary utilities into a single one, namely :

$$\Pi(V_r = T \mid pa(V_r)) = \min(\overline{u}(pa(V_r)), \underline{u}(pa(V_r))). \tag{6.11}$$

and

$$\Pi(V_r = F \mid pa(V_r)) = 1. \tag{6.12}$$

In this case the following proposition is available :

Proposition 6.3 *The optimal instantiation of the decision D_i maximizing the binary utility in a possibilistic network is determined as follows :*

$$PU^*(D_i, E) = \max_{D_i} \Pi(V_r = T|D_i, E). \tag{6.13}$$

Example 6.3 *Let us transform the possibilistic influence diagram in Figure 6.1 into a possibilistic network in Figure 6.4.*

The decision nodes D1 and D2 are transformed into a chance nodes. The possibility distributions relative to these chance nodes are represented in Table 6.7 *and* 6.8*. The value*

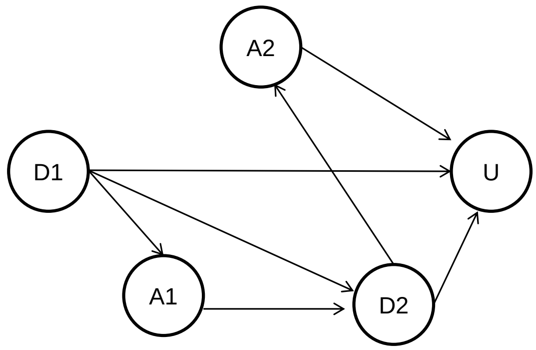

FIGURE 6.4 – Obtained possibilistic network from the transformation of the influence diagram in Figure 6.1

$D1$	$\Pi(D1)$
T	1
F	1

TABLE 6.7 – The conditional possibility for $D1$

$D2$	$D1$	$A1$	$\Pi(D2 \mid D1, A1)$
T	T	T	1
T	T	F	1
T	F	T	1
T	F	F	1
F	T	T	1
F	T	F	1
F	F	T	1
F	F	F	1

TABLE 6.8 – The conditional possibility for $D2$

node U is transformed into a chance node with possibility distributions represented in Table 6.9.

The possibilistic network in Figure 6.4 is multiply connected, so the possibilistic adaptation of junction trees propagation will be used to make inference in this network and to compute the optimal strategy w.r.t optimistic utility.

$D1$	$D2$	$A2$	$\Pi(U = T \mid D1, D2, A2)$	$\Pi(U = F \mid D1, D2, A2)$
T	T	T	0.2	1
T	T	F	0.3	1
T	F	T	0.4	1
T	F	F	0.6	1
F	T	T	1	1
F	T	F	0	1
F	F	T	0.1	1
F	F	F	0.7	1

TABLE 6.9 – The conditional possibility for U

For finding the optimal strategy in the possibilistic network in Figure 6.4, we will start by $D_i = D_2$ *and* $E = \emptyset$. *We compute* $\max_{D_2} \Pi(U = T \mid D_2)$ *we have :*

- $\Pi(U = T \mid D_2 = T) = 0.3$ *and*
- $\Pi(U = T \mid D_2 = F) = 0.4$.

So the best decision for D_2 *is* $D_2 = F$.

Let us now determine the best decision for D_1 *where* $E = (D_2 = F)$, *we have :*

- $\Pi(U = T \mid D_1 = T, D_2 = F) = 0.4$ *and*
- $\Pi(U = T \mid D_1 = F, D_2 = F) = 0.4$.

So, we have two optimal strategies $\Delta^* = (D_1 = T, D_2 = F)$ *and* $(D_1 = F, D_2 = F)$ *with* $U^*_{opt} = 0.4$.

6.4 Conclusion

In this chapter, we have developed possibilistic influence diagrams where possibilistic decision criteria (seen in Chapter 2) are used. We have proposed evaluation algorithms for these graphical models using possibilistic decision trees (detailed in Chapter 4) or possibilistic networks according to the possibilistic decision criterion.

More precisely, if the decision criterion satisfies the monotonicity property then the possibilistic influence diagram can be transformed into a possibilistic decision tree and dynamic programming can be applied to find optimal strategy or it can be transformed into a possibilistic network and possibilistic versions of propagation algorithms should be applied according to the nature of their DAGs (singly or multiply connected). For these types of possibilistic decision criteria, the use of the two indirect methods of evaluation is possible and the choice between them depends on the size of the influence diagram. In fact, for great size it is better to use possibilistic networks as a secondary structure since they are more compact representations.

For possibilistic decision criteria that do not satisfy the monotonicity property, only the transformation into a decision tree is allowed and Branch and Bound algorithm can be applied to find the optimal strategy.

Note that if the decision criterion does not satisfy the monotonicity property and the decision problem contains several variables then the determination of the optimal strategy via its transformation into a possibilistic network is impossible and it cannot be evaluated. In addition, if we proceed by transforming the influence diagram into a decision tree then we obtain a huge tree.

An interesting future work concerns the evaluation of possibilistic influence diagrams with possibilistic Choquet integrals in the case of huge decision problems.

General Conclusion

We have proposed in this thesis a contribution for *possibilistic decision theory* in both single and sequential decision problems.

We have first developed classical decision theories and existing possibilistic decision criteria by giving their axiomatic systems in the style of Von Neumann and Morgenstern and in the style of Savage. Then, we have proposed possibilistic Choquet integrals in order to benefit from possibility theory, to represent qualitative uncertainty, and from Choquet integrals to represent different decision makers behaviors. In fact, we have developed necessity-based Choquet integrals for cautious decision makers and possibility-based Choquet integrals for adventurous decision makers.

Another contribution of this work concerns graphical decision models to deal with sequential decision making, more precisely we have developed *possibilistic decision trees* with different possibilistic decision criterion presented in the first part of our thesis.

More precisely we have proposed a complexity study of decision making in possibilistic decision trees which showed that the strategy optimization problem in possibilistic decision trees is only NP-hard in the case of possibilistic Choquet integrals (Ch_N and Ch_Π) which is not the case of optimistic and pessimistic utility (U_{opt} and U_{pes}), binary utility (PU), possibilistic likely dominance (LN and $L\Pi$) and order of magnitude expected utility ($OMEU$) where this problem is polynomial since they satisfy the monotonicity property.

These results allow us to propose appropriate evaluation algorithms since we show that the *dynamic programming* can be applied in the case of U_{opt}, U_{pes}, PU, LN, $L\Pi$ and $OMEU$ contrarily to the case of Ch_N and Ch_Π where it can provide sub-optimal strategies. For this particular case we have proposed a *Branch and Bound* algorithm that proceeds by implicit enumeration to find the optimal strategy. Then, we have defined three particular classes of possibilistic Choquet integrals that satisfy the monotonicity property and where the polynomial dynamic programming can be applied.

We also proposed an experimental study aiming to compare results of the two algorithms on a synthetic benchmark. This study shows that dynamic programming, even if it generates sub-optimal strategies, allows to have values which are close to those obtained by the Branch and Bound algorithm for small decision trees.

Finally, we have proposed *possibilistic influence diagrams* to deal with huge decision problems where decision trees cannot be generated. More precisely, we have identified several types of possibilistic influence diagrams depending on the quantification of chance and value nodes. To evaluate possibilistic influence diagrams, we have proposed two indirect methods based on their transformation into a secondary structure. The first one transforms possibilistic influence diagrams into possibilistic decision trees and the second one transforms them into possibilistic networks. It is important to note that in the case of possibilistic Choquet integrals, possibilistic influence diagrams cannot be transformed into possibilistic networks since propagation algorithms are a form of dynamic programming This means that for this particular case it is more appropriate to transform the influence diagram into a decision tree and to evaluate it via the Branch and Bound algorithm.

As future work, we can first distinguish direct evaluation of possibilistic influence diagrams in the case of possibilistic Choquet integrals using variable elimination in order to find the optimal strategy in the case of huge decision problems where the transformation into possibilistic decision trees cannot be applied.

Another line of research will be the development of possibilistic unconstrained influence diagrams (UID) [48] to deal with problems where the ordering of the decisions are unspecified. In fact, an anytime algorithm has been proposed in [51] for solving a UID by an indirect method which transforms the UID into a decision tree and performs a search in this tree guided by a heuristic function.

Then, it will be interesting to develop possibilistic hybrid influence diagrams containing a mix of discrete and continuous chance nodes by exploring the solving method based on the approximation of a hybrid influence diagram with a discrete one by discretizing the continuous chance nodes [17].

Bibliographie

[1] M. Allais. Le comportement de l'homme rationnel devant le risque : critique des postulats et axiomes de l'*é*cole am*é*ricaine. Econometrica, Vol 21, pages 503-546, 1953.

[2] R. Bellman, Dynamic programming, Princeton Univ. Press, Princeton, N.J., 1957.

[3] C. Bielza, J.A. Fernández del Pozo and P. Lucas. Explaining clinical decisions by extracting regularity patterns. Decision Support Systems, Vol 44, pages 397408, 2008.

[4] D. Bernoulli. Exposition of a New Theory on the Measurement of Risk. Econometrica 22, Vol 1, pages 2236, 1738.

[5] N. Ben Amor. Qualitative Possibilistic Graphical Models : From Independence to Propagation Algorithms. Ph.D. dissertation, 2002.

[6] N. Ben Amor, S.Benferhat and K.Mellouli. Anytime propagation algorithm for min-based possibilistic graphs. Soft Computing a fusion of foundations methodologies and applications, Springer Verlag, Vol 8, pages 150-161, 2001.

[7] N. Ben Amor, H. Fargier and W. Guezguez. Necessity-based Choquet integrals for sequential decision making under uncertainty, IPMU'2010, pages 222-236, 2010.

[8] N. Ben Amor, H. Fargier, W. Guezguez. Les intégrales de Choquet basées sur les mesures de nécessité pour la prise de décision séquentielle sous incertitude, Logique Floue et ses Applications (LFA), 2010.

[9] N. Ben Amor, H. Fargier, W. Guezguez. On the Complexity of Decision Making in Possibilistic Decision Trees, Uncertainty in Artificial Intelligence, Vol 27, pages 203-211, 2011.

[10] C. Borgelt, J. Gebhardt, and Rudolf Kruse. Possibilistic graphical models. In Proc. of ISSEK98 (Udine, Italy),1998.

[11] J. Busemeyer, R. Hastie and L. Douglas. Decision making from a cognitive perspective. Academic Press, pages 180-190, 1995.

[12] J. Buckley and TJ. Dudley, DBA. How Gerber Used a Decision Tree in Strategic Decision-Making, Graziadio Business Review, Volume 2 Issue 3, 1999.

[13] G. Choquet. Th*é*orie des capacit*é*s. Annales de l'institut Fourier (Grenoble), Vol 5, pages 131-295, 1953.

[14] G.F. Cooper. A method for using belief networks as influence diagrams. Fourth workshop on uncertainty in artificial intelligence, 1988.

[15] G.F. Cooper. Computational complexity of probabilistic inference using Bayesian belief networks. Artificial Intelligence 42, pages 393-405, 1990.

[16] M. Crowley. Evaluating Influence Diagrams. 2004.

[17] B. R. Cobb and P. P. Shenoy. Decision making with hybrid influence diagrams using mixtures of truncated exponentials. European Journal of Operational Research, 186(1), pages 261275, 2008.

[18] D. Dubois and H. Prade. Possibility theory, an approach to computerized processing of uncertainty. Plenum Press, New York, NY, 1988.

[19] D. Dubois and H. Prade. Representing partial ignorance. IEEE Systems Machine and Cybernetic 26, pages 361-377, 1996.

[20] D. Dubois and H. Prade. An introductory survey of possibility theory and its recent developments, 1998.

[21] D. Dubois, L. Gogo, H. Prade and A. Zapico. Making decision in a qualitative setting : from decision under uncertainty to case-based decision.

[22] D. Dubois and H. Prade. Epistemic entrenchment and possibilistic logic. Artificial Intelligence, vol 50, pages 223-239, 1991.

[23] D. Dubois, Hélène Fargier, Henri Prade and Patrice Perny. Qualitative decision theory : from savage's axioms to nonmonotonic reasoning. ACM 4, volume 49, pages 455-495, 2002.

[24] D. Dubois and L. Godo and H. Prade and A. Zapico. Making decision in a qualitative setting : from decision under uncertainty to case-based decision. KR'06, Trento, 1998.

[25] D. Dubois and H. Prade. Possibility theory as a basis for qualitative decision theory. IJCAI'95, Montreal, pages 1924-1930, 1995.

[26] D. Dubois, H. Prade and R. Sabbadin. Qualitative decision theory with Sugeno integrals. 14th conference on Uncertainty in Artificial Intelligence (UAI), Madison, WI, USA, pages 121-128, 1998.

[27] D. Dubois, H. Prade and R. Sabbadin. Towards axiomatic foundations for decsion under qualitative uncertainty. 7th World Congress of the Inter. Fuzzy Systems Association (IFSA'97), pages 441-446, 1997.

[28] D. Dubois, H. Prade and R. Sabbadin. Decision-theoretic foundations of qualitative possibility theory. EJOR 128, pages 459-478, 2001.

[29] D. Dubois and H. Fargier and P. Perny. Qualitative decision theory with preference relations and comparative uncertainty : An axiomatic approach. Artificial Intelligence, volume 148, number 1-2, pages 219-260, 2003.

[30] D. Ellsberg. Risk, ambiguity and and the Savage axioms. The quarterly journal of economics, Vol 75, pages 643-669, 1961.

[31] H. Fargier and P. Perny. Qualitative Models for Decision Under Uncertainty without the Commensurability Assumption, UAI, pages 188-195, 1999.

[32] P. Fonck. Propagating uncertainty in a directed acyclic graph. In IPMU92, pages 1720, 1992.

[33] L. Garcia et R. Sabbadin. Possibilistic influence diagrams, ECAI, pages 372-376, 2006.

[34] L. Garcia et R. Sabbadin. Complexity results and algorithms for Possibilistic Influence Diagrams. Artificial Intelligence, Volume 172, pages 1018-1044, 2008.

[35] P.H. Giang and P.P. Shenoy. A Qualitative Linear Utility Theory for Spohns Theory of Epistemic Beliefs, Uncertainty in Artificial Intelligence (UAI), C. Boutilier and M. Goldszmidt (eds.), Vol. 16, pages 220-229, Morgan Kaufmann, San Francisco, CA, 2000.

[36] P.H. Giang and P.P. Shenoy. Two axiomatic approaches to decision making using possibility theory. European journal of operational research, vol 162, No.2, pages 450-467, 2005.

[37] P. H. Giang and P. P. Shenoy. A Comparison of Axiomatic Approaches to Qualitative Decision Making Using Possibility Theory. UAI, pages 162-170, 2001.

[38] I. Gilboa. Maxmin expected utility with non unique prior. Journal of mathematical economics 18, pages 141-153, 1988.

[39] J. Gildas and O. Spanjaard Rank-dependent drobability weighting in sequential decision problems under uncertainty. Association for the advancement of artificial intelligence, 2008.

[40] M. Gómez and all. A Graphical Decision-Theoretic Model for Neonatal Jaundice. Medical Decision Making, Vol 27, 3, pages 250-265, 2007.

[41] W. Guezguez and N. Ben Amor. Possibilistic influence diagrams using information fusion. 12th International conference, Information Processing and Management of Uncertainty in Knowledge Based Systems (IPMU), pages 345-352, Torremolinos (Malaga), June 22-27, 2008.

[42] W. Guezguez, N. Ben Amor and Khaled Mellouli. Qualitative possibilistic influence diagrams based on qualitative possibilistic utilities. European Journal of Operational Research 195, pages 223-238, 2009.

[43] R.A. Howard and J.E. Matheson. Influence diagrams. In the principles and applications of decision analysis, vol II, R.A Howard and J.E Matheson (eds). Strategic decisions group, Menlo Park, Calif, 1984.

[44] L. Hurwicz. Some Specification Problems and Applications to Econometric Models, Econometrica, Vol. 19, pages 343-44, 1951.

[45] J.Y. Jaffray. Implementing resolute choice under uncertainty. Fourtheenth conference of uncertainty in artificial intelligence, pages 282-288, 1998.

[46] F.V. Jensen. An introduction to Bayesian Networks. Taylor and Francis, London, United Kingdom, 1996.

[47] F.V. Jensen. Bayesian networks and decision graphs. Springer, statistics for engineering and information science, 2002.

[48] F. V. Jensen and M. Vomlelova. Unconstrained influence diagrams. In Proceedings of the 18th Annual Conference on Uncertainty in Artificial Intelligence (UAI'02), pages 234-241, San Francisco, CA, 2002. Morgan Kaufmann.

[49] J. Kim and J. Pearl. Convice ; a conversational inference consolidation engine. IEEE Trans. on Systems, Man and Cybernetics, 17-120132, 1987.

[50] R. D. Luce and H. Raiffa. Games and Decision. John Wiley and Sons, 1957.

[51] M. Luque, T. D. Nielsen and F. V. Jensen. An anytime algorithm for evaluating unconstrained influence diagrams. The fourth European Workshop on Probabilistic Graphical Models (PGM08), 2008.

[52] SE. MacNeill and PA. Lichtenberg. The MacNeill-Lichtenberg Decision Tree : a unique method of triaging mental health problems in older medical rehabilitation patients, Arch Phys Med Rehabil 81(5),pages 618-622, 2000.

[53] R. Marinesc and N. Wilson. Order-of-Magnitude Influence Diagrams, UAI'2011, 2011.

[54] E.F. McClennen. Rationality and Dynamic Choice Foundational Explorations, Cambridge University Press, Cambridge, 1990.

[55] A. B. Meijer. A Methodology for Automated Extraction of the Optimal Pathways from Influence Diagrams. Lecture notes in computer sciences, Vol 4594/2007, pages 351-358, 2007.

[56] W. Nesrine Zemirli and Lynda Tamine-Lechani and Mohand Boughanem, A Personalized Retrieval Model based on Influence Diagrams, Context-Based Information Retrieval (CIR), Vol 326, 2008.

[57] J.V. Neumann and O. Morgenstern. Theory of games and economic behavior. Princeton University Press, 1948.

[58] D.N Nielson and J.Y Jaffray. Dynamic decision making without expected utility : an operationnal approach. European journal of operational research 169, pages 226-246, 2004.

[59] Z. Pawlak. Rough sets. Auszug aus : Kunstliche Intelligent, Heft 3/01, 1991.

[60] J. Pearl. Fusion propagation and structuring in belief networks. Artificial Intelligence, 29, pages 241-288, 1986.

[61] J. Pearl. Probabilistic reasoning in intelligent systems networks of plausible inference. Morgan Kaufmann, San Francisco (California), 1988.

[62] J. Pearl. From conditional oughts to qualitative decision theory. UAI'93, pages 12-20, 1993.

[63] J. Pearl. Causality Models, Reasoning and Inference. Cambridge University Press, 2000.

[64] J. Quiggin. A theory of anticipated utility. Journal of economic behavior and organisation, Vol 3, pages 323-343, 1982.

[65] H. Raiffa. Decision Analysis : Introductory lectures on choices under uncertainty. Addison-Wesley, Reading, MA, 1968.

[66] Y. R*ébillé*. Decision making over necessity measures through the Choquet integral criterion. Fuzzy sets and systems 157, pages 3025-3039, 2006.

[67] R. Sabbadin. Possibilistic Markov decision processes. Engineering Applications of Artificial Intelligence 14, pages 287-300, 2001.

[68] L.J. Savage, The foundations of statistics. Dover, New York, 1972.

[69] D. Schmeidler. Subjective probability and expected utility without additivity. Econometrica, Vol 57, pages 517-587, 1989.

[70] R.D. Shachter. Evaluating influence diagrams. Operation Research 34, pages 871-882, 1986.

[71] R.D. Shachter. An ordered examination of influence diagrams. NETWORKS, Vol. 20, pages 535-563, 1990.

[72] G. Shafer. A mathematical theory of evidence. Princeton Univ.Press, Princeton, VJ, 1976.

[73] P.P Shenoy. Valuation based systems : A framework for managing uncertainty in expert systems. Fuzzy Logic for the Management of Uncertainty (L. A. Zadeh and J. Kacprzyk, Eds.), John Wiley and Sons, New York, NY, pages 83-104, 1992.

[74] P.P Shenoy. A comparison of graphical techniques for decision analysis. European Journal of Operational Research, Vol 78, pages 1-21, 1994.

[75] Pierre Simon and Marquis de Laplace. A philosophical essay on probabilities, translated from the sixth French edition by Frederick Wilson Truscott and Frederick Lincoln Emory, 1902.

[76] W. Spohn (1988). Ordinal Conditional Functions, A Dynamic Theory of Epistemic States. In W. L. Harper B. Skyrms (eds.), Causation in Decision, Belief Change, and Statistics, vol. II. Kluwer, 1988.

[77] W. Spohn. A general non probabilistic theory of inductive reasoning, UAI'90 (Elsevier Science), pages315-322, 1990.

[78] M. Sugeno. Theory of fuzzy integral and its applications. PhD thesis, Tokyo institute of technology, Tokyo, 1974.

[79] M. Sugeno. Fuzzy automata and decision processes, chapter fuzzy measures and fuzzy integrals a survey. M.M. Gupta, G.N. Saridis and B.R. Gaines (Eds.), North Holland, Amsterdam, pages 89-102, 1977.

[80] D. Surowik. Leonard Savage's mathematical theory of decision. Studies in logic ; grammar and rhrtoric 5 (18), 2002.

[81] J.A. Tatman and R.D. Shachter. Dynamic programming and influence diagrams. IEEE Transactions on systems, Man and Cybernetics 20, pages 365-379, 1990.

[82] C. Tseng and Y. Lin. Financial Computational Intelligence, Computing in Economics and Finance 2005, Vol 42, 2005.

[83] J. Ullberg, R. Lagerström and M. Ekstedt, A Framework for Interoperability Analysis on the Semantic Web using Architecture Models, International Workshop on Enterprise Interoperability (IWEI), 2008.

[84] P. Walley. Statistical Reasoning with imprecise probabilities. Chapman and Hall, London, 1991.

[85] A.Wald. Statistical decision functions which minimize the maximum risk. The Annals of Mathematics, 46(2), pages 265-280, 1945.

[86] A.Wald. Statistical Decision Functions, John Wiley, NY, 1950.

[87] N. Wilson. An order of magnitude calculus. UAI'95, pages 548-555, 1995.

[88] L.Zadeh, Fuzzy sets, Information and Control 8, pages 338353, 1965.

[89] L. Zadeh. Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems 1, pages 3-28, 1978.

[90] L. Zadeh. A theory of approximate reasoning. Machine Intelligence (J. E. Ayes, D.Mitchie and L. I. Mikulich, Eds.), Ellis Horwood, Chichester, U.K., 9, pages 149194, 1979.