Université de Tunis

Institut Supérieur de Gestion de Tunis

Laboratoire de Recherche: LARODEC

MEMOIRE DE MASTÈRE EN INFORMATIQUE DE GESTION

Possibilistic Influence Diagrams

Réalisé par: Wided Guezguez

Encadré par: Pr. Khaled Mellouli, Professeur (IHEC, Tunis)
Dr. Nahla Ben Amor, Maître assistante (ISG, Tunis)

2005 2006

Abstract

This work proposes a new approach for decision making under uncertainty based on influence diagrams and possibility theory. The so-called *possibilistic influence diagrams* aim to extend standard influence diagrams in order to avoid difficulties during the specification of probability distributions relative to chance nodes and utilities relative to value nodes. In fact, it is easier for experts to quantify dependencies between chance nodes qualitatively via possibility distributions. Moreover, in some situations decision makers can only provide a preferential relation between different consequences which can be modelled by possibility utilities. Different combinations of the quantification between chance and utility nodes offer several kinds of possibilistic influence diagrams. In this work, we will focus on homogeneous ones, i.e. those where both chance and value nodes are quantified similarly. This leads us to treat product-based and min-based influence diagrams. To evaluate these possibilistic influence diagrams, we have opted for an indirect method based on the transformation of the initial network into a causal possibilistic network. We have proposed a possibilistic influence diagrams toolbox (PIDT) implementing different evaluation algorithms developed in this work.

Résumé

Ce travail propose une nouvelle approche pour la prise de décisions dans un environnement incertain basée sur les diagrammes d'influence et la théorie de possibilité. Cette nouvelle approche, appelée *diagrammes d'influence possibilistes*, vise á étendre les diagrammes d'influence standard dans le but d'éviter les difficultés durant la spcification des distributions de probabilité relatives aux noeud d'incertitude et des utilités relatives aux noeuds de valeur. En fait, il est plus facile pour les experts de quantifier les dépendances entre les nuds d'incertitude travers des distributions de possibilit qualitatives. En plus, pour certaines situations les décideurs peuvent seulement établir un ordre de préférence entre les différentes conséquences qui peut être modélisée par des utilités possibilistes. Les différentes combinaisons de quantification entre les noeuds d'incertitude et les noeuds de valeur offrent plusieurs types de diagrammes d'influence possibilistes. Dans ce travail, on s'intéresse ceux qui sont homogènes c'est-á-dire ceux qui contiennent des noeuds d'incertitudes et des noeuds de valeur similairement quantifiés. Ce qui nous emmène á traiter les diagrammes d'influence possibilistes basés sur l'opérateur produit ainsi ceux basés sur l'opérateur min. Pour évaluer ces différents diagrammes d'influence, nous avons opté pour une méthode indirecte basée sur la transformation des réseaux initiaux aux réseaux causaux possibilistes. On a proposé un outil pour les diagrammes d'influence possibilistes implémentant les différents algorithmes d'évaluation développés dans ce travail.

Acknowledgments

I would like to thank my advisor Professor Khaled Mellouli for supporting my first steps in the world of research. Thank you for your encouragements and for all the interested discussion during the elaboration of this work, despite the little time, you had available.

I am deeply indebted to 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.

Thanks to all the members of my laboratory LARODEC in Institut Suprieur de Gestion de Tunis and to all my friends for the good company.

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 for their steadfast support, their love and patience. To them I dedicate this work.

Thanks "Dad" for all good things you have done for me, I would have been lost without you.

Contents

Introduction **1**

1 Influence Diagrams **4**

1.1 Introduction 4

1.2 Formal definitions of influence diagrams 5

1.2.1 Graphical component 5

1.2.2 Numerical component 10

1.3 Evaluation of influence diagrams 15

1.3.1 Bayesian networks 16

1.3.2 Cooper's method 17

1.3.3 Other indirect methods 19

1.4 Conclusion 22

2 Possibility theory **23**

2.1 Introduction 23

2.2 Notations 24

2.3 Basics of possibility theory 24

2.3.1 Possibility distribution 24

2.3.2 Possibility and necessity measures 26

2.4 Qualitative possibility theory 27

2.5 Combination of possibility distributions 28
2.5.1 The conjunctive fusion 29
2.5.2 The disjunctive fusion 30
2.5.3 Possibilistic mixture 30
2.6 Possibilistic conditioning 30
2.7 Qualitative possibilistic utility 32
2.7.1 Notations 32
2.7.2 Pessimistic and optimistic utilities 33
2.7.3 Limits of pessimistic and optimistic utilities 38
2.7.4 Binary possibilistic utility 39
2.8 Conclusion 44

3 Possibilistic Influence Diagrams 45

3.1 Introduction 45
3.2 Definition of possibilistic influence diagrams 46
3.2.1 Product-based possibilistic influence diagrams 48
3.2.2 Min-based possibilistic influence diagrams 49
3.3 Evaluation of possibilistic influence diagrams 50
3.3.1 Decision nodes transformation 51
3.3.2 Value node transformation 51
3.4 Evaluation of product-based possibilistic influence diagrams 52
3.5 Evaluation of min-based possibilistic influence diagram 55
3.5.1 Evaluation of ΠID^{O}_{min} 55
3.5.2 Evaluation of $\Pi ID^{\Pi O}_{min}$ 58
3.6 Conclusion 62

4 Implementation and Experimentations **63**

4.1 Introduction 63

4.2 Implementation 64

4.2.1 Major variables of PIDT 64

4.2.2 Principal programs of PIDT 65

4.3 Illustrative example: Project risk analysis decision problem 67

4.4 Conclusion 77

Conclusion **79**

List of Figures

1.1 Possible cases of arcs in influence diagrams 6

1.2 Graphical component of the influence diagram (Example 1) 7

1.3 Possible connections 8

1.4 An example of influence diagram 9

1.5 Obtained Bayesian network 18

3.1 Influence Diagram 48

3.2 Obtained possibilistic network 53

4.1 PIDT architecture 66

4.2 An Influence Diagram for project risk analysis 69

4.3 Result causal network for project risk analysis 75

List of Tables

1.1 A priori and conditional probabilities 13

1.2 Physician's utilities 14

1.3 A priori probability distribution for Tr 19

1.4 Conditional probability distribution for V 19

2.1 Conventions for possibility distribution π_ω 25

2.2 Possibility measure Π 26

2.3 Necessity measure N 27

2.4 The possibility distributions π_1 and π_2 29

2.5 Possibilistic conditioning 31

2.6 Possibilistic lotteries 36

2.7 The utility assignment 36

2.8 The function n and h 37

3.1 Different variants of possibilistic influence diagrams 47

3.2 A priori and conditional possibility distributions for chance nodes 49

3.3 Cardinal utilities 49

3.4 The possibility distribution of the node D 54

3.5 A priori and conditional possibility distribution of the value node V 54

3.6 Results of propagation 55

3.7 Numerical utilities 57

3.8 The possibility distribution of the value node 58

3.9 Binary utilities and their transformation using k^{-1} 60

3.10 New possibility distribution π' 60

3.11 Normalized possibility distribution for V 61

4.1 A priori possibility distributions concerning Site Selection decision 70

4.2 A priori possibility distributions concerning New Technology decision 70

4.3 A priori possibility distributions concerning Credit decision 70

4.4 A priori possibility distributions for chance nodes ER and TSU 70

4.5 Conditional possibility distributions concerning Site Selection decision . . . 71

4.6 Conditional possibility distributions concerning New Technology decision . 72

4.7 Conditional possibility distribution concerning Fund decision 72

Introduction

Decision analysis represents an important area of research since it helps private and public organizations in making strategic and tactical decisions.

Indeed, several approaches have been developed to assists decision makers in their tasks such that dynamic programming, simulation techniques and graphical decision models.

In our work, we are interested in graphical decision models which provide an efficient tool for decision analysis. Indeed, they represent a compact, intuitive and simple representation of decision problems. Moreover, they are a powerful tools for communication, elicitation and representation of human knowledge.

Within existing graphical models, we are in particular interested in influence diagrams (IDs), initially developed by Howard and Matheson in 1981 [20]. Influence diagrams are a popular framework representing a decision maker's belief and preferences about a sequence of decisions to be made under uncertainty [20].

Nevertheless, in most real problems it is not always obvious for experts to provide probability distributions representing dependences between uncertain variables. Indeed, it is generally more easier for them to express their uncertainty qualitatively by ranking different states of the world.

Moreover, decision makers encounter several difficulties when expressing their utilities and it is generally more easier for them to provide a preferential relation between different consequences rather than exact numerical values.

In such situations, standard influence diagrams cannot be applied. Hence, our idea to extend them using possibility theory for specifying the numerical component of influence diagrams. More precisely, dependencies between chance nodes will be expressed via pos-

sibility distributions and value nodes will be quantified using ordinal [30] and qualitative possibilistic utilities [17].

This non-classical theory of uncertainty, initially proposed by Zadeh [55] and developed by Dubois and Prade [9, 11], offers a natural and simple framework to handle imperfect data. Possibility theory can be interepted in two ways: qualitatively, if the handled values reflect only an ordering between the different states of the world or quantitatively, if these values make sense in the ranking scale.

This work proposes a new approach, so-called *possibilistic influence diagrams*, presenting a powerful tool for decision making under uncertainty.

Different combinations of quantification between chance and value nodes offer several kinds of possibilistic influence diagrams which can be summarized as follows:

- *Product-based possibilistic influence diagrams* where both dependences between chance nodes and value nodes are quantified in a genuine numerical setting.
- *Min-based possibilistic influence diagrams* 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.
- *Mixed possibilistic influence diagrams* where dependencies between chance nodes and value nodes are not quantified in the same setting.

In this work, we will focus on homogeneous ones, i.e. those where both chance and value nodes are quantified similarly. This leads us to treat product-based and min-based influence diagrams.

Once the influence diagram constructed, it can be used to identify optimal decision strategies via evaluation methods. Standard influence diagrams can be evaluated through direct methods [42, 50] or indirect ones [6, 44, 53, 52, 8, 40] based on a transformation of the initial structure into a secondary one corresponding to a Bayesian network.

The success of indirect methods motivates us to develop an indirect method to evaluate possibilistic influence diagrams. This allows us to avoid heavy computations of direct methods and to benefit from the recent developments of possibilistic causal networks and their propagation algorithms [3].

This report is organized as follows: *Chapter 1* presents influence diagrams, it details how these graphical decision models are used to structure and resolve decision problems. *Chapter 2* presents theorical aspects regarding possibility theory. *Chapter 3* develops a possibilistic counterpart of standard influence diagrams based on qualitative and quantitative aspects of possibility theory. Moreover, in this chapter we present evaluation methods of possibilistic influence diagrams. Finally, *Chapter 4* deals with the implementation of the possibilistic influence diagram toolbox with Matlab 6.3 allowing the modelisation and the evaluation of different kinds of possibilistic influence diagrams.

Chapter 1

Influence Diagrams

1.1 Introduction

The decision making process is composed of three principal steps, the first one consists on data collection and problem identification. In the second step, a model is formulated to represent the problem. Finally, a solution maximizing the satisfaction of the decision maker is generated.

We can distinguish decision making under certainty which assumes that complete knowledge is available and that the decision maker (DM) knows exact information about the problem. Whereas, decision making under uncertainty assumes that the DM does not know or cannot estimate exact informations about the actual problem.

Several models manipulating decision problems under uncertainty have been proposed, we can mention decision trees [38], influence diagrams [20] and valuation based systems [46].

The major weaknesses of decision trees is the combinatorial explosiveness of their representations for problems with many variables and their inflexibility in representing probability models and information constraints.

Valuation based systems (VBS) [46] are general tools since uncertainty may be represented by probabilily distributions, possibility distribution or belief function. However, possibilistic VBS are not appropriate for the qualitative interpretation of possibility distribution because they use the product operator as a combination operator.

The ID formalism presents a powerful tool for representing decision problems, it was used in several real applications such that probabilistic risk analysis in the aerospace sector [2], formulating risk into research and engineering projects [27], medical diagnosis [28] and military domain [51].

This chapter proposes an overview of influence diagrams, we will focus on the basis of influence diagram (ID). First, section 1.2 gives some useful definitions. Section 1.3 presents different methods of influence diagram's evaluation.

1.2 Formal definitions of influence diagrams

Influence diagrams (IDs) (Howard and Matheson, 1981) are a popular framework representing a decision maker's belief and preferences about a sequence of decisions to be made under uncertainty [20].
Influence diagrams are composed of a graphical component and a numerical one detailed in what follows.

1.2.1 Graphical component

The graphical component (or qualitative component) of IDs is composed of 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 subsets C, D and V such that:

- $C = \{C_1...C_n\}$ is a set of chance nodes which represent relevant uncertain factors for decision problem. Chance nodes are represented by circles.
- $D = \{D_1...D_m\}$ is a set of decision nodes which depict decision options, they have a temporal order. Decision nodes are represented by rectangles.
- $V = \{V_1...V_k\}$ is a set of value nodes which represent utilities to be maximized, they are represented by lozenges.

In what follows c_{ij} (resp. d_{ij}, v_{ij}) denotes an instance of the variable C_i (resp. D_i, V_i).

Let x be a node pertaining to N (chance, decision or value one) there:

- Ω_x denotes a set of possible values of the node x (the frame of x).
- S(x) denotes the set of direct successors of x.
- Pa(x) denotes the set of parents of x and pa(x) is an instance of Pa(x).

The set **A** consists of a set of arcs in the graph, these arcs have different meanings according to their targets. We can distinguish two types of arcs:

- *conditional arcs* (into chance and value nodes), those that have as target chance nodes represent probabilistic dependence.
- *informational arcs* (into decision nodes) which imply time precedence.

Possible types of arcs are presented by Figure 1.1.

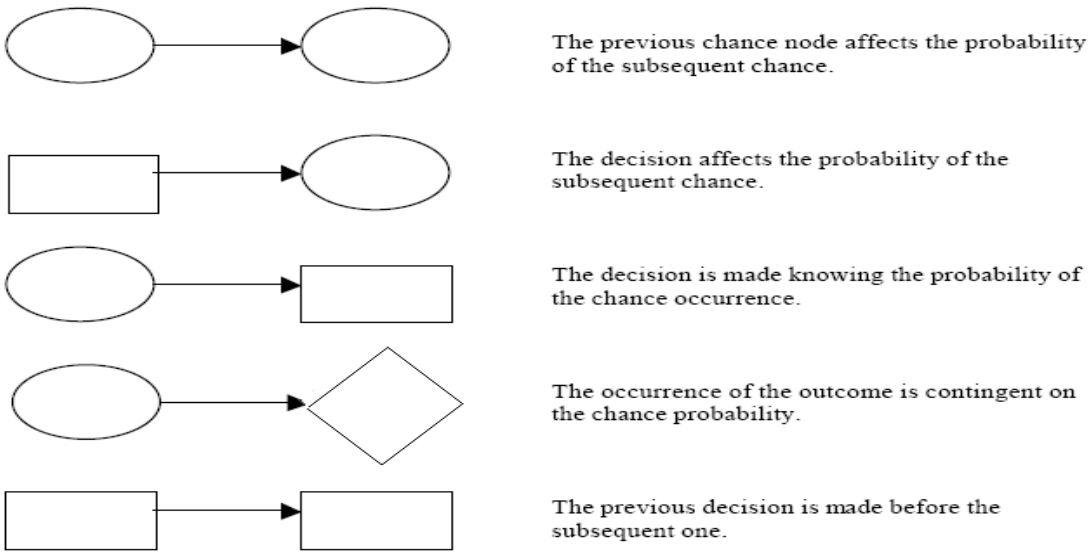

Figure 1.1: Possible cases of arcs in influence diagrams

Example 1.1 *Let us state a simple problem of Medical Diagnosis [47]: 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 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 1.2 presents an influence diagram for the medical diagnosis problem, it contains 3 chance nodes, 1 decision node and 1 value node.

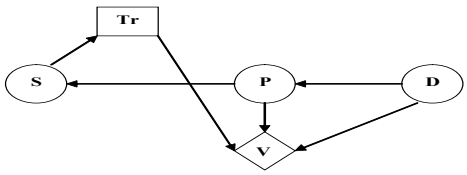

Figure 1.2: Graphical component of the influence diagram (Example 1)

The graphical component of IDs encodes different conditional independences between chance nodes. Jensen [23] has investigated the problem of determining exactly what independences are implied by the DAG structure in IDs, he affirms that utility nodes and links into decision nodes must be ignored.
According to when it is observed, the set of chance nodes C is partitioned into three subsets:

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 is taken and before the decision D_{i+1} is taken.
3. SC_n is the set of chance nodes never observed or observed too late to have an impact on any decision.

Decision nodes act differently from remaining nodes, since they are not quantified. In fact, when we have an arc from a decision node D_i to a chance node C_i, the question is whether an action from D_i will have influence on C_i. So, links from decision to chance nodes reflect causality.

Indeed, Jensen [23] proposed three types of connection between three variables X, Y and Z in causal networks:

1. **Serial connection** X has a causal influence on Y that has a causal influence on Z (see Figure 1.3(a)). Information may be transmitted through a serial connection, unless the state of the variable Y in the connection is known.

2. **Diverging connection** Y has a causal influence on both X and Z (see Figure 1.3(b)). Information may be transmitted through a diverging connection unless, the state of the variable Y in the connection is known.

3. **Converging connection** Both X and Z has a causal influence on Y (see Figure 1.3(c)). Information may only be transmitted through a converging connection if either information about the state of the variable Y in the connection or its descendants is available.

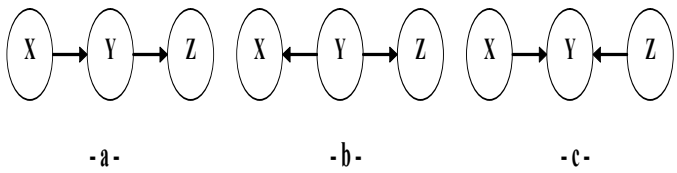

Figure 1.3: Possible connections

Using the rule of evidence's transmission over converging connection, Jensen gives the proposition 1.1 presenting the d-separation criterion for IDs.
In fact, d-separation for influence diagrams is performed slightly differently from the way it is done for Bayesian network [24, 22]. More precisely, utility nodes and links into decision nodes are ignored.

Proposition 1.1 *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:

$$P(C_l \mid D_j) = P(C_l) \tag{1.1}$$

(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:

$$P(C_l \mid D_j, W) = P(C_l \mid W) \tag{1.2}$$

Proof 1.1 *(i) D_j has no parents because arcs into decision nodes are ignored. So, any impact from D_j must follow the direction of a link from D_j. The only way the impact can*

start going in the opposite direction of a link is if it meets a converging connection at a chance node C_b, and then it can only to do so if either C_b or one of its children C_c have received evidence. Since D_j is the only evidence, C_b or one of its children C_c cannot having evidence. Hence, if we suppose that D_j and C_a are not d-separated there must be a directed path from D_j to C_a. Since $C_a < D_j$ in the temporel ordering, there is a directed path from C_a to D_j. There cannot be a directed path from D_j to C_a because the graph is acyclic.

(ii) The same proof in (i) is used. By following directions of links from D_j, we can only start going opposite to the direction by meeting evidence. Since, all evidence is prior in the temporal ordering, we know from (i) that we cannot meet it.

Example 1.2 *Given the influence diagram represented by the DAG in Figure 1.4, we have $SC_0 = \emptyset$, $SC_1 = \{T\}$ and $SC_2 = \{A, B, C\}$.*
A and T are d-separated from D_2, this means that any action from D_2 has no impact on T.

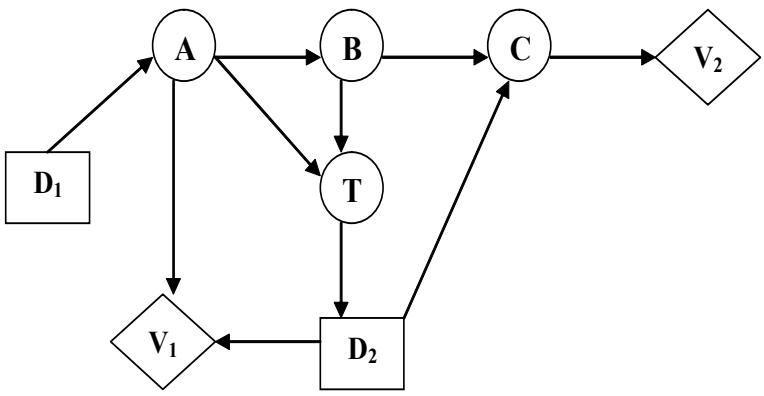

Figure 1.4: An example of influence diagram

We can distinguish some particular cases of influence diagram as follows:

Definition 1.1 *An ID is said to be **proper**, if it is an unambiguous representation of a single decision maker's view of the world. Then, it is said to be **oriented**, if it contains a value node [42].*

Definition 1.2 *An ID is said to be **regular**, if it satisfies the following conditions:*

1. *The directed graph has no cycles,*
2. *The value node, if present, has no successors,*
3. *There is a directed path that contains all of the decision nodes.*

Example 1.3 *Let us consider the ID represented in Figure 1.2. This ID does not contain any cycles, the value node V has no successors and it contains only one decision node T consequently, it is regular and proper.*

In our work, we are only interested in proper, oriented and regular IDs.

1.2.2 Numerical component

The numerical component (or quantitative component) of IDs consists in evaluating 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. 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:

$$\sum_{c_{ij} \in \Omega_{C_i}} P(c_{ij}) = 1 \tag{1.3}$$

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

$$\sum_{c_{ij} \in \Omega_{C_i}} P(c_{ij} \mid Pa(C_i)) = 1 \tag{1.4}$$

Chance nodes represent uncertain variables characterizing decision problem. Each decision's alternative may have several consequences according to uncertain 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 is quantified by a utility function, denoted U, in the context of its parents.

In the literature, two kinds of utility have been defined:

1. *Cardinal utility* proposed by Walras, Menger and Jevons in late 1870s [1]. In this case, we suppose that decision makers are able to give a cardinal number representing their satisfaction. This cardinal approach supposes that the following axioms should be satisfied [1]:

   - Axiom 1. **Measurability**: the decision maker is supposed to be able to assign to each consequence a numerical value reflecting his satisfaction.
   - Axiom 2. **Independence**: the utility of a consequence A is independent of another consequence B.
   - Axiom 3. **Additivity**: for two consequences A and B, total utility U(A,B) is the sum of utilities of each consequence, namely $U(A, B) = U(A) + U(B)$.

   The problem with cardinal utility functions comes from the difficulty in expressing utilities by numerical values, finding the appropriate measurement index (metric) and assuming the independence axiom which is irrational.

2. *Ordinal utility* proposed by Von Neumann and Morgenstern [30]. In this case, we suppose that decision makers can only define a total preference's relation on the set of consequences. Let A, B and C be three possible consequences and $\succeq$ be the relative preference relation then:

   - $A \succeq B$ means that A is more preferred than B.
   - $A \preceq B$ means that A is less preferred than B.
   - $A \sim B$ means that A and B are equally preferred.

   Von Neumann and Morgenstern [30] have defined the following axiom's system for any preference relation:

   - Axiom 1. **Completeness (Orderability)**: it is always possible to state either that $A \succeq B$ or $B \succeq A$ or $A \sim B$.
   - Axiom 2. **Reflexivility**: any consequence A is always at least as preferred as itself: $A \succeq A$.
   - Axiom 3. **Transitivity**: if $A \succeq B$ and $B \succeq C$ then $A \succeq C$.

- Axiom 4. **Continuity**: if B is between A and C in preference, then there is a probability p for which the rational agent (DM) will be indifferent between consequence B and the lottery in which A comes with probability p, C with probability (1-P).

$$(A \succeq B \succeq C) \rightarrow \exists p\ [p : A; (1-p) : C] \sim B$$

- Axiom 5. **Substitutability**: if a decision maker is indifferent between two lotteries, A and B, then there is a more complex lottery in which A can be substitued with B.

$$(A \sim B) \rightarrow \exists p\ [p : A; (1-p) : C] \sim [p : B; (1-p) : C]$$

- Axiom 6. **Monotonicity**: if a decision maker prefers A to B, then he must prefer the lottery in which A occurs with a higher probability.

$$(A \succeq B) \rightarrow \exists (p > q \Leftrightarrow [p : A; (1-p) : B] \succ [q : A(1-q) : B])$$

- Axiom 7. **Decomposability**: compound lotteries can be reduced to simpler lotteries using the laws of probability.

$$[p : A; (1-p) : [q : B; (1-q) : C]] \Rightarrow [p : A; (1-p)q : B; (1-p)(1-q) : C]$$

Definition 1.3 *If a preference relation $\succeq$ satisfies Axiom 1, 2 and 3 then there exists a utility function U such that:*

$$\begin{aligned} A \succeq B \quad iff \quad U(A) \geq U(B) \\ A \sim B \quad iff \quad U(A) = U(B) \end{aligned} \tag{1.5}$$

In this case, utility levels are an ordinal concept, for instance if U(A) = 6 and U(B) = 2 then the consequence A is preferred to B, but A is not preferred three times as much as is B.

Generally, each decision d_{ij} induces a probability distribution, denoted by $P_{d_{ij}}$ over the set of consequences denoted X. Namely, $P_{d_{ij}}(X)$ is the probability of having the set of outcomes X under decision d_{ij}, then, the *expected utility* of decision d_{ij} is defined as follows [30]:

$$EU(d_{ij}) = \sum_{x \in X} P_{d_{ij}}(x).U(x) \tag{1.6}$$

The principle of Maximum Expected Utility (MEU) states that the optimal decision under uncertainty is the one having the greatest expected utility.

In 1950s, Leonard Savage [41] merges Von Neumann and Morgenstern model of expected utility with de Finetti's model of subjective probability [1] to form axiomatic bases for statistics and decision analysis.
Savage's axioms are closer to those of VNM's, but they require other special axioms in order to achieve subjective probability from subjective utility. For more details see [41, 49].

In the literature, the quantification of value nodes in IDs have been made using cardinal or ordinal utility, based on VNM's axiomatic system, according to available informations during the structuring of decision problems.

Example 1.4 *Let us consider again the ID defined in Example 1. Table 1.1 represents a priori and conditional probabilities for chance nodes S, P and D. Table 1.2 represents the utility function (ordinal utilities) for the value node V, in the context of its parents (Tr, P and D).*

Table 1.1: A priori and conditional probabilities

D	$P(D)$	P	D	$P(P \mid D)$	S	P	$P(S \mid P)$
d	.10	p	d	.8	s	p	.7
$\tilde{d}$	.9	p	$\tilde{d}$	.15	s	$\tilde{p}$	.2
		$\tilde{p}$	d	.2	$\tilde{s}$	p	.3
		$\tilde{p}$	$\tilde{d}$	.85	$\tilde{s}$	$\tilde{p}$	.8

[1] Probability is a subjective degree of conviction which could be attributed to any event. The sum of probability of event A and its opposite is not equal to one.

Table 1.2: Physician's utilities

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

As mentioned before, 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 [36], 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 [24, 22], we can infer the following chain rule relative to IDs [23]:

$$P(C \mid D) = \Pi_{C_i \in C} P(C_i \mid Pa(C_i)) \tag{1.7}$$

Example 1.5 *Let us consider again the influence diagram presented in Figure 1.6, using Bayes's rule and (1.7) we obtain:*
$P(C, T, B, A \mid D_1, D_2) = P(C \mid T, B, A, D_1, D_2) P(T, B, A \mid D_1, D_2)$
$= P(C \mid T, B, A, D_1, D_2) P(T \mid B, A, D_1, D_2) P(B \mid A, D_1, D_2) P(A \mid D_1, D_2)$
C is d-separated from A, T and D_1 given B and D_2, so, we have:
$P(C \mid B, A, D_1, D_2) = P(C \mid B, D_2)$
$P(T \mid B, A, D_1, D_2) = P(T \mid B, A)$
$P(B \mid A, D_1, D_2) = P(B \mid A)$

$P(A \mid D_1, D_2) = P(A \mid D_1)$
So, $P(C, T, B, A \mid D_1, D_2) = P(C \mid B, D_2)P(T \mid B, A)P(B \mid A)P(A \mid D_1)$

1.3 Evaluation of influence diagrams

The construction of influence diagrams requires five general principles such that:

1. Start at the value node and work back to the decision nodes,
2. Draw the arcs in the direction that makes the probabilities easiest to assess,
3. Use informational arcs to specify which events will have been observed at the time each decision is made,
4. Ensure that missing arcs reflect intentional assertions about conditional independence and the timing of observations,
5. Ensure that there are no cycles in the influence diagram.

After specifying the structure of an ID, we should define a priori and conditional probabilty distributions relative to chance nodes which can be provided by experts or learned from data bases. In addition, decision makers should quantify value nodes to express their utilities.

Once the ID constructed it can be used to identify the optimal policy, this can be ensured via evaluation algorithms which allow to generate the best strategy yielding to the highest expected utility.

More formally, a strategy is a set of policies which are lists of decision rules, denoted by $\Delta = (\delta_i, .., \delta k)$, where δ_i is a decision rule relative to the decision node D_i which is a mapping $\delta_i : \Omega_{Pa_{D_i}} \rightarrow \Omega_{D_i}$.

In 1990, Cooper has shown that the problem of evaluation of ID is NP-hard [7].
ID evaluation algorithms can be divided into exhaustive and focused ones.
Exhaustive algorithms compute the expected utility of each possible decision strategy contrary to focused ones which only identify optimal strategy. Many anytime algorithms were proposed for evaluating influence diagrams.

Ramoni [39] proposed an anytime algorithm for solving IDs that incrementally refines the confidence in the expected utility of decisions. Horsch and Poole [18] developed another anytime algorithm that constructs a tree structure for each decision node in the ID. Each of the constructed tree structures represents a decision function that is improved incrementally.

Within evaluation algorithms, we can distinguish direct methods [42, 50] which operate directly on influence diagrams or indirect methods [6, 42, 53] which transform them into Bayesian networks [22, 34] and reduce the ID evaluation problem into a Bayesian networks inference one.

Direct evaluation requires a lot of probabilistic calculations which justify the great development of indirect methods initiated by Cooper [46] for the particular case of influence diagrams with a unique value node.

The key idea of indirect methods is to transform decision and value nodes into chance nodes to obtain a Bayesian network.

In what follows, we will only focus on indirect methods and especially on the algorithm proposed by Cooper [6].

We will first give a brief overview of Bayesian network.

1.3.1 Bayesian networks

Bayesian networks, introduced by Pearl [24], are a formalism of probabilistic reasoning used increasingly in several fields as decision aid, diagnosis and complex systems control [22, 29, 35].
A Bayesian network over a set of variables (nodes) consists of two components:

- A *graphical component* composed of a directed acyclic graph (DAG) G reflecting the causal relations relative to the modeled domain. BNs are IDs with only chance nodes.
- A *numerical component* consisting in a quantification of different links in the DAG by a conditional probability distribution of each node A in the context of its parents Pa(A).

One of the interesting tasks that we can perform using BNs is the propagation which consists in determing the impact of a certain information on the remaining ones. Several propagation algorithms was developed, they can be divided into two classes:

- *Exact algorithms*: The fundamental algorithm for exact probability is the one proposed by Kim and Pearl [25] and Pearl [33, 34]. In this algorithm, the impact of each new piece of evidence is viewed as a perturbation that propagates in parallel through the network, via a message mechanism, between neighboring variables. A centralized version of this algorithm was proposed by Poet and Shachter [37] and converges in two iterations.

- *Approximate algorithms*: Approximation techniques give result approaching the exact solutions. Unfortunately, approximate inference is NP-complete [7], but we can nonetheless come up with approximations which often work well in practice. As an example of approximate methods, we quote, sampling (Monte Carlo) methods [5], loopy belief propagation [26], which entails applying Pearl's algorithm to the original graph.

1.3.2 Cooper's method

Most indirect methods are based on this method [6], defined to transform ID with a unique value node into a Bayesian network. This algorithm is based on three steps, presented as follows:

1. Transform each decision node into a chance node characterized by an equi-probability, as follows:
$$P(D_i \mid Pa(D_i)) = \frac{1}{\mid dom(D_i) \mid} \tag{1.8}$$

2. Transform the value node into a chance node,

3. Convert the utility function into a probability distribution as it will be detailed in what follows.

The value node, V is transformed into a new binary chance node which has two values False (F) and True (T). In the obtained BN, the utility function is transformed into a probability function as follows, $\forall pa(V) \in Pa(V)$:

$$P(v = T \mid pa(V)) = \frac{U(pa(V)) + K_2}{k_1} \tag{1.9}$$

where

$$K_1 = U_{max} - U_{min} \tag{1.10}$$

and

$$K_2 = -U_{min} \tag{1.11}$$

Where U_{max} and U_{min} be the maximal utility level and the minimal utility level, 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. The maximal expected utility is computed begining by the last decision D_m to the first one D_1, by considering for each node a set of evidence E updated in the light of the previous step.

More formally, for each decision node D_i, the maximal expected utility is computed as follows:

$$MEV(D_i, E) = K_1 * max_{d_{ij}}[P(v = T \mid d_{ij}, E)] - K_2 \tag{1.12}$$

Example 1.6 *Let us continue with the Medical Diagnosis's example. After the transformation of the ID, the decision node Tr (treat for disease) will become an chance node Tr, its a priori possibility distribution is presented in Table 1.3. The new chance node V is characterized by a conditional probability distribution detailed in Table 1.4. Figure 1.5 presents the obtained Bayesian network.*

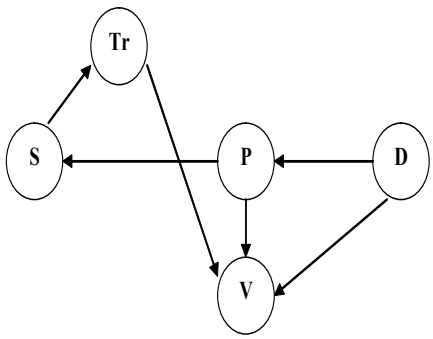

Figure 1.5: Obtained Bayesian network

Table 1.3: A priori probability distribution for Tr

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

Table 1.4: Conditional probability distribution for V

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

1.3.3 Other indirect methods

Shachter and Poet's method

In his algorithm [44], Shachter uses Cooper's transformation to convert the value node into a chance node. Then, each decision node D_i is converted into a chance node with an equi-probability distribution using (1.8).
Before the conversion of the ID into a BN, a simple linear-time algorithm can examine the ID to determine which nodes can be omitted and of those remaining, those for which probability distribution are not needed [44].
The simple linear-time algorithm was defined by Shachter in 1989 in his work called "The relational pastime for determining irrelevance and requisite information in BN and ID elaborated previously".

Within the diagram, it can be recognized that only a subset of the non evidence variables observed at the time of decision are needed to make the best choice. We say that this set

(denoted Ri) contains relevant information for decision node and the optimal policy can be written as a function from the possible values of Ri to the alternative choices for decision node.

Shachter's algorithm of reduction is based on the operation used to remove nodes presented by Olmested in 1983. The application of this algorithm, will reduce enormously the number of numerical computations, because all irrelevant nodes are removed from ID.

Zhang's method

This method was based on previous advances made by Zhang and Poole regarding the notion of stepwise decomposable IDs and Bayesian inference techniques, Zhang proposed a recursive algorithm based on decomposition theorem.

If D_i is the tail decision node in the ID (the last decision node in the ordered list of all decision nodes), then the ID can be decomposed into two components, called tail and body, such that an optimal decision rule for D_i can be found in the tail and optimal decision rules for all other decision nodes can be found in the body. The body is again an ID and hence the decomposition can be repeated in the body [53].
In his algorithm, Zhang uses a procedure for simplifying computations in the body and in the tail. To transform decision nodes and value nodes to chance nodes, Cooper's transformations is used.

Previous methods always included the entire graph for calculation the MEU (Maximal Expected Utility), though several nodes have no impact on optimal decision. The decomposition of ID and the simplification of the graph reduce greatly the computational task for finding the optimal decision.

Xiang and Ye's method

It is a simple method to evaluate IDs by first converting them into BNs and then by applying a junction tree (JT) based inference algorithm [52]. This method convert ID into BN as follows :

- All existing chance nodes remain the same.

- Convert each decision node (D_i) into chance node, it has the same space Ω_{D_i} and the same set of parent nodes $Pa(D_i)$ and it is associated with a uniform distribution $P(D_i \mid Pa(D_i)) = const$, where const stands for a normalizing constant such that $\Sigma_j P(d_{ij} \mid Pa(D_i)) = 1$ for each configuration of $Pa(D_i)$.
- Convert the value node into a binary chance node V with a space (T,F), the probability distribution of the node V is $P(v = T \mid Pa(V)) = U(Pa(V))$.

Once the BN obtained, it will be converted into a junction tree which is a tree where each node is labeled by a subset of nodes called a cluster. The conditional probability distributions in the BN are converted into the distribution over each cluster. Computation of posterior distributions over each cluster can be performed effectively by message passing between clusters, called belief propagation [52].

Mark Crowley's method

It is an indirect algorithm which has been proposed informally by David Poole and follows from several intuitive observations [8]. This method transforms ID into BN by the standard way as described by Cooper and it uses the variable elimination algorithm presented by Zhang [53].

Daniel Garcia-Sanchez and Marek J.Druzdzel's method

Another work elaborated by Daniel Garcia-Sanchez and Marek J.Druzdzel describe an efficient sampling algorithm for solving influence diagrams that achieves its efficiency by reusing samples for each of the decision strategies [40]. The algorithm is exhaustive in the sense of computing the expected utility of each of the possible decision strategies.
The method indicates how sampling algorithms can be used within an indirect evaluation method. Evaluation all decision strategies on the same set of random samples allows to save a significant amount of computation and also produces high quality anytime behavior. To transform an ID into a BN, Cooper's transformation is used, then we iterate through nodes in the sample sets. The simulation can be interrupted and the best estimate of the expected utilities will be returned. The algorithm can be based on any simulation algorithm for BN, it allows for taking advantages of these developments.

1.4 Conclusion

In this chapter, we have presented the basis of influence diagrams, its quantitative and qualitative components. Several evaluation's algorithm are presented, which are divided into two principal categories: direct evaluation's algorithm and indirect ones.

An important task in ID's construction is the specification of its quantitative component, precisely conditional probability distributions for chance nodes and utility functions for value nodes.

Nevertheless, the specification of this numerical component is not always obvious. In fact, experts may be enable to provide exact conditional probabilities. Moreover, it is generally more easier for them to provide a preferential relation between different consequences to express their utilities rather than numerical values. In addition, the probability theory models the state of total ignorance by uniform probabilities which express randomness rather than ignorance.

To overcome these limitations, we are interested in using a non probabilistic theory to handle uncertainty and imprecision during the construction of IDs.

Indeed, possibility theory offers a natural and simple model to deal with uncertainty qualitatively and quantitatively. The basis of this no-classical theory, based on Fuzzy sets, will be presented in the next chapter.

Chapter 2

Possibility Theory

2.1 Introduction

There are many practical problems characterized by a large number of interrelated uncertain alternatives. Information and uncertainty are two sides of the same coin, any information contains inherent uncertainty which manifests itself in different forms: ambiguity, vagueness, incomplete and imprecise knowledge. Uncertainty is a serious problem that must be treated suitably. Imprecision and uncertainty can be considered as two complementary aspects of a single reality, that of imperfect information [9]. Imprecision relates to the content of an information (its value) while uncertainty relates to its truth, understood as its conformity to a reality.

Traditionally, probability theory was the most known model of representing imperfect information. It represents a tried and tested mathematical theory, it deals with all kinds of imperfect information. However, Dubois et al. [10] proved that this theory is not appropriate in the case of total ignorance because uniform probability distribution on finite sets better express randomness than ignorance. Then, probability theory is not concerned with the modeling of imprecision but rather it captures the notion of random variability.

Recent advances in cognitive sciences and uncertainty theory have brought forth several novel mathematical approaches to deal with different kinds of imperfect information such as fuzzy sets theory [54], possibility theory [9], evidence theory [45] and rough set theory [32].

This chapter presents basic elements of possibility theory. It is organized as follows: Section 2.2 represents some useful notations. The next section introduces the basics of possibility theory, the qualitative setting of this theory is detailed in Section 2.4. Section 2.5 presents the different methods to combine possibility distributions. In Section 2.6 we focus in possibilistic conditioning, qualitative possibilistic utility was defined in Section 2.7.

2.2 Notations

Let us give some useful notations.

- Ω represents the universe of discourse, it represents the set of all possible situations, $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$.
- Each element $\omega \in \Omega$ is called an interpretation, a possible world or a state of Ω.
- The set of all subsets of Ω is called the power set and it is denoted by 2^{Ω}.
- Let us denote by ψ an element of the power set ($\psi \in 2^{\Omega}$).

2.3 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 [9, 55, 56].

2.3.1 Possibility distribution

Possibility theory, issued from Fuzzy sets theory, was introduced by Zadeh [55, 56] and further developed by Dubois and Prade [9].

The basic building block in the possibility theory is the notion of possibility distribution denoted by π, it is a mapping from Ω to a bounded linearly ordered scale L exemplified by the unit interval [0,1]. A possibility degree is the value from the interval [0,1] associated to each element ω of Ω.

Table 2.1: Conventions for possibility distribution π_ω

$\pi_\omega(u) = 0$	$u = \omega$ is rejected as impossible
$\pi_\omega(u) = 1$	$u = \omega$ is totally possible (plausible)
$\pi_\omega(u) > \pi_\omega(u')$	$u = \omega$ is preferred to $u' = \omega$ (or is more plausible)

The function π_ω represents a flexible restriction of the values of ω with the conventions represented in Table 2.1.

The quantity $\pi_\omega(u)$ represents the degree of possibility of the assignment $u = \omega$.
An important property relative to possibility distribution is the **normalization**.
Namely, at least one element of Ω should be fully possible as the value of ω. More formally:

$$\exists u \in \Omega \ s.t \ \pi_\omega(u) = 1 \quad (normalization) \tag{2.1}$$

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:

$$\exists u_0, \pi_\omega(u_0) = 1 \qquad and \qquad \forall u \neq u_0, \pi_\omega(u) = 0 \tag{2.2}$$

- Complete ignorance:

$$\forall u \in \Omega, \pi_\omega(u) = 1 \tag{2.3}$$

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 as follows:*
$\Omega = \{Cenozoic(Ceno), Mesozoic(Meso), Paleozoic(Paleo)\}$
Suppose that a geologist must determine the geological era (E) of a fossil (F). He gave his opinion in the form of a possibility distribution π_1 defined as follows:

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

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 F dates from the Mesozoic era.*

2.3.2 Possibility and necessity measures

In probability theory, $P(\neg\psi) = 1 - P(\psi)$. The expression *It is not probable that "not* ψ*"* means that *It is probable that* ψ. Conversely, *it is possible that* ψ does not entail anything about the possibility nor the impossibility 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 $N(\psi) = 1 - \Pi(\neg\psi)$.
These two dual measures are defined as follows:

Possibility measure

Given a possibility distribution π, the possibility measure can be defined as follows:

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

$\Pi(\psi)$ is called the possibility degree of ψ, it corresponds to the possibility degree to have one of the models of ψ as the real world [11].
Table 2.2 gives main properties of possibility measures (case of normalized π).

Table 2.2: Possibility measure Π

$\Pi(\psi) = 1$ 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 a prior 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

Necessity measure

This measure represents the dual of the possibility measure:

$$N(\psi) = 1 - \Pi(\neg\psi) = min_{\omega \notin \psi}(1 - \pi(\omega)) \quad \forall \psi \subseteq \Omega \tag{2.5}$$

$N(\psi)$ is called the necessity degree of ψ. It 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 necessity measure.

Table 2.3: Necessity measure N

$N(\psi) = 1$ 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

Sometimes, the expert may be enable to express his opinions with exact values, but he can provide a preferential relation relative to each possible situation of the universe of discourse Ω. This preferential relation is the base of a new form of uncertainty expression which is the qualitative one. It corresponds to a plausibility relation allowing to equip Ω with a total pre-order[1] denoted by $\geq_\pi$.
Let us present how uncertainty is represented qualitatively in possibility theory.

2.4 Qualitative possibility theory

The universe of discourse is equipped with a finite and totally ordered scale denoted by $\mathcal{L}$ $= \{\alpha_0 = 1, \alpha_1, .., \alpha_n, \alpha_{n+1} = 0\}$ such that $\alpha_0 = 1 > \alpha_1 > ..\alpha_n > \alpha_{n+1} = 0$ to allow the determination of the qualitative possibility distribution (QPD), also denoted by π. Indeed,

[1] A relation $\geq$ on Ω is a total pre-order if $\geq$ is complete ($\forall \omega_1, \omega_2$ we have either $\omega_1 \geq \omega_2$ *or* $\omega_1 \leq \omega_2$, reflexive and transitive.

a qualitative possibility distribution is a mapping $\pi : \Omega \rightarrow \mathcal{L}$, enabling us to express that some states are more plausible than others without referring to any numerical value as in the quantitative setting. We can have three possible cases:

- $\omega =_\pi \omega'$ means that ω is as plausible as ω'.
- $\omega >_\pi \omega'$ means that ω is more plausible than ω'.
- $\omega <_\pi \omega'$ means that ω is less plausible than ω'.

Generally, the ordered scale $\mathcal{L}$ is represented by the numerical scale [0,1], but the handled values reflect only on ordering between the different states of the world.

Example 2.2 *Let us continue with the last example. The plausibility relation $>_{\pi_1}$ on Ω is given by: $E = Meso >_{\pi_1} E = Ceno >_{\pi_1} E = paleo$. It can be represented qualitatively by the QPD π'_1 as follows:*

$$\pi'_1(E = Ceno) = \alpha_1, \quad \pi'_1(E = Meso) = \alpha_0, \quad \pi'_1(E = Paleo) = \alpha_2.$$

Thus, the Mesozoic era is more plausible than the Cenozoic era which is in turn more plausible than the Paleozoic era, since $\alpha_0 > \alpha_1 > \alpha_2$. We can assign any numerical values to α_i, but its ranking is very important. Hence, an infinity of quantitative possibility distributions can be derived from a qualitative one.

2.5 Combination of possibility distributions

In the possibilistic approach, there are different combination modes to assure the fusion of informations. Hence, we should choose the appropriate combination mode, this choice is strongly related to the reliability of informations to be combined, namely to the reliability of information's sources.

The most known combination operators are the symmetric ones, namely the conjunctive and the disjunctive operators which will be detailed in what follows.

2.5.1 The conjunctive fusion

If all sources are fully reliable, then we can combine them using the intersection, the conjunctive operator $\bigotimes$ is defined as follows:
Let π_i be the possibility distribution supplied by source i.

$$\forall \omega \in \Omega, \quad \pi_\wedge(\omega) = \bigotimes_{i=1..n} \pi_i(\omega) \tag{2.6}$$

$\bigotimes$ is a t-norms such that minimum or product or linear product according to the uncertainty scale's interpretation. Indeed, the min operator is supported by both quantitative and qualitative possibility distributions. However, the use of the product operator assumes that possibility degrees are numerical.

After the conjunction of possibility distributions, we may obtain a subnormalized possibility distribution, i.e., it may happen that $\forall \omega \in \Omega,\ \pi_\wedge(\omega) < 1$. Such cases express that there is a conflict between the combined sources.
The result given by $\pi_\wedge$ must be normalized by the following operator:

$$\forall \omega \in \Omega,\ \pi(\omega) = \frac{\pi_\wedge(\omega)}{h(\pi_1, \pi_2)} \tag{2.7}$$

Where $\pi_\wedge(\omega) = \pi_1(\omega) \otimes \pi_2(\omega)$ and $h(\pi_1, \pi 2) = sup_{\omega \in \Omega}(\pi_1(\omega) \otimes \pi_2(\omega)) = sup\ \pi_\wedge$. We must have $h(\pi_1, \pi 2) \neq 0$

Example 2.3 *Let us continue with the same example, suppose that another geologist G_2 providing his opinion about the geologic era of the founded fossil in the form of a possibility distribution π_2.*
The two possibility distributions are presented in Table 2.4.

Table 2.4: The possibility distributions π_1 and π_2

$\pi_1(E = Ceno) = 0.4$	$\pi_2(E = Ceno) = 0.6$
$\pi_1(E = Meso) = 1$	$\pi_2(E = Meso) = 0.1$
$\pi_1(E = Paleo) = 0.3$	$\pi_2(E = Paleo) = 1$

If we assume that geologists are reliable and $\otimes = min$, we will obtain the following result $\pi_\wedge$ for the conjunctive fusion:

$$\pi_\wedge(E = Ceno) = 0.4, \quad \pi_\wedge(E = Meso) = 0.1, \quad \pi_\wedge(E = Paleo) = 0.3.$$

$\pi_\wedge$ is subnormalized, the normalized possibility distribution $\pi_\wedge$ will be as follows ($h(\pi_1, \pi_2) = 0.4$):

$$\pi_\wedge(E = Ceno) = 1, \quad \pi_\wedge(E = Meso) = \frac{0.1}{0.4} = 0.25, \quad \pi_\wedge(E = Paleo) = \frac{0.3}{0.4} = 0.75.$$

The qualitative normalization consists in moving to 1 the highest possibility degrees of the obtained distribution $\pi_\wedge$.

2.5.2 The disjunctive fusion

This mode of combination is applied when it is known for sure that at least one of the sources is reliable but it is not known which one. The disjunctive operator $\bigoplus$ is defined as follows:

$$\pi_\vee(\omega) = \bigoplus_{i=1..n} \pi_i(\omega) \quad \forall \omega \in \Omega \tag{2.8}$$

$\bigoplus$ is a t-conorms such that maximum or probabilistic sum or Lukasievicz according to the uncertainty scale's interpretation. Indeed, all of these t-conorms can be applied in the quantitative setting, but in the qualitative setting, the maximum operator is applied.

2.5.3 Possibilistic mixture

In the litterature, we can distinguish an other form of fusion inspired from the probabilistic mixture. The possibilistic mixture [12] aim to combine m possibility distributions $\pi_1, .., \pi_m$ with possibility degrees $\lambda_1, .., \lambda_m$ such that $max_{i \in \{1,..,m\}}\{\lambda_i\} = 1$, as follows:

$$(\lambda_1/\pi_1, .., \lambda_m/\pi_m)(x) = max_{i \in \{1,..,m\}} min(\lambda_i, \pi_i(x)) \tag{2.9}$$

2.6 Possibilistic conditioning

The conditioning may be considered as a special case of informations fusion. Indeed, it 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

$$\pi(\omega|_m\psi) = \begin{cases} 1 & if\ \pi(\omega) = \Pi(\psi)\ \ and\ \ \omega \in \psi \\ \pi(\omega) & if\ \pi(\omega) < \Pi(\psi)\ \ and\ \ \omega \in \psi \\ 0 & otherwise. \end{cases} \tag{2.10}$$

- product-based conditioning relative to the numerical setting

$$\pi(\omega|_p\psi) = \begin{cases} \frac{\pi(\omega)}{\Pi(\psi)} & if\ \omega \in \psi \\ 0 & otherwise. \end{cases} \tag{2.11}$$

Example 2.4 *Let us consider again the problem of fossil's dating. Suppose that the geologist make 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 (2.10) and (2.11), new distributions are presented in Table 2.5.*

Table 2.5: Possibilistic conditioning

Era	Breed	$\pi(Era \wedge Breed)$	$\pi(Era \wedge Breed\ \|_p\ \psi)$	$\pi(Era \wedge Breed\ \|_m\ \psi)$
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

2.7 Qualitative possibilistic utility

Classical approaches for decision making under uncertainty are based on maximal expected utility (c.f section 1.2). However, the computation of this measure is not always obvious especially in complex decision problems when probabilities and numerical utilities are difficult to assess.

To overcome this limitation many researches aim to introduce non-probabilistic decision theory. Indeed, many efforts have been made to create an axiomatic basis for possibility-based decision theory. We can distinguish those following Von Neumann and Morgenstern axiomatic system [30] and those following Savage axiomatic system [41] (c.f section 1.2). More precisely, Dubois et al. proposed pessimistic and optimistic qualitative utilities [13] [12] and Giang et al. introduced the concept of binary qualitative utility [17] as a possibilistic counterpart to VNM's axiomatic system. Moreover, Dubois et al. [14] and Fargier et al. [16] proposed qualitative possibilistic utility based on Savage's axiomatic system.

In our work, we are only interested on possibilistic qualitative utility following VNM's axiomatic system. The two principal approaches in this direction are detailed in what follows.

2.7.1 Notations

- Let S be a finite set of situations (states).
- Let X be a finite set of consequences (outcomes, prizes) with at least two elements.
- Let V be a finite ordered scale assumed to be a set of points in [0,1] and representing uncertainty.
- Let d be a decision, represented by a function $d : S \rightarrow X$, called an act.
- Let $\succeq$ be a preference relation among consequences in the style of Von Neumann and Morgenstern [30] (c.f section 1.2).
- Within the set X, we can distinguish two particular elements:
    1. The best consequence, denoted by $x_\top$, s.t $\forall x \in X$, $x_\top \succeq x$,

2. The worst consequence, denoted by $x_\perp$, s.t $\forall x \in X$, $x \succeq x_\perp$.

- Let UT be an ordered utility scale (follow the same order criterion than V).

2.7.2 Pessimistic and optimistic utilities

Dubois et al. [13] defined qualitative possibilistic decision theory based on the uncertainty-aversion decision maker's behaviour which corresponds to a pessimistic view of what may happen and on the uncertainty-attraction behaviour which corresponds to an optimistic view. Several improvements of this proposal are proposed in [12].

In this work, uncertainty in decision problems is represented by two possibility distributions:

1. the first one, denoted by π, encodes our knowledge about the set of situations (S) by a mapping $\pi : S \rightarrow V$.
2. the second possibility distribution, denoted by μ, represents the decision maker's preference on X by a mapping $\mu : X \rightarrow UT$.

The decision d induces a possibility distribution π_d on X s.t:

$$\pi_d(x) = \pi(s) \quad \forall s \in S \quad and \quad d(s) = x \tag{2.12}$$

Later on, $\pi_d(x)$ and $\pi(x)$ are used interchangeably, namely the possibility distribution π can be defined as a mapping from X to V ($\pi : X \rightarrow V$). To the decision maker, the value of decision d is the same as the value of the induced lottery π_d. Thus a preference relation between decisions is the same as a preference relation between lotteries. Hence, the utility of d is evaluated by mixing the two possibility distributions $\pi(x)$ and $\mu(x)$ in a suitable way using (2.9) $\forall x \in X$.

Therefore, to mix μ and π a commensurateness assumption should be made between the uncertainty scale V and the utility scale UT. Hence, we need an order preserving mapping $h : V \rightarrow UT$ such that, let $card_V$ be the cardinality of the set V:

$$h(V(i)) = \begin{cases} 1 & if \quad V(i) = 1 \\ 0 & if \quad V(i) = 0 \\ UT(i) \quad if \quad V(i) \neq 0, \quad V(i) \neq 1 \quad and \quad i \leq card_V \end{cases}$$

A possibility distribution on X is called *qualitative possibilistic lottery* and it can be represented by $[\pi(x_1)/x_1, .., \pi(x_n)/x_n]$ such that impossible consequences are omitted. The set of lotteries is denoted Π_X.

Note that each consequence $x \in X$ can be written in the form of a lottery π_x as follows: $\pi_x(x) = 1 \;\; and \;\; \pi_x(x') = 0 \;\; \forall \;\; x' \neq x$.
A particular form of lottery is the canonical lottery that occurs either in the best or in the worst consequence. The set of best or worst lotteries are called canonical lotteries and denoted by Πc_X, such that:

$$\Pi c_X = \{[\lambda/x_\top, \mu/x_\bot] \;\; | \; \lambda, \mu \in V, max(\lambda, \mu) = 1\} \tag{2.13}$$

In what follows, we detail pessimistic and optimistic utilities which offer two frameworks to establish an order of preference between the set of lotteries on X.

Pessimistic utility

The pessimistic criterion was originally proposed by Whalen [12], it supposes that the decision maker considers that all bad consequences are hardly plausible.
A pessimistic qualitative utility function $QU^- : \Pi_X \rightarrow UT$ is defined by:

$$QU^-(\pi) = min_{x \in X} max(n(\pi^*(x)), \mu(x)) \;\; \forall \pi \in \Pi_X \tag{2.14}$$

Where $\pi^*(x) = h(\pi(x))$ and $n : UT \rightarrow UT$ is the order reversing involution on UT such that, let $card_{UT}$ be the cardinality of the set UT:

$n(UT_i) = Ut(card_{UT} - i + 1) \;\; \forall \;\; i \leq card_{UT}$

Particular values for QU^- are as follows:

- $QU^-(x_\bot) = QU^-([1/x_\top, 1/x_\bot]) = 0$
- $QU^-(x_\top) = 1$
- $QU^-([1/x_\top, \lambda/x_\bot]) = nh(\lambda)$ (where nh is a composition of n and h).

Let $\succeq$ be a preference relation on Π_X, it should satisfy the following axiomatic system S_p to be represented by a pessimistic qualitative utilitity function.

- Axiom 1^-. **Total pre-order**: $\succeq$ is reflexive, transitive and complete.
- Axiom 2^-. **Uncertainty aversion**: if $\pi' \geq \pi$ then $\pi \succeq \pi'$.
- Axiom 3^-. **Substitutability**: if $\pi_1 \sim \pi_2$ then $[\lambda/\pi_1, \mu/\pi] \sim [\lambda/\pi_2, \mu/\pi]$.
- Axiom 4^-. **Continuity**: $\forall \pi \in \Pi_X, \exists \lambda \in V \ \pi \sim [1/x_\top, \lambda/x_\bot]$.

Theorem 2.1 *A preference relation $\succeq$ on Π_X satisfies the axiomatic system S_P iff there exists a pessimistic utility function QU^- such that:*

$$\pi \succeq \pi' \quad iff \quad QU^-(\pi) \geq QU^-(\pi') \tag{2.15}$$

Optimistic utility

The optimistic criterion was originally defined by Yager [12], it captures the optimistic behavior of the decision maker that makes at least one of the good consequence highly plausible.
An optimistic qualitative utility function $QU^+ : \Pi_X \to UT$ is defined by:

$$QU^+(\pi) = max_{x \in X} min(\pi^*(x), \mu(x)) \quad \forall \pi \in \Pi_X \tag{2.16}$$

Particular values for QU^+ are as follows:

- $QU^+(x_\top) = QU^+([1/x_\top, 1/x_\bot]) = 1$
- $QU^+(x_\bot) = 0$
- $QU^+([\lambda/x_\top, 1/x_\bot]) = h(\lambda)$.

The system S_O of axioms that characterized QU^+ is obtained from S_P by substituting axioms Axiom 2^- and Axiom 4^- by their diametrical counterparts Axiom 2^+ and Axiom 4^+, respectively:

- Axiom 2^+. **Uncertainty attraction**: if $\pi \geq \pi'$ then $\pi \succeq \pi'$.
- Axiom 4^+. **Continuity**: $\forall \pi \in \Pi_X, \exists \lambda \in V \ \pi \sim [\lambda/x_\top, 1/x_\bot]$.

In order to decide on a preference relation between lotteries, we should compute qualitative utility (QU) relative to each lottery according to the decision maker's behaviour (optimistic or pessimistic). The computation of the qualitative utility requires the determination of $\mu(x)$ $(\forall x \in X)$.

Since each consequence x can be represented by a canonical lottery as illustrated by the axiom of continuity (Axiom 4^- or Axiom 4^+). $\mu(x)$ is equal to the qualitative utility of the canonical lottery equivalent to x. Namely:
$\mu(x) = QU^-([1/x_\top, \lambda/x_\bot]) = nh(\lambda)$ (pessimistic behaviour).
$\mu(x) = QU^+([\lambda/x_\top, 1/x_\bot]) = h(\lambda)$ (optimistic behaviour).

The following example illustrates the use of pessimistic and optimistic utilities.

Example 2.5 *Let* $X = \{x_\top, x_2, x_3, x_\bot\}$ *s.t* $x_\top \succeq x_2 \succeq x_3 \succeq x_\bot$, $V = \{1, 0.7, 0.5, 0\}$ *and* $UT = \{1, 0.5, 0.3, 0\}$.
Suppose that the decision maker is considering three lotteries $\pi_1, \pi_2, \pi_3 \in \Pi_X$ *represented in Table 2.6 and the utility assignment is represented in Table 2.7.* n *and* h *are represented in Table 2.8.*

Table 2.6: Possibilistic lotteries

X	$x_\top$	x_2	x_3	$x_\bot$
π_1	0.7	1	0.5	0.5
π_2	0.5	0.7	0	1
π_3	1	1	1	1

Table 2.7: The utility assignment

X	$x_\top$	x_2	x_3	$x_\bot$
μ	1	0.5	0.3	0

Table 2.8: The function n and h

UT	1	0.5	0.3	0	V	1	0.7	0.5	0
n	0	0.3	0.5	1	h	1	0.5	0.3	0

Case 1: The pessimistic case

For a pessimistic decision maker, the indifference set derived from the Axiom 4^- *is defined by:*

$$x_\top \sim [1/x_\top, 0/x_\bot]$$
$$x_2 \sim [1/x_\top, 0.5/x_\bot]$$
$$x_3 \sim [1/x_\top, 0.7/x_\bot]$$
$$x_\bot \sim [1/x_\top, 1/x_\bot]$$

Then, the pessimistic qualitative utility function relative to the lottery π_1 *is defined as follows:*

$$QU^-(\pi_1) = min \begin{cases} max(nh(0.7), U(x_\top)) \\ max(nh(1.0), U(x_2)) \\ max(nh(0.5), U(x_3)) \\ max(nh(0.5), U(x_\bot)) \end{cases} = min \begin{cases} max(0.3, 1.0) \\ max(0.0, 0.5) \\ max(0.5, 0.3) \\ max(0.5, 0.0) \end{cases} = 0.5.$$

Following the same calculation, we have $QU^-(\pi_2) = 0$ *and* $QU^-(\pi_3) = 0$.

Then, the preference relation between these lotteries is: $\pi_1 \succeq \pi_2 \sim \pi_3$.

This means that a pessimistic decision maker prefers the lottery that offers the highest possibility degree for the best consequence and the lower one for the worst consequence.

Case 2: The optimistic case

For a optimistic decision maker, the indifference set derived from the Axiom 4^+ *is defined by:*

$$x_\top \sim [1/x_\top, 1/x_\bot]$$
$$x_2 \sim [0.7/x_\top, 1/x_\bot]$$
$$x_3 \sim [0.5/x_\top, 1/x_\bot]$$
$$x_\bot \sim [0/x_\top, 1/x_\bot]$$

Then, the optimistic qualitative utility function relative to the lottery π_1 is defined as follows:

$$QU^+(\pi_1) = max \begin{cases} min(h(0.7), U(x_\top)) \\ min(h(1.0), U(x_2)) \\ min(h(0.5), U(x_3)) \\ min(h(0.5), U(x_\bot)) \end{cases} = max \begin{cases} min(0.5, 1.0) \\ min(1.0, 0.5) \\ min(0.3, 0.3) \\ min(0.3, 0.0) \end{cases} = 0.5.$$

Following the same calculation, we have $QU^+(\pi_2) = 0.5$ and $QU^+(\pi_3) = 1$.
Then, the preference relation between these lotteries is: $\pi_3 \succeq \pi_2 \sim \pi_1$.
This means that an optimistic decision maker will prefer the lottery π_3 since in this lottery $x_\top$ is totally possible even he is in form of a total ignorance. This behaviour is clearly different from the pessimistic one when π_3 is remoted.

2.7.3 Limits of pessimistic and optimistic utilities

Giang and Shenoy [17] criticized pessimistic and optimistic utilities presented by Dubois et al. [12], their arguments are based on the fact that proposed frameworks for possibilistic utilities are based on axioms (i.e Axiom 2^- and Axiom 2^+) 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 as shown by the following example:

Example 2.6 *Let us consider the following decision problem with two lotteries $\pi_1 = [1/x_\top, 1/x_\bot]$ and $\pi_2 = x_\bot$.*
π_2 is more specific than π_1, since it describes a situation of complete knowledge.
Let us compute $QU^-(\pi_1)$ assuming that n,h,U, UT and V are those given in Example 2.5.
Then, we obtain $QU^-(\pi_1) = QU^-(\pi_2) = 0$
Therefore, axiom $A2^-$ will force the decision maker to consider that $\pi_2 \sim \pi_1$, namely he will be indifferent between a worst prize for sure (π_2) and the case that the best prize is as likely as the worst one (π_1).

2.7.4 Binary possibilistic utility

Shenoy and Giang [17] investigated an axiomatic system for preference relation for the set of lotteries, based on the fact that lotteries that realize in the best prize $x_\top$ or in the worst prize $x_\bot$ play an important role in decision making (as demonstrates in Example 2.5). The system of axioms (denoted by S) is defined as follows:

- Axiom 1^S. **Total pre-order**: $\succeq$ is reflexive, transitive and complete.
- Axiom 2^S. **Qualitative monotonicity** $\succeq$ restricted to C satisfies the following condition:

$$[\lambda/x_\top, \mu/x_\bot] \succeq [\lambda'/x_\top, \mu'/x_\bot] \;\; if \;\; \begin{cases} (1 \geq \lambda \geq \lambda' \;\; and \;\; \mu = \mu' = 1) \;\; or \\ (\lambda = 1 \;\; and \;\; \lambda' < 1) \;\; or \\ (\lambda = \lambda' = 1 \;\; and \;\; \mu' \geq \mu) \end{cases} \tag{2.17}$$

- Axiom 3^S. **Substitutability**: if $\pi_1 \sim \pi_2$ then $[\lambda/\pi_1, \mu/\pi] \sim [\lambda/\pi_2, \mu/\pi]$.
- Axiom 4^S. **Continuity**: $\forall x \in X$, $\exists \sigma \in C$ s.t $x \sim \sigma$.

Let $\succeq$ be a preference relation on Π_X that satisfies the axiomatic system S. For each $\pi \in \Pi_X$, there exists a canonical lottery $\sigma \in C$ such that $\pi \sim \sigma$.

Proof 2.1 *[17] Let* $\sigma \in$ C

$\pi = [\pi(x_1)/x_1, \pi(x_2)/x_2, .., \pi(x_n)/x_n]$ *(possibilistic lottery)*
$\forall i \in \{1, 2, .., n\}$ $\sigma_i = [\lambda_i/x_\top, \mu_i/x_\bot]$, $x_i \sim \sigma_i$ *(from Axiom* 4^S*)*
$\pi \sim [\pi(x_1)/\sigma_1, \pi(x_2)/\sigma_2, .., \pi(x_n)/\sigma_n]$ *(from Axiom* 3^S*)*
$\pi \sim [\pi(x_1)/\sigma_1, \pi(x_2)/\sigma_2, .., \pi(x_n)/\sigma_n] = [\lambda/x_\top, \mu/x_\top]$
where $\lambda = max_{i\in\{1,2,..n\}} min(\pi(x_i), \lambda_i)$ *and* $\mu = max_{i\in\{1,2,..n\}} min(\pi(x_i), \mu_i)$
Thus, $\pi \sim [\lambda/x_\top, \mu/x_\bot]$.

To represent a canonical lottery in the probabilistic case, we need only a simple number which is the probability p of getting the best prize. For the worst case, this probability is $1 - p$.

In possibility theory, we need two measures to represent a canonical lottery because the knowledge that the best prize is fully possible could imply nothing about the possibility of getting the worst prize.

The concept of binary utility was introduced to represent canonical lotteries in the possibility theory [17].

Definition 2.1 *Let W be a linearly ordered scale. U_W is a set of pairs of elements in W such that one of them is equal to 1. It is defined as follows:*

$$U_W = \{< a, b > \quad | \, a, b \in W \quad and \quad max(a, b) = 1\} \tag{2.18}$$

Let $\gg$ be a complete order on U_W defined by:

$$< a, b > \quad \gg \quad < a', b' > \quad iff \quad (a > a') \quad or \quad (b' > b) \tag{2.19}$$

A binary utility scale *based on W is the set U_W equipped with the order $\gg$.*

Since indices in U_W pertains to the linearly ordered scale W, then the min and max operators can be extended for pairs as follows. Let $\alpha, \beta, \gamma, \delta \in W$, then:

$$min(< \alpha, \beta >, < \gamma, \delta >) =< min(\alpha, \gamma), min(\beta, \delta) > \tag{2.20}$$

$$max(< \alpha, \beta >, < \gamma, \delta >) =< max(\alpha, \gamma), max(\beta, \delta) > \tag{2.21}$$

Let us call a A function $BU : X \rightarrow U_W$, that assigns utility for each prize in X, *a basic utility assessment.* BU is said to be consistent with a preference relation $\succeq$ if for any pair x,y $\in$ X,

$$x \succeq y \; iff \; BU(x) \;\gg\; BU(y) \tag{2.22}$$

Let BU be a basic utility assessment, then a possibilistic utility function PU relative to each lottery π is a mapping from Π_X to U_W defined as follows:

$$PU(\pi) = max_{x \in X} min(k(\pi(x)), BU(x)) \tag{2.23}$$

where k is a binary vector function defined as follows:

Definition 2.2 *Let $k =< k_1, k_2 >$ such that $\forall r = 1, 2$, k_r is an order preserving mapping from V to W, such that $k_r(1) = 1$ and $k_r(0) = 0$.*

$\forall x \in X$, $BU^L(x)$ and $BU^R(x)$ are defined as follows:

- *$BU^L(x)$ is the left value of a pair BU(x),*

- *$BU^R(x)$ is the right value of a pair BU(x).*

The binary vector function k is relative to the type of the decision maker [17]. More precisely:

- *If the DM is* pessimistic *then* $k_1(0) = 0$, $\forall v \in \{V - 0\}$ $k_1(v) = 1$ *and* $\forall v \in V$ $k_2(v) = v$,
- *If the DM is* optimistic *then* $\forall v \in V$ $k_1(v) = v$, $k_2(0) = 0$ *and* $\forall v \in \{V - 0\}$ $k_2(v) = 1$,
- *If the DM is* neutral *then* $\forall v \in V$ $k_1(v) = k_2(v) = v$.

The possibilistic utility of the lottery π denoted by $PU(\pi)$ is considered as the *expected qualitative utility* of Π. This means that the expected qualitative utility of the consequence x is PU(x), since each consequence can be written in the form of a lottery π_x.

Proof 2.2 *Let π_x be a lottery representing the consequence x. The expected qualitative utility of this lottery is computed as follows:*
$PU(\pi_x) = max_{y \in V} min(k(\pi_x(y)), BU(y))$
$PU(\pi_x) = max\{\{min(< 0, 0 >, BU(y)) | y \neq x)\}, min(< 1, 1 >, BU(x))\}$
$PU(\pi_x) = BU(x)$

The main representation theorem that summarizes the utility theory for the possibility theory is the following:

Theorem 2.2 *Let $\succeq$ be a preference relation on Π_X. $\succeq$ satisfies axioms Axiom 1^S through Axiom 4^S iff there exist a utility function PU satisfying (2.22) s.t:*

$$\pi \succeq \pi' \; iff \; PU(\pi) \gg PU(\pi') \; \forall \pi, \pi' \in \Pi_X \tag{2.24}$$

Obviously, a rational decision maker (optimistic, pessimistic or neutral) will choose the best lottery denoted by π^* and defined by:

$$\pi^* = argsup_{i \in \{1,..,n\}} PU(\pi_i) \tag{2.25}$$

Example 2.7 *Let us reconsider the sets X, V and the lotteries π_1, π_2, π_3 given in Example 2.5 with the following binary utility scale:*
$U_V = \{< 0,1 >; < 0.5,1 >; < 0.7,1 >; < 1,1 >; < 1,0.7 >; < 1,0.5 >; < 1,0 >\}$.
We can distinguish three cases of indifference relation between consequences and canonical lotteries depending on the decision maker's behaviour.

Case 1: The pessimistic case

Utility assignment for consequences is: $BU(x_\top) =< 1,0 >$, $BU(x_2) =< 1,0.5 >$, $BU(x_3) =< 1,0.7 >$ *and* $BU(x_\bot) =< 1,1 >$.
The binary vector function $k =< k_1, k_2 >$ *is defined by:*
$k_1(0) = 0$ *and* $\forall v \in \{V - 0\}$ $k_1(v) = 1$, $\forall v \in V$ $k_2(v) = v$.
Then, the expected qualitative utility for each lottery is computed as follows:

$$PU(\pi_1) = max \begin{cases} min(k(0.7), BU(x_\top)) \\ min(k(1.0), BU(x_2)) \\ min(k(0.5), BU(x_3)) \\ min(k(0.5), BU(x_\bot)) \end{cases} = max \begin{cases} min(< 1.0, 0.7 >, < 1.0, 0.0 >) \\ min(< 1.0, 1.0 >, < 1.0, 0.5 >) \\ min(< 1.0, 0.5 >, < 1.0, 0.7 >) \\ min(< 1.0, 0.5 >, < 1.0, 1.0 >) \end{cases}$$

$$= max \begin{cases} < 1.0, 0.0 > \\ < 1.0, 0.5 > \\ < 1.0, 0.5 > \\ < 1.0, 0.5 > \end{cases} =< 1, 0.5 >.$$

Similarly, $PU(\pi_2) =< 1,1 >$ *and* $PU(\pi_3) =< 1,1 >$, *then* $\pi_1 \succeq \pi_2 \sim \pi_3$.

Case 2: The optimistic case

Utility assignment for consequences is: $BU(x_\top) =< 1,1 >$, $BU(x_2) =< 0.7,1 >$, $BU(x_3) =< 0.5,1 >$ *and* $BU(x_\bot) =< 0,1 >$
The binary vector function $k =< k_1, k_2 >$ *is defined by:*
$\forall v \in V$ $k_1(v) = v$ *and* $k_2(0) = 0$, $\forall v \in \{V - 0\}$ $k_2(v) = 1$
The expected qualitative utility for each lottery is computed as follows:

$$PU(\pi_1) = max \begin{cases} min(k(0.7), BU(x_\top)) \\ min(k(1.0), BU(x_2)) \\ min(k(0.5), BU(x_3)) \\ min(k(0.5), BU(x_\perp)) \end{cases} = max \begin{cases} min(< 0.7, 1.0 >, < 1.0, 1.0 >) \\ min(< 1.0, 1.0 >, < 0.7, 1.0 >) \\ min(< 0.5, 1.0 >, < 0.5, 1.0 >) \\ min(< 0.5, 1.0 >, < 0.0, 1.0 >) \end{cases}$$

$$= max \begin{cases} < 0.7, 1.0 > \\ < 0.7, 1.0 > \\ < 0.5, 1.0 > \\ < 0, 1.0.0 > \end{cases} =< 0.7, 1 >.$$

Similarly, $PU(\pi_2) =< 0.7, 1 >$ *and* $PU(\pi_3) =< 1, 1 >$*, then* $\pi_3 \succeq \pi_1 \sim \pi_2$.

Case 3: The neutral case

Utility assignment for consequences is: $U(x_\top) =< 1, 0 >$, $U(x_2) =< 1, 0.5 >$, $U(x_3) =< 0.7, 1 >$ *and* $U(x_\perp) =< 0, 1 >$

The binary vector function $k =< k_1, k_2 >$ *is defined by:*

$\forall v \in V \;\; k_1(v) = k_2(v) = v$

The expected qualitative utility for each lottery is computed as follows:

$$PU(\pi_1) = max \begin{cases} min(k(0.7), BU(x_\top)) \\ min(k(1.0), BU(x_2)) \\ min(k(0.5), BU(x_3)) \\ min(k(0.5), BU(x_\perp)) \end{cases} = max \begin{cases} min(< 0.7, 0.7 >, < 1.0, 0.0 >) \\ min(< 1.0, 1.0 >, < 1.0, 0.5 >) \\ min(< 0.5, 0.5 >, < 0.7, 1.0 >) \\ min(< 0.5, 0.5 >, < 0.0, 1.0 >) \end{cases}$$

$$= max \begin{cases} < 0.7, 0 > \\ < 1, 0.5 > \\ < 0.5, 0.5 > \\ < 0, 0.5 > \end{cases} =< 1, 0.5 >.$$

Similarly, $PU(\pi_2) =< 0.7, 1 >$ *and* $PU(\pi_3) =< 1, 1 >$*, then* $\pi_1 \succeq \pi_3 \succeq \pi_2$.

Note that the preference relation changes according to the decision maker's behaviour.

The axiomatic system for decision making with possibility theory, proposed by Giang et al. [17], unifies and generalizes the pessimistic and optimistic systems of axioms (S_P, S_O) proposed by Dubois et al. [12]. The two informational axioms expressing pessimism and optimism are substituted by the monotonicity axiom (Axiom 2^S). Then, the continuity axiom is generalized by affecting to each prize an equivalent canonical lottery as defined by (Axiom 4^S).

2.8 Conclusion

In this chapter we have presented the basics of possibility theory, the non classical theory of uncertainty that offers an appropriate framework to handle uncertainty qualitatively and quantitatively.

Possibility theory allows avoiding limitations of probability theory used for the quantification of influence diagrams. To profit from the benefits of this theory and of the simplicity of influence diagrams, we are interested in developing the possibilistic counterpart of standard IDs based on the possibility theory to specify the numerical component of this graphical decision model.

Our new approach for decision making under uncertainty, called possibilistic influence diagrams, will be detailed in next chapter.

Chapter 3

Possibilistic Influence Diagrams

3.1 Introduction

As detailed in Chapter 1, influence diagrams are a powerful tool for representing and analyzing decision problems under uncertainty. However, they work under the assumption that exact conditional probability tables related to chance nodes are well known and that decision makers are able to specify their utility functions.

This is not always obvious, especially when experts can only provide a preferential relation between different states of the world and when decision makers can only express their utilities qualitatively.

Our idea is to use possibility theory which offers an appropriate framework to deal with imprecision and qualitative uncertainty and to benefit from the simplicity of influence diagrams to develop a new approach for decision making under uncertainty, called *Possibilistic Influence Diagrams*.

This chapter is organized as follows: Section 3.2 will focus on the definition of possibilistic influence diagrams. Fundamental steps for the evaluation of possibilistic influence diagrams are presented in section 3.3. Section 3.4 and 3.5 will detail evaluation methods of product-based and min-based possibilistic influence diagrams, respectively.

3.2 Definition of possibilistic influence diagrams

Possibilistic influence diagrams, as standard influence diagrams, are composed of two components:

1. **A graphical component** defined by a directed acyclic graph (DAG), denoted by G(N,A), where N is the set of chance, decision and value nodes and A is the set of arcs in the directed graph.

2. **A numerical component** evaluating different dependences between chance nodes and utilities for value nodes.

   - For each chance node C_i, we should provide conditional possibility degree $\Pi(c_{ij} \mid pa(C_i))$ 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:

     $$max_{c_{ij}} \Pi(c_{ij} \mid_{\diamondsuit} pa(C_i)) = 1, \ \forall pa(C_i) \tag{3.1}$$

     Where $\diamondsuit$ is the min-based or the product-based conditioning according to the interpretation of the possibilistic scale.
     Note that for root chance nodes (i.e $Pa(C_i) = \emptyset$), (3.1) corresponds to $max_{c_{ij}} \Pi(c_{ij}) = 1$.

   - For each value node V_i, there are three ways to represent decision maker's preferences on the set of consequences, namely, using cardinal, ordinal or qualitative binary utilities.

   - Likewise standard influence diagrams, decision nodes in possibilistic IDs are not quantified.

To develop possibilistic chain rule, utility nodes and links into decision nodes in possibilistic influence diagrams must be ignored, as in standard influence diagrams.

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 $\Pi(c_{ij} \mid_{\diamondsuit} do(d_{ij}))$ where $do(d_{ij})$ is the particular operator

defined by Pearl [36]. When iterating this reasoning we can bunch the whole decision nodes together and express the joint possiblity 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 possibilistic causal network [3] representing $\Pi(C \mid_{\diamond} do(d))$ i.e the joint possibility relative to C, in the context of the decision's configuration d.
Using the chain rule relative to possibilistic causal networks [3], the following chain rule for possibilistic IDs can be inferred:

$$\pi_{\diamond}(C \mid_{\diamond} D) = \bigotimes_{C_i \in C} \Pi(C_i \mid_{\diamond} Pa(C_i)) \tag{3.2}$$

In other words, possibilistic influence diagrams is 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 offer several kinds of possibilistic influence diagrams summarized in Table 3.1.

Table 3.1: Different variants of possibilistic influence diagrams

Value nodes Chance nodes	Cardinal utility	Ordinal utility	Qualitative possibilistic utility (ΠO)
Product-based conditioning (*)	ΠID_*^C	ΠID_*^O	$\Pi ID_*^{\Pi O}$
Min-based conditioning (min)	ΠID_{min}^C	ΠID_{min}^O	$\Pi ID_{min}^{\Pi O}$

These different kinds of possibilistic IDs can be grouped into three principal classes:

- *Product-based possibilistic influence diagrams* where both dependences between chance nodes and value nodes are quantified in a genuine numerical setting (i.e ΠID_*^C).
- *Min-based possibilistic influence diagrams* where both dependences between chance nodes and value nodes are quantified in a qualitative setting used for encoding an ordering between different states of the world (i.e ΠID_{min}^O and $\Pi ID_{min}^{\Pi O}$).
- *Mixed possibilistic influence diagrams* where dependences between chance nodes and value nodes are not quantified in the same setting (i.e ΠID_{min}^C, ΠID_*^O and $\Pi ID_*^{\Pi O}$).

In this work, we are only interested in homogeneous possibilistic influence diagrams, namely product-based and min-based ones.

3.2.1 Product-based possibilistic influence diagrams

Product-based possibilistic influence diagrams are appropriate for a numerical interpretation of the possibilistic scale. In these models, conditionals between chance nodes are defined using product-based conditioning expressed by (2.11) and value nodes are characterized by cardinal utilities (detailed in section 1.2). In fact, the decision maker is supposed to be able to provide exact numerical values reflecting his preferences relative to the set of consequences.

The product-based chain rule relative to product-based possibilistic influence diagrams, denoted by ΠID_*^C, is defined by (3.2) where $\diamondsuit$ corresponds to the product-based conditioning expressed by (2.11) and $\bigotimes$ corresponds to the product operator which is the most appropriate conjunctive operator for the numerical setting, namely:

$$\pi_p(C \mid_p D) = \prod_{C_i \in C} \Pi(C_i \mid_p Pa(C_i)) \tag{3.3}$$

Example 3.1 *Let us consider the ΠID_*^C presented in Figure 3.1 and quantified by Table 3.2 for chance nodes and Table 3.3 for value nodes.*

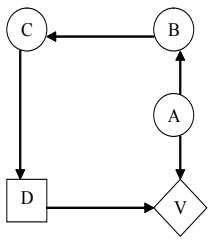

Figure 3.1: Influence Diagram

Table 3.2: A priori and conditional possibility distributions for chance nodes

A	$\Pi(A)$	A	B	$\Pi(B\|A)$	B	C	$\Pi(C\|B)$
T	1	T	T	0.9	T	T	1
F	0.6	F	T	0.2	F	T	0.3
		T	F	1	T	F	0.2
		F	F	1	F	F	1

Table 3.3: Cardinal utilities

A	D	U(A,D)
T	Act1	45
F	Act1	19
T	Act2	8
F	Act2	70

3.2.2 Min-based possibilistic influence diagrams

Min-based possibilistic influence diagrams are appropriate for a qualitative interpretation of the possibilistic scale. In these models, conditionals between chance nodes are defined using min-based conditioning expressed by (2.10). This means that $\diamondsuit$ in (3.1) corresponds to this type of conditioning.

To quantify value nodes, we propose to use two kinds of utility which respect VNM axiomatic system :

1. Ordinal utility (see section 1.2),
2. Qualitative binary utilities (see section 2.7).

The min-based chain rule relative to min-based possibilistic influence diagrams, denoted by ΠID_{min}, is defined by (3.2) where $\diamondsuit$ corresponds to the min-based conditioning expressed by (2.11) and $\bigotimes$ corresponds to the min operator which is the most appropriate conjunctive operator for the qualitative setting, namely:

$$\pi_m(C \mid_m D) = min_{C_i \in C} \Pi(C_i \mid_m Pa(C_i)) \tag{3.4}$$

In min-based possibilistic influence diagrams, value nodes are quantified using ordinal utilities or qualitative binary utilities.

Min-based possibilistic IDs with ordinal utilities

Min-based possibilistic IDs with ordinal utilities, denoted by ΠID^{O}_{min}, are min-based possibilistic IDs in which value nodes are characterized by an ordinal utility. This means that decision maker should provide a complete preferential relation, denoted by $\succeq$, between different consequences.

Example 3.2 *Let the ΠID^{O}_{min} represented in Figure 3.1. Possibility distributions for chance nodes A, B and C are the same as those for ΠID^{C}_{*}, presented in Table 3.2 but the possibility scale is considered as a qualitative one.*
The decision maker provides the following preferential relation between different consequences:
$(D = Act2 \wedge A = F) \succeq (D = Act1 \wedge A = T) \succeq (D = Act1 \wedge A = F) \succeq (D = Act2 \wedge A = T)$.

Min-based possibilistic IDs with qualitative binary utilities

Min-based possibilistic IDs with binary utilities, denoted by $\Pi ID^{\Pi O}_{min}$, are min-based possibilistic IDs in which value nodes are characterized by binary utility. In that case, the decision maker should provide a complete preferential relation, denoted by$\succeq$, between different consequences and a utility scale. Obviously, he should precise his behavior (i.e pessimistic, optimistic or neutral).

Example 3.3 *Let us consider the $\Pi ID^{\Pi O}_{min}$ represented in Figure 3.1. A priori and conditional possibility distributions of chance nodes A, B and C are presented in Table 3.2.*
Let us consider an optimistic decision maker providing the following preferential relation:
$(D = Act1 \wedge A = T) \succeq (D = Act2 \wedge A = F) \succeq (D = Act2 \wedge A = T) \succeq (D = Act1 \wedge A = F)$ *and an utility scale* $W = \{1, 0.7, 0.5, 0\}$

3.3 Evaluation of possibilistic influence diagrams

Once the possibilistic ID is constructed, it must be evaluated to determine optimal decisions. As we have seen in chapter 1, there are two approaches to evaluate standard influence

diagrams, namely, the direct and the indirect ones.

Direct evaluation methods [42, 50] require heavy computations since they are based on arc reversal and node deletion, contrary to indirect ones which are based on the transformation of influence diagrams into Bayesian networks. This explains the great development of indirect methods in the probabilistic case [6, 44, 53, 52, 8, 40].

This motivates us to develop an indirect evaluation method for possibilistic influence diagrams. Our choice is reinforced by the fact that a possibilistic counterpart of Bayesian networks has been recently developed as well as their propagation algorithms [3].

More precisely, we will develop a possibilistic counterpart of Cooper's method for the particular case of influence diagram with a unique value node, since it represents the basis of all other indirect methods.

The key idea of the proposed algorithm is to transform decision and value nodes into chance nodes in order to obtain possibilistic causal networks and then, perform propagation in this secondary structure.

3.3.1 Decision nodes transformation

Each decision node D_i in the possibilistic influence diagram is transformed into a chance node which will be quantified by the following conditional possibility distribution in the context of its parents:

$$\Pi(d_{ij} \mid_{\diamond} pa(D_i)) = 1, \quad \forall d_{ij}, pa(D_i) \tag{3.5}$$

This distribution expresses the state of total ignorance with respect to the new chance node D_i. Note that this quantification is more appropriate than the one used in standard influence diagram (expressed by (1.8)), since equi-probability represents randomness rather than total ignorance.

3.3.2 Value node transformation

The value node V is converted into a new binary chance node having two values: False (F) and True (T). This new chance node is characterized by a possibility distribution expressed by:

$$\Pi(v \mid_{\diamond} pa(V)), \quad \forall v \in \{T, F\} \;\; and \;\; pa(V) \in Pa(V) \tag{3.6}$$

Note that different steps of the quantification of the new chance node V depend on the type of utility used to specify decision maker's preferences on the set of consequences in possibilistic IDs. Different evaluation's methods are detailed in what follows.

3.4 Evaluation of product-based possibilistic influence diagrams

In this type of possibilistic influence diagram, decision nodes will be transformed into chance nodes using (3.5) where $\diamondsuit$ corresponds to the product-based conditioning expressed by (2.11). Moreover, the possibility distribution attached to the new chance node issued from the value node and expressed by (3.6) will be computed by rescaling the set of numerical utilities into the unit interval [0,1] as follows, $\forall pa(V) \in Pa(V)$:

$$\Pi(v = T \mid_p pa(V)) = \frac{U(pa(V)) - U_{min}}{U_{max} - U_{min}} \tag{3.7}$$

$$\Pi(v = F \mid_p pa(V)) = \frac{U(pa(V)) - U_{max}}{U_{min} - U_{max}} \tag{3.8}$$

Where U_{max} and U_{min} are the maximal utility level and the minimal utility level, respectively.

Proof 3.1 *We will prove that (3.7) and (3.8) transform U into [0,1] and save the same order induced initially by U. We have* $\forall pa(V) \in Pa(V)$*:*

$U_{min} \leq U(pa(V)) \leq U_{max}$,

$\Rightarrow 0 \leq U(pa(V)) - U_{min} \leq U_{max} - U_{min}$,

$\Rightarrow 0 \leq \dfrac{U(pa(V)) - U_{min}}{U_{max} - U_{min}} \leq 1$,

Thus, $\dfrac{U(pa(V)) - U_{min}}{U_{max} - U_{min}} \in [0, 1]$,

Let f(Pa(V)) $= \frac{U(Pa(V)) - U_{min}}{U_{max} - U_{min}}$. *The function f and U have the same derivative since* U_{min} *and* U_{max} *are constant. Hence, f and U have the same direction of variation in [0,1].*

The proof for (3.8) can be done in the same manner.

Note that the resulting possibility distribution relative to the new chance node V is subnormalized. In order to satisfy the normalization constraint, the obtained possibility distribution should be transformed as follows: $\forall pa(V) \in Pa(V)$, $\forall v \in \{T, F\}$

$$\Pi(v|_p pa(V)) = \begin{cases} 1 & if\ max(\Pi(v|_p pa(V)), \Pi(\neg v|_p pa(V))) \ = \ \Pi(v|_p pa(V)) \\ \frac{\Pi(v|_p pa(V))}{\Pi(\neg v|_p pa(V))} & otherwise. \end{cases} \tag{3.9}$$

Once the product-based possibilistic causal network is constructed, the Maximal Expected Utility (MEU) is computed to generate the optimal strategy. The computation of MEU begins from the last decision D_m to the first one D_1 by considering for each decision a set of evidence E updated in the light of the previous step.
More formally, for each decision D_i:

$$MEU(D_i, E) = max_{d_{ij}}[max_{pa'(V)}(U(Pa(V))\Pi(pa'(V) \mid d_{ij}, E))] \tag{3.10}$$

Note that $\Pi(pa'(V) \mid d_{ij}, E))$ is computed via the product-based propagation algorithm in product-based causal networks. Indeed, two product-based propagation algorithms have been defined according to the nature of the DAG in the possibilistic causal network [3]. Namely, the possibilistic adaptation of the centralized version of Pearl's algorithm is used when the DAG is singly connected, and the possibilistic adaptation of junction trees propagation are appropriate for multiply connected DAGs.

Example 3.4 *Let us reconsider the same example of possibilistic ID represented in Figure 3.1. After the transformation, the obtained possibilistic network is presented in Figure 3.2.*

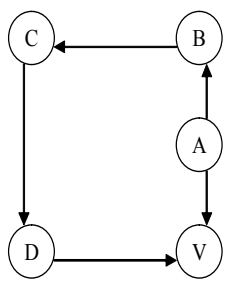

Figure 3.2: Obtained possibilistic network

- *The possibility distribution relative to the node D in context of C (using (3.5)), is presented in Table 3.4.*

- *Table 3.5 gives conditional possibility distribution $\pi(V \mid A, D)$ (using (3.7) and (3.8)) on the set of numerical utilities presented in Table (3.3).*
  For instance, $\Pi(v = T \mid A = T, D = Act1) = (45 - 8)/(70 - 8) = 0.596$
  This conditional possibility distribution is normalized using (3.9).

Table 3.4: The possibility distribution of the node D

C	D	$\Pi(D \mid C)$
T	Act1	1
F	Act2	1
T	Act2	1
F	Act1	1

Table 3.5: A priori and conditional possibility distribution of the value node V

V	D	A	$\Pi(V\|A, D)$	normalized $\Pi(V\|A, D)$
T	Act1	T	0.59	1
T	Act1	F	0.17	0.20
T	Act2	T	0	0
T	Act2	F	1	1
F	Act1	T	0.41	0.69
F	Act1	F	0.83	1
F	Act2	T	1	1
F	Act2	F	0	0

Suppose that we receive a certain information saying that the variable C takes the value T, then $\Pi(pa'(V) \mid d_{ij}, E))$ is computed using product-based possibilistic propagation in Junction Trees since the obtained causal network is multiply connected. Obtained values are represented in Table 3.6.

Thus, the maximal expected utility is obtained as follows: $U(A = T, D = Act1) * \Pi(A = T \mid D = Act1, C = T) = 0.9 * 45 = 0.405$,
$U(A = F, D = Act1) * \Pi(A = F \mid D = Act1, C = T) = 0.18 * 19 = 0.034$,
$U(A = T, D = Act2) * \Pi(A = T \mid D = Act2, C = T) = 0.18 * 8 = 1.44$,

Table 3.6: Results of propagation

A	D	C	$\Pi(A\|D,C)$
T	Act1	T	0.9
F	Act1	T	0.18
T	Act2	T	0.18
F	Act2	T	0.18

$U(A = F, D = Act2) * \Pi(A = F \mid D = Act2, C = T) = 0.18 * 70 = 12.6$,
Thus, $MEU(D, C = T) = max(0.405, 0.034, 1.44, 12.6) = 12.6$, *which means that the optimal decision is* $D^* = Act_2$.

3.5 Evaluation of min-based possibilistic influence diagram

In this type of possibilistic influence diagrams, decision nodes will be transformed into chance nodes using (3.5) where $\diamondsuit$ corresponds here to the min-based conditioning, expressed by (2.10). The definition of the possibility distribution relative to the chance node issued from the value node depends on the used utility i.e ordinal utility or qualitative binary one. Therefore, we can distinguish two evaluation methods for min-based possibilistic influence diagrams detailed in what follows.

3.5.1 Evaluation of ΠID^O_{min}

In this case, the value node is transformed into a new chance node V characterized by a possibility distribution deduced from the original preferential relation between different consequences.

The following proposition gives a transformation function of the preferential relation between different consequences, which satisfies VNM axioms, into numerical utilities.

Proposition 3.1 *Let* $\succeq$ *be a preference relation between different consequences. Then, the order induced by* $\succeq$ *can be transformed into a numerical scale as follows:*

$$U(pa(V)) = card(Pa(V)) - rank(pa(V)) + 1 \tag{3.11}$$

where *card(Pa(V))* is a function that determines the number of combinations of $pa(V) \in Pa(V)$ and *rank(pa(V))* is the rank of pa(V) in the preference order.

Proof 3.2 *We will prove that the proposed transformation method verifies the fundamental theorem of VNM (presented in section 1.2). Let* $pa_1(V)$ *and* $pa_2(V)$ *be two consequences such that* $pa_1(V) \succeq pa_2(V)$*, we have:*
$rank(pa_1(V)) \leq rank(pa_2(V))$
$\Rightarrow -rank(pa_1(V)) + 1 \geq -rank(pa_2(V)) + 1$
$\Rightarrow card(Pa(V)) - rank(pa_1(V)) + 1 \geq card(Pa(V)) - rank(pa_2(V)) + 1$
Thus, $U(pa_1(V)) \geq U(pa_2(V))$.

The set of utilities assigned to the different consequences according to the preference relation will be transformed into a possibility distribution using (3.7) and (3.8) where the conditioning is a min-based one.

The obtained qualitative possibilistic distribution $\Pi(V \mid Pa(V))$ is sub-normalized. Then, to satisfy the normalization constraint, the qualitative possibilistic distribution should be transformed as follows: $\forall pa(V) \in Pa(V)$ *and* $v \in \{T, F\}$,

$$\Pi(v|_m pa(V)) = \begin{cases} 1 & if\ max(\Pi(v|_m pa(V)), \Pi(\neg v|_m pa(V))) \ = \ \Pi(v|_m pa(V)) \\ \Pi(v|_m pa(V)) & otherwise. \end{cases} \tag{3.12}$$

After the transformation phase, the Maximal Expected Utility relative to each decision node D_i is computed in order to generate the optimal decision strategy. The maximal expected utility is computed beginning by the last decision D_m to the first one D_1 with respect to the set of evidence E which should be updated at each step considering the best decision in the previous step. More formally, for each decision D_i, the maximal expected utility is as follows:

$$MEU(D_i, E) = max_{d_{ij}}[max_{pa'(V)} min(\Pi(v = T \mid pa(V)), \Pi(pa'(V) \mid d_{ij}, E))] \tag{3.13}$$

where Pa'(V) denotes the set of chance nodes in Pa(V) and pa'(V) is an instance of Pa'(V).

Note that $\Pi(v = T \mid pa(V))$ is a transformation of U(Pa(V)) into the unit interval [0,1] using (3.7) in order to let $U(Pa(V))$ and $\Pi(pa'(V) \mid d_{ij}, E)$ commensurable since, the min operator is used.

To compute $\Pi(pa'(V) \mid d_{ij}, E)$ and $\Pi(v = T \mid pa(V))$, a min-based propagation algorithm for qualitative possibilistic networks can be applied. Indeed, two min-based propa-

gation algorithms have been defined according to the nature of the DAG in the possibilistic causal network [3]. Namely, the qualitative possibilistic adaptation of the centralized version of Pearl's algorithm is used when the DAG is singly connected, and the qualitative possibilistic adaptation of junction trees propagation are appropriate for multiply connected DAGs. In the case that these two algorithms are blocked, the anytime algorithm [4] can be used.

Example 3.5 *Let us consider the graphical component presented in Figure 3.1 with the following preferential relation relative to the value node V is as follows:*
$(D = Act2 \wedge A = F) \succeq (D = Act1 \wedge A = T) \succeq (D = Act1 \wedge A = F) \succeq (D = Act2 \wedge A = T)$.

- *The possibility distribution relative to the node D in context of C (computed using (3.5)), is presented in Table 3.4.*
- *This preference relation will be transformed into numerical utilities using (3.11) as presented in Table 3.7.*
  Then the possibility degrees relative to the value node are computed via (3.7) and (3.8) and are presented in Table 3.8. Normalized values are computed using (3.12).

Table 3.7: Numerical utilities

A	D	U(A,D)
T	Act1	3
F	Act2	4
T	Act2	1
F	Act1	2

Suppose that we receive a certain information saying that the variable C takes the value T, then $\Pi(pa'(V) \mid d_{ij}, E))$ *is computed using a min-based propagation algorithm in Junction Trees.*
We have $min(\Pi(v = T \mid A = T, D = Act1), \Pi(A = T \mid D = Act1, C = T)) = min(0.666, 0.666)$,
$min(\Pi(v = T \mid A = F, D = Act1), \Pi(A = F \mid D = Act1, C = T)) = min(0.6, 0.3)$,
$min(\Pi(v = T \mid A = T, D = Act2), \Pi(A = T \mid D = Act2, C = T)) = min(0, 0.9)$,

Table 3.8: The possibility distribution of the value node

V	A	D	$\Pi(V \mid A, D)$	$\Pi(V \mid A, D)$ normalized
T	T	Act1	2/3	1
T	F	Act2	1	1
T	T	Act2	0	0
T	F	Act1	1/3	1/3
F	T	Act1	1/3	1/3
F	F	Act2	0	0
F	T	Act2	1	1
F	F	Act1	2/3	1

$min(\Pi(v = T \mid A = F, D = Act2), \Pi(A = F \mid D = Act2, C = T)) = min(0.333, 0.3)$,
Thus $MEU(D, C = T) = 0.666$ *and* $D^* = Act_1$.

3.5.2 Evaluation of $\Pi ID^{\Pi O}_{min}$

For this type of min-based possibilistic ID, the decision maker gives his utility scale and the preferential relation on the set of consequences and defines his behavior. During the transformation phase, the value node V is transformed into a new chance node V quantified through the following steps:

1. The binary utility scale is constructed from the utility scale provided by the decision maker using (2.18).

2. According to the decision maker's behavior and the preference relation, the binary utilities are assigned to each consequence.

3. In order to transform binary utilities into $\Pi(V \mid Pa(V))$, we propose to define reciprocal function of the binary vector function k, proposed by Giang and Shenoy [17] (c.f Definition 2.7).

   This reciprocal function, denoted by k^{-1} is a mapping from U_W to W, and it is defined independently of the decision maker behavior. More formally, $\forall w \in W$:

$$k^{-1}(BU(w)) = min(BU^R(w), BU^L(w)) \tag{3.14}$$

Where $BU^R(w)$ and $BU^L(w)$ are respectively, the right and the left element of the binary utility (detailed in section 2.7).

Proof 3.3 *We will prove that whatever the behavior of the decision maker is, we have* $\forall w \in W$, $k^{-1}(k(w)) = w$.

- *If the DM is* pessimistic *then* $k^{-1}(k(0)) = k^{-1}(< 0, 0 >) = 0$, *and* $k^{-1}(k(w)) = k^{-1}(< 1, w >) = w$,
- *If the DM is* optimistic *then* $k^{-1}(k(0)) = k^{-1}(< 0, 0 >) = 0$, *and* $k^{-1}(k(w)) = k^{-1}(< w, 1 >) = w$,
- *If the DM is* neutral *then* $k^{-1}(k(w)) = k^{-1}(< w, w >) = w$.

The possibility distribution of the new chance node V obtained by the transformation (3.14) is subnormalized.

To insure this constraint, we should first compute $\Pi(v = F \mid Pa(V))$. However, the computation of the exact value of this degree is not possible from available data. Since we are in an ordinal setting, our idea is to recompute $\Pi'(v = T \mid Pa(V))$ and $\Pi'(v = F \mid Pa(V))$ using (3.7) and (3.8) just to know the maximal value between them. This allows us to normalize the distribution issued from (3.14) as follows, $\forall pa(V) \in Pa(V)$:

$$\Pi(v = T|pa(V)) = \begin{cases} \Pi(v = T|pa(V)) & if\ \Pi'(v = T|pa(V)) < \Pi'(v = F|pa(V)) \\ 1 & otherwise. \end{cases} \tag{3.15}$$

For the computation of the maximal expected utility, we will use the fact that the expected qualitative utility of a lottery π is $PU(\pi)$. Since each consequence x can be written in the form of a lottery (see section 2.7), then we can conclude that the expected utility of each consequence $x \in X$ can be computed as follows (see section 2.7):

$$EU(x) = PU(x) = U(x) \tag{3.16}$$

Hence, to generate the Maximal Expected Utility relative to each decision node D_i is computed in order to generate the optimal decision strategy. The maximal expected utility is computed beginning by the last decision D_m to the first one D_1 with respect to the set of evidence E which should be updated at each step considering the best decision in the previous step, more formally:

$$MEU(D_i, E) = max_{d_{ij}} U(Pa(V)) = max_{d_{ij}} \Pi(v = T \mid Pa(V), d_{ij}, E) \tag{3.17}$$

Example 3.6 *Let us continue with Example 3.3.*
The binary utility scale formed from the utility scale $W = \{1, 0.7, 0.5, 0\}$ *using (2.18) is:*
$U_W = \{< 0, 1 >; < 0.5, 1 >; < 0.7, 1 >; < 1, 1 >; < 1, 0.7 >; < 1, 0.5 >; < 1, 0 >\}$.
Table 3.9 presents binary utilities for the value node V, for different decision makers and their transformation using (3.14).

The new possibility distribution π' *is presented in Table 3.10. The normalized possibility distribution that characterizes the new chance node V is presented in Table 3.11.*

Table 3.9: Binary utilities and their transformation using k^{-1}

		Pessimistic DM		Optimistic DM		Neutral DM	
A	D	U(A,D)	$k^{-1}(U(A,D))$	U(A,D)	$k^{-1}(U(A,D))$	U(A,D)	$k^{-1}(U(A,D))$
T	Act1	$< 1, 0 >$	0	$< 1, 1 >$	1	$< 1, 0 >$	0
F	Act2	$< 1, 0.5 >$	0.5	$< 0.7, 1 >$	0.7	$< 1, 0.5 >$	0.5
T	Act2	$< 1, 0.7 >$	0.7	$< 0.5, 1 >$	0.5	$< 0.7, 1 >$	0.7
F	Act1	$< 1, 1 >$	1	$< 0, 1 >$	0	$< 0, 1 >$	0

Table 3.10: New possibility distribution π'

V	A	D	Pessimistic DM $\Pi'(V \mid A, D)$	Optimistic DM $\Pi'(V \mid A, D)$	Neutral DM $\Pi'(V \mid A, D)$
T	T	Act1	0	1	0
T	F	Act2	0.5	0.7	0.5
T	T	Act2	0.7	0.5	0.7
T	F	Act1	1	0	0
F	T	Act1	1	0	1
F	F	Act2	0.5	0.3	0.5
F	T	Act2	0.3	0.5	0.3
F	F	Act1	0	1	1

Suppose that we receive a certain information saying that the variable C takes the value T, then $\Pi(pa'(V) \mid d_{ij}, E))$ *is computed using a min-based propagation algorithm in Junction*

Table 3.11: Normalized possibility distribution for V

V	A	D	Pessimistic DM $\Pi(V \mid A, D)$	Optimistic DM $\Pi(V \mid A, D)$	Neutral DM $\Pi(V \mid A, D)$
T	T	Act1	0	1	0
T	F	Act2	1	1	1
T	T	Act2	1	1	1
T	F	Act1	1	0	0
F	T	Act1	1	0	1
F	F	Act2	0.5	0.3	0.5
F	T	Act2	0.3	0.5	0.3
F	F	Act1	0	1	1

Trees.

If the decision maker is pessimistic then we have:

$\Pi(v = T \mid A = T, D = Act1, C = T) = 0,$

$\Pi(v = T \mid A = T, D = Act2, C = T) = 0.9,$

$\Pi(v = T \mid A = F, D = Act1, C = T) = 0.3,$

$\Pi(v = T \mid A = F, D = Act2, C = T) = 0.3,$

Thus, $MEU(D, C = T) = 0.9$ *and* $D^* = Act_2$.

If the decision maker is optimistic then we have:

$\Pi(v = T \mid A = T, D = Act1, C = T) = 0.9,$

$\Pi(v = T \mid A = T, D = Act2, C = T) = 0.5,$

$\Pi(v = T \mid A = F, D = Act1, C = T) = 0.3,$

$\Pi(v = T \mid A = F, D = Act2, C = T) = 0,$

Thus, $MEU(D, C = T) = 0.9$ *and* $D^* = Act_1$.

If the decision maker is neutral then we have:

$\Pi(v = T \mid A = T, D = Act1, C = T) = 0,$

$\Pi(v = T \mid A = T, D = Act2, C = T) = 0.9,$

$\Pi(v = T \mid A = F, D = Act1, C = T) = 0.3,$

$\Pi(v = T \mid A = F, D = Act2, C = T) = 0,$

Thus, $MEU(D, C = T) = 0.9$ *and* $D^* = Act_2$.

3.6 Conclusion

In this chapter, we have proposed a new approach for graphical decision models in the possibilistic setting. This approach based on influence diagrams offers to decision makers the possibility to express their utilities in different manners. Namely, we have treated the case where the value node in possibilistic influence diagrams are quantified using cardinal, ordinal or qualitative binary utilities.

Moreover, we have defined evaluation methods for homogeneous possibilistic influence diagrams, i.e those that represent decision problems when experts and decision makers are coherent on their specification of possibility distributions relative to chance nodes and of utilities relative to value nodes.

In the next chapter, we will deal with the implementation and the experimentation of possibilistic influence diagrams.

Chapter 4

Implementation and Experimentations

4.1 Introduction

In this chapter, we present the implementation of possibilistic influence diagrams. Indeed, we propose a new toolbox named "Possibilistic Influence Diagrams" (PIDT) implemented with Matlab 6.5 in order to offer a mean for decision making when uncertainty is presented in the possibilistic framework.

Our toolbox has two purposes. The first one concerns the construction of different kinds of possibilistic influence diagrams, presented in chapter 3, based on the Bayes Net Toolbox (BNT) which is an open source package in Matlab for directed graphical models. The second one, is the implementation of the evaluation methods of the possibilistic ID built. These methods consist on the transformation of the possibilistic ID into a possibilistic network and the use of the Possibistic Networks Toolbox (PNT) to construct this secondary model and make inference on it.

This chapter is composed of two parts. The first one deals with the implementation of the PIDT, different variables and programs used in PIDT are detailed in section 4.2. In the second part, we gives an example of a real decision problem as an illustration of our software in section 4.3.

4.2 Implementation

As we have mentioned before, PIDT offers different kinds of possibilistic influence diagrams as well as their evaluation methods.

Hence, our software has different parameters of the decision problem as inputs and the best decision strategy as well as the maximal expected utility as outputs.

4.2.1 Major variables of PIDT

Major variables used in our programs to implement necessary modules of possibilistic influence diagram toolbox are the following:

- *PID*: a record presenting possibilistic influence diagram, defined as follows:
  - *.cnodes*: double arrays contains chance nodes in a topological order (i.e. ancestors before descendants),
  - *.dnodes*: double arrays contains decision nodes in a topological order,
  - *.unodes*: double arrays contains value nodes in a topological order,
  - *.node_sizes*: double arrays contains cardinalities of the set of nodes. $Node_sizes(i)$ is the cardinality of the node i,
  - *.dag*: N_by_N matrix which represents different links in the PID, we have dag(i,j) = 1 if and only if the node i is a parent of the node J,
  - *.CPD*: 1_by_N cell array of matrices, contains:
    - a *tabular_cpd* object for chance nodes, such that for each chance node i, the vector $.CPD\{i\}.CPT$ contains the initial conditional possibility distribution of the node i in the context of its parents,
    - a *tabular_decision_node* object for decision nodes,
    - a *tabular_utility_node* object for value nodes, such that for each value node i, the vector $.CPD\{i\}.T$ contains utilities of the node i in the context of its parents,
  - *.DM_behaviour*: the type of the decision maker (pessimistic, optimistic or neutral),

  - $.DM_utility_scale$: a vector containing different values of the decision maker utility scale,

- $PNET$: a record presenting possibilistic network, defined as follows:
  - $.nodes$: double arrays contains nodes in a topological order,
  - $.node_sizes$: double arrays contains cardinalities of the set of nodes. $Node_sizes(i)$ is the cardinality of the node i,
  - $.dag$: N_by_N matrix which represents different links in the PID, we have dag(i,j) = 1 if and only if the node i is a parent of the node J,
  - $.CPD$: 1_by_N cell array of matrices, such that each cell $.CPD\{i\}$ contains a $tabular_cpd$ object and the vector $.CPD\{i\}.CPT$ contains the initial conditional possibility distribution of node i in the context of its parents.

- *Other variables*:
  - evid: the set of evidences provided by the decision maker,
  - MEU: the value of the maximal expected utility,
  - strat: the best strategy presented by a vector which has as length the number of decision nodes in PID such that $strat\{i\}$ is the best value of the decision node i.

4.2.2 Principal programs of PIDT

To construct our software, we have developed several programs regrouped by principal task. In fact, for any type of possibilistic influence diagram (min-based or product-based) there are two principal procedures that should be implemented:

1. **The construction procedure** used to built the possibilistic influence diagram (PID) (i.e Built_PID).

2. **The evaluation procedure** used to evaluate the possibilistic ID (i.e Eval_PID). This procedure is composed of two sub_programs as follows:

   - **The transformation procedure** used to transform the possibilistic influence diagram (PID) into a possibilistic causal network (PNET) (i.e Trans_PID).

- **The inference procedure** used to construct the obtained possibilistic causal network (PNET) and apply the appropriate propagation algorithm on it (i.e Inf_pnet).

The possibilistic influence diagram toolbox's architecture is illustrated in Figure 4.1.

More precisely, the min-based possibilistic influence diagram specification's file should

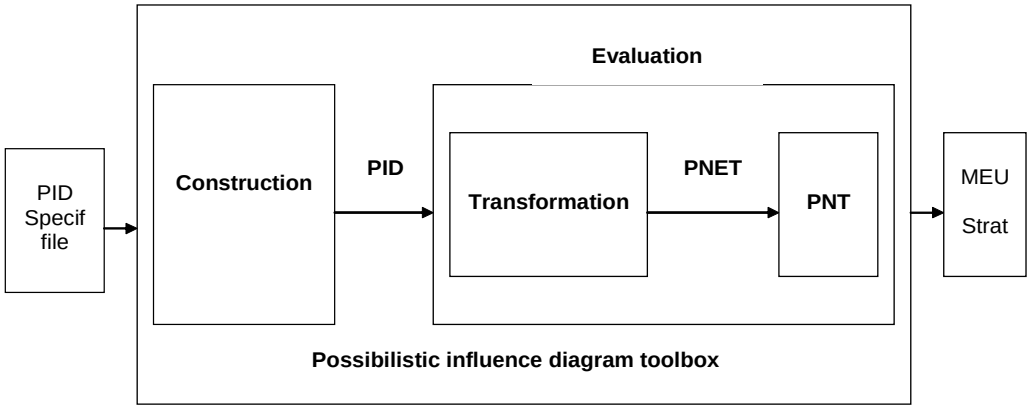

Figure 4.1: PIDT architecture

contain several variables forming the PID record detailed as follows:

- The number of nodes N,
- The topological order between the set of nodes,
- The DAG representing different dependences between the set of nodes, it is a N-by-N matrix containing 0 or 1. Where the value 1 in the line i column j means that the variable i is a parent of the variable j, where $i\ and\ j \in [1, N]$,
- The set of chance nodes,
- The set of decision nodes,
- The set of value nodes,
- The vector specifying sizes of different nodes,

- The conditional possibility distributions relative to chance nodes and stored in multidimentional arrays (PID.CPD), where $PID.CPD\{i\}.CPT$ is the conditional possibility distribution of the chance node i. For instance, if a binary chance node A has two binary parents B and C then, we should provide $[\Pi(a1 \mid b1, c1)\ \Pi(a1 \mid b1, c2)\ \Pi(a1 \mid b2, c1)\ \Pi(a1 \mid b2, c2)\ \Pi(a2 \mid b1, c1)\ \Pi(a2 \mid b1, c2)\ \Pi(a2 \mid b2, c1)\ \Pi(a2 \mid b2, c2)]$,
- The decision maker should provide a preferential relation between different consequences. This preferential relation will be encoded by a vector of ranks such that, the value in the position i of this vector is the rank of the consequence i in the preferential relation,
- The utlity scale of the decision maker,
- The behaviour of the decision maker.

4.3 Illustrative example: Project risk analysis decision problem

As an illustration of the possibilistic influence diagram toolbox, we will simulate a risk analysis project decision problem.

In fact, in business economists often spend a lot of time trying to predict best decisions concerning the implantation of a new project, given available informations. To achieve a successful project risk analysis, the influence diagram model will be used to formulate this problem providing an evaluation of project alternatives.

We will consider the case that a project team will decide about a new project **P**, they should decide to **fund** F or not $\tilde{F}$ in the new project in order to maximize their profit.

Main decisions and critical data concerning **P** will be detailed in what follows:

- **Site selection** (denoted by SS): Suppose that the new project can be established in two sites such that *site A* and *site B*, the project team should choose the most appropriate site for the new project according to the project requirements and to the site properties. Each site is characterized by the following properties:

- Market conditions (denoted by MC), such that $D_{MC} = \{Good, Bad\}$. This means that MC can has two values, Good and Bad.
- Infrastructure of the site (denoted by I), such that $D_I = \{Good, Bad\}$. It is affected by:
  - Electrical power (denoted by EP), such that $D_{EP} = \{High, Low\}$
  - Aircraft (denoted by Air), such that $D_{Air} = \{Good, Bad\}$
  - Rail road (denoted by RR), such that $D_{RR} = \{Good, Bad\}$.
- Tax incentives (denoted by TI), such that $D_{TI} = \{Good, Bad\}$.
- Political and social climate (denoted by PSC) such that $D_{PSC} = \{Yes, No\}$ affected by:
  - Political state (denoted by PS), such that $D_{PS} = \{Stable, Disturbing\}$
  - Skill level (denoted by SL), such that $D_{SL} = \{High, Low\}$.
- Raw material sources (denoted by RMS), such that $D_{RMS} = \{Beneficial,$ B̃eneficial$\}$ affected by:
  - Quality (denoted by Qual), such that $D_{Qual} = \{Good, Bad\}$.
  - Long term availability (denoted by LTA), such that $D_{LTA} = \{High, Low\}$.

- **New technology** (denoted by NT), such that $D_{NT} = \{T1, T2\}$. There are two technologies, *T1* and *T2* that may be used in the new project. Each technology is characterized by two principal properties:
  - Energy consumption (denoted by EC), such that $D_{EC} = \{High, Low\}$.
  - Degree of automation (denoted by DA), such that $D_{DA} = \{High, Low\}$..
- **Credit obtention** (denoted by C), such that $D_C = \{Yes, No\}$.: The project team will decide to have a credit for the foundation of the new project or not. This decision depends on:
  - Availability of the capital for the project (denoted by AC), such that $D_{AC} = \{Yes, No\}$.
  - Bank facilities (denoted by BF), such that $D_{BF} = \{Yes, No\}$.
  - Project budjet (denoted by PB), such that $D_{PB} = \{High, Low\}$..

Figure 4.2 represents an influence diagram for analysing project risk, it illustrates a global approach that could be extended and used in real time to support economists in their decisions.

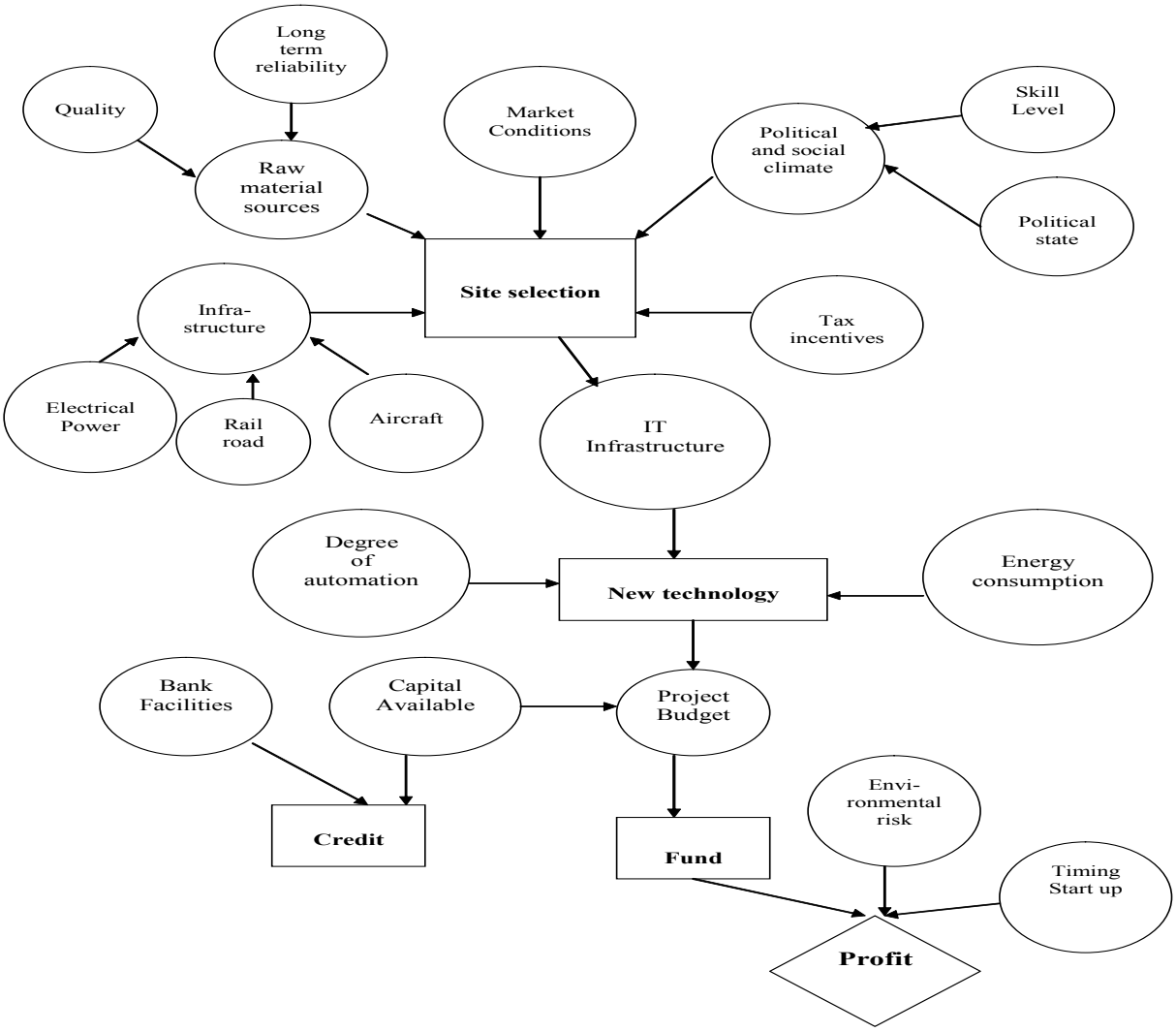

Figure 4.2: An Influence Diagram for project risk analysis

The quantification of incertain variables relative to the project risk analysis problem is presented by following Tables 4.1-4.7 grouped by decision.

Table 4.1: A priori possibility distributions concerning Site Selection decision

LTR	$\Pi(LTR)$	Qual	$\Pi(Qual)$	MC	$\Pi(MC)$
Yes	1	Good	0.2	Good	0.2
No	0.4	Bad	1	Bad	1

SL	$\Pi(SL)$	PS	$\Pi(PS)$	EP	$\Pi(EP)$
High	0	Stable	0.4	High	0.1
Low	1	Disturbing	1	Low	1

RR	$\Pi(RR)$	Air	$\Pi(Air)$	TI	$\Pi(TI)$
Good	0.2	Good	0	Good	0.6
Bad	1	Bad	1	Bad	1

Table 4.2: A priori possibility distributions concerning New Technology decision

DA	$\Pi(DA)$	EC	$\Pi(EC)$
High	0.35	High	0.76
Low	1	Low	1

Table 4.3: A priori possibility distributions concerning Credit decision

BF	$\Pi(BF)$	CA	$\Pi(CA)$
Yes	1	Yes	1
No	0.56	No	0.75

Table 4.4: A priori possibility distributions for chance nodes ER and TSU

ER	$\Pi(ER)$
High	1
Low	0.75

TSU	$\Pi(TSU)$
Yes	1
No	0.3

Table 4.5: Conditional possibility distributions concerning Site Selection decision

RMS	LTR	Qual	$\Pi(RMS \mid Qual, LTR)$
beneficial	Yes	good	0.9
$\tilde{\text{b}}$eneficial	Yes	good	1
beneficial	No	good	1
$\tilde{\text{b}}$eneficial	No	good	0.3
beneficial	Yes	Bad	1
$\tilde{\text{b}}$eneficial	Yes	Bad	1
beneficial	No	Bad	0
$\tilde{\text{b}}$eneficial	No	Bad	1

PSC	PS	SL	$\Pi(PSC \mid PL, SL)$
Yes	Stable	High	0.9
NO	Stable	High	1
Yes	Disturbing	High	1
No	Disturbing	High	0.7
Yes	Stable	Low	1
No	Stable	Low	0.55
Yes	Disturbing	Low	1
No	Disturbing	Low	0.7

I	EP	RR	Air	$\Pi(I \mid EP, RR, Air)$
Good	High	Good	Good	0.9
Bad	High	Good	Good	1
Good	High	Bad	Bad	0.4
Bad	High	Bad	Bad	1
Good	Low	Good	Good	0.8
Bad	Low	Good	Good	1
Good	High	Bad	Good	1
Bad	High	Bad	Good	1
Good	High	Good	Bad	0.4
Bad	High	Good	Bad	1
Good	Low	Bad	Good	0.2
Bad	Low	Bad	Good	1
Good	Low	Good	Bad	0.2
Bad	Low	Good	Bad	1
Good	Low	Bad	Bad	0.4
Bad	Low	Bad	Bad	1

Table 4.6: Conditional possibility distributions concerning New Technology decision

ITI	SS	$\Pi(ITI \mid SS)$
Yes	site A	0.4
No	site A	1
Yes	site B	1
NO	site B	0.2

Table 4.7: Conditional possibility distribution concerning Fund decision

PB	CA	NI	$\Pi(PB \mid CA, NI)$
High	Yes	T1	0.6
Low	Yes	T1	1
High	No	T1	0.3
Low	NO	T1	1
High	Yes	T2	0.1
Low	Yes	T2	1
High	No	T2	0
Low	NO	T2	1

For the formulation of a project risk analysis decision problem, a min-based influence diagram with binary qualitative utilities will be used, according to the nature of decision problem and to available informations.

In fact, decision maker can only provide his utility scale and his preferential relation between different consequences as follows:
$W = \{1, 0.9, 0.85, 0.7, 0.5, 0.3, 0.2, 0\}$ and
$(Fund = F, ER = Low, TSU = Yes) \succ (Fund = \tilde{F}, ER = High, TSU = No) \succ$
$(Fund = \tilde{F}, ER = High, TSU = Yes) \succ (Fund = F, ER = Low, TSU = No) \succ$
$(Fund = \tilde{F}, ER = Low, TSU = No) \succ (Fund = \tilde{F}, ER = Low, TSU = Yes) \succ$
$(Fund = F, ER = High, TSU = No) \succ (Fund = F, ER = High, TSU = Yes)$.

PID specification:

To construct the relative PID, the procedure *Built_PID* will have as input a file specifying necessary parameters of PID.

To encode such possibilistic influence diagram, we should first define the topological order which can be:
[LTR Qual RMS MC SL PS PSC EP RR Air I TI SS ITI DA EC NT BF CA PB C F ER TSU Profit] i.e. $LTR = 1$, $Qual = 2$, $RMS = 3$, $MC = 4$, $SL = 5$, $PS = 6$, $PSC = 7$, $EP = 8$, $RR = 9$, $Air = 10$, $I = 11$, $TI = 12$, $SS = 13$, $ITI = 14$, $DA = 15$, $EC = 16$, $NT = 17$, $BF = 18$, $CA = 19$, $PB = 20$, $C = 21$, $F = 22$, $ER = 23$, $TSU = 24$ and $Profit = 25$.

Then, the **PID** record takes the following values:

- **PID.N**=25,
- **PID.dag**= 25-by-25 matrix containing 0 or 1,
- **PID.node_sizes**= [2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2],
- **PID.cnodes**=[LTR Qual RMS MC SL PS PSC EP RR Air I TI ITI DA EC BF CA PB ER TSU],
- **PID.dnodes**=[SS NT C F],
- **PID.unodes**=[Profit],
- The conditional possibility distributions are stored as multidimentional arrays (PID.CPD).

  - PID.CPD{1}.CPT = [1 0.4],
  - PID.CPD{2}.CPT = [0.2 1],
  - PID.CPD{3}.CPT = [0.9 1 1 0.3 1 1 0 1],
  - PID.CPD{4}.CPT = [0.2 1],
  - PID.CPD{5}.CPT = [0 1],
  - PID.CPD{6}.CPT = [0.4 1],
  - PID.CPD{7}.CPT = [0.9 1 1 0.7 1 0.55 1 0.7],
  - PID.CPD{8}.CPT = [0.1 1],
  - PID.CPD{9}.CPT = [0.2 1],
  - PID.CPD{10}.CPT = [0 1],
  - PID.CPD{11}.CPT = [0.9 1 0.4 1 0.8 1 1 1 0.4 1 0.2 1 0.2 1 0.4 1],
  - PID.CPD{12}.CPT = [0.6 1],
  - PID.CPD{14}.CPT = [0.4 1 1 0.2],
  - PID.CPD{15}.CPT = [0.35 1],
  - PID.CPD{16}.CPT = [0.76 1],
  - PID.CPD{18}.CPT = [1 0.56],
  - PID.CPD{19}.CPT = [1 0.75],
  - PID.CPD{20}.CPT = [0.6 1 0.3 1 0.1 1 0 1],
  - PID.CPD{23}.CPT = [1 0.75],
  - PID.CPD{24}.CPT = [1 0.3],

- PID.DM_utlity_scale = [1 0.9 0.85 0.7 0.5 0.3 0.2 0],
- PID.CPD{25}.T = [8 7 1 4 3 2 6 5],
- PID.DM_behaviour can take three different values which are pessimistic, optimistic or neutral. In this example, we will study these three different cases.

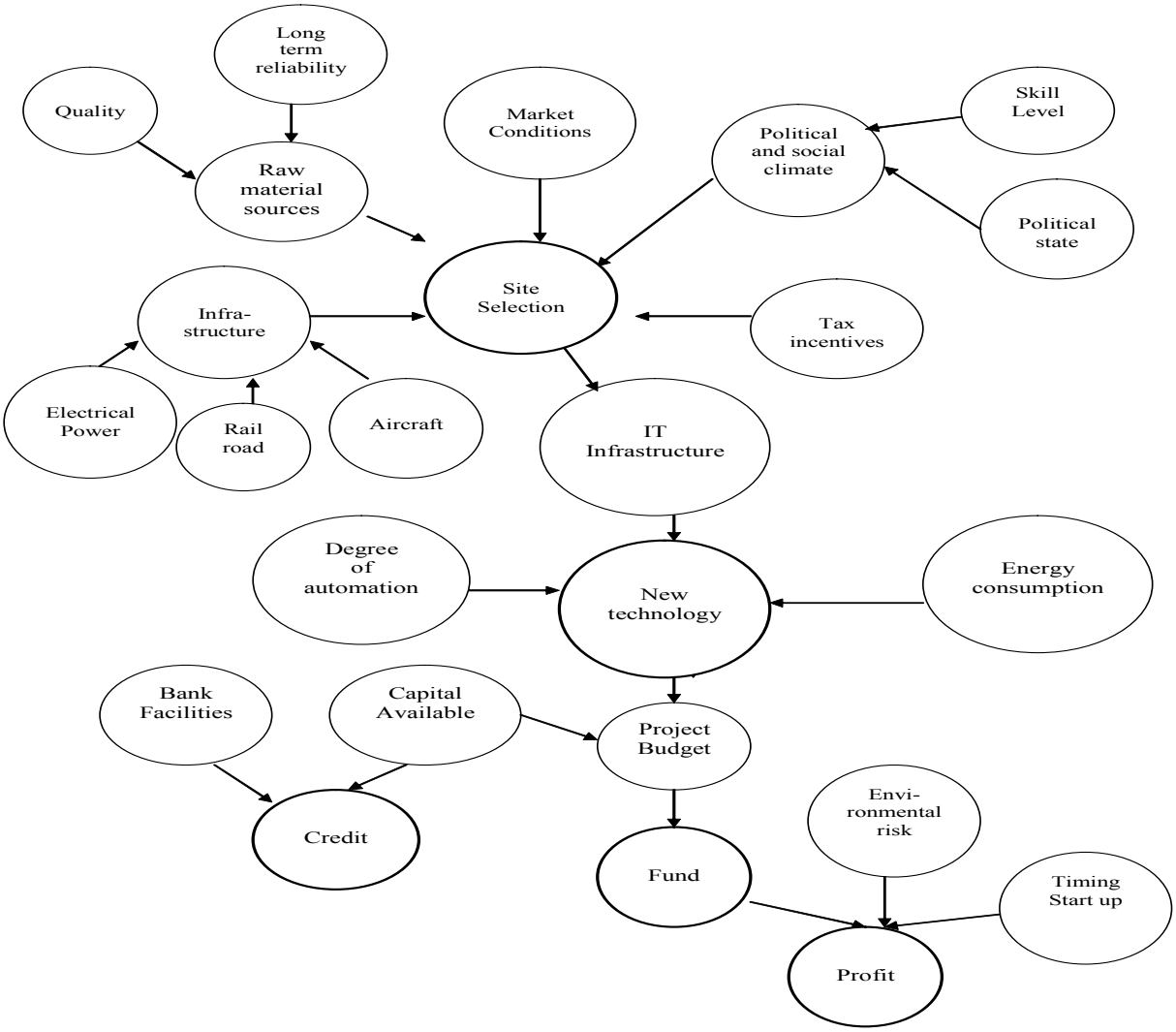

Figure 4.3: Result causal network for project risk analysis

PID evaluation:

Once the influence diagram is constructed, $Eval_PID$ procedure is used to determine best decision strategy. $Trans_PID$ procedure allows to obtain the causal network $PNET$ represented in Figure 4.3.

The PNET record contains following variables:

- $PNET.N = PID.N$,

- $PNET.nodes = [PID.cnodes \;\; PID.dnodes \;\; PID.unodes]$,
- $PNET.dag = PID.dag$,
- $PNET.node_sizes = PID.node_sizes$,
- $PNET.CPD\{i\}.CPT = PID.CPD\{i\}.CPT$
  $\forall i \in \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 18, 19, 20, 23, 24\}$,
- $PNET.CPD\{22\}.CPT$ =[1 1 1 1],
- $PNET.CPD\{17\}.CPT$ =[1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1],
- $PNET.CPD\{21\}.CPT$ =[1 1 1 1 1 1 1 1],
- $PNET.CPD\{13\}.CPT$ =[1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

Thus, in order to maximize his profit a pessimistic decision maker shouldn't fund in the new project P.

- If the deccision maker is **optimistic**: MEU= 0.14 and strat=[2 2 2 2]. More precisely,

  1. strat{4}=2 means that for the fourth decision (Fund), the best alternative is not fund,
  2. strat{3}=2 means that for the third decision (Credit), the best alternative is not obtain a credit,
  3. strat{2}=2 means that for the second decision (New Technology), the best alternative is to adopt T2,
  4. strat{1}=2 means that for the first decision (Site Selection), the best alternative is to choose site B for the new project,

  Thus, in order to maximize his profit an optimistic decision maker shouldn't fund in the new project P.

- If the deccision maker is **neutral**: MEU= 0.7 and strat=[1 2 1 1]. More precisely,

  1. strat{4}=1 means that for the fourth decision (Fund), the best alternative is to fund,
  2. strat{3}=2 means that for the third decision (Credit), the best alternative is not obtain a credit,
  3. strat{2}=1 means that for the second decision (New Technology), the best alternative is to adopt T1,
  4. strat{1}=1 means that for the first decision (Site Selection), the best alternative is to choose site A for the new project,

  Thus, in order to maximize his profit a natural decision maker should fund in the new project P.

4.4 Conclusion

In this chapter, we have detailed major variables and main programs that we have used in order to implement different kinds of possibilistic influence diagrams as well as their

evaluation methods.

Then, we have presented a real decision problem as an illustration of our new toolbox.

As a future work, we aim to develop a user interface of our toolbox.

Conclusion

In this report we have proposed a new approach for decision making under uncertainty using influence diagrams when both chance and value nodes are quantified using possibility theory.

The possibility theory encodes uncertainty in two ways, quantitatively using numerical values in the unit interval or qualitatively using a total preferential relation between different states of the world. These two aspects leads to several kinds of *possibilistic influence diagrams* organized in three principal classes. Namely, product-based, min-based and mixed possibilistic influence diagrams.

In this work, we were only interested in product-based and min-based possibilistic ones corresponding to the case when experts and decision makers are coherent and express their uncertainty in the same setting. More precisely, we have developed three kinds of possibilistic influence diagrams:

- product-based possibilistic influence diagrams with cardinal utilities,
- min-based possibilistic influence diagrams with ordinal utilities,
- min-based possibilistic influence diagrams with qualitative binary utilities.

To evaluate these possibilistic influence diagrams we have opted for an indirect method based on the transformation of the initial network into a causal possibilistic network [3].

Our developments show that product-based possibilistic influence diagrams are very close to standard ones, in fact they share same operators (i.e. product operator). This is not the case for min-based possibilistic influence diagrams, since the min operator has different properties than the product one which leads to a deep study of normalization constraints.

In addition, value node transformation procedure in product-based possibilistic influence diagrams is quiet similar to the probabilistic one used for standard influence diagrams, whereas it is different in min-based influence diagrams.

As future works, we can first distinguish direct improvements of our proposals. In fact, we aim to optimize proposed evaluation methods to deal with possibilistic influence diagrams with more than one value node in order to treat multi-objective decision problems. Moreover, it will be useful to enrich our toolbox with a convivial user-interface.

Another line of research will be to extend our work to mixed influence diagrams in order to treat the case when experts and decision makers are heterogeneous. Our idea is to propose a recursive evaluation method which decomposes the mixed influence diagrams into several homogeneous parts which will be evaluated separately.

Finally, it will be interesting to develop possibilistic influence diagrams to deal with asymmetric decision problems. In fact, until 2005, the only graphical decision model treating asymmetric decision problems were decision trees [38]. Recently, hybrid influence diagrams [48] were proposed as a new approach which combines decision trees and influence diagrams in order to benefit from the strengths of these two models and to deal with asymmetric decision problems. Thus, our idea is to develop the possibilistic counterpart of these diagrams.

Bibliography

[1] A. Akari. Microeconomie du consommateur et du producteur. Etigraph editor, 2000.

[2] T. Bedford and R. Cooke. Probabilistic risk analysis: foundations and methods. Cambridge university Press, 2001.

[3] N. Ben Amor. Qualitative Possibilistic Graphical Models: From Independence to Propagation Algorithms. Ph.D. dissertation, 2002.

[4] 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.

[5] R.M. Chavez and G.F. Cooper. A randomized approximation algorithm for probabilistic inference on Bayesian belief networks. Artificial Intelligence 20, pages 661-685, 1990

[6] G.F. Cooper. A method for using belief networks as influence diagrams. Fourth workshop on uncertainty in artificial intelligence, 1988.

[7] G.F. Cooper. Computational complexity of probabilistic inference using Bayesian belief networks. Artificial Intelligence 42, pages 393-405, 1990.

[8] M. Crowley. Evaluating Influence Diagrams. 2004.

[9] D. Dubois and H. Prade. Possibility theory, an approach to computerized processing of uncertainty. Plenum Press, New York, NY, 1988.

[10] D. Dubois and H. Prade. Representing partial ignorance. IEEE Systems Machine and Cybernetic 26, pages 361-377, 1996.

[11] D. Dubois and H. Prade. An introductory survey of possibility theory and its recent developments, 1998.

[12] D. Dubois, L. Gogo, H. Prade and A. Zapico. Making decision in a qualitative setting: from decision under uncertainty to case-based decision, 1998.

[13] D. Dubois and H. Prade. Possibility theory as a basis for qualitative decision theory. 14th Int. Joint Conf. on Artificial Intelligence (IJCAI'95), Montreal, pages 1924-1930, 1995.

[14] D. Dubois, H. Prade and R. Sabbadin. Decision theoretic foundation of qualitative possibility theory. European Journal of Operational Research 128, pages 478-495, 2001.

[15] J. Dyer and R. Sarin. Relative risk aversion. Management Science 28, pages 875-886, 1982.

[16] H. Fargier and R.Sabbadin. Qualitative decision under uncertainty: back to expected utility. Artificial Intelligence, 2004.

[17] 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.

[18] M.C. Horsch and D. Poole. Flexible policy construction by information refinement. Proceedings of the Twelfth Annual Conference on Uncertainty in Artificial Intelligence (UAI 96), pages 315-324, San Francisco, CA. Morgan Kaufmann Publishers, 1996.

[19] R.A. Howard. An assessment of decision analysis. Operation Research, pages 4-27, 1980.

[20] 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.

[21] I. Jenhani. Qualitative Possibilistic Option Decision Trees. Master dissertation, 2005.

[22] F.V. Jensen. An introduction to Bayesian Networks. Taylor and Francis, London, United Kingdom, 1996.

[23] F.V. Jensen. Bayesian networks and decision graphs. Springer, statistics for engineering and information science, 2002.

[24] J. Kim and J. Pearl. Convice; a conversational inference consolidation engine. IEEE Trans. on Systems, Man and Cybernetics 17, pages 120-132, 1987.

[25] J.H. Kim and J. Pearl. A computational model for causal and diagnostic reasoning in inference systems. Proceedings of International Joint Conference on Artificial Intelligence (IJCAI'83), pages 190-193, Karlsruhe (Germany), 1983.

[26] K.P. Murphy, Y. Weiss and M.I. Jordan. Loopy belief propagation for approximate inference, An empirical study. Proceeding of the Fifteenth Conference on Uncertainty in Artificial Intelligence (UAI'99), pages 467-475, San Francisco, CA, 1999. Morgan Kaufmann Publishers.

[27] B.L. Mckinney and D.R. Engfer. Formulating risk into research and engineering projects. Proceedings in Crystal Ball User conference, 2004.

[28] J. Meyer, M.H. Phillips, P.S Cho, I. Kalet and N. Jason. Application of ID to prostate intensity modulated radiation therapy plan selection. Physics in medicine and biology 49, pages 1637-1653, 2004.

[29] P. Naim, P.H. Wuillemin, P. Leray, O. Pourret, and A. Becker. Rseaux baysiens. Eyrolles, Paris, 2004.

[30] J.V. Neumann and O. Morgenstern. Theory of games and economic behavior. Princeton University Press, 1948.

[31] D.K. Owens, R.D. Shachter, J. Nease. Representation and analysis of medical decision problems with influence diagrams. Medical Decision Making 17, pages 241-262, 1997.

[32] Z. Pawlak. Rough sets. Auszug aus: Kunstliche Intelligent, Heft 3/01, 1991.

[33] J. Pearl. Fusion propagation and structuring in belief networks. Artificial Intelligence, 29, pages 241-288, 1986.

[34] J. Pearl. Probabilistic reasoning in intelligent systems networks of plausible inference. Morgan Kaufmann, San Francisco (California), 1988.

[35] J. Pearl. Graphical models for probabilistic and causal reasoning. In D. M. Gabbay and P. Smets, editors, Handbook of Defeasible Reasoning and Uncertainty Management Systems, Volume 1: Quantified Representation of Uncertainty and Imprecision, pages 367-389. Kluwer Academic Publishers, Dordrecht, 1998.

[36] J. Pearl. Causality Models, Reasoning and Inference. Cambridge University Press, 2000.

[37] M.A. Poet and R.D. Shachter. Fusion and propagation with multiple observations in belief networks. Artificial Intelligence 48, pages 299-318, 1991.

[38] H. Raiffa. Decision Analysis: Introductory Lectures on Choices Under Uncertainty. Addison Wesley, Reading, MA,1968.

[39] M. Ramoni. Anytime influence diagrams. Working notes of the IJCAI 95 workshop on "Anytime Algorithms and Deliberation Scheduling", pages 55-62, Montreal Canada, 1995.

[40] D.G. Sanchez and M.J. Druzdzel. An efficient sampling Algorithm for influence diagrams. Proceedings of the second European Workshop on probabilistic graphical models (PGM), pages 97-104, The Netherlands, 2004.

[41] L.J. Savage, The foundations of statistics. Dover, New York, 1972.

[42] R.D. Shachter. Evaluating influence diagrams. Operation Research 34, pages 871-882, 1986.

[43] R.D. Shachter and D.E. Heckerman. A backwards view for assessment. Lemmer JF, Kanal LN (eds), Uncertainty in Artificial Intelligence 2, Amsterdam, pages 317-324, 1988.

[44] R.D. Shachter and M.A. Poet. Decision Making using probabilistic inference methods. Proceedings of 8th Conference on Uncertainty in Artificial Intelligence, pages 276-283, 1992.

[45] G. Shafer. A mathematical theory of evidence. Princeton Univ.Press, Princeton, VJ, 1976.

[46] PP. 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.

[47] PP. Shenoy. A comparison of graphical techniques for decision analysis. European Journal of Operational Research, vol 78, pages 1-21, 1994.

[48] P.P. Shenoy, Finn .V Jensen Thomas D. Nielson. Sequential influence diagrams: A unfied asymmetry framework. Proceedings of the second European Workshop on probabilistic graphical models, pages 121-128, 2004.

[49] D. Surowik, Leonard Savage's mathematical theory of decision. Studies in logic; grammar and rhrtoric 5 (18), 2002.

[50] J.A. Tatman and R.D. Shachter. Dynamic programming and influence diagrams. IEEE Transactions on systems, Man and Cybernetics 20, pages 365-379, 1990.

[51] K. Virtanen, T. Raivio and R.P. Hamalmainen. An Influence Diagram Approach to One-on-One Air Combat. 10th International Symposium on dynamic games and applications, 2002.

[52] Y.Xiang and C.Ye. A simple Method to Evaluate Influence diagrams. Published in third international conference on cognitive science, 2001.

[53] N.L Zhang. Probabilistic inference in influence diagrams. Proceedings of 14th Conference on Uncertaity in Artificial Intelligence, pages 514-522, 1998.

[54] L.Zadeh, Fuzzy sets, Information and Control 8, pages 338-353, 1965.

[55] L. Zadeh. Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems 1, pages 3-28, 1978.

[56] 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 149-194, 1979.