

N°Ordre : 04/2007 – D/IN

Université des Sciences et de la Technologie Houari Boumediene
Faculté d'Électronique et d'Informatique

Université d'Artois
Faculté Jean Perrin Lens

THESE EN COTUTELLE

présentée pour l'obtention du diplôme de doctorat en:
INFORMATIQUE

Spécialité:
Intelligence Artificielle et Bases de Données avancées

Par:
Hadja Faiza KHELLAF-HANED

*Autour des différents modes de Représentation et de
Raisonnement
en théorie des possibilités quantitative*

Soutenue publiquement le 30 Juin 2007, devant le jury composé de:

Professeur	Mohamed	BETTAZ	INI	Président
Professeur	Aicha	AISSANI	USTHB	Directrice de thèse
Professeur	Salem	BENFERHAT	Université d'Artois	Co-directeur de thèse
Professeur	Amar	AISSANI	USTHB	Examinateur
Professeur	Choukri	BEN-YELLES	Université de Grenoble 2	Examinateur
Professeur	Malika	BOUKALA	USTHB	Examinatrice
Professeur	Odile	PAPINI	Université de la Méditerranée	Examinatrice
Professeur	Lakhdar	SAIS	Université d'Artois	Examinateur

*On mesure l'intelligence d'un individu
à la quantité d'incertitude
qu'il est capable de supporter*
EMMANUEL KANT (1724-1804)

*A la mémoire de mon père
A ma famille
A ma fleur Yasmine*

Remerciements

Je remercie le professeur Mohamed Bettaz qui me fait l'honneur de présider ce jury et d'avoir bien voulu juger ce travail.

J'adresse mes remerciements au professeur Choukri Bey Ben-Yelles d'avoir répondu à notre demande pour faire partie de ce jury en tant que rapporteur.

Je remercie chaleureusement le professeur Odile Papini de m'avoir fait l'honneur d'accepter d'être le rapporteur de ce travail.

Mes remerciements vont également au professeur Amar Aissani d'avoir bien voulu participer au jury de soutenance.

Je remercie vivement le professeur Malika Boukala-Ioualalen d'avoir bien voulu accepter d'examiner ce travail. Je la remercie vivement pour ces encouragements.

Que le professeur Lakhdar Sais trouve ma gratitude d'avoir accepté de faire partie de ce jury.

Je tiens à exprimer ma gratitude au professeur Aicha Aissani-Mokhtari de m'avoir accordé sa confiance en voulant bien être ma directrice de thèse et de m'avoir intégré dans son équipe de recherche. Ce fût avec un grand plaisir que j'ai travaillé avec elle durant ces années.

Je suis infiniment reconnaissante au professeur Salem Benferhat de m'avoir proposé ce sujet de thèse, d'avoir dirigé cette thèse et de m'avoir initié à la théorie des possibilités. J'espère que j'ai été à la hauteur de ses attentes malgré tout.

Je remercie très fortement les professeurs Daniel Kayser et Aicha Aissani d'une part et les professeurs Amar Aissani et Salem Benferhat d'autres part, de m'avoir fait bénéficier, dans le cadre des différents accords CMEP, de stages de courtes durées qui m'ont permis d'avancer dans mon travail.

Enfin, je tiens à remercier les membres de ma famille pour le soutien moral qu'ils m'ont prodigué durant ces longues années de labeur et plus particulièrement ma chère mère ainsi que mon époux pour ses encouragements et pour également son soutien logistique.

Table des matières

Table des matières	1
Table des figures	7
Liste des tables	9
Introduction Générale	11
I État de l’art	17
1 Théories de l’incertain et de l’imprécision	21
1.1 Introduction	21
1.2 Quelques notations	21
1.3 La théorie des probabilités	22
1.3.1 Introduction	22
1.3.2 Principes de base	22
1.3.2.1 Mesure de probabilité	22
1.3.2.2 Conditionnement probabiliste	24
1.3.2.3 Théorème de Bayes	25
1.3.3 Indépendance et Indépendance conditionnelle	26
1.3.4 Interprétations des probabilités	27
1.3.5 conclusion	27
1.4 La théorie des fonctions de croyance	28
1.4.1 Fonction de croyance ou de crédibilité	29
1.4.2 Fonction de plausibilité	30
1.4.3 La règle de combinaison de Dempster	31
1.4.4 Le conditionnement	32
1.4.5 Les fonctions de croyances et les probabilités	32
1.4.6 Conclusion sur les fonctions de croyances	33
1.5 La Théorie des possibilités	34
1.5.1 Distribution de possibilités	35
1.5.2 Les mesures de possibilités et de nécessité	35
1.5.3 Conditionnement possibiliste	36

1.5.4	Mesures de probabilités et Mesures de possibilités	37
1.5.5	Mesures de possibilités et les fonctions de croyance	38
1.5.6	Conclusion	39
1.6	Théorie des ensembles flous	39
1.6.1	Fonctions d'appartenance	40
1.6.2	Les caractéristiques d'un ensemble flou	41
1.6.3	Les coupes de niveau	41
1.6.4	Les opérations ensemblistes	42
1.6.4.1	Égalité et inclusion des ensembles flous	42
1.6.4.2	Intersection et Union des ensembles flous	42
1.6.4.3	Complément d'un ensemble flou	43
1.6.5	Mesures de possibilité et de nécessité	43
1.6.6	Raisonnement approché	45
1.6.6.1	L'implication floue	46
1.6.6.2	Le modus ponens généralisé	46
1.7	Conclusion	47
2	Modèles Logiques de l'incertain	49
2.1	Introduction	49
2.2	La logique possibiliste qualitative	49
2.2.1	Base possibiliste standard	50
2.2.2	Distribution de possibilités associée à une base possibiliste standard	50
2.3	La logique possibiliste quantitative	51
2.4	La logique des pénalités	52
2.4.1	Coût d'une interprétation	52
2.4.2	Coût d'une formule	53
2.5	Les logiques connexes	54
2.5.1	La logique possibiliste et la logique probabiliste	54
2.5.2	La logique possibiliste et la logique modale	55
2.5.3	Raisonnement par défaut utilisant les logiques de l'incertain	55
2.5.4	La logique des pénalités et la logique possibiliste	56
2.5.5	La logique des pénalités et les fonctions de croyance	57
2.5.6	La logique des pénalités et les fonctions conditionnelles ordinales	57
2.6	Conclusion	58
3	Modèles Graphiques de l'incertain	59
3.1	Introduction	59
3.2	Réseaux causaux Bayésiens	60
3.2.1	Les concepts de base	60
3.2.1.1	Quelques définitions	61
3.2.1.2	Les réseaux causaux et le critère de D-Séparation	63
3.2.1.3	L'indépendance conditionnelle dans les DAGs	67
3.2.1.4	Modélisation d'un réseau Bayésien	68
3.2.2	Algorithmes de propagation dans les réseaux causaux bayésiens	70

3.2.2.1	Introduction	70
3.2.2.2	La propagation dans les Polyarbres	71
3.2.2.3	La propagation dans les graphes à connexions multiples	76
3.3	Réseaux causaux possibilistes	87
3.3.1	Définition d'un réseau causal possibiliste	88
3.3.1.1	Les réseaux causaux possibilistes basés sur le produit :	88
3.3.1.2	Les réseaux causaux possibilistes basés sur le minimum :	89
3.3.2	Propagation dans les réseaux causaux possibilistes basés sur le minimum	90
3.3.2.1	Propagation basée sur le minimum dans les polyarbres	90
3.3.2.2	Propagation basée sur le minimum dans les graphes à connexions multiples	93
3.4	Conclusion	97
4	Rétrospective sur les transformations en théorie des possibilités	99
4.1	Introduction	99
4.2	Codification logique d'un réseau causal possibiliste basé sur le minimum	100
4.3	Codification logique d'un réseau causal possibiliste basé sur le produit	102
4.3.1	Principe de la transformation	102
4.4	D'une base possibiliste qualitative vers un réseau causal basé sur le minimum	104
4.4.1	Mise sous forme clausale des bases et suppression des tautologies	104
4.4.2	Construction du graphe	105
4.4.3	Calcul des distributions conditionnelles locales	106
4.5	D'une base possibiliste qualitative vers un réseau causal basé sur le produit	106
4.5.1	Calcul de la base marginalisée	107
4.5.1.1	Décomposition de π	107
4.5.1.2	Marginalisation	107
4.5.1.3	Calcul effectif de Σ_C	108
4.5.2	Détermination des parents de A_i	108
4.5.3	Calcul des degrés de possibilités conditionnelles	109
4.6	D'une base possibiliste quantitative vers une base de connaissances qualitative	110
4.7	conclusion	110
II	Transformations entre modes de représentations des connaissances en théorie des possibilités quantitative	113
5	Transformations en théorie des possibilités quantitative	117
5.1	Introduction	117
5.2	Rappel sur la représentation des connaissances en théorie des possibilités quantitative	119
5.2.1	Représentation logique	119

5.2.2	Représentation graphique	119
5.2.3	Relation entre le conditionnement et l'implication matérielle . . .	121
5.3	D'un réseau causal basé sur le produit vers une base possibiliste quantitative	122
5.4	D'une base possibiliste quantitative vers un réseau causal basé sur le produit	126
5.4.1	Réductions et extensions des bases possibilistes quantitatives . .	126
5.4.2	Construction du DAG	130
5.4.3	Calcul des degrés de possibilités conditionnelles	135
5.5	Conclusion	138
6	La théorie des possibilités quantitative, les OCF et la logique des pénalités	141
6.1	Introduction	141
6.2	La logique possibiliste quantitative et la logique des pénalités	142
6.2.1	D'une base de pénalités vers une base possibiliste quantitative . .	143
6.2.2	D'une base possibiliste quantitative vers une base à pénalités . .	144
6.3	Les fonctions conditionnelles ordinales	144
6.3.1	concepts de base	144
6.3.2	Les OCF, la théorie des possibilités et les probabilités infinitésimales	145
6.3.3	Les réseaux causaux basés sur les OCF	145
6.4	D'un réseau causal basé sur les OCF vers un réseau causal possibiliste basés sur le produit	146
6.5	D'un réseau causal possibiliste basé sur le produit vers un réseau causal basé sur les OCF	148
6.6	Relations entre les réseaux causaux basés sur les OCF et les bases de pénalités	148
6.7	Conclusion	150
III	Algorithmes et Expérimentations	151
7	Algorithmes de propagation pour les réseaux causaux possibilistes basés sur le produit	155
7.1	Introduction	155
7.2	Réseaux causaux basés sur le produit	155
7.3	Propagation basée sur le produit dans les polyarbres	157
7.3.1	Algorithme de propagation :	158
7.4	Propagation basée sur le produit dans les réseaux causaux possibilistes à connexions multiples	160
7.4.1	Algorithme de la propagation	160
7.5	conclusion	164

8 Algorithmes d'inférence pour la logique possibiliste quantitative	165
8.1 Introduction	165
8.2 L'inférence en logique possibiliste standard (SPL)	166
8.3 Un nouvel algorithme pour raisonner en logique des possibilités quantitative	169
8.3.1 Le problème WMAXSAT	169
8.3.2 Principe de l'algorithme	170
8.3.3 Complexité de l'algorithme d'inférence en logique des pénalités	175
8.4 Conclusion	175
9 Implémentation et résultats expérimentaux	177
9.1 Introduction	177
9.2 Implémentation	177
9.2.1 Mise en œuvre du processus de propagation dans le mode graphique	178
9.2.2 Mise en œuvre du processus de transformation des connaissances	179
9.2.3 Mise en œuvre de l'algorithme d'inférence	180
9.2.4 Formats des fichiers utilisés	180
9.2.4.1 Les connaissances	180
9.2.4.2 Les fichiers résultats	182
9.3 Etude expérimentale	183
9.3.1 Comportements des algorithmes de raisonnement	184
9.3.1.1 Cas des polyarbres	184
9.3.1.2 Cas des graphes de complexité moyenne	184
9.3.1.3 Cas des graphes complexes	185
9.4 Conclusion	187
Conclusion Générale	195
Bibliographie	206

Table des figures

1.1	Somme orthogonale entre m_1 et m_2	32
1.2	Intervalle de croyance d'un élément focal A	34
1.3	Relations entre les modèles de l'incertitude illustrées par un exemple	39
1.4	Exemple d'ensemble flou "grandes tailles"	40
1.5	Distribution de possibilités $\mu_{age(Jean)}$	44
1.6	Mesure de possibilité	45
1.7	Tableau récapitulatif sur les mesures de l'incertain	48
1.8	Relation entre les modèles de l'incertain	48
3.1	Réseau causal correspondant à l'exemple 3.1	61
3.2	Différents types de graphes	63
3.3	Les types de liaisons dans un DAG	64
3.4	Exemple de liaisons en série	64
3.5	Exemple de liaisons divergentes	65
3.6	Exemple de liaisons convergentes	65
3.7	Réseau bayésien représentant les dépendances entre cinq variables	67
3.8	Topologies où les nœuds A et C sont conditionnellement indépendants étant donné B	67
3.9	Exemple d'un réseau Bayésien	69
3.10	Cautionnement de l'évidence E en sous-ensembles associés aux parents et aux enfants d'un nœud typique A	72
3.11	Exemple d'un polyarbre	72
3.12	Exemple de propagation dans un DAG simplement connecté	75
3.13	Un graphe acyclique dirigé (DAG)	78
3.14	Graphe moral associé au DAG de la figure 3.13	78
3.15	Triangulation du graphe moral de la figure 3.14	81
3.16	Arbre de jonction associé au DAG de la figure 3.13	82
3.17	Le diagramme de l'algorithme de propagation sans évidence	82
3.18	Passage de messages durant la propagation globale associé à l'arbre de jonction de la figure 3.16	86
3.19	Exemple de DAG	90
4.1	Exemple d'un DAG	101

5.1	DAG correspondant à l'exemple 5.2	121
5.2	DAG correspondant à l'exemple 5.3	123
5.3	DAG associé à la base possibiliste de l'exemple 5.3	136
6.1	Exemple d'un DAG	146
7.1	Exemple d'un DAG	156
9.1	Exemple d'un DAG	181
9.2	Temps d'exécution dans le cas des polyarbres avec une évidence	185
9.3	Temps d'exécution dans le cas des polyarbres avec deux évidences	186
9.4	Temps d'exécution dans le cas des graphes moyennement denses (nbpar- max=3)	187
9.5	Temps d'exécution dans le cas de graphes denses avec une évidence	188
9.6	Comportement des algorithmes de raisonnement (nbparmax=7)	189
9.7	Temps d'exécution dans le cas des graphes à connexion multiples avec une évidence	192
9.8	Temps d'exécution dans le cas des graphes à connexions multiples avec deux évidences	193

Liste des tableaux

1.1	Distributions initiales	37
1.2	Conditionnement basé sur le minimum	37
1.3	Conditionnement basé sur le produit	37
1.4	Comparaison entre la théorie des probabilités et la théorie des possibilités	38
1.5	Exemples de T-normes	42
1.6	Exemples de Co-normes	43
1.7	Principales implications floues	46
2.1	Distribution de possibilités associée à la base de connaissances de l'exemple	
2.1	2.1	51
2.2	Distribution de possibilités associée à la base de connaissances de l'exemple	
2.2	2.2	52
2.3	Coûts d'interprétation de l'exemple 2.4	53
3.1	Les probabilités a priori $P(B)$ et $P(C)$	69
3.2	Les probabilités conditionnelles $P(A \mid B \wedge C)$	70
3.3	Les probabilités conditionnelles $P(D \mid A)$	70
3.4	L'ordre d'élimination des variables	80
3.5	Choix des séparateurs appropriés	83
3.6	Distributions initiales	90
3.7	Distributions de possibilités jointes utilisant la règle de chaînage basée sur le minimum	91
4.1	Les degrés de possibilités a priori $\Pi(A)$	101
4.2	Les degrés de possibilités conditionnelles $\Pi(B \mid A)$	101
4.3	Les degrés de possibilités conditionnelles $\Pi(C \mid A)$	102
4.4	Les degrés de possibilités conditionnelles $\Pi(D \mid B \wedge C)$	102
4.5	Les degrés de possibilités conditionnelles $\Pi(E \mid D)$	103
5.1	Distributions de possibilités associée à la base possibiliste quantitative Σ de l'exemple 5.1	120
5.2	Distributions de possibilités a priori associées au graphe de l'exemple 5.2	120
5.3	Distributions de possibilités conditionnelles associées au graphe de l'exemple 5.2	121

5.4	Distributions de possibilités associée au graphe de l'exemple 5.2	122
5.5	Distributions de possibilités conditionnelles initiales de $\Pi(A)$, $\Pi(B A)$ et $\Pi(C A)$	123
5.6	Distributions de possibilités conditionnelles initiales de $\Pi(D B \wedge C)$. .	124
5.7	Distributions de possibilités jointes utilisant la règle de chaînage basée sur le produit	124
5.8	Distributions de possibilités associée à la base π_{Σ_G} de l'exemple 5.4 . . .	125
5.9	Distributions de possibilités locales conditionnelles finales de $\Pi(A CFH)$	137
5.10	Distributions de possibilités conditionnelles finales de $\Pi(B E)$	137
5.11	Distributions de possibilités locales conditionnelles finales de $\Pi(C DFH)$	137
5.12	Distributions de possibilités locales conditionnelles finales de $\Pi(D FGH)$	138
5.13	Distributions de possibilités locales conditionnelles finales de $\Pi(E G)$.	138
5.14	Distributions de possibilités locales conditionnelles finales de $\Pi(F GH)$	139
5.15	Distributions de possibilités locales conditionnelles finales de $\Pi(G H)$ et $\Pi(H)$	139
6.1	Fonction coût associée à la base de pénalités PK de l'exemple 6.1	142
6.2	Les distributions jointes associées à PK et à Σ de l'exemple 6.2	144
6.3	Distributions d'OCF $\kappa(A)$	146
6.4	Distributions d'OCF $\kappa(B A)$	146
6.5	Distributions d'OCF $\kappa(C A)$	147
6.6	Distributions d'OCF $\kappa(D BC)$	147
6.7	Distributions de possibilités conditionnelles $\Pi(A)$	148
6.8	Distributions de possibilités conditionnelles $\Pi(B A)$	148
6.9	Distributions de possibilités conditionnelles $\Pi(C A)$	149
6.10	Distributions de possibilités conditionnelles $\Pi(D BC)$	149
7.1	Distributions initiales	156
7.2	Distributions de possibilités jointes utilisant la règle de chaînage basée sur le produit	157
9.1	Distributions a priori associées au graphe 9.1	181
9.2	Distributions conditionnelles associées au graphe 9.1	182
9.3	Description des formats des connaissances	190
9.4	prototype du fichier résultat	191
9.5	prototype du fichier résultat dans le cas où la base n'infère pas la variable d'intérêt	191

Introduction Générale

L'intelligence artificielle a pour vocation de développer des outils pour représenter les connaissances et de fournir des mécanismes d'inférence pour élaborer des conclusions à partir des connaissances disponibles [1].

C'est la capacité de transmettre des représentations qui a permis l'extraordinaire accélération des possibilités humaines [2]. Quant à l'inférence, c'est l'élément de base de toute description de la cognition [2].

Néanmoins, les connaissances du monde dont nous disposons sont imparfaites. Ceci est dû à plusieurs facteurs liés soit à la manière dont nous récoltons ces connaissances soit à la nature même de ces connaissances.

L'imperfection est perçue soit comme une incertitude soit comme une imprécision.

La notion d'incertitude est une forme subjective de l'ignorance [3]. Elle apparaît quand un agent construit une opinion subjective sur une proposition qui n'est pas définitivement établie.

L'imprécision est une forme objective de l'ignorance. Elle couvre les cas où un agent attribue une valeur à une variable mais sans la précision adéquate. Pour illustrer la différence entre imprécision et incertitude, considérons les deux situations suivantes [4] :

1. John a au plus deux enfants et j'en suis sûr.
2. John a trois enfants mais je n'en suis pas sûr.

Dans le premier cas, le nombre d'enfants est imprécis mais certain. Dans le second cas, le nombre d'enfants est précis mais incertain. Ces deux aspects (imprécision et incertitude) peuvent coexister mais ils sont distincts.

Durant plusieurs décennies, les différents aspects de l'imperfection de l'incertitude ont été modélisés par la théorie des probabilités. La notion de probabilité n'est cependant pas adaptée à la représentation des incertitudes de nature psychologiques [5], telles que celles liées à la fiabilité d'un informateur. Ainsi, durant ces trente années, plusieurs modèles ont émergé, reflétant les nouveaux concepts prenant en compte les différents aspects de l'imperfection, confirmant ainsi que la théorie des probabilités, aussi répandue qu'elle l'est, n'est pas le seul modèle normatif pouvant capturer l'ensemble de ces aspects [4].

La théorie de l'évidence [6] permet également de manipuler les degrés de confiance qu'un observateur attribue à la validité des faits, en n'imposant plus la condition d'additivité fondamentale de la théorie des probabilités.

Afin de modéliser la notion d'imprécision, Zadeh a proposé de modéliser le mécanisme de la pensée humaine par un raisonnement approximatif basé sur des variables linguistiques. Il a introduit alors les ensembles flous [7]. Et afin de tenir compte du caractère incertain et non probabilisé des connaissances, Zadeh [8] a introduit une théorie non probabiliste représentée par la théorie des possibilités, qui fût par la suite développée par Dubois et Prade [9].

Au fil de l'évolution des recherches, d'autres types de représentation ont vu le jour comme les représentations graphiques. Les réseaux bayésiens probabilistes [10, 11, 12] figurent parmi les modèles graphiques les plus répandus. Ils mettent en exergue les relations d'indépendances qui existent entre les variables.

La représentation et le raisonnement sur des connaissances incertaines en théorie des possibilités, peuvent être abordés dans des cadres symbolique ou numérique.

Dans le cadre symbolique, une relation d'ordre est définie entre les formules constituant les connaissances. Nous parlons alors de la théorie des possibilités qualitative.

Par contre dans le contexte numérique, un degré numérique, évaluant l'incertitude, est associé à chaque parcelle de connaissances. Ce cas correspond à la théorie des possibilités quantitative.

Pour chaque contexte, il existe deux façons de codifier les connaissances :

- soit en utilisant le mode logique où les connaissances incertaines sont représentées sous forme de bases de connaissances formées par un ensemble de formules propositionnelles associées à des pondérations mesurant l'incertitude,
- soit en utilisant le mode graphique où les connaissances sont représentées par un certain nombre de variables représentées par des nœuds et un ensemble d'arcs représentant les liens causaux qui existent entre les différentes variables. Au niveau de chaque nœud est attribuée une valuation exprimant l'influence des incertitudes entre les variables dépendantes.

Problématique

Les principaux travaux réalisés en théorie des possibilités se sont focalisés sur la logique possibiliste standard ou qualitative. En effet, il a été proposé :

- des transformations équivalentes entre une base possibiliste qualitative ou standard et un réseau causal possibiliste basé sur le minimum [13],
- des transformations équivalentes entre une base possibiliste qualitative et un réseau causal possibiliste basé sur le produit [14],
- la transformation d'une base possibiliste basée sur le produit vers une base possibiliste basée sur le minimum [15]. La transformation inverse n'a pas été développée car elle ne présente pas d'intérêts vu qu'il existe des algorithmes d'inférence performants opérant sur les bases de connaissances qualitatives.

A partir des deux premières transformations équivalentes, il est possible de codifier une base de connaissances possibiliste quantitative en un réseau causal possibiliste basé sur le produit et ceci via une base de connaissances qualitative. Ceci peut être réalisé :

- en transformant d’abord la base de connaissances possibiliste basée sur le produit en une base de connaissances possibiliste basée sur le minimum,
- puis en codifiant la base de connaissances possibiliste basée sur le minimum en un réseau causal possibiliste basé sur le produit.

Néanmoins, en pratique, cette codification indirecte n’est pas optimale vu qu’elle génère en général des DAGs (Directed Acyclic Graphs) complets. Plus précisément, n’importe quel DAG associé à une base de connaissances basée sur le produit est tel qu’il existe au moins une variable A , ayant l’ensemble des autres variables comme parents. Ceci est dû au fait que la transformation d’une base possibiliste basée sur le produit vers une base possibiliste basée sur le minimum provoque le rajout de toutes les disjonctions, ce qui amène à ajouter toutes les dépendances entre les variables (même si elles sont indépendantes).

Il est important de signaler que la transformation d’un graphe possibiliste basé sur le produit vers une base possibiliste quantitative ne peut être déduite à partir des transformations déjà réalisées dans le cadre de la théorie des possibilités.

Principalement, notre travail s’inscrit dans la théorie des possibilités quantitative où nous avons élaboré des procédures de transformations réciproques entre les deux modes de représentation des connaissances, à savoir la logique possibiliste quantitative et les réseaux de croyances possibilistes quantitatifs .

Ces translations sont très utiles à plus d’un titre et peuvent être exploitées de diverses manières :

- Dans le contexte de multi-expertises ou de multi-agents par exemple, un agent peut indépendamment exprimer ses connaissances sous forme d’une base de connaissances logique ou encore sous forme d’un DAG représentant les relations d’influence entre les variables, associé à des distributions conditionnelles.
- Il est également possible de tirer profit des avantages offerts par les deux modes de représentation aussi bien au niveau de la représentation que de l’inférence.
- Elles peuvent être également vues comme un rapprochement entre la théorie des possibilités et la théorie des probabilités.

Principales contributions

Nos travaux se situent dans l’axe de la théorie des possibilités quantitative. Ils complètent le treillis des transformations réalisées dans la théorie des possibilités. De plus, ils effectuent un parallèle avec d’autres modes de représentation de connaissances et ils apportent un nouveau mode de raisonnement pour la logique possibiliste quantitative.

Du point de vue de la représentation des connaissances :

- nous avons développé dans [16] une procédure de translation plus directe d’un graphe causal possibiliste basé sur le produit vers une base possibiliste quantitative ainsi qu’une procédure de transformation d’une base possibiliste quantitative vers un graphe causal possibiliste basé sur le produit.
- puis nous nous sommes intéressé dans [17] à mettre en exergue les liens qu’entretient la théorie des possibilités quantitative avec d’autres modes de représentation.

Ainsi, à partir de l'équivalence sémantique entre les distributions générées par une base de pénalités et une base de possibilités quantitatives [18], nous avons établi des équivalences syntaxiques entre ces deux types de logiques pondérées, à savoir la logique des pénalités et la logique des possibilités quantitative.

- Nous nous sommes également intéressé dans [17], pour le mode graphique, aux relations qui existent entre les réseaux causaux possibilistes basés sur le produit et les réseaux causaux basés sur les Fonctions conditionnelles Ordinales.
- Nous avons complété, dans [19], ces différentes transformations en apportant des preuves formelles sur l'ensemble des phases de transformation et ce pour chaque procédure.

Du point de vue du raisonnement en mode graphique, nous avons exploité et adapté les procédures de propagation dans les réseaux causaux possibilistes développées dans [20], qui sont des adaptations des noyaux présentés dans [21] pour les réseaux causaux bayésiens.

Du point de vue du raisonnement en mode logique, nous avons développé un algorithme d'inférence sur une base de pénalités basé sur un prouveur des plus performants de type WMAXSAT. Cet algorithme exploite les relations syntaxiques, révélées dans ce travail, entre la logique des pénalités et la logique des possibilités quantitative.

Sur le plan expérimental, nous avons mené quelques expériences exploitant une partie des procédures développées. Ces résultats sont totalement cohérents avec cette étude et montrent que le choix d'un mode de représentation (logique ou graphique) dépend de la structure du graphe généré. Le message reflété par ces expériences est clair :

- Dans le cas où les connaissances des agents à formaliser induisent des graphes de complexité faible ou moyenne, il serait plus judicieux de basculer au mode graphique pour la représentation et le raisonnement.
- Dans le cas où les connaissances des agents à formaliser induisent des graphes complexes, il est conseillé alors d'opter pour le mode logique, car les algorithmes d'inférence deviennent plus performants au delà d'un certain seuil de complexité.

Lecture de la thèse

D'une manière plus détaillée, le manuscrit se compose de trois parties.

La première partie est consacrée à faire un tour d'horizon sur les principaux concepts et travaux réalisés dans le domaine étudié. Dans le premier chapitre nous aborderons les principales théories de l'incertain et de l'imprécision en commençant par le précurseur des théories de l'incertain qui est la théorie des probabilités. Nous présenterons par la suite la théorie de l'évidence qui sera suivie par la théorie des possibilités. La suite du chapitre sera consacrée à présenter les concepts de base de la théorie des ensembles flous pour modéliser les connaissances imprécises.

Le deuxième chapitre s'intéresse aux modes logiques de l'incertain et principalement à la logique possibiliste et à la logique des pénalités. Dans cette dernière, à chaque formule

est associée une pondération qui représente le prix à payer pour violer cette formule. Cette logique présente des caractéristiques qui seront utilisées dans l'élaboration des algorithmes d'inférence.

Dans le troisième chapitre, nous nous intéresserons aux modes de représentations graphiques où nous détaillerons le mode de représentation et de propagation dans les réseaux causaux bayésiens et dans les réseaux causaux possibilistes qualitatifs.

Dans le quatrième chapitre, nous nous pencherons sur les différents travaux qui ont été effectués sur les transformations entre les modes graphiques et les modes logiques en théorie des possibilités. Les deux parties suivantes s'étaleront sur notre contribution.

La seconde partie sera consacrée à présenter les transformations développées dans le cadre de cette thèse. D'une manière détaillée, dans le cinquième chapitre, nous présenterons de nouvelles transformations en théorie des possibilités quantitative. Nous commencerons par le passage immédiat d'un réseau causal possibiliste basé sur le produit vers une base de connaissances possibiliste. Dans le sixième chapitre nous allons nous intéresser à d'autres modes de représentations des connaissances, à savoir à la logique des pénalités [22, 23], ainsi qu'aux réseaux causaux basés sur les OCF (Ordinal Conditional Functions) [24]. Une fonction conditionnelle ordinale (OCF) ou fonction Kappa définie par Sphon [25, 26] permet de représenter les connaissances incertaines.

La troisième partie sera consacrée à présenter les algorithmes d'inférence ainsi que les expérimentations. Ainsi, le septième chapitre s'étalera sur les deux algorithmes de propagation dans les réseaux causaux possibilistes basés sur le produit présentés dans [20]. Dans le huitième chapitre, nous présenterons un nouvel algorithme d'inférence pour la logique des possibilités quantitative.

Pour finir, nous présenterons quelques résultats expérimentaux, en comparant les deux modes de raisonnement à savoir la propagation dans les réseaux causaux possibilistes basés sur le produit et l'inférence logique dans une base de pénalités, ceci en exploitant les différentes procédures de transformations présentées dans les deux derniers chapitres de la deuxième partie.

Première partie

État de l'art

Introduction de la partie I

Le développement de différents modes de représentation des connaissances a toujours été au cœur des recherches en intelligence artificielle. Vu que les connaissances manipulées sont souvent imparfaites, un certain nombre de théories modélisant les notions d'imperfection telles que l'incertitude et l'imprécision sont apparues.

Le premier chapitre donne un aperçu sur les principales théories de l'incertain et du flou. De ces théories ont découlé deux principaux modes de représentation des connaissances, le mode logique et le mode graphique. Les deux chapitres suivants donnent un aperçu sur ces deux types de modèles, présentant chacun des atouts majeurs.

La théorie des possibilités est une des théories de l'incertain, offrant la possibilité d'exploiter les deux modes de représentation des connaissances, soit sous forme de bases de connaissances logiques soit sous forme de graphes causaux possibilistes et ceci d'une manière numérique ou d'une manière ordinale selon la nature de connaissances à traiter. Vu cette richesse sur le plan de la représentation, plusieurs travaux se sont penchés sur les différentes passerelles qui existent entre ces deux modes de représentations (logique et graphique) dans les deux contextes (quantitatif et qualitatif). Ces transformations permettront ainsi, une grande liberté d'expression des connaissances aux experts.

A cet effet, le quatrième chapitre retrace l'ensemble des procédures de transformations développées, principalement, autour de la théorie des possibilités qualitative.

Chapitre 1

Théories de l'incertain et de l'imprécision

1.1 Introduction

Plusieurs modèles numériques ont été établis pour traiter ces différents types d'ignorance. La théorie des probabilités fût la première à s'être intéressée à la modélisation de l'incertitude. En effet, les incertitudes ont été identifiées depuis fort longtemps par l'intermédiaire de la théorie des probabilités [5]. Elle permet, par exemple, de quantifier la croyance pour qu'une proposition soit vraie. Cependant, elle est assez limitée pour exprimer certaines formes d'incertitudes [3].

Par la suite, les travaux de Dempster et Shafer sur la théorie des fonctions de croyance sont apparus afin de modéliser la notion de crédibilité issue des connaissances du type "Mon degré de croyance que le cancer X est dû à un virus". Vu que les connaissances incertaines peuvent ne pas être de nature probabiliste (numérique), la théorie des possibilités fût introduite afin de représenter des possibilités épistémiques du type "la possibilité que la hauteur de Paul soit autour de 175 cm".

Le concept de l'imprécision fût modélisé par les ensembles flous [27].

Nous allons nous focaliser, dans ce chapitre, sur ces différentes approches numériques proposées pour représenter les connaissances incertaines et imprécises. La section 3 présente les concepts de base de la théorie des probabilités. Puis nous exposerons les théories non probabilistes. Ainsi, la section 4 s'intéresse aux fondements de base de la théorie des fonctions de croyance. Quant à la section 5, elle est dédiée à la théorie des possibilités. La section 6 relate les principaux fondements de la théorie des ensembles flous qui traitent les connaissances imprécises.

1.2 Quelques notations

Nous allons dans cette section présenter les conventions de notations utilisées dans cette thèse.

Soit \mathcal{L} le langage propositionnel. Ω représente l'ensemble des interprétations propo-

sitionnelles. ω représente une interprétation de Ω . ϕ, ψ, \dots représentent des formules propositionnelles. A, B, C, ... représentent des variables. \top représente une tautologie, \perp représente une antilogie et \neg représente la négation, \wedge représente la conjonction et \vee représente la disjonction.

Soit $\omega \in \Omega$ une interprétation et soit ϕ une formule propositionnelle. $\omega \models \phi$ signifie que l'interprétation ω est un modèle de ϕ dans le sens de la logique propositionnelle. $\omega \not\models \phi$ signifie que l'interprétation ω n'est pas un modèle de ϕ .

1.3 La théorie des probabilités

1.3.1 Introduction

De toutes les théories qui représentent l'incertitude, la théorie des probabilités [10] est la plus ancienne et la plus répandue. Elle constitue une base solide pour le raisonnement logique avec incertitude car elle permet de modéliser quantitativement l'incertitude associée à un événement.

Considérons par exemple l'événement N "*il neigera sur les hauts plateaux d'ici la fin de semaine*". L'occurrence de N est difficile à prédire. Cette incertitude est quantifiée par un nombre $P(N)$, appelé la probabilité associée à N. Cette probabilité est non négative et ne doit pas excéder 1.

1.3.2 Principes de base

Les bases de la formalisation axiomatique de la théorie des probabilités sont les suivantes :

1.3.2.1 Mesure de probabilité

La notion de base de la théorie des probabilités est la mesure de probabilités définie dans un espace de probabilité.

Définition 1.1 (Espace de probabilités) (Kolmogoroff, 1933) *Un espace de probabilités formel est défini par trois entités (S, F, P) où :*

1. S est l'espace de réalisation,
2. F est l'ensemble des événements en S ,
3. P est une mesure de probabilité définie sur les éléments de F .

Sur l'espace de probabilités S , sont définies des opérations algébriques entre événements :

- \bar{A} représente l'événement complémentaire de A,
- $A \cup B$ représente l'union de deux événements A et B,
- $A \cap B$ représente l'intersection de deux événements A et B,
- $A \setminus B$ représente la différence entre deux événements A et B.

Considérons deux événements A et B. Deux cas particuliers sont à noter :

- A et B sont dits exhaustifs si $A \cup B = S$.

– A et B sont dits exclusifs si $A \cap B = \emptyset$ où \emptyset est l'événement impossible.
La mesure de probabilités P satisfait les trois axiomes de base suivants :

1. $P(\emptyset)=0$;
2. $P(S)=1$;
3. Si A_1, A_2, \dots, A_n sont des événements exclusifs ($A_i \cap A_j = \emptyset \forall i \neq j$), alors :

$$P(\cup_{h=1}^n A_h) = \sum_{h=1}^n P(A_h).$$

Le troisième axiome est l'axiome d'additivité propre au modèle probabiliste dans le cas où les A_i sont mutuellement exclusifs. Il atteste que la croyance d'un événement est la somme des croyances assignées à ces composants exclusifs.

A partir de ces axiomes sur la fonction de probabilité P, les propriétés élémentaires suivantes sont dérivées :

- Propriété 1.1** 1. $\forall A \in S, 0 \leq P(A) \leq 1$. Les deux valeurs extrêmes sont interprétées comme suit : $P(A)=0$ signifie que l'événement A est impossible et $P(A)=1$ signifie que A est certain.
2. $\forall A \in S, P(A)+P(\bar{A})=1$. C'est une équation qui permet de déduire la probabilité d'un événement dès que la probabilité de l'événement contraire est connue. Ceci implique que quand nous augmentons la croyance d'un événement, nous diminuons la croyance de l'événement contraire.
3. Soient A et B deux événements, la réalisation de A et B quand A et B sont impossibles est un événement impossible :

$$A \cap B = \emptyset \Rightarrow P(A, B) = 0$$

4. Soient A et B deux événements.
Si $A \subseteq B$ Alors $P(A) \leq P(B)$.
Si $A \subset B$ alors $P(B \setminus A) = P(B) - P(A)$.
5. Soient A et B deux événements.

$$P(A \cup B) + P(A, B) = P(A) + P(B) \Leftrightarrow P(A \cup B) = P(A) + P(B) - P(A, B)$$

Lorsque A et B sont exclusifs auquel cas, $P(A \cap B) = 0$, alors, $P(A \cup B) = P(A) + P(B)$.
Plus généralement, si tous les A_i sont incompatibles deux à deux alors :

$$P(\cup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$$

1.3.2.2 Conditionnement probabiliste

La notion du conditionnement est inhérente à la théorie des probabilités. Elle permet de modéliser les liens de dépendance entre deux événements. Plus précisément, elle évalue l'impact de l'arrivée d'une information certaine sur les connaissances du monde.

Exemple 1.1 Soit l'expérience de jeter un dé. Soit X un jet, F l'événement $\{X=6\}$ et E l'événement $\{X>4\}$. Ainsi, $P(F)=1/6$. Supposons qu'un dé a été jeté et que l'événement E s'est produit générant ainsi deux possibilités : (5 ou 6). Intuitivement, en l'absence d'autres informations : $P(F | E)=1/2$.

$P(F | E)$ est appelée probabilité conditionnelle de F sachant E .

D'une manière formelle, le conditionnement est donc matérialisé par la définition suivante :

Définition 1.2 (Probabilité Conditionnelle) Soient A et B deux événements avec $P(B)>0$. La probabilité conditionnelle de A sachant B est définie par :

$$P(A | B) = \frac{P(A, B)}{P(B)} \quad (1.1)$$

$P(B)$ est dite probabilité marginale.

Exemple 1.2 Si on reprend l'exemple précédent du dé avec F l'événement $\{X=6\}$ et E l'événement $\{X>4\}$. L'événement $E \cap F$ est l'événement F d'où :

$$P(F | E) = \frac{P(F \cap E)}{P(E)} = \frac{P(F)}{P(E)} = \frac{1/6}{1/3} = 1/2. \text{ Ce qui correspond au résultat obtenu précédemment.}$$

Propriété 1.2 1. Si $A \supset B \Rightarrow P(A | B)=1$ (car $A \cap B=B$).

2. Si $A \subset B \Rightarrow P(A | B) \geq P(A)$ (car $A \cap B=A$).

3. Les probabilités conditionnelles satisfont les axiomes de la mesure de probabilité.

A partir de la définition de la probabilité conditionnelle, est déduite la probabilité de l'intersection de deux événements, appelée probabilité jointe, en termes de probabilité conditionnelle et de probabilité marginale comme suit :

$$P(A \cap B) = P(A | B) * P(B) = P(B | A) * P(A) \quad (1.2)$$

Cette règle peut être appliquée à plusieurs d'événements produisant ainsi la règle de multiplication ou la règle de chaînage.

Définition 1.3 (Règle de chaînage) Soient A_1, A_2, \dots, A_n n événements, la probabilité jointe associée aux événements (A_1, A_2, \dots, A_n) est définie comme le produit des probabilités conditionnelles comme suit :

$$P(A_1, A_2, \dots, A_n) = P(A_1 | A_2, \dots, A_{n-1}, A_n) * \dots * P(A_{n-1} | A_n) * P(A_n) \quad (1.3)$$

A partir de l'équation des probabilités jointes et de l'axiome 3 de la mesure des probabilités, la notion de probabilité totale est définie par : Soit $\{A_i\}$ une partition de S , c'est à dire :

- $A_i \cap A_j = \emptyset, i \neq j$
- $S = \bigcup_{i=1, \dots, N} A_i$

Alors, pour tout événement B de S ,

$$P(B) = \sum_{i=1}^N P(B, A_i) = \sum_{i=1}^N P(B | A_i) * P(A_i)$$

Cette définition exprime la probabilité marginale d'un événement B comme une moyenne pondérée des probabilités conditionnelles.

La règle de chaînage et le théorème de la probabilité totale peuvent être étendus aux probabilités conditionnelles :

$$P(A_1, A_2, \dots, A_n | C) = P(A_n | A_{n-1}, A_{n-2}, \dots, A_1, C) * \dots * P(A_2 | A_1, C) * P(A_1 | C) \quad (1.4)$$

$$P(B | C) = \sum_{i=1}^N P(B | A_i, C) * P(A_i, C) \quad (1.5)$$

1.3.2.3 Théorème de Bayes

Le théorème de Bayes est une formulation mathématique du conditionnement modélisant la modification des croyances après l'arrivée d'une nouvelle information liant les probabilités conditionnelles de deux événements A et B . Il stipule :

$$P(A | B) = \frac{P(B | A) * P(A)}{P(B)} \quad (1.6)$$

avec $P(A) \neq 0$ et $P(B) \neq 0$.

Le théorème de Bayes peut être vu comme une règle qui permet de mettre à jour une probabilité initiale $P(A)$ appelée aussi probabilité a priori donnant ainsi lieu à une probabilité révisée $P(A | B)$, appelée probabilité a posteriori.

La règle de Bayes joue un rôle important pour toutes sortes d'applications comme pour le diagnostic médical, car elle permet de calculer la probabilité d'un diagnostic à partir de la probabilité de la cause :

$$P(\text{cause} | \text{effet}) = \frac{P(\text{effet} | \text{cause})P(\text{cause})}{P(\text{effet})} \quad (1.7)$$

Cependant, la difficulté d'un problème augmente infiniment lorsque les variables ne sont pas indépendantes. Ainsi, l'utilisation du théorème de Bayes, nécessite la prise en compte des indépendances conditionnelles.

1.3.3 Indépendance et Indépendance conditionnelle

Deux événements A et B sont indépendants si leurs probabilités ne changent pas quand nous les conditionnons l'un à l'autre : $P(A | B) = P(A)$ et $P(B | A) = P(B)$ avec $P(A) > 0$ et $P(B) > 0$. Cette indépendance peut être également formalisée par :

Définition 1.4 (Indépendance) Deux événements A et B sont dits indépendants, avec $P(A) > 0$ et $P(B) > 0$, si $P(A, B) = P(A) * P(B)$

En d'autres termes, la probabilité jointe de ces deux événements correspond au produit des probabilités marginales. Si A et B sont indépendants, alors (A et \bar{B}), (\bar{A} et B) et (\bar{A} et \bar{B}) le sont également.

La notion d'indépendance peut être étendue à un ensemble fini d'événements A_1, A_2, \dots, A_n .

Définition 1.5 Un ensemble d'événements A_1, A_2, \dots, A_n est dit indépendant si pour chaque sous ensemble A_i, A_j, \dots, A_m de ces événements, nous avons :

$$P(A_i, A_j, \dots, A_m) = P(A_i) * P(A_j) * \dots * P(A_m)$$

Remarque 1.1 Il est important de remarquer que la notion d'indépendance ne coïncide pas avec celle d'événements mutuellement exclusifs. En effet, si A et B sont mutuellement exclusifs, alors leur intersection est vide et donc $P(A, B) = 0$, ce qui entraîne $P(A|B) = 0$. Ainsi, si B se produit, l'occurrence de A est impossible.

Définition 1.6 (Indépendance conditionnelle) Deux événements A et B, avec $P(A) > 0$ et $P(B) > 0$ sont dits indépendants dans le contexte C si :

$$P(A, B | C) = P(A | C) * P(B | C)$$

Il est important de remarquer que l'indépendance conditionnelle n'implique pas l'indépendance. En effet, deux événements A et B peuvent être indépendants dans le contexte C mais si nous les considérons à part, ils ne seront pas indépendants.

Remarque 1.2 Il est important à ce niveau de souligner l'impact de la notion d'indépendance sur la complexité du raisonnement probabiliste. En effet, le calcul des probabilités requiert :

- Soit la connaissance d'un nombre important de probabilités conditionnelles du type $P(A|B)$,
- Soit l'exploitation des indépendances pour un certain nombre de couples d'événements (A, B).

Ainsi, l'utilisation de l'indépendance probabiliste entre certains événements permet de réduire notablement la complexité des calculs pour le raisonnement probabiliste dans l'incertain [28]. En effet, la difficulté d'un problème augmente considérablement lorsque les variables ne sont pas indépendantes.

1.3.4 Interprétations des probabilités

Le modèle probabiliste peut être interprété de plusieurs manières. Essentiellement, nous parlons des probabilités fréquentistes ou des probabilités subjectives. Selon l'interprétation fréquentiste, les probabilités sont obtenues à partir de la notion de "fréquence relative d'apparition lors de la répétition d'un grand nombre de fois d'une expérience aléatoire".

Exemple 1.3 Soit la situation où il s'agit de tirer une boule au hasard dans un ensemble de n boules formé de r boules rouges et de j boules jaunes. La probabilité que la proposition $A =$ "La boule tirée est rouge" correspond à la fréquence relative avec laquelle nous nous attendons à tirer une boule rouge. $P(A) = r/n$.

L'interprétation fréquentiste ne permet pas de raisonner sur l'incertain si les données statistiques disponibles ne sont pas suffisantes, ou si le cadre d'une expérience aléatoire répétée dans des conditions identiques n'est pas respecté [28].

L'approche subjective peut être définie par le slogan "La probabilité est un degré de croyance". En effet, il y a des situations pour lesquelles il n'y a pas d'interprétations fréquentistes objectives comme par exemple la probabilité d'avoir une bonne note à un examen.

1.3.5 conclusion

Le modèle probabiliste reste le modèle le plus utilisé. Néanmoins, il présente des limites :

- La théorie des probabilités ne permet pas de distinguer entre l'ignorance et l'incertitude, qui cependant, dans le sens commun, ne recèlent pas la même quantité de savoir. En effet, soient A et B deux événements. Affirmer que $P(A) = P(B) = 0.5$ peut désigner l'équiprobabilité entre les deux événements ou encore la totale ignorance sur les événements A et B . En effet, l'axiome d'additivité de la théorie des probabilités entraîne que si nous ne savons rien sur un événement A alors

$$P(A) = P(\neg A) = 1/2.$$

Le problème majeur qui en découle repose sur le fait que d'augmenter les croyances sur un événement entraîne la diminution sur la croyance de l'événement contraire, ce qui n'est pas toujours souhaitable.

- De plus, il est difficile en théorie des probabilités de représenter l'ignorance totale lorsqu'on se trouve en présence de plus de deux alternatives. En effet, considérons le paradoxe suivant [29].

Exemple 1.4 Soient les propositions $P_1 =$ "La vie existe sur Mars", $P_2 = \neg P_1 \wedge P_3 =$ "La vie sur Mars est seulement végétale", $P_4 =$ "La vie sur Mars n'est pas seulement végétale"; Il est clair que $P_3 \wedge P_4 = \emptyset$ et que $P_1 = \neg P_3 \vee P_4$. Quelle que soit la répartition des probabilités (par exemple $P_2 = 1/2$, $P_3 = P_4 = 1/4$ ou $P_2 = P_3 = P_4 = 1/3$) utilisée pour rendre compte de notre ignorance a priori, certaines propositions apparaissent plus probables que d'autres, ce qui est paradoxal vu qu'on est dans le cas d'une ignorance totale.

Ces problèmes découlent de l'axiome d'additivité.

- Un autre inconvénient du calcul des probabilités réside dans la nécessité d'énumérer un ensemble exhaustif d'alternatives mutuellement exclusives [30]. En pratique, l'ensemble des faits fournis par un expert ne répondent pas forcément à ces exigences.
- L'utilisation du calcul des probabilités conduit souvent, en pratique, à poser des axiomes d'indépendance entre les faits manipulés [30].
- Du point de vue du calcul, les règles de calcul de la théorie des probabilités favorisent l'augmentation des erreurs (il y a des sommes et des produits). Si bien qu'une longue chaîne de propagation des coefficients d'incertitude peut réduire à néant la signification de ces derniers [30].

Pour tout cela, des travaux de recherche ont conduit à d'autres formalismes pour représenter l'incertain.

1.4 La théorie des fonctions de croyance

Pour pallier aux problèmes posés par le modèle probabiliste, notamment ceux liés à l'axiome d'additivité, Dempster et Shafer [6] ont défini la notion de fonctions de croyances qui ont pour objectif de modéliser et de quantifier la notion de degrés de croyance pour représenter l'incertitude. Dans la théorie des croyances [31, 32, 33], les distributions de probabilités sont construites non pas sur Ω mais sur l'ensemble des parties de Ω . Ω est dit cadre de discernement représentant l'ensemble exhaustif et exclusif des états connus du monde. 2^Ω est l'ensemble des parties de Ω . La distribution représente la masse de probabilités m définie par :

$$m : 2^\Omega \rightarrow [0,1]$$

Les propriétés de la fonction de masse sont :

$$\sum_{A \subseteq \Omega} m(A) = 1 \quad (1.8)$$

$$m(\emptyset) = 0. \quad (1.9)$$

L'égalité (1.8) représente la condition de normalisation.

Tout sous-ensemble de Ω vérifiant $m(A) \neq 0$ est dit élément focal de m .

Le couple (Ω, m) est appelé corps de la croyance.

Si $m(A)=1$ et $\forall B \neq A \in 2^\Omega, m(B)=0$, alors A est certain dans le sens où un de ces éléments a la valeur cherchée. Dans ce cas nous ne savons pas quel est l'élément concerné sauf si A est un singleton.

Si $A=\Omega$, alors l'information $m(A)=1$ correspond à une situation d'ignorance totale.

Exemple 1.5 [34] *Imaginons que nous nous intéressons à l'évolution sur 24 heures du cours d'une action en bourse. Ainsi, au bout de 24 heures, soit le cours aura changé (événement a_1), soit il restera stable (événement a_2). Il faut attribuer une masse à*

chacun des sous-ensembles possibles : $m(\emptyset), m(\{a_1\}), m(\{a_2\}), m(\{a_1, a_2\})$. Supposons que selon une analyse financière, le cours restera stable, mais sans grande certitude. Nous en déduisons que $m(\{a_2\})=0.3$. D'un autre côté, le fait que le marché boursier est très volatil, permet de dire que le cours aura tendance à fluctuer. Nous en déduisons que $m(\{a_1\})=0.2$. La masse restante sera attribuée à $\{a_1, a_2\}$ qui se traduit par "Nous ne pouvons pas nous prononcer". Il s'agit de la masse restante qui n'a pas encore été attribuée. En résumé, notre crédibilité sera partagée entre une partie de connaissance et une partie d'ignorance. Nous aurons :

- $m(\emptyset)=0$ par définition,
- $m(\{a_1\})=0.2$,
- $m(\{a_2\})=0.3$,
- $m(\{a_1, a_2\})=0.5$.

L'élément focal correspond à un événement en lequel croit un observateur au minimum un peu. L'ensemble des éléments focaux constitue le noyau : $N_\Omega = \{A \in 2^\Omega : m(A) \geq 0\}$.

1.4.1 Fonction de croyance ou de crédibilité

A la fonction de masse m est associée la notion de fonction de croyance Bel définie comme suit : la fonction de croyance évalue l'intensité de l'implication d'une formule à partir des croyances disponibles. Ainsi, si A est un ensemble dans un cadre de discernement Ω , $Bel(A)$ représente la somme de toutes les masses attribuées aux éléments contenus dans A , plus la masse attribuée à A lui-même. D'une manière plus formelle, la fonction de croyance est définie par : $Bel : 2^\Omega \rightarrow [0,1]$ telle que :

$$Bel(A) = \sum_{B \subseteq A} m(B) \quad (1.10)$$

La fonction de crédibilité vérifient les propriétés suivantes :

- $Bel(\emptyset)=0$
- $Bel(\Omega)=1$
- Règle de sous-additivité :

$$Bel(A) + Bel(\bar{A}) \leq 1$$

Cette propriété permet d'augmenter la croyance d'un événement sans pour autant diminuer celle de l'événement contraire. Elle offre une souplesse qui fait défaut à la théorie des probabilités à cause de l'axiome d'additivité qui est contraignant.

$$Bel(A \cup B) \geq Bel(A) + Bel(B) - Bel(A \cap B).$$

Concernant cette propriété, la distinction est importante. En effet, pour la théorie des probabilités, le degré de probabilité accordé à l'union de deux événements mutuellement exclusifs est égal à la somme des probabilités allouées à chaque événement.

– Règle de monotonie :

$$\text{Si } A \subseteq B \Rightarrow \text{Bel}(A) \leq \text{Bel}(B).$$

Exemple 1.6 Si nous reprenons l'exemple 1.5, nous avons les masses suivantes : $m(\emptyset)=0$, $m(\{a_1\})=0.2$, $m(\{a_2\})=0.3$, $m(\{a_1, a_2\})=0.5$. Les degrés de crédibilité qui en résultent sont :

- $\text{Bel}(\emptyset)=0$,
- $\text{Bel}(\{a_1\})=0.2$,
- $\text{Bel}(\{a_2\})=0.3$,
- $\text{Bel}(\{a_1, a_2\})=1$ car la somme des masses de tous les sous-ensembles du cadre de discernement est toujours égale à 1 ($\Omega=\{a_1, a_2\}$).

La fonction de crédibilité est également considérée comme un élément focal. L'ensemble de tous les éléments focaux d'une fonction de crédibilité s'appelle le cœur de la fonction de crédibilité.

1.4.2 Fonction de plausibilité

La fonction de plausibilité évalue dans quelle mesure un événement est consistant avec les croyances disponibles. Elle est définie par : $Pl : 2^\Omega \rightarrow [0,1]$

$$Pl(A) = \sum_{A \cap B \neq \emptyset} m(B) \quad (1.11)$$

La plausibilité d'un ensemble d'éléments A dans un cadre de discernement Ω est la somme des croyances qui ne contredisent pas A.

La fonction de plausibilité peut également être définie à travers la relation suivante :

$$Pl(A) = 1 - \text{Bel}(\bar{A})$$

En effet, le degré de plausibilité de A représente tout le degré de crédibilité qui n'est pas attribué à \bar{A} . En d'autres termes, A n'est quelque peu crédible que dans la mesure où l'événement opposé n'est pas très plausible. La fonction de plausibilité satisfait les relations suivantes :

- $Pl(\emptyset)=0$
- $Pl(\Omega)=1$
- $Pl(A)=1-\text{Bel}(\bar{A})$
- $Pl(A)+Pl(\neg A) \geq 1$
- Règle de monotonie : Si $A \subseteq B$ Alors $Pl(A) \leq Pl(B)$
- $Pl(A \cap B) \leq Pl(A)+Pl(B)-Pl(A \cup B)$

Exemple 1.7 Si nous reprenons l'exemple 1.5, nous avons les masses suivantes : $m(\emptyset)=0$, $m(\{a_1\})=0.2$, $m(\{a_2\})=0.3$, $m(\{a_1, a_2\})=0.5$. Les degrés de plausibilité qui en résultent sont :

- $Pl(\emptyset)=0$,

- $Pl(a_1) = m(\{a_1\}) + m(\{a_1, a_2\}) = 0.7$,
- $Pl(a_2) = m(\{a_2\}) + m(\{a_1, a_2\}) = 0.8$,
- $Pl(\{a_1, a_2\}) = Pl(\Omega) = 1 - Bel(\emptyset) = 1$ en utilisant la relation entre la fonction de crédibilité et la fonction de plausibilité.

La fonction de crédibilité et la fonction de plausibilité sont reliées par plusieurs points :

- Si nous comparons la fonction de crédibilité avec la fonction de plausibilité, il apparaît que la fonction de plausibilité domine la fonction de crédibilité. Ceci est illustré par l'équation suivante :

$$\forall A \subset B, Bel(A) \leq Pl(A)$$

- $\forall A \subseteq \Omega, Bel(A) + Bel(\bar{A}) \leq 1$ et $Pl(A) + Pl(\bar{A}) \geq 1$.

1.4.3 La règle de combinaison de Dempster

La règle de combinaison de Dempster permet de modifier les croyances après l'arrivée d'une nouvelle croyance. Cette règle permet de calculer, à partir de plusieurs fonctions de crédibilité sur le même ensemble de discernement Ω , une nouvelle fonction de crédibilité dite somme orthogonale en combinant les croyances.

D'une manière formelle, soient m_1, m_2 deux distributions de masses définies sur un référentiel Ω . Soient A_1, A_2, \dots, A_n les éléments focaux de m_1 et B_1, B_2, \dots, B_l les éléments focaux de m_2 . La distribution de masse $m = m_1 \oplus m_2$ est définie par :

$$m(\emptyset) = 0$$

$$\forall A \neq \emptyset, m(A) = \frac{1}{k} \sum_{A_i \cap B_j = A} m_1(A_i) * m_2(B_j) \quad (1.12)$$

où

$$k = 1 - \sum_{A_i \cap B_j = \emptyset} m_1(A_i) * m_2(B_j)$$

k est une constante de normalisation et la quantité :

$$\sum_{A_i \cap B_j = \emptyset} m_1(A_i) * m_2(B_j)$$

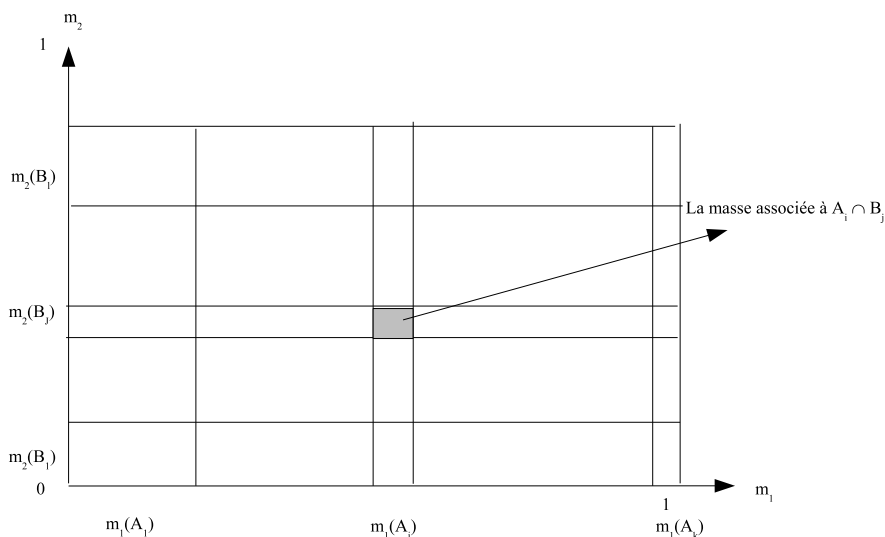
mesure le degré de conflits entre les deux distributions de masse m_1 et m_2 .

La figure 1.1 [6] illustre la combinaison orthogonale de deux masses m_1 et m_2 .

Si $\sum_{A_i \cap B_j = \emptyset} m_1(A_i) * m_2(B_j) = 1$ alors les deux distributions de masses sont contradictoires et la combinaison est impossible.

La masse associée à un sous-ensemble P de Ω est donnée par :

$$m(P) = \sum_{A_i \cap B_j = P} m_1(A_i) * m_2(B_j)$$

FIG. 1.1 – Somme orthogonale entre m_1 et m_2

1.4.4 Le conditionnement

Différentes définitions ont été développées pour le conditionnement dans la théorie des fonctions de croyance. Parmi lesquelles nous trouvons la règle de Dempster :

$$Pl(A | B) = \frac{Pl(A)}{Pl(B)} \quad (1.13)$$

1.4.5 Les fonctions de croyances et les probabilités

Nous pouvons faire le parallèle entre fonction de croyance et probabilités en plusieurs points :

- Si tous les éléments focaux associés à une croyance sont des singletons alors la mesure de probabilité est à la fois une mesure de croyance et une mesure de plausibilité.
- L'axiome d'additivité des probabilités est remplacé par l'inégalité suivante :

$$Bel(A \cup B) \geq Bel(A) + Bel(B) - Bel(A \cap B)$$

- La théorie des fonctions de croyance permet de distinguer l'ignorance de l'incertitude, contrairement à la théorie des probabilités. En effet, pour la théorie des fonctions de croyance, l'ignorance est représentée par le fait que deux événements peuvent être tous les deux plausibles sans être pour autant crédibles. Plus formellement, nous pouvons avoir :

$$Pl(A) = Pl(\neg A) = 1$$

et

$$Bel(A) = Bel(\neg A) = 0$$

Ainsi, la théorie des fonctions de croyance permet de distinguer entre l'ignorance et l'incertitude, contrairement à la théorie des probabilités.

- La théorie des probabilités est, sous certaines conditions, un cas particulier de la théorie des fonctions de croyance. En effet, les mesures de croyance et de plausibilité peuvent être vues respectivement comme étant la borne inférieure et la borne supérieure de la probabilité [35] :

$$\forall A \in \Omega, Bel(A) \leq P(A) \leq Pl(A)$$

Ainsi, l'intervalle $[Bel(A), Pl(A)]$ est interprété comme un encadrement de la probabilité réelle de A [31].

- Dans le cas où la distribution de masse m est définie sur les singletons du référentiel Ω et non pas sur les sous ensembles de Ω , elle devient équivalente à une distribution de probabilités. Ceci induit une égalité entre la mesure de probabilités, la mesure de crédibilité et la mesure de plausibilité. Ainsi,

$$\forall A \in \Omega, Bel(\{A\}) = Pl(\{A\}) = P(A) \quad (1.14)$$

L'équation de la théorie de Dempster-Shafer

$$Pl(A) = 1 - Bel(\neg A)$$

devient alors

$$P(A) + P(\neg A) = 1$$

C'est la propriété caractéristique de la théorie des probabilités.

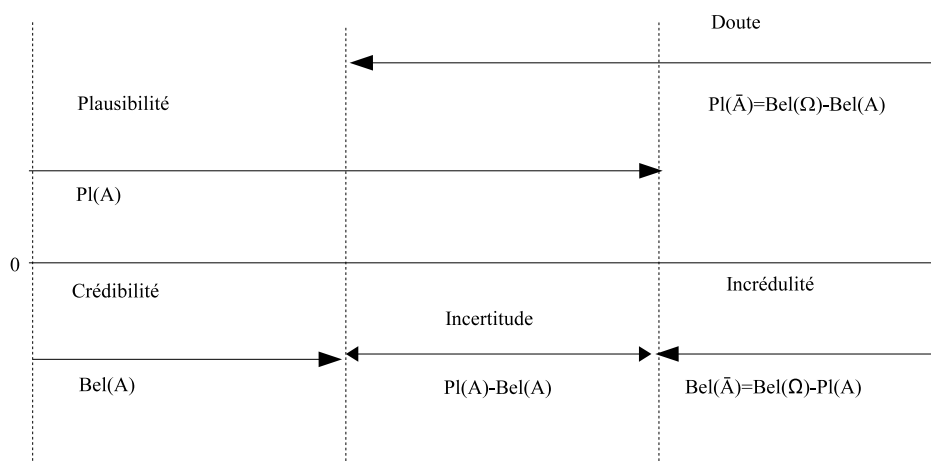
Les différentes relations entre les mesures relatives à la théorie des croyances peuvent être illustrées sous forme graphique appelée intervalle de croyance (voir la figure 1.2). Nous pouvons souligner l'existence des fonctions symétriques de la crédibilité et de la plausibilité à savoir l'incrédibilité (crédibilité de la négation) et le doute (la plausibilité de la négation).

1.4.6 Conclusion sur les fonctions de croyances

Les fonctions de croyances permettent la modélisation des connaissances incertaines incluant le caractère de "non monotonie" du raisonnement où une connaissance peut être revue après l'avènement d'une croyance grâce à la règle de conditionnement. Cette théorie fût populaire [36] et le modèle de base a été étendu dans de nombreuses directions [37, 38]. Entre autre, dans [39], l'auteur a développé une théorie de l'incertain non probabiliste appelé le modèle de croyance transférable qui est une interprétation particulière de la théorie des fonctions de croyance où la représentation s'effectue sur la base des informations partielles. Dans cette approche, le concept d'additivité est remplacé par l'inégalité suivante :

$$Bel(A \cup B) \geq Bel(A) + Bel(B) - Bel(A \cap B)$$

Cependant, la théorie des fonctions de croyances ne permet pas de prendre en compte les informations contradictoires.

FIG. 1.2 – Intervalle de croyance d'un élément focal A

1.5 La Théorie des possibilités

Une manière d'évaluer le degré de confiance attribué à un événement incertain était de déterminer le degré de certitude de l'événement contraire comme pour la théorie des probabilités. Une autre approche consiste à évaluer l'incertitude accordée à un événement par deux mesures distinctes, comme le cas de la théorie des possibilités.

La théorie des possibilités proposée par Zadeh [8] puis développée par Dubois et Prade [9] est une théorie de l'incertain, ayant pour vocation de manipuler des connaissances incomplètes [40] décrites en termes d'ensembles flous. Elle diffère de la théorie des probabilités vu qu'elle manipule deux mesures duales (mesure de possibilités et mesure de nécessité). De plus, c'est une théorie non additive permettant de prendre en compte la notion de structures ordinales. Cette théorie a été développée dans deux directions : qualitative et quantitative. Ces deux approches partagent l'axiome de *maxitivité* mais diffèrent dans leurs définitions des concepts de conditionnement et d'indépendance.

Les mesures de possibilités qualitatives permettent d'évaluer les connaissances en utilisant la notion d'ordre. La logique possibiliste est issue de la théorie des possibilités. C'est une extension de la logique classique offrant ainsi un contexte sémantique pour le raisonnement non monotone dont la complexité reste proche à celle de la logique classique. Les échelles ne sont pas numériques mais ordinales : les connaissances sont organisées en niveaux (strates) selon leurs degrés de certitude. Ainsi, les formules de même degré occupent la même strate.

La théorie des possibilités quantitative possède diverses branches : le degré de possibilité peut être vu comme la borne supérieure d'une probabilité. Une distribution de possibilités peut être considérée comme une fonction *likelihood* et peut codifier une distribution de probabilités avec des valeurs extrêmes [40].

La théorie des possibilités repose sur les concepts décrits dans les sous-sections suivantes :

1.5.1 Distribution de possibilités

La théorie des possibilités est basée sur la notion de distribution de possibilités qui représente une fonction de Ω vers $[0, 1]$. Une distribution de possibilités π représente un état sur les connaissances permettant de distinguer ce qui est plausible à partir des connaissances les moins plausibles.

La distribution de possibilités présente les propriétés suivantes :

- $\pi(\omega) > 0$: ω est quelque peu possible,
- $\pi(\omega) = 0$: ω est impossible,
- $\pi(\omega) = 1$: ω est totalement possible,
- $\pi(\omega) > \pi(\omega')$: ω est plus plausible que ω' .

π est dite normalisée s'il existe une interprétation ω telle que $\pi(\omega)=1$. La condition de normalisation de la distribution de possibilités reflète la cohérence des connaissances représentées par π . Dans ce document, les distributions de possibilités considérées sont normalisées.

Soient π et π' deux distributions de possibilités. Si $\forall \omega, \pi(\omega) < \pi'(\omega)$, π est dite plus spécifique que π' ou que π' est plus générale ou moins informative.

Si $\forall \omega \neq \omega_0, \pi(\omega)=0$ et $\pi(\omega_0) = 1$ (toutes les interprétations de Ω sont impossibles sauf une), alors la connaissance du monde est complète.

Ces distributions sont les plus spécifiques.

1.5.2 Les mesures de possibilités et de nécessité

La théorie des possibilités a été définie afin de représenter des connaissances incomplètes, imprécises et incertaines. Comme pour la théorie des probabilités, l'évaluation des degrés de confiance d'un événement est numérique. Le degré de possibilité est défini par :

$$\Pi(\phi) = \max\{\pi(\omega) : \omega \models \phi\} \quad (1.15)$$

$\Pi(\phi)$ permet d'évaluer la cohérence de ϕ avec les connaissances disponibles représentées par π .

Exemple 1.8 *Dans le domaine des prévisions météorologiques, si le mois en cours est Août, alors il est totalement possible que le ciel soit dégagé vu qu'il n'y a pas d'informations qui contredisent ce fait (exemple prévisions d'orages).*

Le degré de possibilité Π satisfait les propriétés suivantes :

- $\Pi(\perp)=0$,
- $\Pi(\top)=1$,
- $\Pi(\phi)=0$: ϕ est impossible,
- $\Pi(\phi)=1$: ϕ est complètement possible,
- $\Pi(\phi)=\Pi(\neg \phi)=1$: ϕ et $\neg \phi$ sont totalement possibles (cas de l'ignorance totale),
- $\max(\Pi(\phi), \Pi(\neg \phi))=1$: ϕ ou $\neg \phi$ doit être possible (critère de normalisation),
- $\forall \phi, \forall \psi \Pi(\phi \vee \psi)=\max(\Pi(\phi), \Pi(\psi))$.

Les relations qui lient $\Pi(\phi)$ et $\Pi(\neg \phi)$ induisent la prise en compte de ces deux quantificateurs pour la description de la croyance sur l'occurrence d'un événement. $\Pi(\neg \phi)$ informe sur la possibilité de $\neg \phi$, donc sur la certitude ou la nécessité de l'occurrence de $\neg \phi$ vu que si $\neg \phi$ est impossible alors ϕ est certaine. Cette dualité permet de définir le degré de nécessité comme suit :

$$N(\phi) = 1 - \Pi(\neg \phi) = \min\{1 - \pi(\omega), \omega \models \neg \phi\}$$

Cette relation duale entre Π et N exprime le fait que plus ϕ est certaine moins $\neg \phi$ devient cohérente avec les croyances disponibles.

Par exemple si un étudiant veut faire de la recherche en intelligence artificielle, il est certain qu'il a des bases solides en logique.

Le degré de nécessité présente les propriétés suivantes :

- $N(\phi)=0$: ϕ n'est pas certaine, néanmoins ϕ peut être possible,
- $N(\phi)=1$: ϕ est certainement vraie,
- $N(\phi) > 0 \Rightarrow N(\neg \phi)=0$: si ϕ est quelque peu certaine alors son contraire n'est pas du tout certain,
- $\Pi(\phi) < 1 \Rightarrow N(\phi)=0$: si ϕ n'est pas complètement possible alors ϕ est complètement incertaine,
- $\forall \phi, \forall \psi N(\phi \wedge \psi) = \min(N(\phi), N(\psi))$.

Remarque 1.3 *Le domaine de ces différentes mesures est fini.*

1.5.3 Conditionnement possibiliste

Le conditionnement est un concept fondamental pour la mise à jour des pondérations initiales attribuées aux connaissances représentées par la distribution des possibilités après l'arrivée d'une information certaine, notée ϕ .

Soit π la distribution de possibilités initiales. Vu que la théorie des possibilités offre deux directions (qualitative et quantitative), le conditionnement possibiliste admet alors deux définitions selon le contexte ordinal ou numérique (pour plus de détails, voir [41, 42, 43]).

- Dans le contexte ordinal ou qualitatif, nous assignons au meilleur modèle de ϕ le degré maximal de possibilités comme suit :

$$\Pi(\omega \mid \phi) = \begin{cases} 1 & \text{si } \pi(\omega) = \Pi(\phi) \text{ et } \omega \models \phi \\ \pi(\omega) & \text{si } \pi(\omega) < \Pi(\phi) \text{ et } \omega \models \phi \\ 0 & \text{sinon} \end{cases} \quad (1.16)$$

Cette définition correspond au conditionnement basé sur le minimum.

- Dans le contexte numérique, tous les éléments de ϕ sont augmentés proportionnellement :

$$\Pi(\omega \mid \phi) = \begin{cases} \frac{\pi(\omega)}{\Pi(\phi)} & \text{si } \omega \models \phi \\ 0 & \text{sinon} \end{cases} \quad (1.17)$$

Cette définition correspond au conditionnement basé sur le produit.

Les deux définitions du conditionnement satisfont l'équation suivante qui rappelle l'équation de Bayes :

$$\forall \omega, \pi(\omega) = \pi(\omega \mid \phi) \otimes \Pi(\phi)$$

où \otimes représente le minimum (équation 1.16) ou le produit (équation 1.17).

Exemple 1.9 *Considérons deux variables binaires, relatives à la dégradation de la couche d'ozone (DO) et la pollution industrielle (PI) telles que : $D_{Do} = \text{Oui}(o), \text{Non}(\neg o)$ et $D_{Pi} = \text{Faible}(f), \text{Élevée}(\neg f)$. La table 1.1 donne les distributions de possibilités initiales. Supposons qu'une nouvelle information certaine arrive relative au fait que la pollution soit élevée. La croyance est représentée par ϕ . $\Pi(\phi) = \max(0.2, 0.5) = 0.5$.*

TAB. 1.1 – Distributions initiales

Do	Pi	$\Pi(\text{Do} \wedge \text{Pi})$
<i>o</i>	<i>f</i>	0.6
<i>o</i>	$\neg f$	0.2
$\neg o$	<i>f</i>	1
$\neg o$	$\neg f$	0.5

En utilisant l'équation du conditionnement basé sur le minimum, nous obtenons les distributions données par la table 1.2. En utilisant l'équation du conditionnement basé sur

TAB. 1.2 – Conditionnement basé sur le minimum

Do	Pi	$\Pi(\text{Do} \wedge \text{Pi} \mid \phi)$
<i>o</i>	<i>f</i>	0
<i>o</i>	$\neg f$	0.2
$\neg o$	<i>f</i>	0
$\neg o$	$\neg f$	1

le produit, nous obtenons les distributions données par la table 1.3.

TAB. 1.3 – Conditionnement basé sur le produit

Do	Pi	$\Pi(\text{Do} \wedge \text{Pi} \mid \phi)$
<i>o</i>	<i>f</i>	0
<i>o</i>	$\neg f$	0.4
$\neg o$	<i>f</i>	0
$\neg o$	$\neg f$	1

1.5.4 Mesures de probabilités et Mesures de possibilités

La théorie des possibilités est utilisée pour le raisonnement sous incertitude basé sur le concept d'ordre (qualitatif) aussi bien que sur le concept numérique (quantitatif),

contrairement à la théorie des probabilités qui est essentiellement numérique.

Dans le contexte subjectif [3], la théorie des possibilités quantitative concurrence la théorie des probabilités. Le terme subjectif signifie que les probabilités et les théories numériques proposées pour représenter l'incertitude sont considérées comme des outils permettant de quantifier les croyances d'un agent sans tenir compte de la nature aléatoire ou fréquentiste des événements.

La table 1.4 présente une comparaison entre les deux théories par rapport à leurs concepts respectifs.

TAB. 1.4 – Comparaison entre la théorie des probabilités et la théorie des possibilités

Probabilités	Possibilités
Probabilité P Distribution p	Mesures Π et N Distribution π
Normalisation : $\sum_{\omega \in \Omega} P(\omega) = 1$	Normalisation : $\max_{\omega \in \Omega} \pi(\omega) = 1$
Marginalisation : $P(A) = \sum_{\omega} p(\omega)$	Marginalisation $\Pi(A) = \sup_{\omega} \pi(\omega)$
$P(A \cup B) = P(A) + P(B)$ si $A \cap B = \emptyset$	$\Pi(A \cup B) = \max(\Pi(A), \Pi(B))$
$P(A \cap B) = P(A) * P(B)$ si A et B sont indépendants	$N(A \cap B) = \min(N(A), N(B))$
$P(A) + P(\bar{A}) = 1$	$\max(\Pi(A), \Pi(\bar{A})) = 1$ $\Pi(A) + \Pi(\bar{A}) \geq 1$ $N(A) + N(\bar{A}) \leq 1$
$\forall \omega \in \Omega, p(\omega) = \frac{1}{ \Omega }$	Ignorance : $\forall \omega \in \Omega, \pi(\omega) = 1$

1.5.5 Mesures de possibilités et les fonctions de croyance

Dans le cas où tous les éléments focaux sont emboîtés, c'est à dire si :

$$A_1 \subseteq A_2 \dots \subseteq A_n$$

alors, les mesures de croyance et de plausibilité deviennent équivalentes aux mesures de nécessité et de possibilité par les relations suivantes :

$$\forall A \subset \Omega, \forall B \subset \Omega, Bel(A \cap B) = \min(Bel(A), Bel(B)) \quad (1.18)$$

$$\forall A \subset \Omega, \forall B \subset \Omega, Pl(A \cup B) = \max(Pl(A), Pl(B)) \quad (1.19)$$

Une famille d'ensembles est dite emboîtée lorsque ses éléments sont ordonnés de telle sorte que chacun soit contenu dans le suivant. En pratique, dans une approche système expert, cette condition signifie que les informations fournies par les experts sont naturellement ordonnées au sens de leur précision car elles correspondent à des ensembles emboîtés de valeurs : $A_1 \subseteq A_2 \subseteq \dots \subseteq A_k$ [30]. Les différentes relations qui existent entre ces différentes mesures de l'incertitude sont illustrées par la figure 1.3.

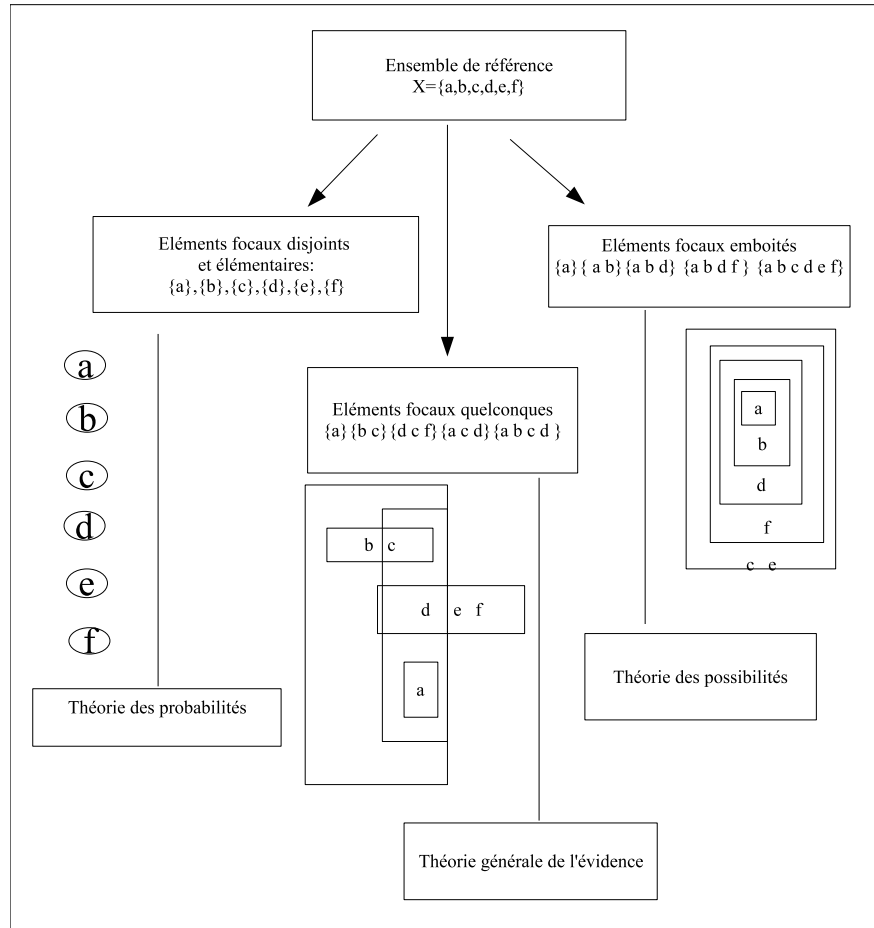


FIG. 1.3 – Relations entre les modèles de l'incertitude illustrées par un exemple

1.5.6 Conclusion

La théorie des possibilités permet une représentation naturelle de la notion d'incertitude. L'incertitude associée à un événement est évaluée par deux mesures duales (le degré de nécessité et le degré de possibilités). De plus, elle offre deux contextes de travail, l'un quantitatif ou numérique et l'autre qualitatif ou ordinal. Ainsi, le choix du mode dépendra, par exemple, de la nature des connaissances incertaines à modéliser.

1.6 Théorie des ensembles flous

La notion d'ensemble flou provient du constat établi par Zadeh que "très souvent, les classes d'objets rencontrés dans le monde physique ne possèdent pas de critères d'appartenance bien définis". La définition des ensembles flous répond au besoin de représenter des connaissances imprécises, floues et vagues qu'offre les langues naturelles

du type "la plupart des suédois mesure plus de 1.75 m", "Mohammed est de grande taille". Dans cette proposition, le qualificatif grand est un prédicat imprécis vu qu'il n'y a pas de transitions marquées entre l'appartenance à un ensemble "grandes tailles" et la non appartenance. De tels ensembles sont dits flous.

Les principales notions de base de cette théorie [7] sont illustrées dans les prochaines sous-sections.

1.6.1 Fonctions d'appartenance

Afin de représenter des concepts flous, Zadeh [7] a défini la fonction caractéristique μ comme suit :

Définition 1.7 *Un ensemble flou A sur un référentiel Ω est caractérisé par une fonction d'appartenance μ_A qui associe à chaque élément x de Ω un nombre réel dans l'intervalle $[0,1]$, c'est à dire :*

$$\mu_A : \Omega \rightarrow [0, 1]$$

La valeur $\mu_A(\omega)$ représente le degré d'appartenance de ω à A . Ainsi, plus la valeur $\mu_A(\omega)$ est proche de 1, plus grand est le degré d'appartenance de ω à A . En particulier :

- $\mu_A(\omega)=0$ représente la non-appartenance absolue,
- $\mu_A(\omega)=1$ représente l'appartenance totale,
- $0 < \mu_A(\omega) < 1$ représente l'appartenance nuancée ou partielle.

Un ensemble flou est dit normalisé s'il existe au moins un élément de son référentiel qui a son degré d'appartenance égale à 1. Il est à noter que la théorie des possibilités fait appel à des ensembles flous normalisés.

Nous illustrons cette notion d'ensemble flou par l'exemple suivant :

Exemple 1.10 *Soit A l'ensemble flou des tailles sur le référentiel des tailles illustré par la figure 1.4. Elle indique :*

- qu'une personne telle que sa taille est inférieure à 1,6 m est de petite taille,
- qu'une personne est d'autant plus grande que la taille se rapproche de 1.90 m,
- qu'une personne est considérée très grande si sa taille dépasse les 1.90 m.

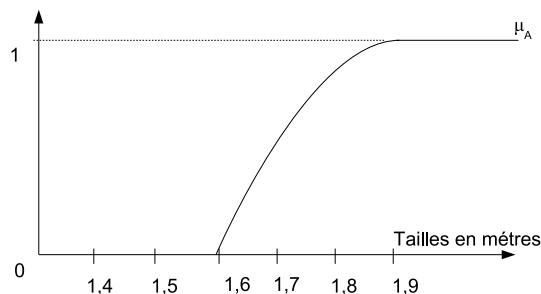


FIG. 1.4 – Exemple d'ensemble flou "grandes tailles"

Remarque 1.4 *La fonction d'appartenance n'a pas un caractère universel, mais dépend du contexte. Si nous prenons l'exemple de l'ensemble flou "grand", il diffère lorsque nous parlons de personnes, d'animaux ou autres.*

1.6.2 Les caractéristiques d'un ensemble flou

Les caractéristiques d'un ensemble flou A sont les suivantes[5].

1. La première caractéristique est le support de A , noté $\text{Supp}(A)$ qui représente l'ensemble des éléments ω qui ont un degré d'appartenance positif. Plus formellement, le support de A est défini par :

$$\text{Supp}(A) = \{\omega \in \Omega : \mu_A(\omega) > 0\}$$

2. La deuxième caractéristique de A est sa hauteur notée $h(A)$ qui représente le plus fort degré avec lequel un élément de X appartient à A . Plus formellement, la hauteur de A est définie par :

$$h(A) = \{\sup_{\omega \in \Omega} \mu_A(\omega)\}$$

Ainsi, un ensemble flou est dit normalisé si sa hauteur est égale à 1.

3. Le noyau de A , noté $\text{Noy}(A)$ est l'ensemble des éléments de Ω pour lesquels la fonction d'appartenance de A vaut 1. Il est défini par :

$$\{\omega \in \Omega : \mu_A(\omega) = 1\}$$

4. Lorsque Ω est fini, l'ensemble flou A de Ω est caractérisé par sa cardinalité, qui indique le degré global avec lequel les éléments de Ω appartiennent à A . Elle est définie par :

$$|A| = \left\{ \sum_{\omega \in \Omega} \mu_A(\omega) \right\}$$

1.6.3 Les coupes de niveau

La représentation d'ensemble flou par une fonction d'appartenance peut être vue comme "une représentation verticale" des ensembles flous. Une autre vue possible, est celle qui consiste à considérer un ensemble flou comme une famille de sous-ensembles classiques emboîtés, à l'aide de la notion de *coupe de niveau* [44]. Pour $\alpha \in]0,1]$, une coupe de niveau α d'un ensemble flou A est définie par :

$$A_\alpha = \{\omega \in \Omega : \mu_A(\omega) \geq \alpha\}$$

Nous pouvons aussi définir une α -coupe stricte par :

$$A_\alpha^- = \{\omega \in \Omega : \mu_A(\omega) > \alpha\}$$

1.6.4 Les opérations ensemblistes

Zadeh a défini [7] des opérations ensemblistes sur les ensembles flous qui sont en fait une extension des opérations sur les ensembles classiques :

1.6.4.1 Égalité et inclusion des ensembles flous

- Deux ensembles flous A et B sont égaux si leurs fonctions d'appartenance prennent les mêmes valeurs pour tout élément.

$$\forall \omega \in \Omega, \mu_A(\omega) = \mu_B(\omega) \quad (1.20)$$

- Un ensemble flou A est dit inclus dans un ensemble flou B si leurs fonctions d'appartenance sont telles que :

$$\forall \omega \in \Omega, \mu_A(\omega) \leq \mu_B(\omega) \quad (1.21)$$

1.6.4.2 Intersection et Union des ensembles flous

- L'intersection de deux ensembles flous A et B est l'ensemble flou C noté $A \cap B$, tel que

$$\forall \omega \in \Omega \mu_C(\omega) = \min(\mu_A(\omega), \mu_B(\omega)) \quad (1.22)$$

- L'union de deux sous ensembles flous A et B est l'ensemble flou D noté $A \cup B$, tel que :

$$\mu_D(\omega) = \max(\mu_A(\omega), \mu_B(\omega)) \quad (1.23)$$

D'une manière générale, l'intersection et l'union de deux ensembles flous peuvent être définies comme suit :

$$\mu_{A \cap B}(\omega) = T(\mu_A(\omega), \mu_B(\omega))$$

$$\mu_{A \cup B}(\omega) = S(\mu_A(\omega), \mu_B(\omega))$$

où T et S désignent respectivement une T-norme et une Co-norme. Les T-normes et les Co-normes les plus utilisées sont résumées dans les tables suivantes :

TAB. 1.5 – Exemples de T-normes

T-normes	Nom
$T_1(a,b) = \min(a,b)$	Minimum
$T_2(a,b) = a*b$	Produit
$T_3(a,b) = \max(a+b-1, 0)$	T-norme de Lukasiewicz

Les différentes T-normes et Co-normes vérifient les relation suivantes :

$$T_3 < T_2 < T_1$$

$$S_1 < S_2 < S_3$$

TAB. 1.6 – Exemples de Co-normes

Co-normes	Nom
$S_1(a,b)=\max(a,b)$	Maximum
$S_2(a,b)=a+b-a*b$	Somme probabiliste
$S_3(a,b)=\min(1,a+b)$	Somme bornée

En pratique, les opérateurs min et max sont les plus utilisés pour modéliser respectivement l'intersection et l'union car ils vérifient la propriété d'idempotence.

1.6.4.3 Complément d'un ensemble flou

Le complément \bar{A} d'un ensemble flou A est défini comme l'ensemble flou ayant comme fonction d'appartenance :

$$\mu_{\bar{A}}(\omega) = 1 - \mu_A(\omega) \quad (1.24)$$

Propriété 1.3 *Comme dans la théorie des ensembles classiques, les opérations citées précédemment vérifient les propriétés suivantes :*

- Les opérations d'union \cup et d'intersection \cap sont commutatives et associatives.
- Les opérations d'intersection et d'union vérifient les lois de De Morgan et de distributivité :

Les lois de De Morgan

$$(A \bar{\cup} B) = \bar{A} \cap \bar{B}$$

$$(A \bar{\cap} B) = \bar{A} \cup \bar{B}$$

Les lois de distributivité

$$C \cap (A \cup B) = (C \cap A) \cup (C \cap B)$$

$$C \cup (A \cap B) = (C \cup A) \cap (C \cup B)$$

- *L'inclusion est reflexive et transitive*

1.6.5 Mesures de possibilité et de nécessité

La théorie des possibilités est fondée sur la théorie des ensembles flous [8]. Elle se base sur la notion de distributions des possibilités et sur les mesures duales de possibilité et de nécessité.

Du point de vue mathématique, les mesures de possibilités et de nécessité se réfèrent aux ensembles flous [7, 8].

Distribution de possibilités Une distribution de possibilité, notée π_ω sur un référentiel U peut être vue comme une fonction caractéristique généralisée d'ensembles flous représentant les valeurs connues comme des valeurs plus ou moins possibles pour la variable considérée [8].

Définition 1.8 [8] Soit A un sous ensemble flou de l'univers du discours Ω caractérisé par sa fonction d'appartenance μ_A . Soit X une variable prenant ses valeurs sur Ω . La proposition " X est A " associée à X la distribution de possibilités π_x définie comme étant numériquement égale à la fonction d'appartenance de A :

$$\forall u \in \Omega, \pi_x(u) = \mu_A(u)$$

Cette égalité numérique n'implique pas l'égalité des deux concepts. En effet, $\pi_x(u)$ évalue la possibilité que $x=u$ étant donné l'état des connaissances incomplet " $x=a$ " alors que $\mu_A(u)$ évalue le degré de compatibilité de l'information $x=u$ avec la déclaration " x est A " [44].

Exemple 1.11 Nous savons de façon certaine que Jean a aux alentours de 30 à 33 ans et qu'en tout état de cause il n'a pas moins de 27 ans ni plus de 38 ans. Cette information imprécise peut être représentée par une distribution de possibilités $\mu_{age(Jean)}$ comme le montre la figure 1.5.

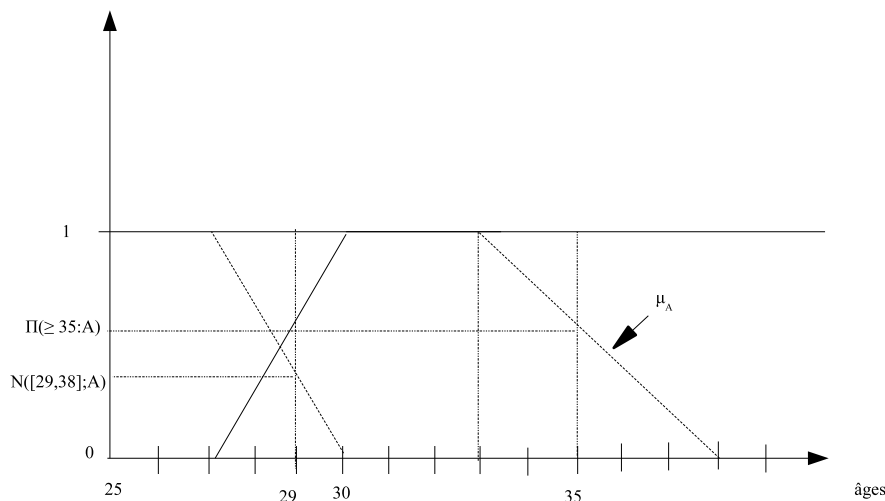


FIG. 1.5 – Distribution de possibilités $\mu_{age(Jean)}$

Exemple 1.12 [3] $\mu_{Tall}(h)$ quantifie l'appartenance d'une personne de taille h à l'ensemble Tall, $\pi_{Tall}(h)$ quantifie la possibilité que la taille d'une personne est h avec l'hypothèse qu'elle appartient à l'ensemble Tall. Le principe possibiliste est traduit par l'égalité

suivante :

$$\pi_{Tall}(h) = \mu_{Tall}(h)$$

avec $h \in H$ et H : ensemble des tailles $[0, +\infty]$

Elle stipule que la possibilité qu'un homme grand a la taille h est numériquement égale au degré d'appartenance d'un homme d'une taille h à l'ensemble des hommes grands.

$$\forall h \in H, \mu(Tall | h) = x \rightarrow \pi(h | Tall) = x, \quad (1.25)$$

Mesure de possibilité Zadeh a proposé de définir la possibilité d'un événement flou de la manière suivante [8] :

Définition 1.9 Soit A un ensemble flou de Ω . Soit π_x une fonction de distribution de possibilité associée à la variable X qui prend ses valeurs sur Ω . La mesure de possibilité, $\Pi(A)$ est définie par :

$$\Pi(A) = \sup_{u \in \Omega} \min(\mu_A(u), \pi_X(u))$$

où μ_A est la fonction d'appartenance de A .

Si A est non flou, nous retrouvons la définition de la possibilité d'un ensemble non flou. La mesure de possibilité peut être vue comme le degré (la hauteur) d'intersection de l'ensemble flou A et l'ensemble flou F_x défini par la fonction de distribution de possibilités π_x (voir figure 1.6).

$$\Pi(A) = 0 \Leftrightarrow \text{supp}(A) \cap \text{supp}(F_x) = \emptyset$$

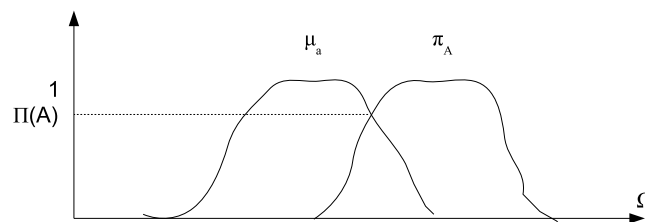


FIG. 1.6 – Mesure de possibilité

1.6.6 Raisonnement approché

En logique classique, le modus ponens constitue le mécanisme de base de l'inférence déductive. Il stipule :

Soit R la règle "Si X est A alors Y est B " et F le fait " X est A ". La conclusion " Y est B " est alors inférée.

En présence de connaissances imprécises, cette règle n'est pas suffisante vu que les

prémises peuvent être juste partiellement vérifiées. Pour pallier à ce problème, Zadeh a défini le modus ponens généralisé [8]. C'est une règle d'inférence compositionnelle permettant de déduire une conclusion imprécise à partir de prémises imprécises et de faits imprécis.

Le modus ponens est représenté par le syllogisme suivant :

Règle : Si X est A alors Y est B

Fait : X est A'

Conclusion : Y est B'

Nous allons décrire dans ce qui suit le modus ponens généralisé qui fait appel aux notions d'implication ainsi qu'au mécanisme de combinaison/projection.

1.6.6.1 L'implication floue

L'implication floue entre deux propositions floues "x est A" et "y est B" est une proposition floue concernant le couple de variable (x,y) dont la valeur de vérité est donnée par la fonction d'appartenance μ_R d'une relation floue entre x et y définie par :

$$\forall (u, v) \in U \times V \mu_R(u, v) = I(\mu_A(u), \mu_B(v)) \quad (1.26)$$

où I est une fonction de $[0,1] \times [0,1] \rightarrow [0,1]$ qui désigne une implication floue et vérifie les propriétés suivantes :

- I_1 : $I(0,0)=I(0,1)=I(1,1)=1$; $I(1,0)=0$,
- I_2 : si $x \leq y$ alors $I(x,y) \geq I(z,y) \forall y \in [0,1]$,
- I_3 : si $y \leq t$ alors $I(x,y) \leq I(x,t) \forall x \in [0,1]$,
- I_4 : $I(0,x)=1 \forall x \in [0,1]$,
- I_5 : $I(x,1)=1 \forall x \in [0,1]$.

Différents types d'implications floues ont été proposées.

La table 1.7 regroupe les implications floues les plus utilisées.

TAB. 1.7 – Principales implications floues

Implication floue	Nom
$I(a,b)=\max(1-a,b)$	Kleene-Dienes
$I(a,b)=1-a+a*b$	Reichenbach
$I(a,b)=\min(1,1-a+b)$	Lukasiewicz
$I(a,b)=1$ Si $a \leq b$ $I(a,b)=b$ sinon	Gödel
$I(a,b)=1$ si $a \leq b$ $I(a,b)=b/a$ sinon	Goguen

1.6.6.2 Le modus ponens généralisé

Considérons une règle floue de la forme "Si X est A alors Y est B" et un fait "X est A'" avec A et A' deux ensembles flous de U et B un ensemble flou de V. Le processus de raisonnement approché consiste à déterminer une conclusion de la forme "Y est B"

à partir de ces deux informations, c'est à dire de quelle façon la différence entre A et A' doit être répercutée sur B pour obtenir B' [44]. La méthode proposée par Zadeh consiste à définir la relation R correspondant à la règle, à la combiner avec A' puis de projeter le résultat sur V. Le calcul de B' s'effectue par le processus de combinaison/projection. La première opération consiste à combiner les différentes distributions de possibilités π_i définies sur U_i associées aux différentes variables comme suit :

$$\pi = \min_{i=1,r} \pi_i$$

La projection de π sur le domaine U_i de x_i est définie de la manière suivante :

$$\forall u_k \in U_k, \pi_i(u_i) = \sup_{u_k, k \rightarrow i} \pi_{x_1, \dots, x_n}(U_1, \dots, U_n)$$

Ainsi, en appliquant le processus de combinaison/projection, nous obtenons :

$$\forall v \in V, \mu_{B'}(v) = \sup_{u \in U} \min(\mu_{A'}(u), \mu_R(u, v))$$

où R est une implication floue I ($\mu_R(u, v) = I(\mu_A(u), \mu_B(v))$).

1.7 Conclusion

Les différents modèles mathématiques de l'incertitude présentés dans ce chapitre à savoir la théorie des probabilités, la théorie des croyances et la théorie des possibilités, sont les modèles de l'incertain les plus couramment utilisés en intelligence artificielle. Ils ont divers mérites dans le domaine de la représentation des connaissances incertaines. En utilisant la théorie des probabilités, l'ignorance totale sur un événement A ($P(A)=0$) entraîne que $\neg A$ doit être considéré comme certain ($P(\neg A)=1$) alors que dans les deux autres modèles, le fait que A ne soit pas vrai n'entraîne pas la croyance totale en l'événement $\neg A$.

Sous un autre angle, ces différentes facettes de l'incertitude sont liées. La figure 1.7 illustre la hiérarchie qui existe entre elles. Nous relevons plusieurs points :

1. La théorie des fonctions de croyances de Dempster-Shafer est un cadre général du raisonnement dans l'incertain qui permet de généraliser, entre autres, les deux cas particuliers que sont la théorie des possibilités et la théorie des probabilités [18] (voir figure 1.8).
2. Les degrés de confiance qui sont à la fois des mesures de possibilités, des mesures de crédibilité, de plausibilité et de probabilités sont des mesures de Dirac : $\exists \omega, \omega \in \Omega$ telle que $d(\omega) = 1$.
3. Toutes ces mesures sont elles mêmes des cas particuliers des mesures floues.

En intelligence artificielle, la logique a été considérée, depuis longtemps, à la base de la représentation et du raisonnement sur les connaissances d'une manière générale. Le chapitre suivant est destiné à exposer deux types de logiques quantitatives, qui nous intéressent particulièrement, modélisant les connaissances incertaines, à savoir la logique possibiliste quantitative et la logique des pénalités.

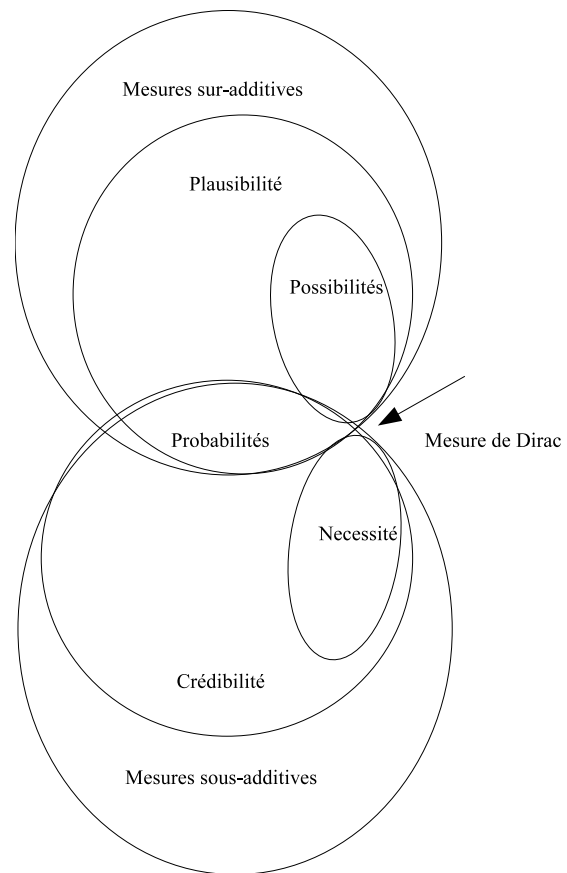


FIG. 1.7 – Tableau récapitulatif sur les mesures de l'incertain

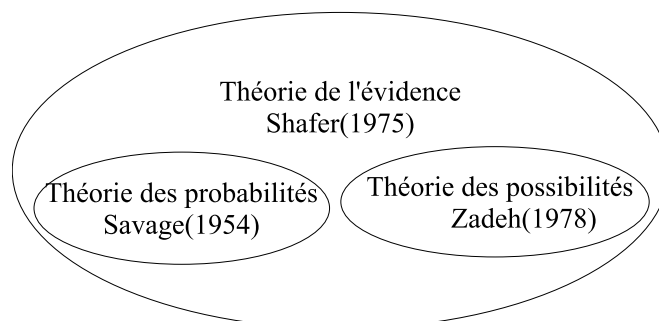


FIG. 1.8 – Relation entre les modèles de l'incertain

Chapitre 2

Modèles Logiques de l'incertain

2.1 Introduction

Le développement des recherches dans le domaine de la logique, vu sa richesse d'expression permettant de spécifier des problèmes complexes, a toujours eu une incidence sur le domaine de l'intelligence artificielle en général et sur la représentation et le raisonnement dans un cadre incertain en particulier.

Dans le contexte de l'incertain, plusieurs logiques ont été développées. Parmi lesquelles nous trouvons :

- les logiques numériques qualitatives ou ordinales telles que la logique possibiliste standard,
- les logiques numériques quantitatives telles que la logique possibiliste quantitative et la logique des pénalités,
- les logiques non numériques comme les logiques non monotones, la logique des défauts, la logique de la circonscription,...etc.

Ce chapitre est consacré à donner un bref aperçu sur ces différentes logiques. La section 2 présente la logique possibiliste standard, tout en mentionnant les relations qui la lie à d'autres logiques. La section 3 est dédiée à la logique possibiliste quantitative. La section 4 expose les concepts de base de la logique des pénalités, tout en invoquant les relations qu'elle entretient avec d'autres approches.

2.2 La logique possibiliste qualitative

Comme nous l'avons vu dans le chapitre précédent, la théorie des possibilités offre deux contextes de travail, l'un ordinal et l'autre numérique. Ainsi, si un expert veut représenter ses connaissances en utilisant la logique, il dispose en fait de deux types de logique possibiliste : la logique possibiliste standard et la logique possibiliste quantitative.

La logique possibiliste standard est une logique de l'incertain conçue pour raisonner avec des connaissances incomplètes et partiellement inconsistantes [45]. Elle est basée

sur l'idée d'ordre (complet) et non pas sur le calcul comme le cas des probabilités.

Au niveau syntaxique, des formules logiques propositionnelles sont manipulées, auxquelles sont attachées des valuations numériques correspondant au seuil minimal des degrés de nécessité.

Au niveau mathématique, les degrés de possibilité et de nécessité se réfèrent aux ensembles flous [7, 8] et la logique possibiliste est alors adaptée pour le raisonnement automatique où les connaissances disponibles sont entachées de flou.

2.2.1 Base possibiliste standard

En logique possibiliste standard, les informations incertaines sont représentées par une base propositionnelle avec priorité, appelée base de croyances possibilistes [46]. Plus formellement, une base possibiliste Σ est constituée d'un ensemble de couples (ϕ_i, α_i) où ϕ_i est une formule propositionnelle et $\alpha_i \in [0,1]$ est le poids associé à chaque formule correspondant au seuil minimal du degré de nécessité. D'une manière formelle, $\Sigma = \{(\phi_i, \alpha_i) : i = 1, n\}$ avec $N(\phi_i) \geq \alpha_i$.

2.2.2 Distribution de possibilités associée à une base possibiliste standard

La base possibiliste est une représentation compacte des distributions de possibilités. Chaque élément d'information (ϕ_i, α_i) d'une base de connaissances peut être vu comme une contrainte qui restreint le degré de possibilité associé aux interprétations [47].

Si une interprétation ω satisfait ϕ_i alors son degré de possibilités est égal à 1 (ω est complètement compatible avec la croyance ϕ_i), sinon, le degré est égal à $1 - \alpha_i$ (plus ϕ_i est certaine, moins ω est possible). En particulier, si $\alpha_i = 1$, alors toute interprétation qui falsifie ϕ_i , est telle que son degré de possibilité est égal à 0, donc est impossible. D'une manière formelle, la distribution de possibilités associée à une formule pondérée (ϕ_i, α_i) est définie par [47] : $\forall \omega \in \Omega$:

$$\pi_{(\phi_i, \alpha_i)}(\omega) = \begin{cases} 1 - \alpha_i & \text{si } \omega \not\models \phi_i \\ 1 & \text{sinon} \end{cases} \quad (2.1)$$

Plus généralement, la distribution de possibilités associée à une base de connaissances possibiliste standard Σ est le résultat de la combinaison, par l'opérateur min, des distributions de possibilités associées à chaque formule pondérée (ϕ_i, α_i) de Σ , à savoir : $\forall \omega \in \Omega$:

$$\pi_{\Sigma}(\omega) = \begin{cases} 1 & \text{si } \forall (\phi_i, \alpha_i) \in \Sigma, \omega \models \phi_i \\ \min\{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma, \omega \not\models \phi_i\} & \text{sinon} \end{cases} \quad (2.2)$$

Exemple 2.1 Soit la base de connaissances possibiliste standard suivante :

$\Sigma = \{(\neg a \vee b, .1), (a \vee b \vee c, .3), (\neg b \vee d, .5), (\neg d, .9)\}$. La distribution de possibilités est donnée dans la table 2.1.

TAB. 2.1 – Distribution de possibilités associée à la base de connaissances de l'exemple 2.1

ω	$\pi_{\Sigma}(\omega)$
$a \wedge b \wedge c \wedge d$.1
$a \wedge b \wedge c \wedge \neg d$.5
$a \wedge b \wedge \neg c \wedge d$.1
$a \wedge b \wedge \neg c \wedge \neg d$.5
$a \wedge \neg b \wedge c \wedge d$.1
$a \wedge \neg b \wedge c \wedge \neg d$.5
$a \wedge \neg b \wedge \neg c \wedge d$.1
$a \wedge \neg b \wedge \neg c \wedge \neg d$.9
$\neg a \wedge b \wedge c \wedge d$.1
$\neg a \wedge b \wedge c \wedge \neg d$.5
$\neg a \wedge b \wedge \neg c \wedge d$.1
$\neg a \wedge b \wedge \neg c \wedge \neg d$.5
$\neg a \wedge \neg b \wedge c \wedge d$.1
$\neg a \wedge \neg b \wedge c \wedge \neg d$	1
$\neg a \wedge \neg b \wedge \neg c \wedge d$.1
$\neg a \wedge \neg b \wedge \neg c \wedge \neg d$.7

2.3 La logique possibiliste quantitative

La logique possibiliste quantitative est une logique numérique. Ainsi, un degré d'incertitude, correspondant au seuil minimal du degré de nécessité est associé à chaque formule ϕ_i de la base de connaissances. Contrairement à la logique possibiliste standard, les quantifications numériques attachées aux formules n'induisent pas une relation d'ordre. En effet, la pondération α_i associée à chaque formule ϕ_i permet juste de quantifier son degré d'incertitude.

Soit $\Sigma = \{(\phi_i, \alpha_i) : i = 1, \dots, n\}$ avec $N(\phi_i) \geq \alpha_i$. La distribution de possibilités associée à une base de connaissances possibiliste quantitative Σ est définie par : $\forall \omega \in \Omega$:

$$\pi_{\Sigma}(\omega) = \begin{cases} 1 & \text{si } \forall (\phi_i, \alpha_i) \in \Sigma, \omega \models \phi_i \\ * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma, \omega \not\models \phi_i\} & \text{sinon} \end{cases} \quad (2.3)$$

Exemple 2.2 Soit Σ la base de connaissances possibiliste quantitative suivante : $\Sigma = \{(a \vee b, .8), (\neg a \vee c, .2), (c \vee \neg d, .5)\}$

La distribution de possibilités associée à Σ est donnée par la table 2.2.

Du côté du mode d'inférence, contrairement à la logique possibiliste standard, il n'y a pas eu de développement d'algorithmes de raisonnement. Pour ce faire, il faudrait tenir compte de la nature des connaissances exprimées dans le contexte de la logique possibiliste quantitative. Nous allons nous y intéresser dans le chapitre 8.

TAB. 2.2 – Distribution de possibilités associée à la base de connaissances de l'exemple 2.2

ω	$\pi_{\Sigma}(\omega)$
$a \wedge b \wedge \neg c \wedge d$.4
$a \wedge b \wedge \neg c \wedge \neg d$.8
$a \wedge \neg b \wedge \neg c \wedge d$.4
$a \wedge \neg b \wedge \neg c \wedge \neg d$.8
$\neg a \wedge b \wedge \neg c \wedge d$.5
$\neg a \wedge \neg b \wedge c \wedge d$.2
$\neg a \wedge \neg b \wedge c \wedge \neg d$.2
$\neg a \wedge \neg b \wedge \neg c \wedge d$.1
$\neg a \wedge \neg b \wedge \neg c \wedge \neg d$.2
Autres interprétations	1

2.4 La logique des pénalités

La logique des pénalités a été proposée par Pinkas [22] et développée par Dupin de Saint-Cyr [48, 49, 18]. Dans cette approche, à chaque formule de la base de connaissances est associée une pénalité, qui correspond au prix à payer pour pouvoir rejeter cette formule. Plus la pénalité est élevée, plus la formule est importante.

D'une manière plus formelle, une base de connaissances à pénalités est définie par :

Définition 2.1 Soit \mathcal{R}^{*+} l'ensemble des réels strictement positifs plus $\{+\infty\}$. Une base de connaissance à pénalités notée PK est un multi-ensembles fini de couples (ϕ_i, α_i) avec $\phi_i \in \mathcal{L}$ et $\alpha_i \in \mathcal{R}^{*+}$, α_i étant la pénalité associée à ϕ_i .

Dans une optique de raisonnement, α_i représente le degré de certitude associé à ϕ_i . Si $\alpha_i = +\infty$ alors ϕ_i est complètement certaine. Si une formule apparaît plusieurs fois dans une base, la pénalité résultante est la somme de l'ensemble des pénalités associées à la formule.

Exemple 2.3 $PK = \{(a, 1), (a, 1)\}$ n'est pas équivalente à $PK' = \{(a, 1)\}$ vu que dans la base PK le coût pour supprimer a est égale à 2 alors que dans la base PK' le coût est égal à 1.

Remarque 2.1 : Si dans une base à pénalités, toutes les pénalités sont infinies, cela implique qu'aucune formule ne peut être violée, ce qui réduit la logique des pénalités à la logique classique.

2.4.1 Coût d'une interprétation

Soit $PK = \{(\phi_i, \alpha_i) : i = 1, \dots, n\}$ une base de connaissances à pénalités.

Définition 2.2 [22]

Le coût d'une interprétation $\omega \in \Omega$ conformément à PK noté k_{PK} est égal à la somme des pénalités des formules de PK violées par ω :

$$\kappa_{PK}(\omega) = \begin{cases} 0 & \text{Si } \forall (\phi_i, \alpha_i) \in PK, \omega \models \phi_i \\ \sum \{\alpha_i : (\phi_i, \alpha_i) \in PK, \omega \not\models \phi_i\} & \text{Sinon} \end{cases} \quad (2.4)$$

Exemple 2.4 Soit $PK = \{(\neg a, 6), (\neg b \vee c, 8), (a \vee c, 3), (\neg c, +\infty)\}$ Les différents coûts d'interprétation sont donnés par la table 2.3.

TAB. 2.3 – Coûts d'interprétation de l'exemple 2.4

ω	k_{PK}
$a \ b \ c$	$+\infty$
$a \ b \ \neg c$	$6+8=14$
$a \ \neg b \ c$	$+\infty$
$a \ \neg b \ \neg c$	6
$\neg a \ b \ c$	$+\infty$
$\neg a \ b \ \neg c$	$3+8=11$
$\neg a \ \neg b \ c$	$+\infty$
$\neg a \ \neg b \ \neg c$	3

Une base à pénalités induit une distribution de pénalités unique $k_{PK}(\omega)$. $k_{PK}(\omega)$ est d'autant plus faible qu'il est plausible que ω soit le monde réel.

2.4.2 Coût d'une formule

Le coût d'une formule ϕ cohérente, noté $K_{PK}(\phi)$ est le coût minimal relativement à PK qui satisfait ϕ :

$$K_{PK}(\phi) = \min\{k_{PK}(\omega) : \omega \models \phi\} \quad (2.5)$$

Le coût k_{PK} satisfait les propriétés suivantes :

Propriété 2.1 $K_{PK}(\perp) = +\infty$

Cette propriété est naturelle, vu qu'il n'est pas possible de rendre K_{PK} cohérent avec \perp .

Propriété 2.2 $K_{PK}(\top) = \min\{k_{PK}(\omega) : \omega \in \Omega\}$

$K_{PK}(\top)$ représente le coût minimal pour rendre K_{PK} cohérente.

Propriété 2.3 $K_{PK}(\top) = +\infty \Leftrightarrow \forall (\phi_i, \alpha_i) \in PK, \alpha_i = +\infty$ et $\{(\phi_i, \alpha_i) \in PK\}$ est incohérente.

Propriété 2.4 $K_{PK}(\top) = 0 \Leftrightarrow PK$ est cohérente.

Propriété 2.5 $\forall \phi, \psi \in \mathcal{L}, (\phi \models \psi) \rightarrow K_{PK}(\phi) \geq K_{PK}(\psi)$

Propriété 2.6 $\forall \phi, \psi \in \mathcal{L},$

1. $K_{PK}(\phi \wedge \psi) \geq \max(K_{PK}(\phi), K_{PK}(\psi))$
2. $K_{PK}(\phi \vee \psi) = \min(K_{PK}(\phi), K_{PK}(\psi))$
3. $K_{PK}(\perp) \geq K_{PK}(\phi) \geq K_{PK}(\top)$

Remarque 2.2 Dans [50], Freund a mis en exergue la relation qui existe entre la logique des pénalités et la logique des défauts. Il a proposé une approche additive pour générer une relation d'inférence non-monotone à partir d'un ensemble de règles de défauts. Cette approche associe d'abord une valuation entière à chaque règle. Ces valuations sont en fait des pénalités puisqu'elles induisent une valuation sur chaque interprétation ω en sommant les poids des règles violées par ω . Ainsi, il y a génération automatique des pénalités à partir d'un ensemble de défauts $\phi_i \rightarrow \psi_i$.

2.5 Les logiques connexes

Nous allons brièvement aborder dans cette section :

- les relations qui existent entre la logique possibiliste standard et certaines logiques numériques telles que la logique probabiliste et certaines logiques non numériques telles que la logique modale et la logique des défauts,
- les relations qu'entretient la logique des pénalités avec la logique possibiliste, les fonctions de croyance et les fonctions conditionnelles ordinales.

2.5.1 La logique possibiliste et la logique probabiliste

La logique probabiliste peut être utilisée pour modéliser le raisonnement automatique. En effet, le mécanisme d'inférence peut être étendu à la logique probabiliste. Ainsi par exemple le principe de résolution s'exprimerait par [42] :

$$P(q \vee r) \geq \max(0, P(p \vee q) + P(\neg p \vee r) - 1)$$

Il est à noter que la borne inférieure est plus petite si nous la comparons à celle utilisant les degrés de nécessité. De plus, l'utilisation de l'axiome d'additivité dans la logique probabiliste entraîne un taux d'erreur croissant des données de départ vers les conclusions, tandis que dans le contexte possibiliste, ce taux est constant. Le critère de normalisation dans la logique probabiliste ($\sum_{i=1..n} P(\omega_i)=1$ où ω_i sont des événements élémentaires) est plus difficile à satisfaire que la règle de normalisation pour les mesures de possibilités ($\max_{i=1..n} \Pi(\omega_i)=1$).

Néanmoins, il s'est avéré que les mesures de possibilités peuvent être interprétées comme une famille spéciale des probabilités inférieures et supérieures [41] et les fonctions ordinales conditionnelles [26] ont permis d'interpréter les mesures de possibilités en termes de probabilités infinitésimales [51].

La logique possibiliste et la logique probabiliste ne sont pas dédiées aux mêmes buts

[47]. La logique probabiliste apparaît mieux adaptée pour traiter les connaissances incomplètes de type statistique. Elle a été également construite comme une théorie de l'induction [52] ou une logique pour les croyances subjectives [53]. La logique probabiliste est plus complexe que la logique possibiliste, spécialement si les probabilités conditionnelles doivent être prises en compte dans le langage. Ceci implique que le raisonnement possibiliste est plus simple. La logique possibiliste quant à elle a pour but de permettre de raisonner avec la partie la plus certaine de la base de connaissances.

2.5.2 La logique possibiliste et la logique modale

Il est également intéressant de discuter des liens qui existent entre la logique des possibilités et la logique modale qui procure une modélisation des modalités telles que la possibilité et la nécessité [42].

La logique modale a été introduite afin de prendre en compte des affaiblissements ou des renforcements d'affirmations présents dans les langues naturelles. Les principales différences entre ces deux approches sont :

- Dans la logique modale, possibilité et nécessité sont des concepts "tout ou rien". Ils peuvent être introduits comme des symboles spéciaux du langage. $\Box p$ correspondrait à "p est nécessaire" et $\Diamond p$ correspondrait à "p est possible". Dans la théorie des possibilités, les notions de possibilité et de nécessité sont des notions graduelles.
- La logique modale propose différents types d'axiomatisation, tandis que les axiomes de la théorie des possibilités sont bien définis et sont uniques. Dans cette optique, il est judicieux de définir la contrepartie qualitative des axiomes de la théorie des possibilités dans le style de la logique modale. Une façon de faire consiste à utiliser les règles de traductions suivantes :
 - $\vdash \Box P$ est traduit en $N(p)=1$
 - $\vdash \Diamond P$ est traduit en $\Pi(p)=1$
 - l'identité classique $\neg \Diamond p = \Box \neg p$ est traduite en $1-\Pi(p)=N(\neg p)$
- De plus, une traduction numérique de l'implication de Lewis $\Box(p \rightarrow q)$ serait $N(p \rightarrow q)=1$ qui implique que $p \rightarrow q$ est vrai dans la logique possibiliste.

Pour plus de détails sur les relations entre la logique possibiliste et la logique modale voir [54].

2.5.3 Raisonnement par défaut utilisant les logiques de l'incertain

La logique probabiliste et la logique possibiliste ont été suggérées comme des approches possibles pour le raisonnement par défaut [55, 56]. L'idée est d'interpréter le poids accordé à une règle "Si-Alors" comme une mesure qui évaluerait à quel point cette règle n'a pas d'exception. Cette interprétation engendre plusieurs problèmes [42] :

- Les règles par défaut n'ont pas toujours une interprétation statistique. En particulier, la typicalité semble être de nature différente. Une règle par défaut est alors modélisée par la connaissance de la quantité $\Pi(q | p)$ définie par la relation :

$$\Pi(p \wedge q) = \min(\Pi(p), \Pi(q | p))$$

Néanmoins, cette notion de conditionnement est très liée à l'implication logique, vu que $\Pi(q | p) = 1 - N(p \rightarrow \neg q)$.

- Ce n'est pas clair que dans la logique des défauts, nous puissions attacher des degrés d'incertitude à une implication logique $p \rightarrow q$.

L'utilisation du conditionnement au lieu de l'implication peut apparaître une approche plus naturelle pour modéliser les relations imparfaites de type "est un".

2.5.4 La logique des pénalités et la logique possibiliste

Le coût d'une formule est une mesure de possibilité non normalisée avec un changement d'intervalle de $[0, +\infty]$ en $[0, 1]$. En effet [18], la logique des pénalités peut être considérée comme une logique multi-sources où chaque formule $\langle \phi_i, \alpha_i \rangle$ d'une base de connaissances à pénalité provient d'une source indépendante et exprime une contrainte sur une fonction coût k_i définie par :

$$k_i(\omega) = \begin{cases} \alpha_i & \text{si } \omega \models \neg\phi_i \\ 0 & \text{sinon} \end{cases} \quad (2.6)$$

En combinant toutes ces fonctions coûts k_i par l'addition, nous obtenons le coût d'une interprétation défini par :

$$k(\omega) = \sum_{i=1}^n k_i(\omega)$$

De manière similaire, la logique possibiliste peut être aussi considérée comme une logique multi-sources : chaque formule $\langle \phi_i, \beta_i \rangle$ d'une base de connaissances possibiliste provient d'une source indépendante et exprime une contrainte sur une distribution de possibilités π_i définie par :

$$\pi_i(\omega) = \begin{cases} 1 - \beta_i & \text{si } \omega \models \neg\phi_i \\ 1 & \text{sinon} \end{cases} \quad (2.7)$$

En combinant toutes ces distributions de possibilités π_i par le minimum, nous obtenons la distribution de possibilités la moins spécifique qui satisfait les contraintes :

$$\pi^*(\omega) = \min_{i=1 \dots n} \pi_i(\omega)$$

Si nous utilisons une autre norme triangulaire comme le produit, nous obtenons :

$$\pi^*(\omega) = \prod_{i=1 \dots n} \pi_i(\omega)$$

Le logarithme de cette distribution de possibilités correspond à l'opposé de la fonction coût induite par la base de connaissances à pénalités où les pondérations α_i sont telles que $\alpha_i = -\log(1 - \beta_i)$.

$$\log(\pi^*(\omega)) = \sum_{\omega \models \neg\phi_i} \log(1 - \beta_i) = -k(\omega) \quad (2.8)$$

2.5.5 La logique des pénalités et les fonctions de croyance

Dans [23], Dupin de Saint-Cyr a montré la relation entre la théorie des fonctions de croyances et la logique des pénalités sous forme de deux points :

- Le coût d'interprétation $k_{PK} : \Omega \rightarrow [0, +\infty]$ induit par une base à pénalités PK composée de n formules pondérées correspond à la fonction de contour Pl : $\Omega \rightarrow [0, 1]$ induite par la règle de combinaison de Dempster sur n fonctions supports (une pour chaque formule ϕ_i).

Soit $PK = \{(\phi_i, \alpha_i), i = 1, \dots, n\}$ une base de connaissances à pénalités. La fonction de masse m_i est définie comme suit :

$$m_i(\phi_i) = 1 - e^{-\alpha_i}$$

$$m_i(\top) = e^{-\alpha_i}$$

avec $e^{-\infty} = 0$ vu que $\alpha_i \in [0, 1]$.

Soit $m = m_1 \otimes \dots \otimes m_n$ le résultat de la combinaison des m_i avec la règle de Dempster [35] sans re-normalisation.

La fonction de contour pl : $\Omega \rightarrow [0, 1]$ associée à m est la restriction de la fonction de plausibilité Pl sur des singletons, c'est à dire que :

$$Pl(\omega) = Pl(\{\omega\}) = \sum_{\omega \models \phi} m(\phi)$$

Après développement, nous obtenons :

$$Pl(\omega) = e^{-K_{PK}(\omega)}$$

ou encore :

$$K_{PK}(\omega) = -\text{Log}(Pl(\omega))$$

Cette équation ne peut être étendue aux formules ; Il n'y a pas d'équivalence entre la fonction de plausibilité et la fonction coût K_{PK} .

- La fonction $K_{PK} : \mathcal{L} \rightarrow [0, 1]$ correspond à une mesure de plausibilité dans le cas d'une version infinitésimale de la théorie de Dempster-Shafer où les masses impliquées sont toutes infiniment proches de 0 ou de 1.

2.5.6 La logique des pénalités et les fonctions conditionnelles ordinales

Les fonctions conditionnelles ordinales [26] associent à chaque interprétation ω un nombre entier $k(\omega)$ qui représente le "degré d'anormalité" de ω . Cette fonction k s'étend aux formules de la même façon que dans la logique des pénalités [23], par :

$$k(\phi) = \min_{\omega \models \phi} k(\omega)$$

L'interprétation de $k(\phi)$ a trait aux probabilités infinitésimales :

- Si $k(\phi)=\alpha$ alors la probabilité de ϕ est de l'ordre de ϵ^α , où ϵ est infiniment petit.
- Si $\alpha =0$ alors $\Pr(\phi) \approx k$ où $k > 0$.

Spohn a défini également le degré d'anormalité de ω sachant ϕ par :

$$k(\omega | \phi) = k(\omega) - k(\phi)$$

puis :

$$k(\psi | \phi) = \min_{\omega \models \phi} k(\omega | \phi)$$

Si ϕ et ψ ne sont pas incompatibles alors $k(\phi \wedge \psi) = k(\phi) + k(\psi | \phi)$.

Le caractère additif de la fonction k n'a pas trait à une somme de coûts issue de la violation de formules (comme le cas de la logique des pénalités) mais à l'application du théorème de Bayes sur des probabilités infinitésimales.

Remarque 2.3 *Pour les fonctions conditionnelles ordinales, il n'existe pas de bases de connaissances sous-jacentes à la construction de la fonction k [50].*

2.6 Conclusion

Les différentes logiques présentées dans ce chapitre permettent de traiter les informations incertaines. Plusieurs auteurs se sont intéressés à faire le parallèle entre elles. La logique de probabilités, la logique des possibilités et les OCF admettent l'existence d'interprétations totalement possibles exprimées par la valeur 1 dans la théorie des possibilités et par 0 dans les OCF. Cette notion n'existe pas lorsque nous utilisons les relations de plausibilité (les fonctions de croyance). Les théories des probabilités et de possibilités admettent l'existence de mondes totalement impossibles exprimée par 0. Dans les autres formalismes, toutes les interprétations sont quelque peu possibles.

Dans ce chapitre, nous nous sommes intéressés, entre autres, à deux types de logique : la logique possibiliste quantitative et la logique des pénalités. Ces modes de représentation des connaissances incertaines ainsi que leurs modes d'inférence feront l'objet de notre étude dans les chapitres ultérieurs.

Nous nous intéressons dans le chapitre suivant aux modes de représentation graphiques et plus précisément

- aux réseaux causaux probabilistes,
- et aux réseaux causaux possibilistes qualitatifs

Les réseaux causaux possibilistes quantitatifs, faisant l'objet de notre étude, seront abordés plus en détails dans la partie consacrée à notre approche.

Chapitre 3

Modèles Graphiques de l'incertain

3.1 Introduction

Les approches logiques de l'incertain permettent d'utiliser un langage formel pour la description des connaissances et le raisonnement automatique. Elles constituent une référence aux autres formalismes surtout pour le raisonnement. Néanmoins, les connaissances ne sont pas structurées.

Plusieurs recherches ont justement permis l'émergence d'un certain nombre de modèles graphiques offrant un cadre de représentation plus structuré. En effet, afin de capturer les différents aspects inhérents au raisonnement incertain, il est primordial de tenir compte des influences pouvant exister entre certains événements et la répercussion de la prise en charge d'une nouvelle connaissance sur le réseau.

Parmi les modèles graphiques les plus répandus, nous distinguons : les réseaux causaux probabilistes [10, 11, 12], les arbres de décision [57], les diagrammes d'influence [58, 59, 60] et les *Valuation Based Systems (VBS)* [61, 62, 63, 64, 65]. L'ensemble de ces modèles graphiques reposent sur la théorie des probabilités.

Pour les mêmes raisons qu'en logique et dans le même ordre d'idée, des chercheurs ont été amenés à adapter les mécanismes de propagation probabilistes en proposant des versions possibilistes. Ainsi, dans [66], les auteurs proposent un algorithme de propagation possibiliste pour les hypergraphes. Dans [67], Fonck a adapté l'algorithme de Pearl [10] dans le contexte possibiliste. Dans [68, 69, 70] les auteurs ont développé une boîte à outils pour la propagation possibiliste pour les réseaux possibilistes indirects et plus récemment dans [20] l'auteur a adapté les algorithmes de propagation des réseaux causaux probabilistes aux réseaux causaux possibilistes basés sur le produit et sur le minimum.

Les VBS ont également été adaptés en théorie des possibilités. En effet, Shenoy [62] a proposé une version possibiliste de l'algorithme de propagation dans les VBS utilisant uniquement l'opérateur produit. Cano, Delgado et Moral [71] ont présenté un système de propagation pour des graphes simplement connectés utilisant des valuations. Une généralisation de cet algorithme au graphe avec boucles a été développé par Fonck [67].

Dans une représentation graphique, nous distinguons deux composantes :

- Une composante graphique matérialisée par des graphes pouvant être de différentes structures. Elle explicite l'ensemble des variables dont il est question ainsi que les liens qui existent entre elles, mettant ainsi en exergue les différentes influences entre les variables.
- Une composante numérique qui quantifie les différents liens qui existent entre les variables.

Les modèles possibilistes, basés sur la théorie des possibilités, offrent deux contextes de travail, permettant soit une modélisation qualitative soit une modélisation quantitative selon l'hypothèse de la problématique à traiter. Ceci permet en fait de définir deux types de réseaux causaux possibilistes : les réseaux causaux possibilistes basés sur le minimum (qualitatif) et les réseaux causaux possibilistes basés sur le produit (quantitatif).

Nous allons aborder dans ce chapitre les différentes notions inhérentes aux principaux modèles graphiques de l'incertain en commençant par les modèles numériques issus de la théorie des probabilités, puis les modèles issus de la théorie des possibilités. Ainsi, la section 2 aborde les réseaux causaux probabilistes et présente les définitions de bases des différents types de structures de graphes ainsi que les algorithmes de propagation adéquats. La section 3 présente les réseaux causaux basés sur le minimum. Par la suite, nous aborderons les algorithmes de propagation pour chaque type de structure d'un réseau causal possibiliste basé sur le minimum, à savoir : les polyarbres et les graphes à connexions multiples. Les réseaux causaux possibilistes basés sur le produit seront présentés ultérieurement.

3.2 Réseaux causaux Bayésiens

Les réseaux causaux bayésiens sont des modèles graphiques dirigés ou orientés. Ils sont issus d'un mariage entre la théorie des probabilités et la théorie des graphes apportant ainsi des outils naturels pour traiter le problème crucial de l'incertitude dans le domaine de l'intelligence artificielle. La théorie des probabilités est à la base de ces modèles graphiques assurant leurs consistences. La théorie des graphes fournit une interface intuitive permettant ainsi à un humain de structurer un problème par un ensemble de composantes interagissant entre elles.

Cette synthèse a pour but de donner un aperçu sur les principaux concepts des réseaux bayésiens et de présenter les principaux algorithmes exacts de la propagation.

3.2.1 Les concepts de base

Soit l'exemple suivant [72] extrait du *frame problem* [73] :

Exemple 3.1 *En début de journée, la voiture ne démarrait pas. Le starter étant en fonction, les causes de la panne peuvent être diverses : Vu que le starter fonctionne, cela induit que la batterie est chargée. La cause la plus probable est que l'essence a dû être dérobé durant la nuit ou encore que les bougies ne sont pas nettoyées. Cela peut être également dû au carburateur non nettoyé, ou encore à quelque chose de plus sérieux.*

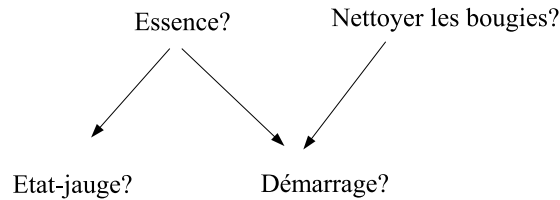


FIG. 3.1 – Réseau causal correspondant à l'exemple 3.1

Afin de localiser la panne, la jauge d'essence est d'abord vérifiée. Celle-ci indiquait la moitié du réservoir. Ainsi, la décision prise est de nettoyer les bougies.

Afin d'implémenter ce type de raisonnement sur machine, il est nécessaire de pouvoir répondre à des questions de type : qu'est ce qui m'a permis d'affirmer que les causes : "Vol d'essence" et "bougies mal nettoyées" sont les plus probables ? Une autre question qui est posée est : qu'est ce qui m'a poussé à vérifier la jauge d'essence ?

Il est alors nécessaire de disposer d'outils afin de représenter le problème ainsi que des méthodes d'inférence.

Une façon de faire, consiste à construire un graphe représentant les relations causales entre les événements. Ainsi, cette situation peut être décrite par un graphe composé par des nœuds et des liens. Les événements sont représentés par les nœuds. Lorsque un nœud A a un impact causal sur le nœud B, alors les deux nœuds seront reliés par un lien.

La figure 3.1 illustre les relations causales de l'exemple 3.1.

3.2.1.1 Quelques définitions

Avant d'entamer les concepts de bases des réseaux causaux bayésiens, nous allons présenter, dans cette section, les préliminaires sur les graphes, nécessaires à la description des algorithmes d'inférence.

1. Généralités

Définition 3.1 (Graphe ou réseau) Soit V un ensemble fini de variables. Soit E une partie de $V \times V$. Un graphe noté $G=(V,E)$ est défini par deux ensembles V et E où V est un ensemble fini de nœuds $V=\{A,B,\dots\}$ et E est un ensemble de liens connectant des paires de nœuds dans V (voir figure 3.2b). Si les liens dans E sont orientés alors, ils sont dits des arcs et le graphe G associé est appelé graphe orienté.

Les figures (3.2a),(3.2b),(3.2c),(3.2d),(3.2e) représentent des graphes.

Définition 3.2 (Chaîne, Chemin) Soit $G=(V,E)$ un graphe. Une séquence de nœuds $[v_0, v_1, \dots, v_n]$ est une chaîne dans G si et seulement si :

$$\forall i, 1 \leq i \leq n, (v_{i-1}, v_i) \in E, (v_i, v_{i+1}) \in E$$

Soit $G=(V,E)$ un graphe orienté. Une séquence de nœuds $[v_0, v_1, \dots, v_n]$ est un chemin dans G si et seulement si :

$$\forall i, 1 \leq i \leq n, (v_{i-1}, v_i) \in E$$

Définition 3.3 (Cycle, Boucle) Soit $G=(V,E)$ un graphe. Un cycle est un chemin simple $[v_0, v_1, \dots, v_0]$.

Une boucle est un cycle indirect.

Définition 3.4 (Graphe Acyclique Orienté) Soit $G=(V,E)$ un graphe orienté. G est dit acyclique si G ne contient aucun cycle orienté.

Définition 3.5 (Parent, Enfant) Soit $G=(V,E)$ un graphe orienté. Si $(u,v) \in A$, alors u est un parent de v , et v est un enfant de u .

Définition 3.6 (Ascendant, Descendant) Soit $G=(V,E)$ un graphe orienté et u, v deux nœuds. S'il existe un chemin orienté de u à v alors u est un ascendant de v , et v est un descendant de u .

Définition 3.7 (Racine, Feuille) Une racine d'un graphe est un nœud sans parent. Une feuille d'un graphe est un nœud sans enfant.

Notations 3.1 – Parents : U_A est l'ensemble des parents du nœud A .

– **Enfants :** Y_A est l'ensemble des enfants du nœud A .

– **Descendants :** X_A est l'ensemble des descendants du nœud A .

– **non-Descendants :** Z_A est l'ensemble des non-descendants du nœud A , tel que $Z_A = V - X_A$.

2. Les types de graphes

Selon la nature des liens qui existent entre les nœuds, plusieurs types de graphes ont été définis :

- **DAG** : (Directed Acyclic Graph) est un graphe orienté acyclique (sans cycle). Les figures (3.2c), (3.2d) et (3.2e) sont des DAG. La figure (3.2a) ne représente pas un DAG.
- **DAG simplement connecté** : c'est un DAG ne contenant pas de boucles. Un DAG est dit un arbre simple si chaque nœud a au plus un parent (voir figure (3.2c)). Autrement, il s'agit d'un polyarbre (voir figure (3.2d)).
- **DAG avec boucle** : est un DAG qui contient au moins une boucle. La figure (3.2e) contient la boucle ACED.

Remarque 3.1 les graphes orientés et non orientés sont très utilisés pour représenter des connaissances. Les modèles à base de graphes non dirigés appelés les modèles de Markov permettent de représenter des relations temporelles ou spatiales [74, 75, 76]. Les graphes dirigés acycliques sont utilisés afin de représenter des relations temporelles ou causales comme dans [77, 78, 79].

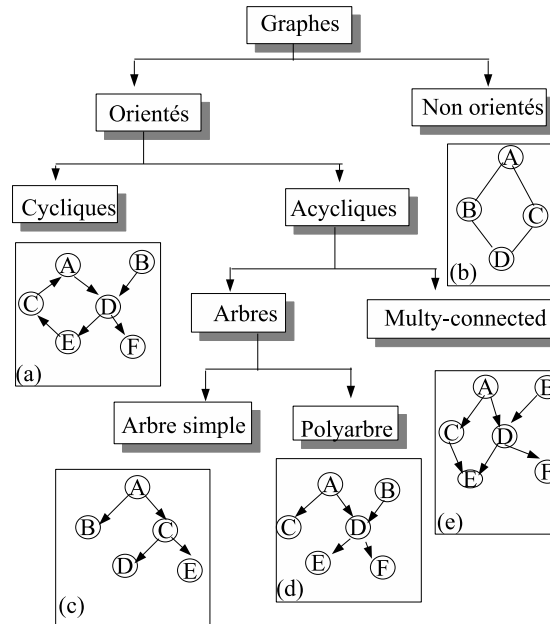


FIG. 3.2 – Différents types de graphes

3.2.1.2 Les réseaux causaux et le critère de D-Séparation

Les réseaux causaux bayésiens sont issus des travaux de [10]. Ils sont fondés sur des Graphes Acycliques Dirigés (DAG) où les nœuds représentent des variables et les arcs résument les liens causaux qui existent entre elles. Les variables représentent des événements (propositions), pouvant avoir un nombre quelconque de valeurs.

Le raisonnement sous incertitude requiert la prise en charge de la retombée de la croyance d'un événement sur les croyances des autres événements du système. Un réseau causal, justement, permet d'étudier l'influence du changement sur le degré de certitude d'une variable sur l'ensemble des variable du réseau. Cette influence est en fonction de la topologie du DAG.

Les topologies des DAG : Il existe trois types de liaisons ou de connexions dans un DAG :

Connexions en série : Considérons la situation de la figure (3.3a). La variable A a une influence sur B, qui à son tour, influe sur C. D'une manière similaire, l'évidence sur C influe sur la certitude de A à travers B. D'autres part, si l'état de B est connu, alors le canal devient bloqué. A et C deviennent alors indépendants. A et C sont alors dits d-séparées étant donné B. Ainsi, l'évidence est transmise à travers la connexion en série jusqu'à ce que l'état de la variable en connexion soit connu.

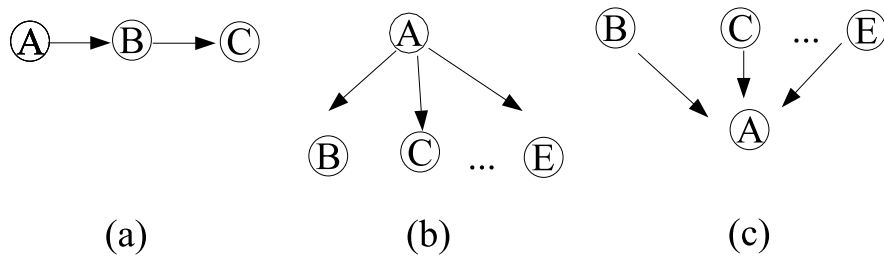
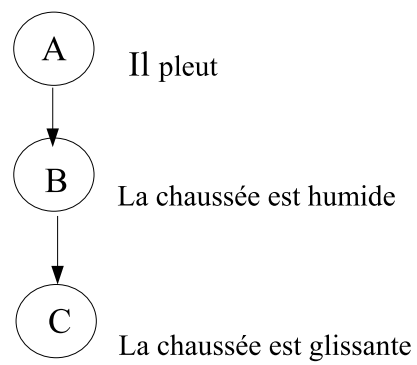
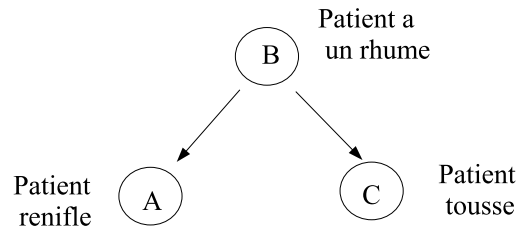


FIG. 3.3 – Les types de liaisons dans un DAG



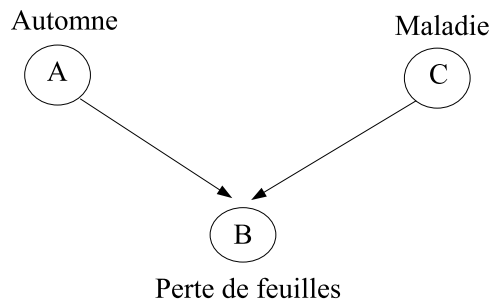
Une fois que B est connue, les variables A et C deviennent indépendantes

FIG. 3.4 – Exemple de liaisons en série



A et C ne sont dépendantes que si B est inconnue

FIG. 3.5 – Exemple de liaisons divergentes



A et C ne sont dépendantes que si B est connue

FIG. 3.6 – Exemple de liaisons convergentes

Connexions divergentes : La situation de la figure (3.3b) est une connexion divergente. L'influence passe entre l'ensemble des fils de A, jusqu'à ce que A soit connue. Ainsi, B,C,...,E sont d-séparées étant donné A. L'évidence peut être transmise à travers des connexions divergentes jusqu'à ce qu'elle soit instanciée.

Connexions convergentes : La figure (3.3c) illustre une connexion convergente. Si nous ne connaissons rien sur A excepté les connaissances inférées de ses parents B,...,E alors ces derniers sont séparés. Si A change alors la communication entre ses parents est alors ouverte.

Indépendance conditionnelle dans les DAGs : le critère de D-séparation La structure d'un DAG encode les relations d'indépendance et ceci sans tenir compte des valuations numériques. Pearl, Verma et Geiger [10, 80, 81, 82] ont examiné le problème de définir avec exactitude les types d'indépendance impliquées par une structure de

DAG dans un réseau causal. Cette notion d'indépendance est à la base du critère de d-séparation. Le critère de d-séparation est une propriété purement graphique permettant de préciser dans quelles condition une information peut être traitée localement sans perturber l'ensemble du graphe.

Ainsi, tout calcul d'incertitude dans une structure causale obéit au principe stipulant que si A et B sont d-séparées, alors une nouvelle information relative à l'une d'entre elles n'influe pas sur la certitude de l'autre. Cette caractéristique est très utile lors de la propagation d'une évidence e à travers un DAG.

La règle générale du critère de d-séparation est la suivante :

Définition 3.8 (Nœuds d-séparés par un ensemble de nœuds) Soit $G=(V,E)$ un DAG. Soient $Z \subseteq V$ et A et B deux nœuds $\in V-Z$. A et B sont dits d-séparés par Z si chaque chemin entre A et B est bloqué par Z. Plus précisément, si :

- La connexion est en série ou elle est divergente et l'état de Z est connu,
 - la connexion est convergente et ni Z ni aucun de ses descendants ne sont connus.
- Cette caractéristique est notée par :

$$\langle A \mid Z \mid B \rangle_G$$

Cette définition peut être étendue aux ensembles :

Définition 3.9 (ensembles) Soit $G=(V,E)$ un DAG. Soient X,Y,Z des sous-ensembles disjoints de V. X et Y sont d-séparés par Z si chaque nœud $X_i \in X$ et $Y_i \in Y$, X_i, Y_i sont d-séparés par Z.

Définition 3.10 (Chemin d-séparé par un ensemble de nœuds) Un chemin p est dit d-séparé par un ensemble Z de nœuds si et seulement si :

1. p contient une séquence en série $i \rightarrow m \rightarrow j$ ou une divergence $i \leftarrow m \rightarrow j$, telle que $m \in Z$, ou bien
2. p contient une convergence $i \rightarrow m \leftarrow j$ telle que $m \notin Z$ et telle qu'aucun descendant de m n'appartienne à Z.

Exemple 3.2 Dans le DAG représenté par la figure 3.7, les relations de d-séparation suivantes sont détectées :

- $X=\{X_2\}$ et $Y=\{X_3\}$ sont d-séparés par $Z=\{X_1\}$ car les deux chemins connectant X_2 à X_3 sont bloqués par Z
- Le chemin $X_2 \leftarrow X_1 \rightarrow X_3$ est bloqué car il s'agit d'une divergence dans laquelle le nœud milieu X_1 appartient à Z.
- Le chemin $X_2 \rightarrow X_4 \leftarrow X_3$ est bloqué car il s'agit d'une convergence dans laquelle le nœud X_4 ainsi que tous ses descendants n'appartiennent pas à Z.

Par contre $Z'=\{X_1, X_5\}$ ne d-sépare pas X_2 et X_3 car le chemin $X_2 \rightarrow X_4 \leftarrow X_3$ n'est pas bloqué par Z' vu que X_5 , qui est un descendant du nœud milieu X_4 , appartient à Z'. Ainsi, le fait de connaître l'effet X_5 rend ses causes X_2 et X_3 dépendantes.

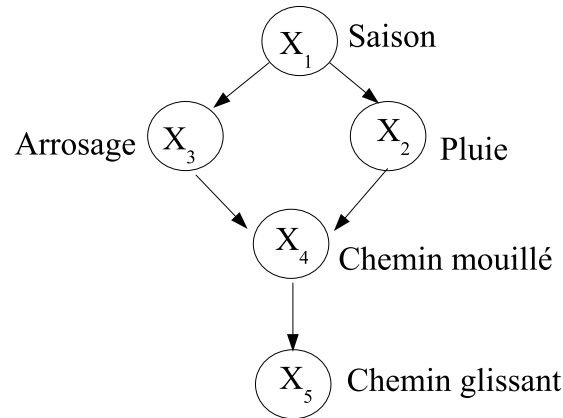


FIG. 3.7 – Réseau bayésien représentant les dépendances entre cinq variables

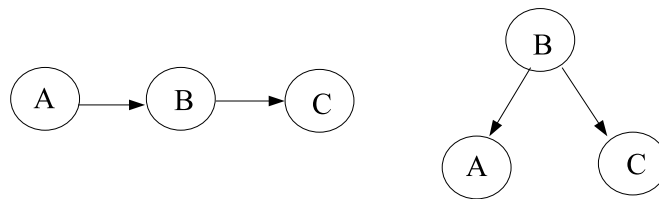


FIG. 3.8 – Topologies où les nœuds A et C sont conditionnellement indépendants étant donné B

Remarque 3.2 Le concept de *d-séparation* est utilisé afin de faciliter la mise à jour des probabilités dans un réseau causal après l'arrivée d'une nouvelle connaissance. Cette opération est aussi connue sous les noms d'*inférence* ou de *propagation*. Néanmoins, comme nous le verrons dans les sections suivantes, même en tenant compte de ce critère, le calcul d'inférence probabiliste est en général NP-Complet.

3.2.1.3 L'indépendance conditionnelle dans les DAGs

Le concept d'indépendance conditionnelle reflète la notion de blocage d'influence entre variables.

Définition 3.11 X et Y sont indépendants conditionnellement de Z , noté $X \perp Y \mid Z$, si :

$$P(X \mid Z, Y) = P(X \mid Z)$$

Remarque 3.3 La notion d'indépendance conditionnelle apparaît dans les cas de connexions en série et de connexions divergentes (voir figure 3.8).

Le lien entre le critère de *d-séparation* et d'indépendance conditionnelle est établi grâce au théorème de Verma et Pearl [83, 84] :

Théorème 3.1 Soit $G=(V,E)$ un DAG. Soient $X \subset V$, $Y \subset V$ et $Z \subset V$, trois sous ensembles de nœuds. Si X et Y sont d -séparés dans G par Z , alors X et Y sont indépendants conditionnellement à Z , c.a.d, $X \perp Y \mid Z$.

3.2.1.4 Modélisation d'un réseau Bayésien

Un réseau Bayésien permet de représenter un ensemble de variables aléatoires pour lesquelles nous connaissons un certain nombre de relations de dépendances et où une distribution de probabilités définie sur l'ensemble des variables est disponible. Formellement :

Définition 3.12 Soit $G=(V,E)$ un DAG. Soit P la distribution de probabilités jointe sur l'ensemble des éléments de V . (G,P) est dit un réseau Bayésien si chaque variable $A \in V$ est conditionnellement indépendante de ses non-descendants Z_A étant donné ses parents (U_A) .

Il est à noter qu'en pratique, la distribution de probabilités jointe sur l'ensemble des variables n'est quasiment jamais spécifiée. Ainsi, ce sont en général des distributions de probabilités locales qui sont utilisées.

Spécification d'un réseau Bayésien : La construction d'un modèle Bayésien se fait en deux phases :

1. **Spécification de la composante qualitative :** Cette phase consiste à définir les relations d'influence pouvant exister entre les variables prises deux à deux. Ces influences seront représentées graphiquement par des arcs.
2. **Spécification de la composante quantitative :** Cette phase consiste à quantifier numériquement les différents liens qui existent dans le DAG par des probabilités conditionnelles de chaque variable A dans le contexte de ses parents (U_A) comme suit :
 - Si $(U_A) = \emptyset$ (A est un nœud racine) alors, il s'agit de spécifier les probabilités a priori relatives aux différentes instances de la variable A , tout en respectant la condition de normalisation qui stipule que :

$$\sum_a P(a) = 1$$

- Si $(U_A) \neq \emptyset$ alors il faudrait spécifier les probabilités conditionnelles des différentes instances a de A dans le contexte des différentes instances de ses parents u_A telles que :

$$\sum_a P(a \mid u_A) = 1$$

Ces spécifications numériques sont souvent données par des experts du domaine modélisé.

Exemple 3.3 Soit le DAG illustré dans la figure (3.9). Les distributions a priori sont données par la table (3.1) et les distributions conditionnelles sont données par les tables (3.2) et (3.3).

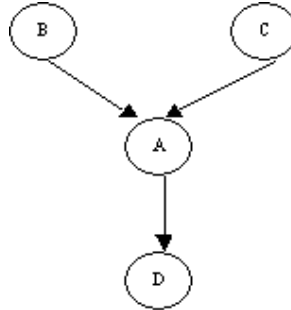


FIG. 3.9 – Exemple d'un réseau Bayésien

TAB. 3.1 – Les probabilités a priori P(B) et P(C)

B	P(B)	C	P(C)
b	.7	c	.5
$\neg b$.3	$\neg c$.5

La règle de chaînage : La règle de chaînage permet une représentation compacte de la distribution de probabilité associée à l'ensemble des variables. En effet, si nous considérons l'ensemble V , la représentation en mémoire de $P(V)$ serait très fastidieuse. Par exemple si l'ensemble V est composé de 30 variables prenant chacune trois valeurs, alors il faudrait mémoriser plus de 10^{14} valeurs.

Afin de palier à cet inconvénient, la règle de chaînage permet, justement, de calculer $P(V)$ à partir des probabilités conditionnelles spécifiées dans le réseau :

Définition 3.13 (Règle de chaînage) *La distribution de probabilités globale jointe à travers un ensemble de variables $V = \{A_1, \dots, A_N\}$ est exprimée par le produit des probabilités conditionnelles via la règle de chaînage suivante :*

$$P(A_1, \dots, A_N) = \prod_{i=1..N} P(A_i | U_{A_i}) \quad (3.1)$$

où U_{A_i} représentent les parents de A_i .

Exemple 3.4 *Soit le réseau causal de la figure 3.9 auquel est associé la table des probabilités a priori (table 3.1) et les tables des probabilités conditionnelles (tables 3.2 et 3.3).*

La distribution des probabilités jointe est définie par :

$$\forall a_i, b_i, c_i, d_i, P(a_i \wedge b_i \wedge c_i \wedge d_i) = P(a_i | b_i \wedge c_i) * P(d_i | a_i) * P(b_i) * P(c_i)$$

où a_i, b_i, c_i, d_i représentent respectivement les instances des variables A, B, C, D .

Par exemple $P(\neg a \wedge b \wedge \neg c \wedge d) = P(\neg a | b \wedge \neg c) * P(d | \neg a) * P(b) * P(\neg c) = .4 * .1 * .7 * .5 = .014$.

TAB. 3.2 – Les probabilités conditionnelles $P(A | B \wedge C)$

A	B	C	$P(A B \wedge C)$	A	B	C	$P(A B \wedge C)$
a	b	c	.8	$\neg a$	b	c	.2
a	b	$\neg c$.6	$\neg a$	b	$\neg c$.4
a	$\neg b$	c	.9	$\neg a$	$\neg b$	c	.1
a	$\neg b$	$\neg c$.3	$\neg a$	$\neg b$	$\neg c$.7

TAB. 3.3 – Les probabilités conditionnelles $P(D | A)$

D	A	$P(D A)$
d	a	.6
d	$\neg a$.1
$\neg d$	a	.4
$\neg d$	$\neg a$.9

3.2.2 Algorithmes de propagation dans les réseaux causaux bayésiens

3.2.2.1 Introduction

Pour les réseaux causaux probabilistes, l'inférence consiste à propager une évidence (une nouvelle information) à travers un réseau et de voir son impact. La propagation entraîne en fait, une mise à jour des distributions de probabilités associées aux variables après l'arrivée de l'évidence.

Il existe principalement deux types d'algorithmes d'inférence :

1. **Algorithmes exacts** : Ces algorithmes calculent les distributions de probabilités associées aux nœuds exactement. L'algorithme exact le plus fondamental est celui proposé par Kim et Pearl [85] et Pearl [10, 86]. L'impact de chaque évidence est perçu comme une perturbation se propageant en parallèle à travers le réseau, via un mécanisme de passage des messages entre les variables voisines. Dans le cas d'un graphe causal simplement connecté, l'algorithme converge rapidement. Néanmoins, pour le cas d'un réseau avec boucles, l'algorithme ne donnera pas les résultats escomptés vu la présence de boucles. Pour cela d'autres algorithmes, tel que celui de l'arbre de jonction [11], ont été développés nécessitant la transformation du graphe initial en un arbre de jonction formé d'un ensemble de groupes de variables.
2. **Algorithmes approximatifs** : Ce type d'algorithme est utilisé dans le cas où il n'existe pas d'algorithmes exacts ou encore lorsque ceux-ci présentent une complexité calculatoire intense. En effet, les algorithmes exacts sont NP-Complet [87] excepté pour les polyarbres.

L'idée de base de ces méthodes est de générer un échantillon à partir de la distribution de probabilités jointe des variables, puis d'utiliser l'échantillon généré pour calculer les valeurs de probabilités d'événement après prise en compte de l'évidence [88].

Les méthodes de propagation approximatives peuvent être classées en deux types :

les méthodes de simulation stochastiques et les méthodes de recherche déterministes. Les méthodes du premier type génèrent l'échantillon à partir de la distribution de probabilités jointe en utilisant des mécanismes aléatoires. Pour le deuxième type de méthodes, l'échantillon est généré d'une manière systématique [88]. Parmi les algorithmes approximatifs, nous distinguons *loopy belief propagation* [89].

Dans ce qui suit, nous présentons les algorithmes exacts les plus utilisés selon la structure du DAG (polyarbre, réseau causal avec boucles).

3.2.2.2 La propagation dans les Polyarbres

Nous allons présenter dans cette section une version centralisée de l'algorithme de Pearl [85, 10] pour la propagation probabiliste dans un polyarbre proposé dans [90]. Cet algorithme a la particularité d'être de complexité polynômiale.

Le principe de propagation : Dans un polyarbre, deux nœuds sont reliés par un chemin unique. Ceci induit que chaque nœud X_i divise le polyarbre en deux parties : l'une incluant les nœuds accessibles à partir des parents du nœud X_i et l'autre inclut les nœuds accessibles à partir des enfants. Afin de faciliter le calcul, l'évidence E est décomposée en deux sous ensembles disjoints E_A^- et E_A^+ (voir la figure (3.10)) correspondants aux deux sous graphes séparés par le nœud A dans le polyarbre. Ces deux ensembles sont définis par :

- $E_A^- = \{E_A, E_{AY_1}^-, \dots, E_{AY_m}^-\}$ où E_A correspond au nœud A et $E_{AY_j}^-$ correspond à l'instanciation des variables en-dessous de l'arc de A à Y_j .
- $E_A^+ = \{E_{U_1A}^+, \dots, E_{U_nA}^+\}$ où $E_{U_iA}^+$ correspond à l'instanciation des variables au-dessus de l'arc de U_i à A .

e_A^+ et e_A^- représentent respectivement l'évidence attachée au nœud dans E_A^+ et E_A^- . e_A dénote l'évidence attachée au nœud A .

Exemple 3.5 *Considérons le DAG représenté par la figure (3.11). Supposons que les variables binaires suivantes sont instanciées : $A = \neg a, B = b, D = \neg d, G = 1, K = \neg k$.*

Alors, $E = \{A, B, D, G, K\}$, $E_A^+ = \{E_{BA}^-, E_{CA}^-, E_{EA}^-\} = \{B, D\}$, $E_A^- = \{E_A, E_{AG}^-, E_{AH}^-\} = \{A, G, K\}$, $e = \{\neg a, b, \neg d, g, \neg k\}$, $e_A^+ = \{e_{BA}^-, e_{CA}^-, e_{EA}^-\} = \{b, \neg d\}$, $e_A^- = \{e_A, e_{AG}^-, e_{AH}^-\} = \{\neg a, g, \neg k\}$

La propagation de l'évidence est effectuée d'une manière efficace en combinant les informations provenant de différents sous-graphes en utilisant la technique de passage de message (calculs locaux) d'un sous graphe vers un autre. Plus précisément, les communications entre les variables se font en utilisant deux types de messages :

- λ message pour transmettre l'information des enfants vers les parents,
- μ message¹ pour transmettre l'information des parents vers les enfants.

D'une manière formelle [10], chaque nœud A sera ainsi caractérisé par sa probabilité conditionnelle en se basant sur une évidence totale e , notée $\text{Bel}(A)$, qui sera fonction

¹Ce message est noté par μ au lieu de π (comme dans la littérature), afin de ne pas confondre avec la distribution de possibilités.

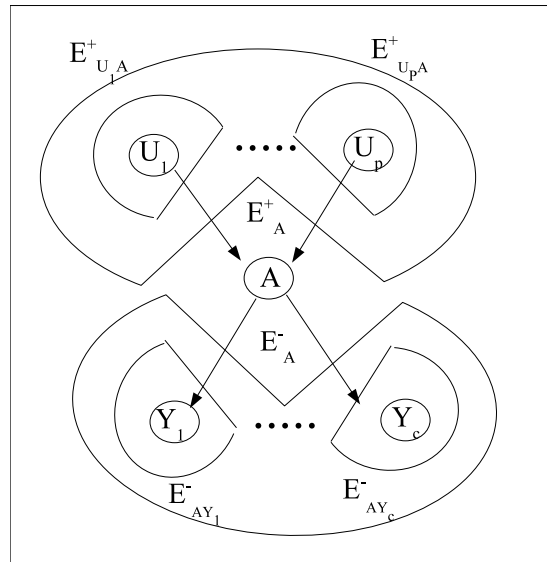


FIG. 3.10 – Cautionnement de l'évidence E en sous-ensembles associés aux parents et aux enfants d'un nœud typique A

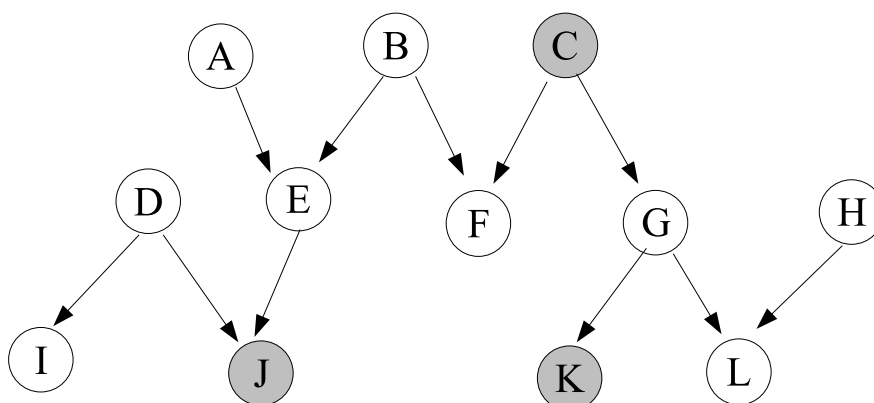


FIG. 3.11 – Exemple d'un polyarbre

des valeurs λ et μ relatives à ce nœud.

La valeur $\lambda(A)$ est calculée à partir des λ messages reçus à partir des enfants du nœud A et la valeurs $\mu(A)$ est calculée à partir des μ messages reçus à partir des parents du nœud A.

Ainsi, $\forall a \in D_A$:

- La mesure de probabilité courante de chaque instance a basée sur l'évidence totale e est définie par :

$$Bel(a) = P(a | e) = \alpha * \lambda(a) * \mu(a) \quad (3.2)$$

où $\alpha = \frac{1}{\sum_a Bel(a)}$ est le facteur de normalisation

- La valeur λ associée à chaque instance $a \in D_A$ est définie par :

$$\lambda(a) = P(a | e_A^-) = \lambda_A(a) * \prod_{j=1}^m \lambda_{Y_j}(a) \quad (3.3)$$

où $\lambda_A(a)$ dénote l'évidence totale relative au nœud A telle que :

$$\lambda_A(a) = \begin{cases} 0 & \text{si } e_A \neq a \text{ (A est instanciée à } a \text{ avec } e_A \neq a) \\ 1 & \text{sinon (A est instanciée à } a \text{ avec } e_A = a \text{ ou A n'est pas instanciée)} \end{cases} \quad (3.4)$$

- La valeur μ associée à chaque instance $a \in D_A$ est définie par :

$$\mu(a) = P(a | e_A^+) = \sum_u P(a | u) * \prod_{i=1}^n \mu_A(u_i) \quad (3.5)$$

- Le message λ de A vers ses parents U_i , ($i \in \{1, \dots, n\}$) où $U_i = u_i$ est défini par :

$$\lambda_A(u_i) = P(e_{U_i A}^- | u_i) = \beta * \sum_a \lambda(a) \left[\sum_{u_k: k \neq i} P(a | u) * \prod_{k \neq i} \mu_A(u_k) \right] \quad (3.6)$$

où β est une constante de normalisation.

Remarque 3.4 Dans le cas où chaque nœud dans le graphe a un parent unique alors la valeur λ est simplifiée comme suit :

$$\lambda_A(u_i) = \beta * \sum_a \lambda(a) * P(a | u)$$

- Le message μ de A à ses enfants Y_j , ($j \in \{1, \dots, m\}$) où $A=a$ est défini par :

$$\mu_{Y_j}(a) = P(a | e_{+AY_j}) = \alpha * \lambda_A(a) * \prod_{i=1 \dots m, i \neq j} \lambda_{Y_i}(a) * \mu(a) \quad (3.7)$$

Algorithme de propagation L'algorithme de propagation de base pour les polyarbres a été proposé par Pearl dans [10]. Néanmoins, Peot et Shachter [90] ont proposé une version révisée de l'algorithme sans modifier sa complexité. Cette version converge au bout de deux itérations.

Algorithme 3.1 : Propagation dans un polyarbre

Début

Choix de la racine de propagation Cette phase consiste à choisir un nœud arbitraire parmi l'ensemble le plus restreint des nœuds connectés contenant les nœuds observés noté \mathcal{S} .

Initialisation Dans cette étape, tous les λ et μ valeurs et messages sont initialisés comme suit :

1. Pour chaque nœud A racine, la valeur $\mu(a)$ est initialisée à $P(a)$, $\forall a \in D_A$.
2. Pour chaque nœud A observé,
 - Si A est instanciée à a , alors $\lambda(a)$ est initialisée à 1,
 - sinon $\lambda(a)$ est initialisée au degré 0.

Collecte de l'évidence Lors de cette phase, pour chaque nœud dans \mathcal{S} :

- ses valeurs λ et μ sont calculées en utilisant respectivement les équations (3.3) et (3.5) permettant ainsi le passage de message à son nœud adjacent dans la direction du pivot.

Si ce nœud est un parent alors

- le message est calculé en utilisant l'équation (3.6) **Sinon**

Le message est calculé en utilisant l'équation (3.7)

- la collecte de l'évidence débute avec le nœud le plus lointain du pivot dans \mathcal{S} .

Distribution de l'évidence Lors de cette phase, chaque nœud :

- calcule ses valeurs λ et μ en utilisant respectivement les équations (3.3) et (3.5)
- Passe les messages aux nœuds adjacents en commençant du pivot lui même jusqu'à atteindre les feuilles.

Le calcul des messages à envoyer aux nœuds adjacents dépend de son type :

Si le nœud traité est un parent alors

- le message est calculé en utilisant l'équation (3.6)

Sinon

Le message est calculé en utilisant l'équation (3.7)

Marginalisation Pour chaque nœud A , $Bel(A) = P(a | e)$ est calculé en utilisant l'équation(3.2).

Fin

Exemple 3.6 Considérons le DAG donné par la figure (3.11). Supposons que de nouvelles informations ont été obtenues pour les nœuds C, J et K . Ainsi, l'ensemble $\mathcal{S} = \{J, E, B, F, C, G, K\}$. Supposons que le nœud pivot est F . La figure (3.12a) représente le diagramme correspondant. Les différents passages de messages sont résumés par la figure (3.12b).

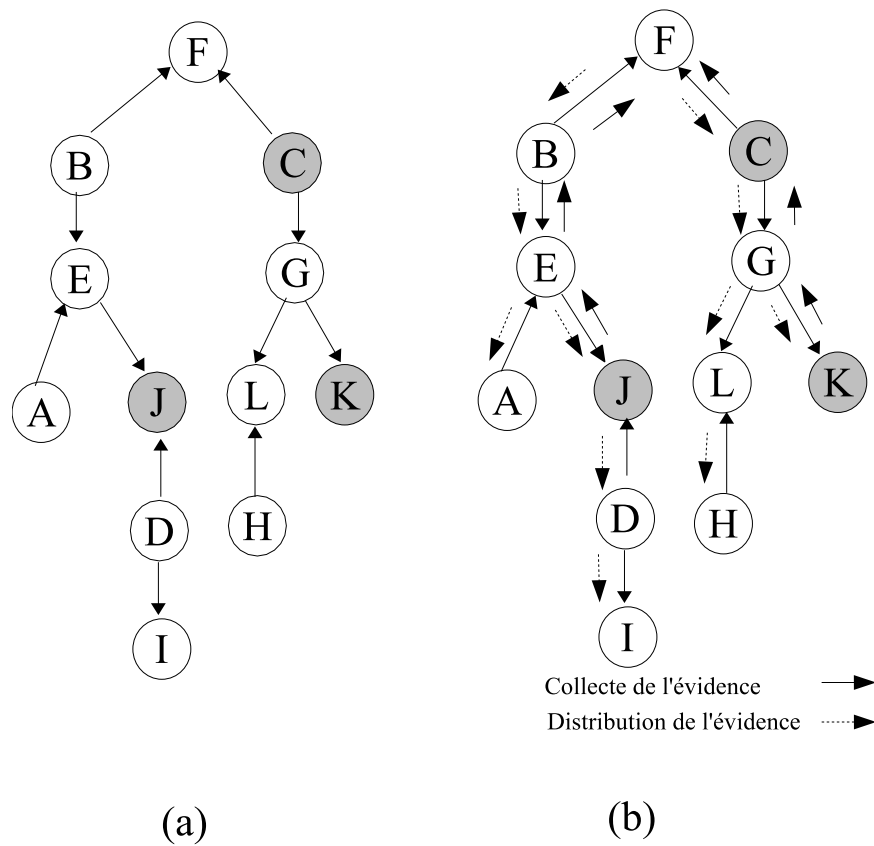


FIG. 3.12 – Exemple de propagation dans un DAG simplement connecté

3.2.2.3 La propagation dans les graphes à connexions multiples

L'algorithme de propagation pour les polyarbres décrit par Pearl [10] n'est pas adapté pour les graphes à connexions multiples. Dans [12], les auteurs proposent un algorithme d'inférence exacte plus efficace et générale (pouvant être appliqué aussi bien pour les polyarbres que pour les graphes à connexions multiples), dit algorithme de l'arbre de jonction, qui consiste à compiler le graphe original en une structure d'arbre secondaire appelée arbre de jonction, puis de réaliser la transmission de message sur cet arbre. Cet algorithme fût amélioré par la suite par Jensen [11]. Ainsi, les experts formalisent leurs connaissances en utilisant les réseaux causaux et la propagation s'effectue sur une structure secondaire (arbre de jonction) caractérisée par :

- Un arbre de jonction (Junction Tree) \mathcal{JT} est formé d'un ensemble de nœuds. Chaque nœud est composé par un ensemble de variables ou *clusters* notés C_i .
- Chaque arc est étiqueté par l'intersection des deux clusters adjacents C_i et C_j appelé séparateur et noté par S_{ij} .
- A chaque cluster C_i (respectivement séparateur S_{ij}) de \mathcal{JT} est affecté une distribution jointe locale relative aux variables du cluster (respectivement au séparateur), appelée potentiel, noté par ψ_{C_i} (respectivement $\psi_{S_{ij}}$).
- Une distribution de probabilités jointe globale unique est associée à un arbre de jonction \mathcal{JT} . Elle est définie par :

Définition 3.14 *La distribution jointe associée à un arbre de jonction \mathcal{JT} est :*

$$P_{\mathcal{JT}}(A_1, \dots, A_N) = \frac{\prod_{i=1}^m \psi_{C_i}}{\prod_{j=1}^{m-1} \psi_{S_{ij}}} \quad (3.8)$$

où m représente le nombre de clusters dans l'arbre de jonction.

- **Définition 3.15 (Lien consistant)** *Soient C_i et C_j deux clusters adjacents dans un arbre de jonction \mathcal{JT} et S_{ij} le séparateur associé. Le lien entre C_i et C_j est dit consistant ou stable si :*

$$\sum_{C_i \setminus S_{ij}} \psi_{C_i} = \psi_{S_{ij}} = \sum_{C_j \setminus S_{ij}} \psi_{C_j} \quad (3.9)$$

Si tous les liens dans l'arbre de jonction sont consistants, alors l'arbre de jonction est dit globalement consistant.

Si un arbre de jonction est globalement consistant, alors le potentiel associé à chaque cluster C_i doit satisfaire :

$$\psi_{C_i} = P(C_i) \quad (3.10)$$

- La distribution de probabilité associée à chaque variable A contenue dans un cluster C_i est obtenue en marginalisant ψ_{C_i} sur A comme suit :

$$P(A) = \sum_{C_i \setminus A} \psi_{C_i} \quad (3.11)$$

Principalement, l'algorithme d'inférence exact pour les graphes à connexions multiples se comporte de la manière suivante :

- **La phase de construction** : Elle nécessite un ensemble de sous-étapes afin de transformer le graphe initial en un arbre de jonction, dont les nœuds sont des **clusters** (regroupements des nœuds du graphe initial). Cette transformation est nécessaire pour :

- Éliminer les boucles du graphe initial,
- Optimiser l'algorithme d'inférence.

Elle s'effectue en trois étapes :

1. Moralisation du DAG initial \mathcal{G} ,
 2. Triangulation du graphe moral,
 3. Création de l'arbre de jonction à partir du graphe moral triangulé.
- **La phase de propagation** : C'est la phase de calcul de l'ensemble des distributions de probabilités du réseau après l'avènement de nouvelles informations. Cette phase s'effectue par le biais de passage de messages entre les nœuds de l'arbre de jonction.

La phase de construction comporte les étapes suivantes :

Les transformations graphiques : Cette étape consiste donc à construire l'arbre de jonction associé au réseau causal original \mathcal{G} .

Étape 1 : Moralisation : Elle consiste à marier deux à deux les parents de chaque variable en les reliant par un arc non-dirigé. A partir d'un graphe \mathcal{G} , le graphe moral associé noté $\mathcal{G}_{\mathcal{M}}$ est obtenu par la procédure suivante :

Algorithme 3.2 : Construction du graphe moral [12, 91]

Début

Associer au DAG initial \mathcal{G} un graphe non dirigé en éliminant les directions de tous les arcs,

Construire le graphe moral $\mathcal{G}_{\mathcal{M}}$ à partir de \mathcal{G} en reliant les parents de chaque nœud en rajoutant des arcs non dirigés.

Fin

Exemple 3.7 La figure 3.14 représente le graphe moral associé au DAG de la figure 3.13. Les arcs en pointillés représentent les arcs qui ont été ajoutés.

Étape 2 : Triangulation du graphe moral :

En général, il existe plusieurs façons de trianguler un graphe moral. La tâche de trouver la triangulation la plus optimale est NP-complet [92, 93, 87]. Néanmoins, plusieurs heuristiques ont été développées afin de réduire le coût induit par l'inférence, parmi lesquelles se trouve le critère de sélection de nœuds qui est une heuristique (induisant une complexité polynômial) produisant une triangulation de haute qualité [94].

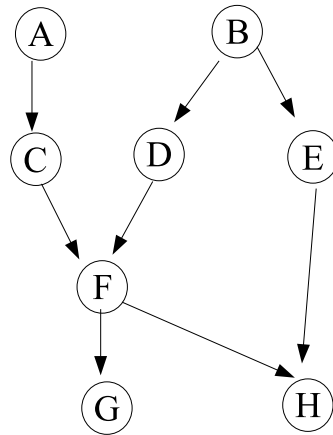


FIG. 3.13 – Un graphe acyclique dirigé (DAG)

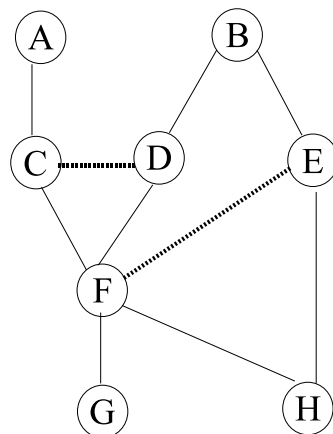


FIG. 3.14 – Graphe moral associé au DAG de la figure 3.13

La procédure de triangulation qui suit permet également d'identifier l'ensemble des clusters ou cliques (noté *cluster-set*) en utilisant l'algorithme de Golumbic [95] qui garantit qu'aucun cluster n'est inclus dans un autre cluster plus grand.

Algorithme 3.3 : Triangulation du graphe moral et identification des clusters [94]

Debut

Cluster-set := \emptyset ;

$\mathcal{G}'_{\mathcal{M}}$:= $\mathcal{G}_{\mathcal{M}}$;

Tantque il existe des nœuds dans $\mathcal{G}'_{\mathcal{M}}$ **Faire**

Si il existe un nœud A tel que tous ses nœuds adjacents sont connectés

Alors

Créer un cluster C contenant A ainsi que ses nœuds adjacents ;

Supprimer A de $\mathcal{G}'_{\mathcal{M}}$

Sinon

Sélectionner un nœud A ayant le nombre de nœuds adjacents le plus petit ;

Ajouter des arcs afin de connecter l'ensemble des nœuds adjacents ;

Pour chaque arc ajouté à $\mathcal{G}'_{\mathcal{M}}$, ajouter l'arc correspondant à $\mathcal{G}_{\mathcal{M}}$;

Créer un cluster C contenant A ainsi que ses nœuds adjacents ;

Supprimer A de $\mathcal{G}'_{\mathcal{M}}$

Si $C \notin \text{cluster-set}$ **Alors** *cluster-set* := *cluster-set* $\cup \{C\}$;

Fin

Exemple 3.8 La figure (3.15) représente le graphe triangulé obtenu à partir du graphe moral de la figure (3.14). L'arc en pointillé représente l'arc ajouté par la procédure de triangulation.

L'ordre d'élimination des variables ainsi que la construction des cliques sont illustrés par la table (3.4).

Étape 3 : Construction de l'arbre de jonction optimal : Afin de construire l'arbre de jonction, il faudrait relier les différents clusters générés par la procédure précédente. Cette tâche est réalisée en insérant, itérativement, des séparateurs entre des paires de cliques, jusqu'à ce que les cliques soient connectées par $m-1$ séparateurs, où m représente le nombre de clusters [96].

La procédure correspondante est donnée en deux parties : la première partie consiste à présenter la construction de l'arbre de jonction et la seconde étape est relative aux choix appropriés des séparateurs.

TAB. 3.4 – L'ordre d'élimination des variables

Nœuds éliminés	Clusters induits	Arcs ajoutés	Cluster-set
A	AC	rien	{AC}
G	GH	rien	{AC, GF}
C	CDF	rien	{AC, GF, CDF}
H	EFH	rien	{AC, GF, CDF, EFH}
B	BDE	(D,E)	{AC, GF, CDF, EFH, BDE}
D	DEF	rien	{AC, GF, CDF, EFH, BDE, DEF}
E	EF	rien	{AC, GF, CDF, EFH, BDE, DEF} (EF \subseteq DEF)
F	F	rien	{AC, GF, CDF, EFH, BDE, DEF} (F \subseteq DEF)

1. **Construction de l'arbre de jonction optimal :** La procedure suivante permet de construire un arbre de jonction optimal à partir d'un ensemble de m clusters (cluster-set) et ceci en les reliant par un ensemble de séparateurs (Separator-set) :

Algorithme 3.4 : Construction d'un arbre de jonction optimal[97]

Debut

Pour $i := 1$ *à* m **Faire**

Separator-set := \emptyset ;

C_i := *Cluster-set*[i] ;

Pour $j := 1$ *à* $(m-1)$ **Faire**

C_j := *Cluster-set*[j] ;

Créer un séparateur candidat S_{ij} pour les deux clusters C_i et C_j ;

Insérer S_{ij} dans *Separator-set* ;

Sélectionner un séparateur S_{ij} à partir de *Separator-set*, en accord avec le critère de selection spécifié ci-après ;

Insérer le séparateur S_{ij} entre le cluster C_i et le cluster C_j

Fin

2. **Choix des séparateurs appropriés :** Dans le but de décrire comment s'effectue la sélection du séparateur, les notions suivantes sont définies :
 - Le poids associé à un nœud A est la taille de son domaine représentée par le nombre de ses valeurs.
 - Le poids associé à un cluster est le produit des poids des variables qui le constituent.
 - La masse d'un séparateur S_{ij} est le nombre de variables qu'il contient ou c'est le nombre de variables contenues dans $C_i \cap C_j$.
 - Le coût d'un séparateur S_{ij} est le produit des poids de ses variables.
Ainsi, le choix du séparateur dans l'algorithme précédent se fait comme suit :

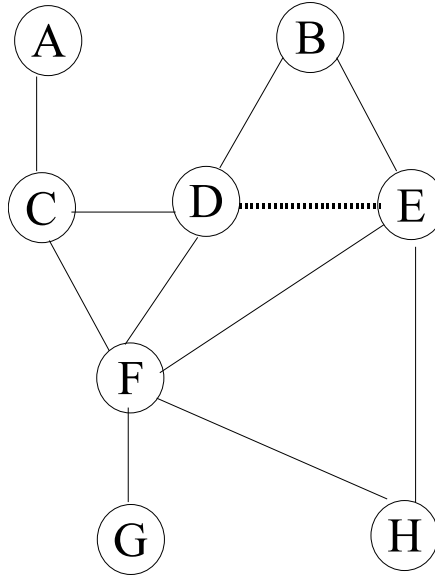


FIG. 3.15 – Triangulation du graphe moral de la figure 3.14

- Le séparateur candidat doit avoir la masse la plus élevée.
- Dans le cas où plusieurs séparateurs ont la même masse, le temps de l'inférence peut être optimisé en choisissant le séparateur induisant le moindre coût.

Exemple 3.9 *Considérons l'ensemble de clusters suivant : $cluster-set = \{AC, GF, CDF, EFH, BDE, DEF\}$ obtenu à partir de la figure (3.15) de l'exemple précédent et supposons que les variables sont binaires. L'application du critère de sélection est résumée dans la table 3.5. Ainsi, l'ensemble des séparateurs associés aux différentes cliques est : $separator-set = \{C, DF, DE, EF, F\}$. L'arbre de jonction associé est représenté par la figure (3.16).*

La phase de propagation : Après avoir construit l'arbre de jonction optimal associé au graphe causal original, l'étape suivante consiste à calculer la composante numérique associée à l'arbre de jonction. Le processus général de propagation dans les graphes à connexions multiples sans variables d'observation est illustré par la figure (3.17). Il consiste à calculer la distribution de probabilités $P(A)$ associée à une variable A dans le contexte d'aucune évidence [97].

Notations 3.2 *Le potentiel associé à un cluster C_i lors de la phase d'initialisation (respectivement lors de consistance globale) est noté $\psi_{C_i}^I$ (respectivement $\psi_{C_i}^C$.)*

D'une manière détaillée, la propagation s'effectue en trois phases :

1. **Initialisation :** Cette étape consiste à quantifier l'arbre de jonction avec les potentiels de croyance. Le résultat est un arbre de jonction inconsistant vu que

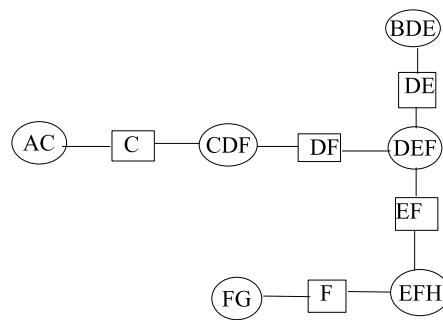


FIG. 3.16 – Arbre de jonction associé au DAG de la figure 3.13

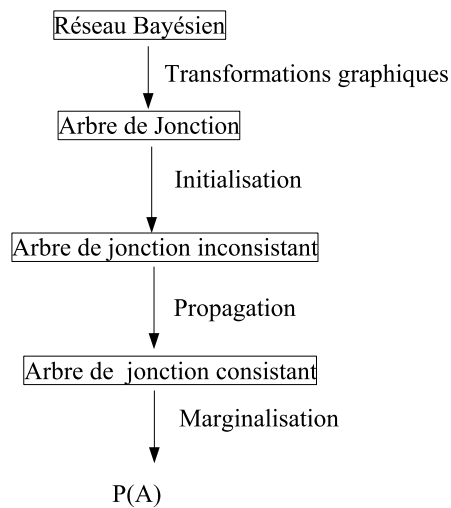


FIG. 3.17 – Le diagramme de l'algorithme de propagation sans évidence

TAB. 3.5 – Choix des séparateurs appropriés

Clusters	Séparateurs candidats	Masse	Coût sélectionnés	Séparateurs
AC,CDF	C	1	2	*
AC,BDE	\emptyset	-	-	
AC,DEF	\emptyset	-	-	
AC,EFH	\emptyset	-	-	
AC,GF	\emptyset	-	-	
CDF,BDE	D	1	2	
CDF,DEF	DF	2	4	*
CDF,EFH	F	1	2	
CDF,GF	F	1	2	
BDE,DEF	DE	2	4	*
BDE,EFH	E	1	2	
BDE,GF	\emptyset	-	-	
DEF,EFH	EF	2	4	*
DED,GF	F	1	2	
GF,EFG	F	1	2	*

les assignations initiales des potentiels ne satisfont pas la propriété de consistance locale formalisée par l'équation (3.9).

2. **Propagation globale :** Elle consiste en une série de passages de messages sur l'arbre de jonction. Ces passages de messages réarrangent les potentiels de l'arbre de jonction de telle sorte à aboutir à une structure consistante. Ainsi, le résultat de cette phase est un arbre de jonction consistant.
3. **Marginalisation :** Elle consiste à calculer, pour chaque variable d'intérêt V , la distribution de probabilités $P(A)$ à partir de l'arbre de jonction consistant. D'une manière détaillée, les étapes sont décrites comme suit :

Initialisation : La procédure suivante initialise les fonctions potentielles associées à l'arbre de jonction en utilisant les probabilités conditionnelles initiales :

Algorithme 3.5 : Initialisation Debut

Pour chaque cluster C_i **Faire** $\psi_{C_i}^I := 1$;

Pour chaque séparateur S_{ij} **Faire** $\psi_{S_{ij}}^I := 1$;

Pour chaque variable A **Faire**

Sélectionner un cluster C_i contenant $\{A\} \cup U_A$;

$\psi_{C_i}^I := \psi_{C_i}^I * P(A | U_A)$;

Fin

Les fonctions potentielles obtenues codifient la distribution probabiliste jointe associée

à l'arbre de jonction, à savoir :

$$P = P_{\mathcal{JT}} \quad (3.12)$$

où P est la distribution de possibilités du réseau causal initial obtenue en utilisant l'équation (3.1) et $P_{\mathcal{JT}}$ représente la distribution de probabilités jointe associée à l'arbre de jonction obtenue en utilisant l'équation (3.8).

Après avoir initialisé les potentiels de l'arbre de jonction, il faudrait s'assurer de la consistance locale en effectuant une propagation globale décrite dans la section suivante :

La propagation globale : Elle induit à chaque cluster de passer des messages à l'ensemble de ses clusters adjacents. Ces messages sont ordonnés afin que chaque passage de message préserve la consistance introduite par le passage de message précédent. Une fois la propagation globale achevée, chaque paire (cluster, séparateur) est consistante et l'arbre de jonction est localement consistant.

Le mécanisme de passages des messages peut être explicité en détaillant la procédure de passage de message entre deux clusters adjacents, puis en explicitant comment se fait la coordination des passages de messages :

1. **Passage de messages simples :** Soient C_i et C_j deux clusters adjacents séparés par un séparateur S_{ij} et soient ψ_{C_i}, ψ_{C_j} et $\psi_{S_{ij}}$ leurs potentiels respectifs. Les différents potentiels sont mis à jour comme suit :

- (a) Préserver le même potentiel pour C_i :

$$\psi_{C_i}^{t+1} \leftarrow \psi_{C_i}^t \quad (3.13)$$

- (b) Assigner un nouveau potentiel à S_{ij} :

$$\psi_{S_{ij}}^{t+1} \leftarrow \sum_{C_i \setminus S_{ij}} \psi_{C_i}^t \quad (3.14)$$

- (c) Assigner un nouveau potentiel à C_j :

$$\psi_{C_j}^{t+1} \leftarrow \psi_{C_j}^t * \frac{\psi_{S_{ij}}^{t+1}}{\psi_{S_{ij}}^t} \quad (3.15)$$

2. **Coordination de messages multiples :** A partir d'un arbre de jonction de n clusters, l'algorithme de propagation globale sélectionne un cluster pivot arbitraire et effectue $2*(n-1)$ passages de messages en deux phases :

- **Collecte de l'évidence :** Chaque cluster passe un message à ses clusters adjacents dans la direction du pivot, en commençant par des clusters les plus éloignés de celui-ci.
- **Distribution de l'évidence :** Chaque cluster envoie des messages à ses clusters adjacents, en s'éloignant de la direction du pivot et ceci en commençant par le pivot lui-même jusqu'à atteindre les feuilles.

Remarque 3.5 Dans le schéma de passages de messages, un cluster ne transmet un message à un voisin dans une direction, qu'après avoir reçu les messages de l'ensemble de ses nœuds (clusters) adjacents. Cette condition permet de garantir la consistance locale de l'arbre de jonction une fois la propagation globale achevée [98, 11].

La procédure de propagation globale peut être résumée comme suit :

Algorithme 3.6 : Propagation globale

Début

Selection de la racine de la propagation

Soit *Pivot*, un cluster arbitraire C_i représentant la racine de la propagation ;

Soit *Postorder*, un vecteur contenant l'ordre dans lequel les messages vont être envoyés dans la phase de la collecte de l'évidence, ;

Soit *Preorder*, un vecteur contenant l'ordre dans lequel les messages vont être envoyés dans la phase de la distribution de l'évidence ;

Collecte de l'évidence

Pour $i := 1$ **jusqu'à** $\text{lenght}(\text{Postorder})-1$ **Faire**

$C_i \leftarrow \text{Postorder}[i]$;

$C_j \leftarrow$ cluster adjacent de C_i dans *Postorder* ;

Envoyer le message de C_i vers C_j en utilisant les équations (3.13), (3.14) et (3.15) ;

Distribution de l'évidence

Pour $i := 1$ **jusqu'à** $\text{lenght}(\text{Preorder})$ **Faire**

$C_i \leftarrow \text{Preorder}[i]$;

Below \leftarrow cluster adjacents de C_i dans *Preorder* ;

Pour $j := 1$ **jusqu'à** $\text{lenght}(\text{Below})$ **Faire**

$C_j \leftarrow \text{Below}[j]$;

Envoyer le message de C_i à C_j en utilisant les équations (3.13), (3.14) et (3.15) ;

Fin

Exemple 3.10 La figure 3.18 illustre le passage de messages dans l'arbre de jonction représenté par la figure 3.16 lors de la procédure de la propagation globale.

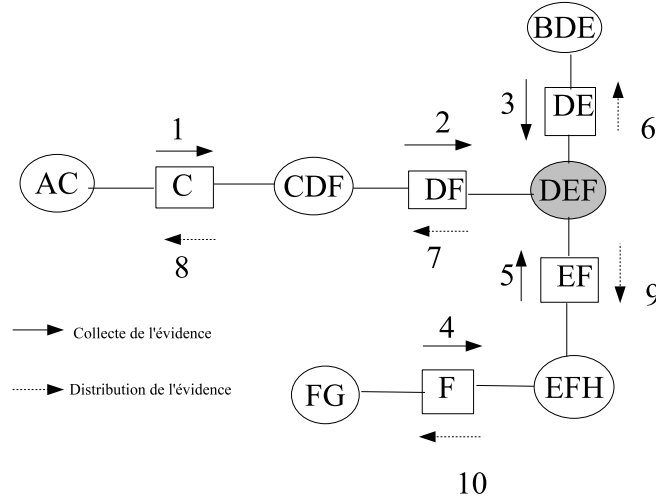


FIG. 3.18 – Passage de messages durant la propagation globale associé à l'arbre de jonction de la figure 3.16

Marginalisation : Une fois avoir obtenu un arbre de jonction consistant, la probabilité P associée à chaque variable d'intérêt A est calculée comme suit :

Algorithme 3.7 : Marginalisation

Debut

Identifier un cluster C_i contenant A ;

Calculer $P(A)$ en marginalisant $\psi_{C_i}^C$ sur A :

$$P(A) \leftarrow \sum_{C_i \setminus A} \psi_{C_i}^C ;$$

Fin

Prise en compte d'une évidence : Les procédures présentées dans la section précédente peuvent être étendues dans le cas général où il s'agirait de calculer la probabilité d'une variable A , étant donnée l'observation d'une évidence totale e , notée $P(A | e)$.

La notion la plus simple de l'évidence est l'observation : Une observation est une expression de la forme $A=a$. Le *likelihood* associé à A , noté Λ_A , est le potentiel associé à A , défini par :

$$\Lambda_A(a) = \begin{cases} 1 & \text{si } A \text{ n'est pas instanciée} \\ 1 & \text{si } A \text{ est instanciée à } a \\ 0 & \text{si } A \text{ est instanciée mais pas à } a \end{cases} \quad (3.16)$$

La procédure de propagation pour les graphes à connexions multiples avec variables d'observation est identique à la procédure précédente (sans variable d'observation) à laquelle des étapes supplémentaires sont ajoutées. Elle se résume comme suit :

1. **Initialisation :**

L'étape d'initialisation est modifiée en introduisant une étape additionnelle qui consiste à initialiser la fonction *likelihood*, associée à chaque variable, à 1.

2. **Prise en charge de l'observation :**

La fonction *likelihood* est incorporée à l'arbre de jonction comme suit :

- (a) Encoder l'observation $A=a$ comme une fonction *likelihood* Λ_A^{new}
- (b) Identifier un cluster C_i contenant A
- (c) Mettre à jour ψ_{C_i} et Λ_A comme suit :

$$\psi_{C_i} \leftarrow \psi_{C_i} * \Lambda_A^{new}$$

$$\Lambda_A \leftarrow \Lambda_A^{new}$$

En introduisant un ensemble d'observations, l'arbre de jonction encode $\mathcal{P}_{\mathcal{JT}}(A \wedge e)$ au lieu de $\mathcal{P}_{\mathcal{JT}}(A)$.

3. **Normalisation :**

La marginalisation d'un potentiel d'un cluster quelconque ψ_{C_i} dans une variable A permet d'obtenir la probabilité de A et e , définie par :

$$P(A, e) = \sum_{C_i \setminus A} \psi_{C_i}$$

Le but est de calculer $P(A | e)$, en normalisant $P(A \wedge e)$ comme suit :

$$P(A | e) = \frac{P(A \wedge e)}{P(e)} = \frac{P(A \wedge e)}{\sum_A P(A \wedge e)} \quad (3.17)$$

3.3 Réseaux causaux possibilistes

Les réseaux causaux possibilistes sont des contreparties possibilistes des réseaux causaux probabilistes. Ils sont issus de la théorie des possibilités. Contrairement à la logique possibiliste (issue également de la théorie des possibilités), il existe peu de travaux relatifs aux réseaux causaux possibilistes dans les domaines de la propagation et de l'apprentissage.

Pour la propagation, dans [66], les auteurs proposent une propagation possibiliste pour les hypergraphes. Dans [67], l'auteur présente une adaptation possibiliste de l'algorithme de Pearl. Pour les graphes possibilistes indirectes, Gebhardt, Kruse et Borgelt ont développé une boîte à outils, appelée POSSINFER [70, 69, 68]. Shenoy [62] propose une version possibiliste de la propagation utilisant les *Valuation-Based Systems*.

Dans le domaine de l'apprentissage des connaissances incertaines, dans [99], les auteurs proposent une méthode d'apprentissage à partir des connaissances incertaines.

En théorie des possibilités deux types de conditionnement ont été définis (voir section 1.5.3) basés sur deux opérateurs différents, (le produit et le minimum). Ceci a conduit à définir deux types de réseaux causaux possibilistes :

- les réseaux causaux possibilistes basés sur le produit qui nous intéresse dans ce travail et qui sont proches du modèle graphique probabiliste, où les interprétations sont numériques,
- les réseaux causaux possibilistes basés sur le minimum, définis dans un contexte ordinal et qui posent le problème de cohérence dû au non-recouvrement des données initiales.

3.3.1 Définition d'un réseau causal possibiliste

Un réseau causal possibiliste, comme le cas d'un réseau causal probabiliste, est défini par :

- Une composante graphique : matérialisée par un Graphe Acyclique Indirect (DAG) \mathcal{G} représentant les relations de dépendances entre les variables.
- Une composante numérique : qui consiste à quantifier les différents liens représentés par le graphe en utilisant les distributions possibilistes conditionnelles de chaque nœud dans le contexte de ses parents. Ces dernières doivent obéir à la condition de normalisation.

Pour chaque variable A :

- Si $U_A = \emptyset$ (A est un nœud racine), alors les possibilités a priori relatives à la variable A doivent satisfaire :

$$\max_a \Pi(a) = 1, \forall a \in D_A,$$

- Si $U_A \neq \emptyset$, alors les distributions conditionnelles de la variable A dans le contexte de ses parents doivent satisfaire :

$$\max_a \Pi(a | a_A) = 1, \forall a \in D_A, u_A \in D_{U_A}$$

La différence de la définition du conditionnement a conduit à définir deux types de réseaux causaux possibilistes :

3.3.1.1 Les réseaux causaux possibilistes basés sur le produit :

Dans le contexte de la théorie des possibilités quantitative, un réseau causal possibiliste basé sur le produit, noté ΠG^* , est un graphe possibiliste où le conditionnement est basé sur le produit :

Définition 3.16 (Conditionnement possibiliste basé sur le produit)

$$\pi(\omega |_p \phi) = \begin{cases} \frac{\pi(\omega)}{\Pi(\phi)} & \text{si } \omega \models \phi \\ 0 & \text{sinon} \end{cases} \quad (3.18)$$

La distribution de possibilités jointe associée à un graphe causal possibiliste basé sur le produit est définie par le biais de la règle de chaînage basée sur le produit :

Définition 3.17 (Règle de chaînage basée sur le produit) *La distribution de possibilités globale jointe d'un réseau causal possibiliste basé sur le produit ΠG_p , à travers un ensemble de variables $V=\{A_1, A_2, \dots, A_N\}$ est exprimée comme le produit des N distributions a priori et des distributions conditionnelles à travers la règle de chaînage suivante :*

$$\pi_p(A_1, \dots, A_N) = \prod_{i=1..N} \Pi(A_i | U_{A_i}) \quad (3.19)$$

Nous reviendrons en détail dans le chapitre 7 sur les réseaux causaux possibilistes basés sur le produit.

3.3.1.2 Les réseaux causaux possibilistes basés sur le minimum :

Dans le contexte de la théorie des possibilités qualitative, un réseau causal possibiliste basé sur le minimum, noté ΠG_m , est un graphe possibiliste où le conditionnement est basé sur le minimum :

$$\pi(\omega |_m \phi) = \begin{cases} 1 & \text{si } \pi(\omega) = \Pi(\phi) \text{ et } \omega \models \phi \\ \pi(\omega) & \text{si } \pi(\omega) < \Pi(\phi) \text{ et } \omega \models \phi \\ 0 & \text{sinon} \end{cases} \quad (3.20)$$

La distribution de possibilités jointe associée à un graphe causal possibiliste basé sur le minimum est définie par le biais de la règle de chaînage basée sur le minimum :

Définition 3.18 (Règle de chaînage basée sur le minimum) *La distribution de possibilités globale jointe d'un réseau causal possibiliste basé sur le minimum ΠG_m , à travers un ensemble de variables $V=\{A_1, A_2, \dots, A_N\}$ est exprimée comme le minimum des N distributions a priori et des distributions conditionnelles à travers la règle de chaînage suivante :*

$$\pi_m(A_1, \dots, A_N) = \min_{i=1..N} \Pi(A_i | U_{A_i}) \quad (3.21)$$

Exemple 3.11 *Considérons le graphe possibiliste basé sur le minimum de la figure (3.19). Les distributions de possibilités conditionnelles sont représentées dans la table 3.6. Les distributions de possibilités jointes, obtenues en utilisant la règle de chaînage basée sur le minimum, sont données par la table(3.7).*

Contrairement aux réseaux causaux basés sur le produit, les réseaux causaux basés sur le minimum ne respectent pas toujours la propriété de recouvrement des données initiales. Ceci influe sur la cohérence du réseau :

Définition 3.19 (Cohérence d'un réseau causal possibiliste basé sur le minimum)

Un réseau causal possibiliste basé sur le minimum ΠG_m est cohérent si l'application de la règle de chaînage correspondante à l'équation(3.21), permet de retrouver les distributions initiales fournies par l'expert : $\Pi_m(a | u_A) = \Pi(a | u_A)$

Ce qui n'est pas systématiquement le cas pour les réseaux causaux possibilistes basés sur le minimum. Néanmoins, il a été montré dans [20] que les données non recouvertes correspondent aux données redondantes, et de ce fait elles peuvent être ignorées sans qu'il n'y ait des répercussions sur la distribution globale jointe.

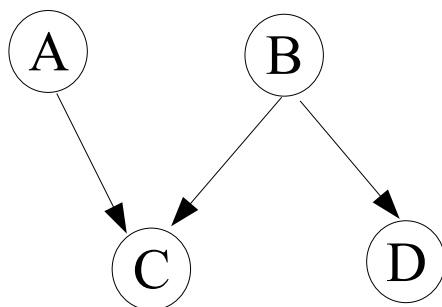


FIG. 3.19 – Exemple de DAG

TAB. 3.6 – Distributions initiales

A	$\Pi(A)$	B	$\Pi(B)$	D	B	$\Pi(D B)$
a	.5	b	1	d	b	1
$\neg a$	1	$\neg b$.4	d	$\neg b$	1
				$\neg d$	b	.3
				$\neg d$	$\neg b$.2

C	A	B	$\Pi(C A \wedge B)$	C	A	B	$\Pi(C A \wedge B)$
c	a	b	.6	$\neg c$	a	b	1
c	a	$\neg b$	1	$\neg c$	a	$\neg b$.8
c	$\neg a$	b	1	$\neg c$	$\neg a$	b	.3
c	$\neg a$	$\neg b$	0	$\neg c$	$\neg a$	$\neg b$	1

3.3.2 Propagation dans les réseaux causaux possibilistes basés sur le minimum

Dans ce qui suit, nous allons présenter les différents algorithmes de propagation dans le réseaux causaux possibilistes basés sur le minimum pour les polyarbres et les graphes à connexions multiples. Ces algorithmes, développés dans [20], sont des adaptations des algorithmes relatifs aux réseaux causaux bayésiens. Plus précisément, nous allons présenter dans cette section une version centralisée de l'algorithme de propagation de Pearl, tout en mettant en exergue le fait que, le problème de non-recouvrement des données initiales n'affecte pas le processus de propagation. Par la suite, une adaptation de l'algorithme de propagation de l'arbre de jonction probabiliste sera donnée.

3.3.2.1 Propagation basée sur le minimum dans les polyarbres

L'algorithme de propagation développé dans [20] est une adaptation de l'algorithme de propagation pour les polyarbre de la section 3.2.2.2. Cette adaptation est légèrement différente de celle proposée pour les réseaux causaux possibilistes basés sur le produit car des étapes supplémentaires sont nécessaires afin de transformer les distributions conditionnelles initiales en des distributions locales jointes.

TAB. 3.7 – Distributions de possibilités jointes utilisant la règle de chaînage basée sur le minimum

A	B	C	D	$\pi_m(A \wedge B \wedge C \wedge D)$	A	B	C	D	$\pi_m(A \wedge B \wedge C \wedge D)$
a	b	c	d	.5	$\neg a$	b	c	d	1
a	b	c	$\neg d$.3	$\neg a$	b	c	$\neg d$.3
a	b	$\neg c$	d	.5	$\neg a$	b	$\neg c$	d	.3
a	b	$\neg c$	$\neg d$.3	$\neg a$	b	$\neg c$	$\neg d$.3
a	$\neg b$	c	d	.4	$\neg a$	$\neg b$	c	d	0
a	$\neg b$	c	$\neg d$.2	$\neg a$	$\neg b$	c	$\neg d$	0
a	$\neg b$	$\neg c$	d	.4	$\neg a$	$\neg b$	$\neg c$	d	.4
a	$\neg b$	$\neg c$	$\neg d$.2	$\neg a$	$\neg b$	$\neg c$	$\neg d$.2

Remarque 3.6 Dans ce qui suit, les mêmes notations que la section 3.2.2.2 vont être utilisées.

Principe de la propagation Le but de la propagation est de calculer, pour chaque nœud A, la distribution conditionnelle associée basée sur une totale évidence, notée $Bel_{Cdt}(a)$ et définie par : $Bel_{Cdt}(a) = \Pi_m(a | e)$.

Cette valeur est obtenue à partir de la distribution jointe, notée $Bel_{joint}(a)$ et définie par :

$$Bel_{joint}(a) = \Pi_m(a \wedge e) = \min(\lambda(a), \mu(a)) \quad (3.22)$$

où $\lambda(a) = \Pi_m(a \wedge \bar{e}_A)$ et $\mu(a) = \Pi_m(a \wedge e_A^+)$.

Ainsi, $\forall a \in D_A$, le calcul de $Bel_{Cdt}(a) = \Pi_m(a | e)$ peut être effectué en utilisant l'équation $Bel_{joint}(a) = \Pi_m(a \wedge e)$ et en appliquant la définition du conditionnement basée sur le minimum (équation 3.20) comme suit :

$$Bel_{Cdt}(a) = \Pi_m(a | e) = \begin{cases} \Pi_m(a \wedge e) & \text{si } \Pi_m(a \wedge e) < \Pi_m(e) = \max_{a \in D_A} \Pi_m(a \wedge e) \\ 1 & \text{sinon} \end{cases} \quad (3.23)$$

L'expression de Bel_{joint} dépend des valeurs λ et μ relatives au nœud A qui dépendent, respectivement, des λ -messages reçus des enfants (notés par $\lambda_{Y_j}(A)$ où λ_{Y_j} est le message reçu par A à partir du fils Y_j) et des μ -messages reçus des parents (notés par $\mu_A(U_i)$) où μ_A est le message que A reçoit de son parent U_i). Formellement, ces deux valeurs sont définies par :

– **Calcul de la valeur** $\lambda \forall a \in D_A$,

$$\lambda(a) = \Pi_m(a \wedge \bar{e}_A) = \min(\lambda_A(a), \min_{j=1}^m \lambda_{Y_j}(a)) \quad (3.24)$$

où $\min_{j=1}^m \lambda_{Y_j}(a)$ correspond au minimum entre le λ -message reçu à partir du fils de A et $\lambda_A(a)$ dénote l'évidence locale de A, telle que :

$$\lambda_A(a) = \begin{cases} 0 & \text{si A est instanciée à } e_A \neq a \\ 1 & \text{si A est instanciée à } e_A = a \text{ ou A n'est pas instanciée} \end{cases} \quad (3.25)$$

- **Calcul de la valeur** $\mu \forall a \in D_A$,

$$\mu(a) = \Pi_m(a \wedge e_A^+) = \max_u \min(\Pi(a \wedge u), \min_{i=1}^n \mu_A(u_i)) \quad (3.26)$$

où $\min_{i=1}^n$ correspond au minimum entre les μ -messages reçus à partir des parents de A.

- **Calcul de λ -message** Le message λ de A à ses parents $U_i, (i \in \{1, \dots, n\})$, lorsque $U_i = u_i$ est défini par :

$$\lambda_A(u_i) = \Pi_m(e_{U_i A}^- \wedge u_i) = \max_{a \in D_A} \min[\lambda(a), \max_{u_k: k \neq i} (\min(\Pi(a \wedge u), \min_{k \neq i} \mu_A(u_k)))] \quad (3.27)$$

- **Calcul de μ -message** Le message μ de A à ses enfants $Y_j, (j \in \{1, \dots, n\})$ lorsque $A=a$ est défini par :

$$\mu_{Y_j}(a) = \Pi_m(a \wedge e_{AY_j}^+) = \min(\lambda_A(a), \min_{i=1 \dots n, i \neq j} \lambda_{Y_i}(a), \mu(a)) \quad (3.28)$$

Algorithme de propagation : Le principe de la propagation dans un réseau causal possibiliste basé sur le minimum est illustré par l'algorithme suivant :

Algorithme 3.12 : Propagation basée sur le minimum dans un polyarbre

Début

Choix de la racine de propagation

- Soit \mathcal{S} le plus petit ensemble connecté incluant les nœuds observés ;
- Soit pivot un nœud arbitraire dans \mathcal{S} représentant la racine de la propagation ;
- Soit Postorder un vecteur contenant l'ordre dans lequel les messages sont envoyés dans la phase de collecte de l'évidence (le dernier nœud étant le pivot) ;
- Soit Preorder un vecteur contenant l'ordre dans lequel les messages sont envoyés dans la phase de collecte de l'évidence (le premier nœud est le pivot)

Initialisation

- Initialiser l'ensemble des valeurs λ et μ messages à 1 ;
- Pour chaque nœud A racine, la valeur $\mu(a)$ est initialisée à $\Pi(a), \forall a \in D_A$.
- Pour chaque nœud A observé, assigner à $\lambda(a)$ la valeur 1 si A est instanciée à a, et lui assigner la valeur 0 dans le cas contraire.

Collecte de l'évidence

Pour $i := 1$ à $\text{length}(\text{Postorder}) - 1$ **Faire**

- $A \leftarrow \text{Postorder}[i]$;
- B \leftarrow un nœud adjacent de A dans Postorder ;
- Calculer $\lambda(A)$ en utilisant l'équation (3.24) ;
- Calculer $\mu(A)$ en utilisant l'équation (3.26) ; **Si** B est un nœud parent de A

alors

Envoyer un message λ de A vers B en utilisant l'équation (3.27)

Sinon

Envoyer un message μ de A à B en utilisant l'équation (3.28)

Distribution de l'évidence**Pour** $i := 1$ à $\text{length}(\text{Preorder}) - 1$ **Faire**

- $A \leftarrow \text{Preorder}[i]$;
- $\text{Below} \leftarrow$ un nœud adjacent de A dans Preorder ;
- Calculer $\lambda(A)$ en utilisant l'équation (3.24) ;
- Calculer $\mu(A)$ en utilisant l'équation (3.26) ;

Pour $j := 1$ à $\text{length}(\text{Below})$ **Faire**

- $B \leftarrow \text{Below}[j]$;
- **Si** B est un nœud parent de A **alors**
Envoyer un message λ de A vers B en utilisant l'équation (3.27) **Sinon**
Envoyer un message μ de A à B en utilisant l'équation (3.28)

Marginalisation Pour chaque nœud A , $\text{Bel}(A) = \Pi_p(a | e)$ est calculé en utilisant l'équation (3.22) ;**Normalisation**

- Calculer $\text{Bel}_{\text{Cdt}}(A)$ à partir de $\text{Bel}_{\text{Joint}}(A)$ en utilisant l'équation (3.23) ;

Fin**3.3.2.2 Propagation basée sur le minimum dans les graphes à connexions multiples**

La méthode de propagation dans ce type de DAG est similaire à celle présentée dans le cas probabiliste, qui consiste d'abord à transformer le réseau initial en un arbre de jonction puis de procéder à la propagation sur la nouvelle structure.

A partir d'un arbre de jonction \mathcal{JT} , l'unique distribution de possibilités globale jointe est définie par :

$$\pi_{\mathcal{JT}}(A_1, \dots, A_N) = \min_{i=1 \dots N} \pi_{C_i} \quad (3.29)$$

où N représente le nombre de clusters dans \mathcal{JT} . Dans le cas où l'arbre de jonction est consistant, le potentiel attaché à chaque cluster C_i doit satisfaire :

$$\pi_{C_i} = \Pi_m(C_i) \quad (3.30)$$

En utilisant cette dernière équation, la distribution de possibilités associée à une variable A dans un arbre de jonction consistant est calculée en utilisant un cluster C_i contenant la variable A , en marginalisant son potentiel sur A comme suit :

$$\Pi_m(A) = \max_{C_i \setminus A} \pi_{C_i} \quad (3.31)$$

Après la phase de transformation graphique, la propagation consiste à calculer le degré de possibilité de n'importe quelle variable $A \in V$, $\Pi_m(A)$. Elle s'effectue en trois phases :

1. **Initialisation** : Cette étape consiste à quantifier l'arbre de jonction, en utilisant les probabilités conditionnelles initiales :

Algorithme 3.13 : Initialisation**Debut**

Pour chaque cluster C_i **Faire** $\pi_{C_i}^I \leftarrow 1$;
Pour chaque séparateur S_{ij} **Faire** $\pi_{S_{ij}}^I \leftarrow 1$;
Pour chaque variable A **Faire**
 Sélectionner un cluster C_i contenant $\{A\} \cup U_A$;
 $\pi_{C_i}^I := \min(\pi_{C_i}^I, \Pi(A | U_A))$;

Fin

L'arbre de jonction, ainsi initialisé, codifie les mêmes distributions que le graphe initial. Ainsi,

$$\pi_m = \pi_{\mathcal{JT}}^I \quad (3.32)$$

où π_m est la distribution de possibilités jointe du réseau causal initial obtenue en utilisant l'équation (3.21) et $\pi_{\mathcal{JT}}^I$ représente la distribution de probabilités jointe associée à l'arbre de jonction obtenue en utilisant l'équation (3.29).

Après avoir initialisé les potentiels de l'arbre de jonction, il faudrait s'assurer de la consistance locale en effectuant la propagation globale décrite ci-après.

2. **La propagation globale** : Elle est réalisée par le mécanisme de passages de messages entre chaque cluster C_i et ses clusters adjacents C_j , comme dans le cas probabiliste.

Si un nœud C_i envoie un message à ses clusters adjacents C_j , les potentiels de C_i , C_j et leurs séparateurs seront mis à jour comme suit :

- (a) Préserver le même potentiel pour C_i :

$$\pi_{C_i}^{t+1} \leftarrow \pi_{C_i}^t \quad (3.33)$$

- (b) Affecter un nouveau potentiel à S_{ij} :

$$\pi_{S_{ij}}^{t+1} \leftarrow \max_{C_i \setminus S_{ij}} \pi_{C_i}^t \quad (3.34)$$

- (c) Affecter un nouveau potentiel à C_j :

$$\pi_{C_j}^{t+1} \leftarrow \min(\pi_{C_j}^t, \pi_{S_{ij}}^{t+1}) \quad (3.35)$$

La procédure de propagation globale peut être résumée comme suit [20] :

Algorithme 3.14 : Propagation globale

Debut

Selection de la racine de la propagation :

Soit *Pivot* un cluster arbitraire C_i représentant la racine de la propagation ;

Soit *Postorder* un vecteur contenant l'ordre dans lequel les messages vont être envoyés dans la phase de la collecte de l'évidence ;

Soit *Preorder* un vecteur contenant l'ordre dans lequel les messages vont être envoyés dans la phase de la distribution de l'évidence ;

Collecte de l'évidence :

Pour $i := 1$ **jusqu'à** $\text{lenght}(\text{Postorder})-1$ **Faire**

$C_i \leftarrow \text{Postorder}[i]$;

$C_j \leftarrow$ cluster adjacent de C_i dans *Postorder* ;

Envoyer le message de C_i vers C_j en utilisant les équations (3.33), (3.34) et (3.35) ;

Distribution de l'évidence :

Pour $i := 1$ **jusqu'à** $\text{lenght}(\text{Preorder})$ **Faire**

$C_i \leftarrow \text{Preorder}[i]$;

Below \leftarrow cluster adjacents de C_i dans *Preorder* ;

Pour $j := 1$ **jusqu'à** $\text{lenght}(\text{Below})$ **Faire**

$C_j \leftarrow \text{Below}[j]$;

Envoyer le message de C_i à C_j en utilisant les équations (3.33), (3.34) et (3.35) ;

Fin

Remarque 3.7 A partir de ces deux procédures, il est important de relever certains points :

- A chaque étape t de la procédure de la propagation globale, l'arbre de jonction codifie la même distribution jointe. Formellement, $\forall t$, de l'étape initiale jusqu'à l'étape finale :

$$\pi_{\mathcal{JT}}^t = \pi_{\mathcal{JT}}^{t+1} \quad (3.36)$$

où $\pi_{\mathcal{JT}}^t$ représente la distribution jointe relative à l'arbre de jonction \mathcal{JT} de l'étape t .

- A partir de équations (3.21) et (3.29), nous déduisons que de la phase d'initialisation jusqu'à la phase de consistance globale, l'arbre de jonction codifie la même distribution jointe. Formellement,

$$\pi_m = \pi_{\mathcal{JT}}^C \quad (3.37)$$

où π_m est la distribution jointe de ΠG_m et $\pi_{\mathcal{JT}}^C$ est la distribution jointe associée à \mathcal{JT} après la procédure de propagation globale.

- La consistance globale est atteinte après les phases de collecte de l'évidence et de distribution de l'évidence.

Marginalisation : A partir de l'arbre de jonction consistant obtenu lors de la phase précédente, la mesure de possibilités associée à chaque variable d'intérêt A , notée $\Pi_m(A)$ peut être calculée par l'algorithme suivant :

Algorithme 3.15 : Marginalisation

Début

Identifier un cluster C_i contenant A ;

Calculer $\Pi_m(A)$ en marginalisant $\pi_{C_i}^C$ sur A :

$$\Pi_m(A) \leftarrow \max_{C_i \setminus A} \pi_{C_i}^C ;$$

Fin

Prise en charge de l'évidence : La prise en charge d'une évidence totale e se fait par la généralisation des différents algorithmes précédents. Il s'agit de calculer $\forall A \in V, \Pi_m(A \wedge e)$. L'évidence peut être codifiée en utilisant la fonction *likelihood* définie par :

$$\Lambda_A(a) = \begin{cases} 1 & \text{si } A \text{ n'est pas instanciée} \\ 1 & \text{si } A \text{ est instanciée à } a \\ 0 & \text{si } A \text{ est instanciée mais pas à } a \end{cases} \quad (3.38)$$

La prise en compte de l'évidence e nécessite l'extension de la procédure d'initialisation afin d'incorporer la connaissance certaine. En effet, l'évidence e doit être codifiée comme une fonction *likelihood*, en utilisant l'équation (3.38), puis elle doit être intégrée dans l'arbre de jonction en ajoutant des étapes supplémentaires à la procédure d'initialisation : Elle se résume comme suit :

- Pour chaque instance de la variable A , codifier l'observation $A=a$ comme une fonction *likelihood*, en utilisant l'équation (3.38).
- Identifier un cluster C_i contenant A :

$$\pi_{C_i}^I \leftarrow \min(\pi_{C_i}^I, \Lambda_A)$$

La marginalisation d'un potentiel d'un cluster quelconque $\pi_{C_i}^C$ dans une variable A , en utilisant l'équation (3.31) permet d'obtenir la mesure de possibilité de A et de e :

$$\Pi_m(A \wedge e) = \max_{C_i \setminus A} \pi_{C_i}^C$$

Le calcul de $\Pi_m(A | e)$ est effectué à partir de $\Pi_m(A \wedge e)$ en appliquant la définition du conditionnement basé sur le produit comme suit :

$$\Pi_m(A | e) = \begin{cases} \Pi_m(A \wedge e) & \text{si } \Pi_m(A \wedge e) < \Pi_m(e) = \max_A \Pi_m(A \wedge e) \\ 1 & \text{sinon} \end{cases} \quad (3.39)$$

3.4 Conclusion

Le traitement de l'incertitude est une issue fondamentale de l'intelligence artificielle. L'approche la plus utilisée de nos jours est sans aucun doute l'approche bayésienne. L'avantage de celle-ci est qu'elle bénéficie des techniques de la théorie des probabilités. De plus, du point de vue philosophique, un agent *rationnel* affecte des incertitudes à des événements de façon à satisfaire les axiomes de probabilités. Elle offre une représentation naturelle des connaissances par des liens entre causes et effets. La représentation des connaissances est graphique ce qui induit sa lisibilité.

Néanmoins, l'approche bayésienne a été largement critiquée car des probabilités subjectives doivent être affectées à chaque événement.

Du point de vue du raisonnement, les différents algorithmes de propagation développés sont complexes dans le cas des réseaux causaux à connexions multiples dans le cadre de la théorie de probabilités ou dans le cadre de la théorie des possibilités. Ces derniers ont été obtenus en adaptant les algorithmes développés dans le contexte possibiliste qui offre un cadre qualitatif et un cadre quantitatif.

Dans les deux derniers chapitres précédents, nous avons présentés deux modes de représentations des connaissances incertaines, à savoir le mode logique et le mode graphique. Après avoir parcouru les modes logiques et les modes graphiques de représentation des connaissances en théorie des possibilités, nous allons, dans le chapitre suivant, faire un tour d'horizon sur les différentes transformations développées dans cette théorie.

Chapitre 4

Rétrospective sur les transformations en théorie des possibilités

4.1 Introduction

Le mode logique et le mode graphique, vus dans les chapitres précédents, constituent les principaux modes de représentation des connaissances incertaines. Sur le plan de la représentation, le mode graphique est plus explicite, vu qu'il exhibe les relations de dépendances qui existent entre les différentes variables. Sur le plan du raisonnement, le mode logique offre une machinerie d'inférence efficace.

La théorie des possibilités est une théorie de l'incertain qui fournit les deux modes de représentation :

- La logique possibiliste offre un cadre général pour représenter les connaissances incertaines, en termes de formules logiques classiques auxquelles sont associées des pondérations appartenant à une échelle linéaire $[0, 1]$.
- Les réseaux causaux possibilistes sont des DAGs où les nœuds représentent des variables (l'âge d'un patient, la résistance d'un matériau,...) et les arcs codifient les liens causaux qui existent entre les variables. L'incertitude est représentée pour chaque nœud et exprime l'intensité des liens entre les différentes variables du réseau.

Comme il a été relaté dans le chapitre 2, la logique possibiliste offre deux contextes de représentation : le contexte ordinal ou qualitatif et le contexte numérique ou quantitatif. D'autres part, comme il a été mentionné dans le chapitre 3, dans le mode graphique, il existe également deux types de règles de chaînage, l'une basée sur le minimum (contexte qualitatif) et l'autre basée sur le produit (contexte quantitatif). Ces deux types de conditionnement induisent deux types de graphes possibilistes : l'un basé sur le minimum et l'autre basé sur le produit.

Ce chapitre est consacré à détailler les différentes transformations réalisées autour de l'ensemble de ces modes de représentation en théorie des possibilités. L'idée principale sur laquelle repose toute transformation est de considérer que :
Deux représentations sont dites équivalentes si elles induisent les mêmes distributions de possibilités.

4.2 Codification logique d'un réseau causal possibiliste basé sur le minimum

Cette section est consacrée à la translation d'un réseau causal possibiliste basé sur le minimum vers une base possibiliste qualitative .

Cette transformation [100, 101, 13] est basée sur l'idée de considérer qu'une base logique possibiliste qualitative associée à un graphe causal possibiliste basé sur le minimum est le résultat de la fusion des bases élémentaires en utilisant l'opérateur minimum. Ces bases élémentaires sont composées de formules auxquelles sont rattachées les degrés de possibilités a priori et les degrés de possibilités conditionnelles associés aux différents nœuds du réseau causal possibiliste.

D'une manière formelle :

- Soit $\{A_1, \dots, A_n\}$ un ensemble de variables binaires d'un graphe causal possibiliste basé sur le minimum et noté ΠG_m . Un réseau causal sera représenté, par souci de simplicité, par un ensemble de triplets : $\Pi G_m = \{(a_i, u_i, \alpha_i) : \alpha_i = \Pi(a_i | u_i) \neq 1\}$,¹ où a_i est une instance de la variable A_i et u_i est un élément du produit cartésien du domaine D_j de la variable $A_j \in Par(A_i)$.
- A chaque triplet (a, u, α) du DAG est associé une formule logique unique $(\neg a \vee \neg u, 1 - \alpha)$, ayant comme distribution de possibilités $\pi_{a,u}$ définie par :

$$\pi_{a,u}(\omega) = \begin{cases} 1 & \text{si } \omega \models \neg a \vee \neg u, \\ \alpha & \text{sinon } (\omega \models a \wedge u). \end{cases}$$

- Soit π_i la distribution de possibilités associée à une formule possibiliste (a_i, u_i, α_i) . La distribution de possibilité jointe calculée à partir du DAG ΠG_m est identique à celle obtenue en combinant les différentes distributions de possibilités π_i , en utilisant l'opérateur minimum. Ainsi,

$$\pi(\omega) = \min\{\alpha_i : (a_i, u_i, \alpha_i) \in \Pi G_m, \omega \models \neg a_i \vee \neg u_i\} = \min_{i=1,n} \pi_i$$

- Soient Σ_1 et Σ_2 deux bases logiques possibilistes. Soient π_1 et π_2 deux distributions de possibilités associées à Σ_1 et Σ_2 . Soit π_{min} la combinaison des deux distributions de possibilités π_1 et π_2 par le minimum. La base logique possibiliste associée à π_{min} est définie par :

$$\Sigma_{min} = \Sigma_1 \cup \Sigma_2.$$

¹Seules les possibilités conditionnelles différentes de 1 sont considérées, vu que 1 est un élément neutre aussi bien pour l'opérateur minimum que pour l'opérateur produit

- Ainsi, la base logique qualitative possibiliste associée à un graphe causal possibiliste basé sur le minimum $\Pi G_m = \{(a_i, u_i, \alpha_i) : \Pi(a_i \mid u_i = \alpha_i) \neq 1\}$ est définie par :

$$\Sigma = \{(\neg a_i \vee \neg u_i, 1 - \alpha_i) : (a_i, u_i, \alpha_i) \in \Pi G_m\}$$

Ce résultat est important car il implique que la machinerie de l'inférence de la logique possibiliste peut être appliquée aux réseaux causaux possibilistes basés sur le minimum.

Exemple 4.1 *Considérons le DAG de la figure 4.1.*

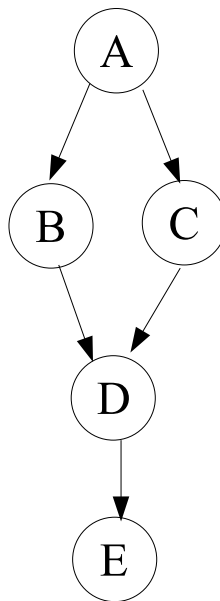


FIG. 4.1 – Exemple d'un DAG

Les différents degrés de possibilités a priori et conditionnelles sont donnés par les tables 4.1 à 4.5.

TAB. 4.1 – Les degrés de possibilités a priori $\Pi(A)$

a	0.4
$\neg a$	1

TAB. 4.2 – Les degrés de possibilités conditionnelles $\Pi(B \mid A)$

B A	a	$\neg a$
b	0.2	1
$\neg b$	1	0.8

TAB. 4.3 – Les degrés de possibilités conditionnelles $\Pi(C | A)$

C A	a	$\neg a$
c	1	0.4
$\neg c$	0.6	1

TAB. 4.4 – Les degrés de possibilités conditionnelles $\Pi(D | B \wedge C)$

D B C	bc	$\neg b c$	ailleurs
d	1	1	1
$\neg d$	1	0.9	1

Ainsi,

$$\Pi G_m = \{(a, \emptyset, 0.4), (b, a, 0.2), (\neg b, \neg a, 0.8), (c, \neg a, 0.4), (\neg c, a, 0.6), (\neg d, \neg bc, 0.9), (\neg e, d, 0.5)\}.$$

La base de connaissances associée au DAG dans le cas où le conditionnement est basé sur le minimum est :

$$\Sigma = \{(\neg a, 0.6), (\neg b \vee \neg a, 0.8), (b \vee a, 0.2), (\neg c \vee a, 0.6), (c \vee \neg a, 0.4), (d \vee b \vee \neg c, 0.1), (e \vee \neg d, 0.5)\}.$$

4.3 Codification logique d'un réseau causal possibiliste basé sur le produit

Le principe de passage d'un graphe causal possibiliste basé sur le produit vers une base logique possibiliste qualitative est identique à celui de la section précédente [100, 101, 13].

4.3.1 Principe de la transformation

Le processus de transformation est résumé par les points suivants :

1. Soit π_i la distribution de possibilités associée à un triplet (a_i, u_i, α_i) . La distribution de possibilités jointe calculée à partir d'un réseau causal possibiliste basé sur le produit ΠG^* est identique à celle obtenue en combinant les différentes distributions de possibilités π_i en utilisant l'opérateur produit.
2. Soient $\Sigma_1 = \{(p_i, \alpha_i) : i \in I\}$ et $\Sigma_2 = \{(q_j, \beta_j) : j \in J\}$. Soient π_1 et π_2 les deux distributions de possibilités associées, respectivement, à Σ_1 et Σ_2 . La base associée à π^* obtenue en combinant π_1 et π_2 par l'opérateur produit est donnée par : $\mathcal{C}^*(\Sigma_1, \Sigma_2) = \Sigma_1 \cup \Sigma_2 \cup \{(p_i \vee q_j, \alpha_i + \beta_j - \alpha_i * \beta_j) : i \in I, j \in J, p_i \vee q_j \neq \top\}$.
3. Soit A un nœud de ΠG et soient $u_1 \neq u_2$ deux instances de $\text{Par}(A)$. Supposons que ΠG contient (a_1, u_1, α_1) et (a_2, u_2, α_2) . Soient $\Sigma_1 = \{(\neg a_1 \vee \neg u_1, \alpha_1)\}$ et $\Sigma_2 = \{(\neg a_2 \vee \neg u_2, \alpha_2)\}$. Alors, $\mathcal{C}^*(\Sigma_1, \Sigma_2) = \Sigma_1 \cup \Sigma_2$.

TAB. 4.5 – Les degrés de possibilités conditionnelles $\Pi(E | D)$

E D	d	$\neg d$
e	1	1
$\neg e$	0.5	1

4. Ainsi, la base de connaissances résultante de la combinaison avec l'opérateur produit est telle que les formules associées aux triplets d'un nœud A est l'union de ces formules :

$$\Sigma_A = \{(\neg a_i \vee \neg u_i, 1 - \alpha_i) : (a_i, u_i, \alpha_i \in \Pi G)\}$$

et la base de connaissances qualitative possibiliste associée à un graphe ΠG^* est obtenue en combinant, avec l'opérateur produit, les bases de connaissances associées à chaque nœud du graphe.

Exemple 4.2 Reconsidérons le graphe causal de l'exemple précédent :

$$\Pi G^* = \{(a, \emptyset, 0.4), (b, a, 0.2), (\neg b, \neg a, 0.8), (c, \neg a, 0.4), (\neg c, a, 0.6), (\neg d, \neg b, c, 0.9), (\neg e, d, 0.5)\}.$$

Les différentes bases de connaissances associées aux différents nœuds du graphe sont :

- $\Sigma_A = \{(\neg a, 0.6)\}$
- $\Sigma_B = \{\neg a \vee \neg b, 0.8), (a \vee b, 0.2)\}$
- $\Sigma_C = \{a \vee \neg c, 0.6), (\neg a \vee c, 0.4)\}$
- $\Sigma_D = \{(b \vee \neg c \vee d, 0.1)\}$
- $\Sigma_E = \{(\neg d \vee e, 0.5)\}$

Les différentes combinaisons de ces différentes bases de connaissances en utilisant \mathcal{C}^* sont :

$$\begin{aligned} - \Sigma_{AB} &= \mathcal{C}^*(\Sigma_A, \Sigma_B) = \Sigma_A \cup \Sigma_B \cup \{(\neg a \vee \neg b, 0.92)\} \\ &= \{(\neg a, 0.6), (a \vee b, 0.2), (\neg a \vee \neg b, 0.92)\} \end{aligned}$$

La formule $(\neg a \vee \neg b, 0.8)$ est supprimée, vu qu'elle est déduite de $(\neg a \vee \neg b, 0.92)$.

$$\begin{aligned} - \Sigma_{ABC} &= \mathcal{C}^*(\Sigma_{AB}, \Sigma_C) = \Sigma_{AB} \cup \Sigma_C \cup \{(\neg a \vee c, 0.76), (a \vee b \vee \neg c, 0.68), (\neg a \vee \neg b \vee c, 0.952)\} \\ &= \{(\neg a, 0.6), (a \vee b, 0.2), (\neg a \vee \neg b, 0.92), (a \vee \neg c, 0.6), (\neg a \vee c, 0.76), (a \vee b \vee \neg c, 0.68), (\neg a \vee \neg b \vee c, 0.952)\} \end{aligned}$$

La formule $(\neg a \vee c, 0.4)$ est supprimée, vu qu'elle est déduite de $(\neg a \vee c, 0.76)$.

$$\begin{aligned} - \Sigma_{ABCD} &= \Sigma_{ABC} \cup \Sigma_D \cup \{(\neg a \vee b \vee \neg c \vee d, 0.64), (a \vee b \vee \neg c \vee d, 0.28), (a \vee b \vee \neg c \vee d, 0.64), (a \vee b \vee \neg c \vee d, 0.712)\} \\ &= \{(\neg a, 0.6), (a \vee b, 0.2), (\neg a \vee \neg b, 0.92), (a \vee \neg c, 0.6), (\neg a \vee c, 0.76), (a \vee b \vee \neg c, 0.68), (\neg a \vee \neg b \vee c, 0.952), (b \vee \neg c \vee d, 0.1), (\neg a \vee b \vee \neg c \vee d, 0.64), (a \vee b \vee \neg c \vee d, 0.712)\} \end{aligned}$$

Les formules $(a \vee b \vee \neg c \vee d, 0.28)$, $(a \vee b \vee \neg c \vee d, 0.64)$ sont supprimées, vu qu'elles sont sous-sommées à partir de la formule $(a \vee b \vee \neg c \vee d, 0.712)$

$$\begin{aligned} - \Sigma_{ABCDE} &= \mathcal{C}^*(\Sigma_{ABCD}, \Sigma_E) \\ &= \Sigma_{ABCD} \cup \Sigma_E \cup \{(\neg a \vee \neg d \vee e, 0.8), (a \vee b \vee \neg d \vee e, 0.6), (\neg a \vee \neg b \vee \neg d \vee e, 0.96), (a \vee \neg c \vee \end{aligned}$$

$$\begin{aligned} & \neg d \vee e, 0.84), (\neg a \vee c \vee \neg d \vee e, 0.88), (a \vee b \vee \neg c \vee \neg d \vee e, 0.84), (\neg a \vee \neg b \vee c \vee \neg d \vee e, 0.976) \} \\ & = \{ (\neg a, 0.6), (a \vee b, 0.2), (\neg a \vee \neg b, 0.92), (a \vee \neg c, 0.6), (\neg a \vee c, 0.76), (a \vee b \vee c, 0.68), (\neg a \vee \\ & \neg b \vee c, 0.952), (b \vee \neg c \vee d, 0.1), (\neg a \vee b \vee \neg c \vee d, 0.64), (a \vee b \vee \neg c \vee d, 0.712), (\neg d \vee \\ & e, 0.5), (\neg a \vee \neg d \vee e, 0.8), (a \vee b \vee \neg d \vee e, 0.6), (\neg a \vee \neg b \vee \neg d \vee e, 0.96), (a \vee \neg c \vee \neg d \vee \\ & e, 0.84), (\neg a \vee c \vee \neg d \vee e, 0.88), (a \vee b \vee \neg c \vee \neg d \vee e, 0.84), (\neg a \vee \neg b \vee c \vee \neg d \vee e, 0.976) \}. \end{aligned}$$

La base de connaissances résultante contient 18 clauses alors que dans le cas du minimum, la base correspondante contient 7 clauses. Ceci illustre le fait que la combinaison avec l'opérateur produit induit une base de connaissances plus grande par rapport à la combinaison avec l'opérateur minimum. En effet, l'opérateur numérique produit n'est pas compatible avec l'échelle ordinale, vu qu'il y a, à chaque étape, des ajouts de formules avec des niveaux intermédiaires qui ne sont pas présents dans l'échelle ordinale initiale.

4.4 D'une base possibiliste qualitative vers un réseau causal basé sur le minimum

Le passage d'une base possibiliste qualitative vers un graphe causal possibiliste basé sur le minimum peut être effectué en utilisant la distribution de possibilités comme une étape intermédiaire [100, 101, 13]. En effet, une base de connaissance induit une distribution de possibilités, à partir de laquelle il est possible de construire un graphe ΠG_m . Cette procédure peut être également appliquée afin de construire un graphe ΠG^* . Néanmoins, cette méthode est coûteuse. Il est plus judicieux de construire le graphe directement à partir de la base de connaissances.

Cependant, cette transformation est moins directe que la précédente et nécessite des étapes de pré-traitement. En effet, il n'y a pas une correspondance directe entre une formule logique et un triplet du graphe.

La construction d'un graphe causal basé sur le minimum à partir d'une base possibiliste est réalisée en trois étapes : la première étape consiste à mettre la base sous forme clausale tout en éliminant les tautologies. La seconde étape consiste à construire le graphe en établissant un ordre arbitraire entre les variables et la troisième étape consiste à calculer les possibilités a priori et conditionnelles associées au graphe.

4.4.1 Mise sous forme clausale des bases et suppression des tautologies

Durant cette phase, la base de connaissances Σ est réécrite sous forme d'une autre base sémantiquement équivalente Σ' comme suit :

Soit $(p, \alpha) \in \Sigma$. Soit $\{c_1, \dots, c_n\}$ un ensemble de clauses codifiant p en logique classique. Σ' est obtenue en remplaçant (p, α) par $\{(c_1, \alpha), \dots, (c_n, \alpha)\}$.

La suppression des tautologies induit également une base de connaissance équivalente à l'initiale. La suppression des tautologies est importante vu qu'elle prévient l'obtention des liens fictifs entre les variables. Par exemple la formule $(\neg a \vee \neg y \vee a, 1)$ peut provoquer la création d'un lien entre la variable A et la variable Y .

4.4.2 Construction du graphe

La construction du graphe revient à définir :

- Les nœuds (variables) du graphe,
- Les parents de chaque nœud (liens).

L'ensemble des variables est simplement constitué par les symboles qui apparaissent dans la base de connaissances. Le but de cette opération est de construire un graphe ΠG_m cohérent, dans le sens où les distributions de possibilités conditionnelles peuvent être retrouvées en utilisant la combinaison à base du minimum. Pour cela :

1. La première phase consiste à ordonner arbitrairement les variables : $\{A_1, A_2, \dots, A_n\}$, de telle sorte que les parents de chaque variable A_i , s'ils existent, sont dans l'ensemble $\{A_{i+1}, \dots, A_n\}$.
2. La deuxième phase consiste à appliquer une procédure, composée de quatre étapes et ce, pour chaque variable $A_i, i \in [1, \dots, n]$, comme suit :

Algorithme 4.1 : Construction du graphe ΠG_m

Début

Pour $i=1, \dots, n$ faire

Début

Determination de la base locale pour A_i

(a) Soit $(a_i \vee p, \alpha)$ une clause de Σ tels que a_i est une instance de A_i et p est construit à partir de $\{A_{i+1}, \dots, A_n\}$

i. **Si** $(a_i \vee p, \alpha)$ est sous-sommée **alors** la supprimer de Σ

ii. **Si** $\Sigma \vdash (p, \alpha)$, **alors** remplacer $(a_i \vee p, \alpha)$ par (p, α) .

(b) Soit K_i l'ensemble de clauses $(a_i \vee p, \alpha)$ dans Σ où p est construit à partir de $\{A_{i+1}, \dots, A_n\}$

(c) Les parents de la variable A_i sont :

$Par(A_i) = \{A_j : \exists c \in K_i \text{ contenant une instance de } A_j\}$

(d) **Calcul de la forme clausale de la forme étendue de K_i**

Remplacer dans Σ, K_i par $E(K_i)$ sous forme clausale. $E(K_i)$ représente la forme étendue de K_i où chaque clause comporte une instance des différentes variables apparaissant dans K_i . Le remplacement de K_i par $E(K_i)$ est justifié comme suit : Soit $(a_i \vee p, \alpha)$ une clause dans Σ .

Si $\Sigma \vdash (p, \alpha)$ **alors** remplacer $(a_i \vee p, \alpha)$ par (p, α)

(e) **Elimination des données incohérentes**

Pour chaque $(a_i \vee p, \alpha)$ dans Σ **faire**

Si $\Sigma \vdash (p, \alpha)$ **alors** remplacer $(a_i \vee p, \alpha)$ par (p, α)

(f) **Production de Σ_i**

Σ_i est alors composée de l'ensemble de clauses $(a_i \vee p, \alpha)$ dans Σ , tel que p est défini à partir de $\{A_{i+1}, \dots, A_n\}$

Fin**Fin**

Les étapes 1 à 3 consistent à déterminer les parents de chaque variable. Les étapes 4 à 6 consistent à réécrire la base de connaissances de telle sorte que :

- (a) L'obtention des distributions de possibilités conditionnelles attachées à chaque variable sera directe et
- (b) Le recouvrement des valeurs originales sera assuré en utilisant la règle de chaînage pour calculer les distributions de possibilités jointes.
En effet, une fois $E(K_i)$ calculée afin d'évaluer $\Pi(a_i | u_i)$ alors :
 - Soit $(\neg a_i \vee \neg u_i, \alpha) \notin E(K_i)$ alors $\Pi(a_i | u_i) = 1$, ou
 - $(\neg a_i \vee \neg u_i, \alpha) \in E(K_i)$, alors,
 - **Si** $\Sigma \vdash (\neg u_i, \alpha)$, **alors** $\Pi(a_i | u_i) = 1$, vu que $\Pi(a_i \wedge u_i) = \Pi(a_i)$. d'où $(\neg a_i \vee u_i, \alpha)$ peut être supprimée de Σ ,
 - **Sinon** $\Pi(a_i | u_i) = 1 - \alpha$

Le résultat de cet algorithme est une partition $\{\Sigma_1, \dots, \Sigma_n\}$, telle que $\Sigma_1 \cup \dots \cup \Sigma_n$ est sémantiquement équivalente à Σ .

4.4.3 Calcul des distributions conditionnelles locales

Une fois le graphe construit, il faudrait calculer les possibilités conditionnelles attachées à chaque nœud du graphe. Le calcul des $\Pi(A_i | Par(A_i))$ est immédiat à partir des différents Σ_i :

Soit $\{\Sigma_1, \dots, \Sigma_n\}$ le résultat obtenu lors de l'étape 6 de l'algorithme précédent. Soit A_i une variable et $Par(A_i) = \{B_1, \dots, B_m\}$ l'ensemble des parents de A_i . Soit a une instance de A_i et $u = b_1 \wedge \dots \wedge b_m$ une instance de $Par(A_i)$. Alors,

$$\Pi(a | u) = \begin{cases} 1 - \alpha_i & \text{si } (\neq a \vee \neg u, \alpha_i) \in \Sigma \\ 1 & \text{sinon} \end{cases}$$

La translation d'une base possibiliste qualitative vers un réseau causal basé sur le minimum a été présentée en établissant au préalable un ordre entre les variables. La simplification de la structure du graphe dépend, justement, de l'optimisation de l'ordre totale entre les variables.

4.5 D'une base possibiliste qualitative vers un réseau causal basé sur le produit

Dans [14, 102], les auteurs ont développé un passage d'une base possibiliste qualitative vers un réseau causal basé sur le produit. Cette approche est différente de celle qui consiste de transformer une base possibiliste qualitative vers un réseau causal basé sur le minimum. Néanmoins, les deux approches ont en commun l'étape préliminaire de transformation de la base sous forme causale ainsi que la suppression des tautologies.

L'idée de base de cette transformation est :

- D'établir un ordre arbitraire entre les variables A_1, \dots, A_n de telle sorte que les parents d'une variable A_i , s'ils existent, doivent se trouver parmi A_{i+1}, \dots, A_n .
- De décomposer la base Σ , associée à la décomposition de la distribution de possibilités :

$$\pi(A_1 \dots A_n) = \pi(A_1 \mid A_2 \dots A_n) * \pi(A_2 \dots A_n)$$

Ainsi, le résultat de chaque étape i est une base de connaissances associée à $\pi(A_i \dots A_n)$. Cette dernière permettra de définir les parents de la variable A_i et de calculer par la suite les possibilités conditionnelles $\pi(A_i \mid Par(A_i))$. Dans ce qui suit, l'étape 1 est considérée pour ($i=1$). Soit Σ_C la base marginalisée courante associée à $\pi(A_2, \dots, A_n)$.

4.5.1 Calcul de la base marginalisée

Le calcul de Σ_c se déroule en trois étapes [14, 102] :

4.5.1.1 Décomposition de π

Cette phase consiste à décomposer la distribution de possibilités π en la restreignant aux deux instances de la variable $A_1(a_1, \neg a_1)$ comme suit :

$$\pi_{a_1}(\omega) = \begin{cases} \pi(\omega) & \text{si } (\omega \models a_1 \\ 0 & \text{sinon} \end{cases}$$

$$\pi_{\neg a_1}(\omega) = \begin{cases} \pi(\omega) & \text{si } (\omega \models \neg a_1 \\ 0 & \text{sinon} \end{cases}$$

La base possibiliste associée à la distribution de possibilités π_{a_1} (respectivement $\pi_{\neg a_1}$) est $\Sigma \cup \{(a_1, 1)\}$ (respectivement $\Sigma \cup \{(\neg a_1, 1)\}$).

Ces deux bases peuvent être simplifiées comme suit : Soit $\Sigma' = \Sigma \cup \{(a_1, 1)\}$.

1. Éliminer toutes les clauses de Σ' de la forme $(a_1 \vee x, \alpha)$, vu que $(a_1 \vee x, \alpha)$ est sous-sommée de $(a_1, 1)$
2. Remplacer dans Σ' chaque clause de la forme $(\neg a_1 \vee x, \alpha)$ par (x, α) car $(a_1, 1)$ et $(\neg a_1 \vee x, \alpha)$ induisent (x, α) qui peut être rajoutée à Σ'_1 . Ainsi, $(\neg a_1 \vee x, \alpha)$ pourra alors être supprimée vu qu'elle sera sous-sommée de (x, α)

4.5.1.2 Marginalisation

Soit Σ_{a_1} (respectivement $\Sigma_{\neg a_1}$) la base obtenue après simplification de $\Sigma \cup \{(a_1, 1)\}$ (respectivement $\Sigma \cup \{(\neg a_1, 1)\}$) de la phase précédente. Ainsi, la seule clause dans Σ_{a_1} qui contient a_1 est $(a_1, 1)$.

Les marginalisations de π_{a_1} et $\pi_{\neg a_1}$ sur $\{A_1, \dots, A_n\}$, notées $\pi_{a_1}^{A_1}$ et $\pi_{\neg a_1}^{A_1}$ sont définies par :

$$\pi_{a_1}^{A_1}(A_2 \dots A_n) = \Pi_{a_1}(A_2 \dots A_n)$$

$$\pi_{\neg a_1}^{A_1}(A_2 \dots A_n) = \Pi_{\neg a_1}(A_2 \dots A_n)$$

où $\Pi_{a_1}(A_2 \dots A_n)$ représente la mesure de possibilité associée avec π_{a_1} définie sur $\{A_1, \dots, A_n\}$. La contrepartie syntaxique associée avec $\pi_{a_1}^{A_1}$ (respectivement $\pi_{\neg a_1}^{A_1}$) est la base $\Sigma_{a_1} - \{(a_1, 1)\}$ (respectivement $\Sigma_{\neg a_1} - \{(\neg a_1, 1)\}$).

4.5.1.3 Calcul effectif de Σ_C

Il est à noter que :

$$\pi(A_2 \dots A_n) = \max(\pi_{a_1}^{A_1}(A_2 \dots A_n), \pi_{\neg a_1}^{A_1}(A_2 \dots A_n))$$

Soient $\Sigma_1 = \Sigma_{a_1} - \{(a_1, 1)\}$ et $\Sigma_2 = \Sigma_{\neg a_1} - \{(\neg a_1, 1)\}$. La base possibiliste associée à $\pi(A_2 \dots A_n)$ est définie comme suit :

$$\Sigma_c = \{(\phi_i \vee \psi_j, \min(\alpha_i, \beta_j)) : (\phi_i, \alpha_i) \in \Sigma_1, (\psi_j, \beta_j) \in \Sigma_2\}$$

Ainsi, le calcul de Σ_C se fait comme suit :

1. Ajouter $(a_1, 1)$ (respectivement $(\neg a_1, 1)$) à Σ ,
2. Supprimer les clauses contenant a_1 (respectivement $\neg a_1$) ayant la forme $(a_1 \vee x, \alpha)$ (respectivement $(\neg a_1 \vee x, \alpha)$),
3. Remplacer les clauses de la forme $(\neg a_1 \vee x, \alpha)$ (respectivement $(a_1 \vee x, \alpha)$) par (x, α)
4. Soit Σ_1 (respectivement Σ_2) le résultat de l'étape précédente. Alors, $\Sigma_c = \{(\phi_i \vee \psi_j, \min(\alpha_i, \beta_j)) : (\phi_i, \alpha_i) \in \Sigma_1, (\psi_j, \beta_j) \in \Sigma_2\}$

Remarque 4.1 Cette phase se base sur l'élimination des variables. Néanmoins, il faut noter que l'élimination d'une variable induit un prix coûteux en terme d'espace car la taille de la base sera multipliée par deux. De ce fait, en terme d'espace, cette méthode n'est pas intéressante.

4.5.2 Détermination des parents de A_i

La détermination des parents de A_i se fait d'une manière incrémentale. En premier lieu, $Par(A_i)$ va être constitué de l'ensemble des variables qui apparaissent dans des clauses contenant soit a_i soit $\neg a_i$. Néanmoins, il peut exister des variables *cachées* pouvant influencer la distribution conditionnelle $\Pi(A_i \mid Par(A_i))$. Le processus est décrit par l'algorithme suivant [14, 102] : **Algorithme 4.2** :

Détermination des parents de A_i

Données : Σ

Résultat : Parents de A_i

Début

Trouver les parents immédiats,

1. $\Sigma_{A_i} \leftarrow \{(\phi_i, \alpha_i) : (\phi_i, \alpha_i) \in \Sigma, \phi_i \text{ contient soit } a_i \text{ soit } \neg a_i\}$,
2. $Par(A_i) \leftarrow \{V : \exists(\phi_i, \alpha_i) \in \Sigma_{A_i} \text{ qui contient une instance de } V\}$,

Rechercher les parents cachés

1. $B \leftarrow \Sigma$, soit (x_1, \dots, x_n) une instance de $Par(A_i)$

2. Supprimer de B chaque clause contenant x_i ,
3. Remplacer dans B chaque clause de la forme $(\neg x_i \vee \neg x_j \dots \vee \neg x_k \vee \phi, \alpha)$ par (ϕ, α)
4. Soit α (respectivement β) le degré de certitude de a_1 (respectivement de $\neg a_1$) de $B \cup \{(x_1, 1), \dots, (x_n, 1)\}$
5. s'il existe $(\phi, \gamma) \in B$ telle que $\gamma \geq \alpha$ (respectivement $\gamma \geq \beta$), alors :
 $Par(A_i) = Par(A_i) \cup \{V : \phi \text{ contient une instance de } V\}$

Aller à l'étape 2

retourner $Par(A_i)$

Fin

Remarque 4.2 Contrairement à la procédure de passage d'une base qualitative vers un réseau causal basé sur le minimum, le nombre de parents induit par cette procédure augmente. Il ne s'agit pas forcément de parents directs.

4.5.3 Calcul des degrés de possibilités conditionnelles

Soit (x_1, \dots, x_n) une instance de $Par(A_1)$ et a_1 une instance de A_1 . Par définition :

$$\Pi(a_1 \mid x_1 x_2 \dots x_n) = \frac{\Pi(a_1 x_1 \dots x_n)}{\Pi(x_1 \dots x_n)}$$

et

$$\Pi(a_1 \mid x_1 x_2 \dots x_n) = 1 \text{ si } \Pi(x_1 \dots x_n) = 0$$

$\Pi(\phi)$ est calculée syntaxiquement comme suit :

$$\Pi(\phi) = 1 - Inc(\Sigma \cup \{(\phi, 1)\})$$

où Inc représente le degré d'incohérence associé à la base.

Ainsi, le calcul de $\Pi(a_1 \mid x_1 \dots x_n)$ se fait comme suit [14, 102] :

1. Ajouter $\{(x_1, 1), \dots, (x_n, 1)\}$ à Σ . Soit Σ' le résultat obtenu.
2. Calculer $h=1-Inc(\Sigma')$. (h représente $\Pi(x_1 \dots x_n)$).
3. Ajouter $\{(a_1, 1)\}$ à Σ' . Soit Σ'' le résultat obtenu.
4. Calculer $h'=1-Inc(\Sigma'')$. (h' représente $\Pi(a_1 x_1 \dots x_n)$). Ainsi,

$$\Pi(a_1 \mid x_1 \dots x_n) = \begin{cases} 1 & \text{si } h=0 \\ \frac{h'}{h} & \text{sinon} \end{cases}$$

4.6 D'une base possibiliste quantitative vers une base de connaissances qualitative

Dans [15], les auteurs ont montré qu'une base de connaissances possibiliste basée sur le produit peut être transformée en une base de connaissances possibiliste basée sur le minimum.

La base possibiliste basée sur le minimum résultante est obtenue en rajoutant toutes les disjonctions possibles de la base possibiliste basée initiale.

Plus précisément, soit Σ une base possibiliste quantitative. La base possibiliste qualitative associée à Σ est définie par :

$$K_{\Sigma} = \{(d(A_j), \alpha(A_j)) : A_j \subseteq \Sigma \text{ et } A_j \neq 0\}$$

où,

- $d(A_j)$ représente la disjonction des formules propositionnelles dans A_j ,
- $\alpha(A_j)$ représente le poids associé à $d(A_j)$ obtenu en combinant les poids de A_j avec l'opérateur associatif $t(a_i, b_j) = a_i + b_j - a_i * b_j$

(Si A_j est un singleton, alors, $\alpha(A_j)$ est simplement le poids de la formule de A_j).

Exemple 4.3 Soit $\Sigma = \{(a \vee b, 0.4), (-c, 0.9)\}$. Alors, $K_{\Sigma} = \{(a \vee b, 0.4), (-c, 0.9), (a \vee b \vee \neg c, 0.94)\}$.

Remarque 4.3 D'une manière générale, la taille de la base possibiliste qualitative K_{Σ} associée à une base possibiliste quantitative Σ est exponentielle et est estimée à 2^n où n est la taille de Σ .

4.7 conclusion

La théorie des possibilités offre quatre formats de représentation de connaissances :

- Représentations logiques
 - Bases possibilistes qualitatives
 - Bases possibilistes quantitatives
- Représentations graphiques
 - Les réseaux causaux possibilistes basés sur le minimum
 - Les réseaux causaux possibilistes basés sur le produit

L'ensemble de ces formats sont des représentations compactes de la distribution de possibilités.

Ce chapitre a été consacré à présenter les différents travaux de transformations effectués autour de ces différents formats de représentation en théorie des possibilités.

En marge de ces transformations, dans [103], les auteurs ont établi des transformations réciproques entre les bases logiques standards et les bases comparatives. En effet, une base de connaissances Σ peut être représentée par une partition ordonnée WOP (Well Ordered Partition), telle que $WOP(\Sigma) = S_1 \cup \dots \cup S_n$ où S_1 contient les formules classiques

les plus certaines de Σ et S_n contient les formules les moins certaines. Ces correspondances induisent une équivalence entre les trois formats de représentation compactes en théorie des possibilités : logique, graphique et comparatif.

Ces correspondances sont très utiles dans le cas où ces différents formats de connaissances doivent être combinés. Chaque format a son impact sur le plan de la communication, aussi bien pour modéliser les connaissances que pour fournir des connaissances à l'utilisateur [102]. De plus, du point de vue de l'inférence, les formats logiques et graphiques offrent des modes de raisonnement intéressants. Pour la logique standard, il a été montré dans [47] que la machinerie possibiliste induit un coût calculatoire raisonnable. Quant au mode graphique, des améliorations successives sont apportées afin de réduire le coût de la propagation.

En pratique, les intérêts de ces transformations ont été mis en évidence dans différents travaux [104, 105].

Dans le même ordre d'idée, la seconde partie de notre thèse relate les principales investigations réalisées dans ce travail en termes de représentation des connaissances.

Afin de compléter le treillis des transformations en théorie des possibilités, nous proposons des transformations équivalentes entre une base possibiliste quantitative et un graphe causal possibiliste basé sur le produit.

Si nous analysons les différents travaux effectués en théorie des possibilités, nous pouvons en déduire le passage d'une base possibiliste quantitative en un réseau causal possibiliste basé sur le produit. Il suffit pour cela :

1. de transformer la base possibiliste quantitative en une base possibiliste qualitative équivalente en utilisant la procédure décrite dans [15],
2. de transformer la base possibiliste qualitative en un réseau causal possibiliste basé sur le produit, en utilisant la procédure décrite dans [102, 14].

Néanmoins, cette méthode n'est pas optimale car la première étape engendre une base qualitative de taille exponentielle ce qui induit un coût élevé.

En ce qui concerne la transformation d'un réseau causal basé sur le produit vers une base possibiliste quantitative, elle ne peut être déduite à partir des différentes transformations proposées.

De plus, nous nous intéressons à d'autres modes de représentation des connaissances incertaines et plus particulièrement à la logique des pénalités et aux réseaux causaux basés sur les OCF.

La deuxième partie de la thèse est consacrée, justement, à présenter une partie de notre contribution. Nous proposons en premier lieu des procédures de passages directes et réciproques entre les deux modes de représentation des connaissances en théorie des possibilités quantitative. En second lieu, nous mettons en exergue les relations qui existent entre les bases possibilistes quantitatives et les bases de pénalités d'une part et entre les réseaux causaux possibilistes basés sur le produit et les réseaux causaux basés sur les OCF.

Deuxième partie

Transformations entre modes de représentations des connaissances en théorie des possibilités quantitative

Introduction de la partie II

En théorie des possibilités, les deux modes de représentation (logique et graphique) sont tous les deux attractifs. En effet,

- d’une manière générale, la logique a toujours été un moyen robuste pour représenter les connaissances et offre des mécanismes d’inférence puissants,
- les atouts majeurs du mode graphique sont dus à leurs capacités de représenter et de manipuler les relations d’indépendance, permettant ainsi des calculs locaux.

Ainsi, le choix d’un des deux modes de représentation ne sera pas limité par leurs caractéristiques.

Nous allons exposer, dans cette partie, notre contribution relative à l’aspect de la représentation des connaissances qui consiste :

- à proposer des procédures de transformations directes entre les deux modes de représentations des connaissances incertaines autour de la théorie des possibilités. Le passage d’une base possibiliste quantitative vers un graphe causal possibiliste basé sur le produit peut se faire en exploitant les différentes procédures présentées dans le chapitre 4. Ce passage s’effectue en deux temps. La première étape consiste à transformer la base de possibilités quantitative en une base base de possibilités qualitative [15]. La seconde étape consiste à transformer la base de possibilités qualitative obtenue en un réseau causal possibiliste basé sur le produit [102, 14]. Néanmoins, cette méthode s’avère fastidieuse car elle génère des graphes complets. La procédure de passage que nous proposons est directe et ne génère pas de graphes complets.

Le passage d’un graphe causal possibiliste basé sur le produit vers une base possibiliste quantitative quant à lui, ne peut être déduit à partir des précédents travaux effectués en théorie des possibilités. La transformation que nous proposons est linéaire.

- à développer les liens syntaxiques qui existent entre la logique des possibilités quantitative et la logique des pénalités, d’une part et de révéler les liens qu’entretiennent les réseaux causaux possibilistes basés sur le produit avec les réseaux causaux basés sur les OCF.

Cette partie se compose de deux chapitres. Dans le cinquième chapitre, nous allons présenter les transformations réciproques entre les réseaux causaux possibilistes basés sur le produit et les base possibilistes quantitatives.

Dans le sixième chapitre, nous allons étendre ces transformations à d’autres modes de représentations. En mode logique, nous allons mettre en exergue les relations réciproques entre les bases possibilistes quantitatives et les bases de pénalités. En mode graphique, nous exhiberons les relations qu’entretiennent les réseaux causaux possibilistes basés sur le produit et les réseaux causaux basés sur les OCF.

Chapitre 5

Transformations en théorie des possibilités quantitative

5.1 Introduction

Comme nous l'avons vu dans le chapitre précédent, les principaux travaux réalisés en théorie des possibilités se sont focalisés sur les bases de connaissances possibilistes basées sur le minimum. En effet, ces travaux ont proposé :

- des transformations équivalentes entre une base possibiliste qualitative ou standard (basée sur le minimum) et un réseau causal possibiliste basé sur le minimum [13],
- des transformation équivalentes entre une base possibiliste qualitative et un réseau causal possibiliste basé sur le produit [14],
- la transformation d'une base possibiliste basée sur le produit vers une base possibiliste basée sur le minimum [15]. La transformation inverse n'a pas été développée car elle ne présente pas d'intérêt vu qu'il existe des algorithmes d'inférence performants opérant sur les bases de connaissances qualitatives.

A partir des deux premières transformations équivalentes, il est possible de codifier une base de connaissances possibiliste quantitative en un réseau causal possibiliste basé sur le produit et ceci via une base de connaissances qualitative. Ceci peut être réalisé :

- en transformant d'abord une base de connaissances possibiliste basée sur le produit en une base de connaissances possibiliste basée sur le minimum,
- puis en codifiant la base de connaissances possibiliste basée sur le minimum en un réseau causal possibiliste basé sur le produit.

Néanmoins, en pratique, cette codification indirecte n'est pas optimale vu que généralement elle engendre des DAGs complets. Plus précisément, n'importe quel DAG associé à une base de connaissances basée sur le produit est tel que il existe au moins une variable A , ayant l'ensemble des autres variables comme parents. Ceci est dû au fait que la transformation d'une base possibiliste basée sur le produit vers une base possibiliste basée sur le minimum provoque le rajout de toutes les disjonctions ce qui amène à ajouter toutes les dépendances entre les variables (même si elles sont indépendantes).

D'une manière plus formelle, si Σ est une base de connaissances basée sur le produit, alors la base de connaissances K_Σ basée sur le minimum qui lui est associée est définie comme suit :

$$K_\Sigma = \{(d(A_j), \alpha(A_j)) : A_j \subseteq \Sigma \text{ et } a_j \neq 0\}$$

où,

- $d(A_j)$ représente la disjonction des formules propositionnelles dans A_j ,
- $\alpha(A_j)$ est le poids associé à $d(A_j)$ obtenu en combinant les poids de A_j avec l'opérateur associatif $(a_i, b_j) = a_i + b_j - a_i * b_j$

(si A_j est un singleton, alors $\alpha(A_j)$ est simplement le poids de la formule dans A_j).

Par exemple, si $\Sigma = \{(a, .7), (b, .5), (c, .6)\}$, alors $K_\Sigma = \{(a, .7), (b, .5), (c, .6), (a \vee b, .85), (a \vee c, .88), (b \vee c, .80)\}$.

Si nous considérons la base Σ , nous déduisons que les variables A et B sont indépendantes. L'application des deux étapes de transformation nous fait alors aboutir à un graphe constitué de trois nœuds A, B et C et ayant les liens suivants :

- soit un lien de A vers B ou un lien de B vers A,
- soit un lien de A vers C ou un lien de C vers A,
- soit un lien de B vers C ou un lien de C vers B,

Le graphe ainsi obtenu est complet. Ceci se justifie par le fait que la transformation de Σ en une base possibiliste qualitative crée de nouvelles formules $(a \vee b, 0.85)$, $(a \vee c, .88)$, $(b \vee c, .80)$.

Nous voyons bien à partir de ce petit exemple que la taille de la base qualitative K_Σ explose par rapport à la taille de la base possibiliste quantitative Σ .

Plus précisément, soit K_Σ une base de connaissances qualitative associée à une base de connaissances basée sur le produit. Il peut être alors vérifié que pour deux variables A et B, il existe une formule $\phi \in K_\Sigma$ telle que ϕ contient une instance de A et une instance de B. Ceci induit que quelque soit l'ordre établi entre les variables, il existe une variable A ayant $V - \{A\}$ comme l'ensemble de ses parents.

Par contre, il est à noter que la transformation d'un graphe possibiliste basé sur le produit vers une base possibiliste quantitative n'est pas aussi évidente et ne peut être déduite à partir des transformations déjà réalisées dans le cadre de la théorie des possibilités.

Ainsi, la matière de ce chapitre provient des travaux que nous avons réalisés tout au long de cette thèse. Les premières ébauches de nos transformations ont été présentées dans [16]. Ces transformations ont été par la suite développées dans [17], en exposant des exemples complets. Puis les preuves formelles de l'ensemble de ces transformations ont été apportées dans [19]. L'un des objectifs de ce travail est donc de compléter le treillis des transformations en théorie des possibilités, en développant des procédures de transformations entre modes de représentations dans le cadre quantitatif. Ainsi, Nous proposons :

- une nouvelle transformation d'une base possibiliste quantitative en un réseau causal basé sur le produit contournant ainsi les limites citées ci dessus. En effet, dans le cadre de ces transformations, l'ensemble de parents de chaque variable A est au plus égal à l'ensemble constitué à partir de la base de connaissances opérant une

influence directe sur celle ci. Aucune dépendance additionnelle n'est rajoutée.

Par exemple, à partir de la base de connaissances possibiliste $\Sigma = \{(a, .7), (b, .5), (c, .6)\}$, nous obtenons un DAG composé de deux nœuds indépendants A et B.

- une transformation linéaire d'un graphe causal possibiliste basé sur le produit vers une base possibiliste quantitative.

Pour cela, nous allons rappeler dans la section suivante les concepts de bases sur la représentation des connaissances en théorie des possibilités quantitative nécessaires afin d'explicitier les transformations.

5.2 Rappel sur la représentation des connaissances en théorie des possibilités quantitative

Comme cela a été abordé dans les chapitres précédents, en théorie des possibilités quantitative, les connaissances peuvent être représentées en mode logique ou en mode graphique.

5.2.1 Représentation logique

Elle repose principalement sur les notions suivantes :

- En mode logique, une base de connaissances possibiliste quantitative est représentée par un ensemble de formules pondérées :

$$\Sigma = \{(\phi_i, \alpha_i) : i = 1, \dots, n\},$$

où ϕ_i est une formule propositionnelle et $\alpha_i \in]0, 1]$ représente le seuil minimal du degré de nécessité de ϕ_i .

- La distribution de possibilités associée à une base de connaissances possibiliste quantitative Σ est définie par : $\forall \omega \in \Omega$:

$$\pi_{\Sigma}(\omega) = \begin{cases} 1 & \text{si } \forall (\phi_i, \alpha_i) \in \Sigma, \omega \models \phi_i \\ * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma, \omega \not\models \phi_i\} & \text{sinon} \end{cases} \quad (5.1)$$

Exemple 5.1 Soit Σ la base possibiliste quantitative suivante : $\Sigma = \{(\neg a \vee b, .6), (b \vee \neg c \vee d, .4), (c \vee \neg d, .3)\}$. La distribution de possibilités associée à Σ est donnée par la table 5.1.

5.2.2 Représentation graphique

Un graphe causal possibiliste, noté Π_G , est un DAG où les nœuds représentent les variables et les arcs matérialisent les liens causaux entre les variables. Lorsqu'il existe un lien d'un nœud A_i vers un A_j , A_i est un parent de A_j . $Par(A_j)$ représente l'ensemble des parents de la variable A_j et u_j est une instance de $Par(A_j)$.

Un graphe causal possibiliste basé sur le produit est caractérisé par :

TAB. 5.1 – Distributions de possibilités associée à la base possibiliste quantitative Σ de l'exemple 5.1

ω	$\pi_{\Sigma}(\omega)$
$ab\neg cd$.7
$a\neg bcd$.4
$a\neg bc\neg d$.24
$a\neg b\neg cd$.28
$\neg ab\neg cd$.7
$\neg a\neg bc\neg d$.6
$\neg a\neg b\neg cd$.7
autres	1

- **Le conditionnement basé sur le produit** : L'incertitude est représentée au niveau de chaque nœud en termes de distributions de possibilités conditionnelles définies comme suit :
 - Pour les nœuds racines A_i , nous spécifions les degrés de possibilités a priori $\Pi(a_i)$ and $\Pi(\neg a_i)$ avec la condition de normalisation $\max(\Pi(a_i), \Pi(\neg a_i)) = 1$.
 - Pour les autres nœuds A_j , nous spécifions pour chaque instance u_j de $Par(A_j)$, les degrés de possibilités conditionnelles de $\Pi(a_j | u_j)$ et $\Pi(\neg a_j | u_j)$ avec la condition $\max(\Pi(a_j | u_j), \Pi(\neg a_j | u_j)) = 1$.
 Le conditionnement basé sur le produit est défini par :

$$\Pi(\psi | \phi) = \begin{cases} \frac{\Pi(\psi \wedge \phi)}{\Pi(\phi)} & \text{si } \Pi(\phi) \neq 0 \\ 1 & \text{sinon} \end{cases} \quad (5.2)$$

- **Distribution de possibilités associée à un DAG** : Chaque réseau causal basé sur le produit Π_G (DAG et distributions de possibilités conditionnelles locales) induit une distribution de possibilités unique jointe en utilisant la règle de chaînage basée sur le produit qui est similaire à celle utilisée dans les réseaux bayésiens probabilistes. La règle de chaînage est définie comme suit :

$$\pi_G(\omega) = * \{ \Pi(x_i | u_i) : \omega \models x_i \wedge u_i, i = 1, \dots, n \} \quad (5.3)$$

où u_i est une instance de parents de A_i .

Exemple 5.2 Soit le réseau possibiliste basé sur le produit dont la composante graphique est représentée par la figure 5.1 et la composante numérique est représentée par les tables 5.2 à 5.3.

TAB. 5.2 – Distributions de possibilités a priori associées au graphe de l'exemple 5.2

A	$\Pi(A)$	B	$\Pi(B)$
a	.5	b	1
$\neg a$	1	$\neg b$.2

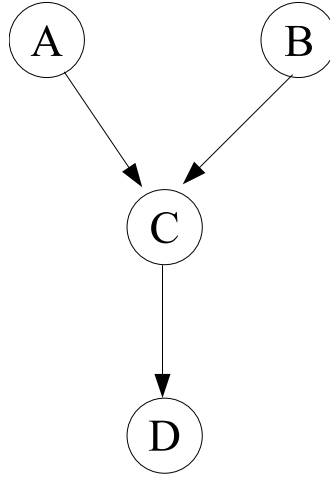


FIG. 5.1 – DAG correspondant à l'exemple 5.2

TAB. 5.3 – Distributions de possibilités conditionnelles associées au graphe de l'exemple 5.2

A B C	$\Pi(C AB)$	D C	$\Pi(D C)$
abc	.2	$c d$	1
$ab\neg c$	1	$c\neg d$.6
$a\neg bc$	1	$\neg cd$.3
$a\neg b\neg c$.4	$\neg c\neg d$	1
$\neg a b c$.5		
$\neg a b\neg c$	1		
$\neg a\neg b c$.3		
$\neg a\neg b\neg c$	1		

5.2.3 Relation entre le conditionnement et l'implication matérielle

Le passage d'un réseau causal basé sur le produit (basé sur le conditionnement) en une base de connaissances quantitative (basé sur l'implication matérielle) nécessite de faire une mise au point concernant la relation qui existe entre le conditionnement et l'implication matérielle en faisant le parallèle entre la théorie des probabilités et la théorie des possibilités.

- En théorie des probabilités, il est clair que le conditionnement est différent de l'implication matérielle. En d'autres termes, en général $P(\neg a \vee b) = P(b | a)$, ne se produit pas. Par exemple, il est facile de montrer que l'ensemble des contraintes $P(a) = r \in [0, 1], P(b | a) = s \in [0, 1], P(b | \neg a) = t \in [0, 1]$, admet toujours une solution.

Néanmoins, comme il a été révélé par Pearl [10], l'ensemble des contraintes $P(a) =$

TAB. 5.4 – Distributions de possibilités associée au graphe de l'exemple 5.2

A	B	C	D	$\pi_G(\omega)$	A	B	C	D	$\pi_G(\omega)$
a	b	c	d	.1	$\neg a$	b	c	d	.5
a	b	c	$\neg d$.06	$\neg a$	b	c	$\neg d$.3
a	b	$\neg c$	d	.15	$\neg a$	b	$\neg c$	d	.3
a	b	$\neg c$	$\neg d$.5	$\neg a$	b	$\neg c$	$\neg d$	1
a	$\neg b$	c	d	.1	$\neg a$	$\neg b$	c	d	0.06
a	$\neg b$	c	$\neg d$.012	$\neg a$	$\neg b$	c	$\neg d$.036
a	$\neg b$	$\neg c$	d	.06	$\neg a$	$\neg b$	$\neg c$	d	.04
a	$\neg b$	$\neg c$	$\neg d$.5	$\neg a$	$\neg b$	$\neg c$	$\neg d$.2

$s \in [0, 1]$ et $P(\neg a \vee b) = t \in [0, 1]$ n'admet pas toujours une solution (par exemple : $P(a) = 0.01$ and $P(\neg a \vee b) = 0.9$).

- En théorie des possibilités, le conditionnement est également différent de l'implication matérielle. A savoir, $\Pi(b | a) \neq \Pi(\neg a \vee b)$. En effet, $\Pi(b | a) = \frac{\Pi(a \wedge b)}{\Pi(a)}$ tandis que $\Pi(\neg a \vee b) = \max(\Pi(a \wedge b), \Pi(\neg a \wedge b), \Pi(\neg a \wedge \neg b))$. En général, $\Pi(\neg a \vee b) \geq \Pi(b | a)$, et nous pouvons avoir une situation où $\Pi(\neg a \vee b) > \Pi(a \wedge b)$ (par exemple $\Pi(\neg a \wedge \neg b) = \Pi(a \wedge \neg b) = 1 > \Pi(a \wedge b) = \Pi(\neg a \wedge b)$). D'une manière similaire, en général, $N(b | a) \neq N(\neg a \vee b)$. En effet, $N(b | a) = 1 - \Pi(\neg b | a)$ et $N(\neg a \vee b) = 1 - \Pi(a \wedge \neg b)$, et en général, $\Pi(\neg b | a) = \frac{\Pi(\neg b \wedge a)}{\Pi(a)} \neq \Pi(\neg b \wedge a)$.

Néanmoins, lorsque $\Pi(a) = 1$ nous pouvons alors vérifier que : $N(b | a) = N(\neg a \vee b)$ puisque $\Pi(\neg b | a) = \Pi(\neg b \wedge a)$.

Même si le conditionnement est en général différent de l'implication matérielle, il est possible de codifier un réseau causal basé sur le produit (basé sur le conditionnement) en une base de connaissances quantitative (basé sur l'implication matérielle).

5.3 D'un réseau causal basé sur le produit vers une base possibiliste quantitative

Dans le but de faciliter la transformation d'un graphe causal possibiliste basé sur le produit vers une base possibiliste quantitative, un réseau causal possibiliste sera représenté par un ensemble de triplets (comme dans [13]) :

$$\mathcal{P}_G = \{(x_i, u_i, \alpha_i) : \Pi(x_i | u_i) = \alpha_i \neq 1 \in \Pi_G\},$$

où x_i est une instance de la variable A_i et u_i est une instance de $Par(A_i)$.

Remarque 5.1 Nous ne considérons que les distributions conditionnelles $\Pi(x_i | u_i) = \alpha_i \neq 1$, car "1" représente un élément neutre pour le produit.

Exemple 5.3 Considérons le réseau causal possibiliste basé sur le produit illustré par la figure 5.2. Les distributions de possibilités locales sont données par les tables 5.5 et 5.6. Les distributions de possibilités jointes associées au graphe sont données par la table 5.7.

Ainsi, en adoptant la représentation explicitée ci-dessus, le graphe sera codifié comme suit :

$$\mathcal{P}_G = \{(a, \emptyset, .3), (\neg b, a, .4), (\neg b, \neg a, .2), (c, a, .8), (\neg c, \neg a, .9), (\neg d, bc, .2), (d, b\neg c, .4), (\neg d, \neg bc, .1), (d, \neg b\neg c, .9)\}.$$

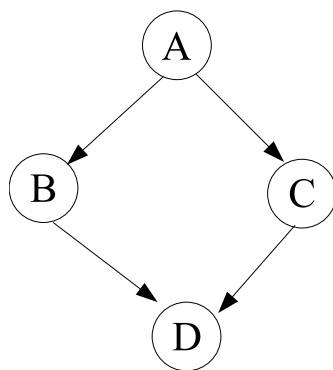


FIG. 5.2 – DAG correspondant à l'exemple 5.3

TAB. 5.5 – Distributions de possibilités conditionnelles initiales de $\Pi(A)$, $\Pi(B | A)$ et $\Pi(C | A)$

A	$\Pi(A)$	AB	$\Pi(B A)$	AC	$\Pi(C A)$
a	.3	ab	1	ac	.8
$\neg a$	1	$a\neg b$.4	$a\neg c$	1
		$\neg ab$	1	$\neg ac$	1
		$\neg a\neg b$.2	$\neg a\neg c$.9

En fait, la construction d'une base de connaissances possibiliste Σ_G associée à un graphe causal basé sur le produit Π_G est même immédiate.

Elle consiste simplement à remplacer chaque triplet (x, u, α) du graphe possibiliste directe \mathcal{P}_G par une formule possibiliste $(\neg x \vee \neg u, 1 - \alpha)$.

La codification, signifie ici que la distribution de possibilités induite à partir de Σ_G (en utilisant l'équation 5.2) est identique à celle induite à partir du graphe G (en utilisant l'équation 5.3).

TAB. 5.6 – Distributions de possibilités conditionnelles initiales de $\Pi(D \mid B \wedge C)$

BCD	$\Pi(D \mid B \wedge C)$
bcd	1
$bc\neg d$.2
$b\neg cd$.4
$b\neg c\neg d$	1
$\neg bcd$	1
$\neg bc\neg d$.1
$\neg b\neg cd$.9
$\neg b\neg c\neg d$	1

TAB. 5.7 – Distributions de possibilités jointes utilisant la règle de chaînage basée sur le produit

A	B	C	D	$\pi_G(A \wedge B \wedge C \wedge D)$	A	B	C	D	$\pi_G(A \wedge B \wedge C \wedge D)$
a	b	c	d	.24	$\neg a$	b	c	d	1
a	b	c	$\neg d$.048	$\neg a$	b	c	$\neg d$.2
a	b	$\neg c$	d	.12	$\neg a$	b	$\neg c$	d	.36
a	b	$\neg c$	$\neg d$.3	$\neg a$	b	$\neg c$	$\neg d$.90
a	$\neg b$	c	d	.096	$\neg a$	$\neg b$	c	d	0.2
a	$\neg b$	c	$\neg d$.0096	$\neg a$	$\neg b$	c	$\neg d$.02
a	$\neg b$	$\neg c$	d	.108	$\neg a$	$\neg b$	$\neg c$	d	.162
a	$\neg b$	$\neg c$	$\neg d$.12	$\neg a$	$\neg b$	$\neg c$	$\neg d$.18

Intuitivement, $(\neg x \vee \neg u, 1 - \alpha)$ est obtenue à partir de $N(\neg x \mid u) = 1 - \alpha \iff \Pi(x \mid u) = \alpha$.

Afin d'illustrer cette codification, considérons une formule pondérée $\Sigma_{x,u} = \{(\neg x \vee \neg u, 1 - \alpha)\}$ associée à $\Pi(x \mid u) = \alpha$. Sa distribution de possibilités $\pi_{x,u}$ est égale à :

$$\pi_{x,u}(\omega) = \begin{cases} 1 & \text{si } \omega \models \neg x \vee \neg u \\ \alpha & \text{sinon} \end{cases}$$

Proposition 1 *Les degrés de possibilités conditionnelles $\Pi(x \mid u) = \alpha$ sont recouverts à partir de $\pi_{x,u}$.*

Preuve 5.1 *En effet, $\Pi_{x,u}(x \wedge u) = \alpha$ se produit vu que chaque interprétation qui satisfait $x \wedge u$ falsifie $(\neg x \vee \neg u, 1 - \alpha)$. De plus, $\Pi_{x,u}(u) = \max(\Pi_{x,u}(x \wedge u), \Pi_{x,u}(\neg x \wedge u))$. Comme chaque interprétation qui satisfait $\neg x \vee u$, satisfait $(\neg x \vee \neg u, 1 - \alpha)$, nous obtenons $\Pi_{x,u}(u) = 1$.*

Ainsi, si nous appliquons le conditionnement basé sur le produit, nous obtenons : $\Pi_{x,u}(x \mid u) = \frac{\Pi_{x,u}(x \wedge u)}{\Pi_{x,u}(u)} = \alpha$.

D'une manière formelle, une base possibiliste Σ_G associée à un graphe causal basé sur le produit Π_G est définie comme suit :

$$\Sigma_G = \{(\neg x_i \vee \neg u_i, 1 - \alpha_i) : (x_i, u_i, \alpha_i) \in \mathcal{P}_G\}. \tag{5.4}$$

Il est néanmoins important de souligner que même si nous remplaçons le conditionnement par l'implication matérielle, ceci ne signifie pas (comme mentionné auparavant) que le conditionnement soit équivalent à l'implication matérielle. Il ne s'agit donc pas de se restreindre aux réseaux causaux pour lesquels $\Pi(x \mid u) = \Pi(x \vee \neg u)$. En effet, tout type de réseau causal peut être utilisé et codifié.

Exemple 5.4 Reconsidérons le graphe causal possibiliste de l'exemple précédent représenté par :

$$\mathcal{P}_G = \{(a, \emptyset, .3), (\neg b, a, .4), (\neg b, \neg a, .2), (c, a, .8), (\neg c, \neg a, .9), (\neg d, bc, .2), (d, b\neg c, .4), (\neg d, \neg bc, .1), (d, \neg b\neg c, .9)\}.$$

La base possibiliste Σ_G associée à Π_G correspond à :

$$\Sigma_G = \{(\neg a, .7), (b \vee \neg a, .6), (b \vee a, .8), (\neg c \vee \neg a, .2), (c \vee a, .1), (d \vee \neg b \vee \neg c, .8), (\neg d \vee \neg b \vee c, .6), (d \vee b \vee \neg c, .9), (\neg d \vee b \vee c, .1)\}.$$

En utilisant l'équation 5.1, nous obtenons les distributions de possibilités associées à la base Σ_G , données par la table 5.8.

TAB. 5.8 – Distributions de possibilités associée à la base π_{Σ_G} de l'exemple 5.4

A	B	C	D	$\pi_{\Sigma_G}(A \wedge B \wedge C \wedge D)$	A	B	C	D	$\pi_G(A \wedge B \wedge C \wedge D)$
a	b	c	d	.24	$\neg a$	b	c	d	1
a	b	c	$\neg d$.048	$\neg a$	b	c	$\neg d$.2
a	b	$\neg c$	d	.12	$\neg a$	b	$\neg c$	d	.36
a	b	$\neg c$	$\neg d$.3	$\neg a$	b	$\neg c$	$\neg d$.90
a	$\neg b$	c	d	.096	$\neg a$	$\neg b$	c	d	0.2
a	$\neg b$	c	$\neg d$.0096	$\neg a$	$\neg b$	c	$\neg d$.02
a	$\neg b$	$\neg c$	d	.108	$\neg a$	$\neg b$	$\neg c$	d	.162
a	$\neg b$	$\neg c$	$\neg d$.12	$\neg a$	$\neg b$	$\neg c$	$\neg d$.18

Dans ce qui suit, nous allons montrer l'équivalence entre la distribution de possibilités associée à la base de connaissances $\pi_{\Sigma_G}(\omega)$ et les distributions de possibilités conditionnelles induites par le réseau causal possibiliste $\pi_G(\omega)$.

Proposition 2 Soit Π_G un graphe causal possibiliste basé sur le produit et soit Σ_G une base de connaissances possibiliste basée sur le produit associée à Π_G en utilisant l'équation (5.4). Nous avons :

$$\forall \omega \in \Omega, \pi_{\Sigma_G}(\omega) = \pi_G(\omega),$$

où π_{Σ_G} est obtenue en utilisant l'équation (5.1)

et π_G est obtenue en utilisant l'équation(5.3).

Preuve 5.2 La preuve se base sur l'idée que le calcul des degrés de possibilités d'une interprétation est en fonction, d'une part, des degrés de possibilités conditionnelles $\Pi(x_i \mid$

u_i) qu'elle satisfait et d'autre part, en fonction des poids des formules $(\neg x_i \vee \neg u_i)$ (associées à $(x_i \mid u_i)$) qu'elle falsifie.

Plus précisément, soit $\omega = (x_1, x_2, \dots, x_n)$ et $\Sigma_G = \{(\neg x_i \vee \neg u_i, 1 - \alpha_i) : \Pi(x_i, u_i, \alpha_i) \in \mathcal{P}_G\}$

En utilisant l'équation(5.2), deux cas se présentent :

- $\omega = (x_1, x_2, \dots, x_n)$ n'est pas un modèle de Σ . Ce qui signifie que : $\pi_{\Sigma_G}(\omega) \neq 1$.
D'où $\pi_{\Sigma_G}(\omega) = * \{1 - (1 - \alpha_i) : (\neg x_i \vee \neg u_i, 1 - \alpha_i) \in \Sigma_G \text{ et } \omega \not\models \neg x_i \vee \neg u_i\}$
 $= * \{\alpha_i : (\neg x_i \vee \neg u_i, 1 - \alpha_i) \in \Sigma_G \text{ and } \omega \models x_i \wedge u_i\}$
 $= * \{\Pi(x_i \mid u_i) : \omega \models x_i \wedge u_i, i = 1 \dots n\}$
 $= \pi_G(\omega)$
- $\omega = (x_1, x_2, \dots, x_n)$ est un mode de Σ c'est à dire que $\pi_{\Sigma_G}(\omega) = 1$.
Ce qui signifie que : $\forall (\neg x_i \vee \neg u_i, 1 - \alpha_i) \in \Sigma_G, \omega \models \neg x_i \vee \neg u_i$.
Ce qui induit également à : $\forall (x_i, u_i, \alpha_i) \in \Pi_G, \omega \not\models (x_i \wedge u_i)$.
Par conséquent, $\forall i = 1 \dots n, \Pi(x_i \mid u_i) = 1$ (les degrés de possibilités conditionnelles non considérés dans \mathcal{P}_G sont égaux à 1). D'où, $\pi_G(\omega) = 1$.

Exemple 5.5 Il peut être facilement vérifié que la distribution de possibilités jointe générée à partir du graphe causal basé sur le produit Π_G de l'exemple 5.3 est identique à celle générée à partir de la base de connaissances de l'exemple 5.4. Donc, $\forall \omega, \pi_G(\omega) = \pi_{\Sigma_G}(\omega)$, comme en témoigne le tableau 5.7 (qui donne les distributions possibilistes associées au graphe de l'exemple 5.3) et le tableau 5.8 (qui donne les distributions possibilistes associées à la base possibiliste quantitative de l'exemple 5.4).

Soit $\omega_1 = \neg a \wedge b \wedge \neg c \wedge d$. En appliquant l'équation (5.3) nous obtenons $\pi_G(\omega_1) = \Pi(\neg a) * \Pi(b \mid \neg a) * \Pi(\neg c \mid \neg a) * \Pi(d \mid b \wedge \neg c) = 1 * 1 * .9 * .4 = .36$

D'autres part, soit Σ_G la base de connaissances possibiliste basée sur le produit correspondant au graphe π_G . En utilisant l'équation (5.1), nous avons : ω_1 falsifie les formules $(c \vee a, .1)$ et $(\neg d \vee \neg b \vee c, .6)$. D'où $\pi_{\Sigma}(\omega_1) = (1-.1) * (1-.6) = .9 * .4 = .36$

5.4 D'une base possibiliste quantitative vers un réseau causal basé sur le produit

La transformation inverse d'une base possibiliste quantitative Σ vers un graphe causal possibiliste basé sur le produit Π_G est moins immédiate.

En effet, nous avons d'abord besoin d'établir un ensemble de lemmes concernant certaines modifications équivalentes de la base de connaissances possibiliste quantitative.

5.4.1 Réductions et extensions des bases possibilistes quantitatives

Nous allons dans cette section définir les lemmes et les propositions nécessaires se rapportant aux différentes transformations qui vont être opérées sur toute base possibiliste quantitative avant la transformation en un graphe causal possibiliste basé sur le produit.

Nous commençons par définir la notion de bases de connaissances équivalentes :

Définition 5.1 Deux bases de connaissances quantitatives Σ et Σ' sont dites équivalentes si elles induisent la même distribution de possibilité, à savoir :

$$\forall \omega \in \Omega, \pi_{\Sigma}(\omega) = \pi_{\Sigma'}(\omega).$$

Le premier lemme indique que les tautologies ainsi que les formules ayant des poids nuls peuvent être supprimées des bases possibilistes quantitatives sans modifier, pour autant, les distributions de possibilités.

Lemme 5.1 Soit Σ une base possibiliste quantitative, Alors : i) Si $(\top, \alpha_i) \in \Sigma$ alors Σ et $\Sigma' = \Sigma - \{(\top, \alpha_i)\}$ sont équivalentes.
ii) Si $(\phi, 0) \in \Sigma$ alors Σ et $\Sigma' = \Sigma - \{(\phi, 0)\}$ sont équivalentes.

La preuve est immédiate. Les tautologies peuvent être supprimées vu que seules les formules qui sont falsifiées par une interprétation donnée sont prises en compte lors du calcul des distributions de possibilités.

La suppression des tautologies est une phase importante vu qu'elle évite de générer des relations de dépendances fictives entre les variables. Par exemple, la tautologie représentée par la formule $(a \vee \neg a \vee b, 1)$ peut induire à l'existence d'un lien entre les variables A et B.

Pour les formules ayant des poids nuls, elles peuvent être supprimées vu que '1' est un élément neutre pour l'opérateur produit.

Les deux lemmes suivants concernent la réduction des bases possibilistes quantitatives.

Lemme 5.2 (Réduction de formules ayant des poids différents)

Soit Σ une base possibiliste quantitative, Alors soit $(x \vee p, \alpha) \in \Sigma$ et $(x \vee \neg p, \beta) \in \Sigma$, avec $\alpha < \beta$. Soit Σ' la base de connaissances quantitative obtenue à partir de Σ en remplaçant $\{(x \vee p, \alpha), (x \vee \neg p, \beta)\}$ par $\{(x, \alpha), (x \vee \neg p, \frac{(1-\alpha)-(1-\beta)}{(1-\alpha)})\}$, à savoir $\Sigma' = \Sigma - \{(x \vee p, \alpha), (x \vee \neg p, \beta)\} \cup \{(x, \alpha), (x \vee \neg p, \frac{(1-\alpha)-(1-\beta)}{(1-\alpha)})\}$. Σ et Σ' sont alors équivalentes.

Preuve 5.3 La preuve est basée sur le fait que si une interprétation satisfait x alors, elle satisfait $x \vee p$ et $x \vee \neg p$ et inversement. et si une interprétation falsifie x alors elle falsifie soit $x \vee p$ ou $x \vee \neg p$ mais pas les deux. Si ω falsifie $x \vee p$, la distribution de possibilités ne va pas changer étant donné que $x \vee \neg p$ est satisfaite. Et si $x \vee \neg p$ est falsifiée, le poids γ associé à $x \vee \neg p$ doit vérifier l'égalité : $(1 - \beta) = (1 - \alpha) * (1 - \gamma)$. A savoir, $\gamma = 1 - \frac{1-\beta}{1-\alpha} = \frac{(1-\alpha)-(1-\beta)}{1-\alpha}$. Par conséquent, ces modifications ne vont pas affecter les distributions de possibilités jointes.

D'une manière plus formelle, soit ω une interprétation possible. Deux cas se présentent :

- ω est un modèle de x . Ce qui signifie que $\omega \models x \vee p$ et $\omega \models x \vee \neg p$.

De nouveau, deux cas sont à considérer :

- ω est un modèle de Σ , à savoir ω satisfait toutes les formules de Σ . Vu que ω satisfait $x \vee p$ et $x \vee \neg p$, alors ω satisfait aussi l'ensemble des formules de Σ' .

Par conséquent, dans ce cas : $\pi_{\Sigma}(\omega) = \pi_{\Sigma'}(\omega) = 1$.

- ω n'est pas un modèle de Σ . En utilisant l'équation (5.1), nous avons :

$$\begin{aligned} \pi_{\Sigma}(\omega) &= * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma \text{ et } \omega \not\models \phi_i\} \\ &= * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma - \{(x \vee p, \alpha), (x \vee \neg p, \beta)\} \text{ et } \omega \not\models \alpha_i\} \text{ (étant donné que } \\ &\omega \models x \vee p \text{ et } \omega \models x \vee \neg p.) \\ &= * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma - \{(x \vee p, \alpha), (x \vee \neg p, \beta)\} \cup \{(x, \alpha), (x \vee \neg p, \frac{(1-\alpha)-(1-\beta)}{(1-\alpha)})\} \\ &\text{ et } \omega \not\models \phi_i\} \\ &\text{(étant donné que } \omega \text{ satisfait } x \text{ et } x \vee \neg p) \\ &= * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma' \text{ et } \omega \not\models \phi_i\} \\ &= \pi_{\Sigma'}(\omega) \end{aligned}$$
- ω n'est pas un modèle de x . Ceci signifie d'abord que ω n'est ni un modèle de Σ ni de Σ' . Nous considérons également à ce niveau deux cas :
 - $\omega \models x \vee \neg p$.
 Ceci implique que $\omega \not\models x \vee p$ (car $\omega \not\models x$)
 En utilisant l'équation (5.1) :

$$\begin{aligned} \pi_{\Sigma}(\omega) &= * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma \text{ et } \omega \not\models \phi_i\} \\ &= (1 - \alpha) * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma - \{(x \vee p, \alpha)\} \text{ et } \omega \not\models \phi_i\} \\ &\text{(vu que } \omega \text{ falsifie } x \vee p) \\ &= (1 - \alpha) * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma - \{(x \vee p, \alpha), (x \vee \neg p, \beta)\} \text{ et } \omega \not\models \phi_i\} \\ &\text{(vu que } \omega \text{ satisfait } x \vee \neg p) \\ &= (1 - \alpha) * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma - \{(x \vee p, \alpha), (x \vee \neg p, \beta)\} \cup \{(x \vee \neg p, \frac{(1-\alpha)-(1-\beta)}{(1-\alpha)})\} \\ &\text{ et } \omega \not\models \phi_i\} \\ &\text{(vu que } \omega \text{ satisfait } x \vee \neg p) \\ &= * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma - \{(x \vee p, \alpha), (x \vee \neg p, \beta)\} \cup \{(x, \alpha), (x \vee \neg p, \frac{(1-\alpha)-(1-\beta)}{(1-\alpha)})\} \\ &\text{ et } \omega \not\models \phi_i\} \\ &\text{(vu que } \omega \text{ falsifie } x) = * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma' \text{ et } \omega \not\models \phi_i\} \\ &= \pi_{\Sigma'}(\omega) \end{aligned}$$
 - $\omega \not\models x \vee \neg p$.
 Ceci implique que $\omega \models x \vee p$ (vu que $\omega \not\models x$)
 En utilisant l'équation (5.1), nous obtenons :

$$\begin{aligned} \pi_{\Sigma}(\omega) &= * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma \text{ et } \omega \not\models \phi_i\} \\ &= * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma - \{(x \vee p, \alpha)\} \text{ et } \omega \not\models \phi_i\} \\ &\text{(vu que } \omega \text{ satisfait } (x \vee p)) \\ &= (1 - \beta) * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma - \{(x \vee p, \alpha), (x \vee \neg p, \beta)\} \text{ et } \omega \not\models \phi_i\} \\ &\text{(vu que } \omega \text{ falsifie } (x \vee \neg p)) \\ &= (1 - \frac{(1-\alpha)-(1-\beta)}{(1-\alpha)}) * (1 - \alpha) * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma - \{(x \vee p, \alpha), (x \vee \neg p, \beta)\} \\ &\text{ et } \omega \not\models \phi_i\} \\ &= (1 - \alpha) * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma - \{(x \vee p, \alpha), (x \vee \neg p, \beta)\} \cup \{(x \vee \neg p, \frac{(1-\alpha)-(1-\beta)}{(1-\alpha)})\} \\ &\text{ et } \omega \not\models \phi_i\} \\ &\text{(vu que } \omega \text{ falsifie } (x \vee \neg p)) \\ &= * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma - \{(x \vee p, \alpha), (x \vee \neg p, \beta)\} \cup \{(x, \alpha), (x \vee \neg p, \frac{(1-\alpha)-(1-\beta)}{(1-\alpha)})\} \end{aligned}$$

$$\begin{aligned}
 & \text{et } \omega \not\models \phi_i \} \\
 & (\text{car } \omega \text{ falsifie } x) \\
 & = * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma' \text{ et } \omega \not\models \phi_i \} \\
 & = \pi_{\Sigma'}(\omega)
 \end{aligned}$$

Exemple 5.6 Soit Σ la base possibiliste quantitative suivante : $\Sigma = \{(\neg a \vee b, .2), (\neg b \vee c \vee \neg d, .5), (\neg b \vee c \vee d, .8), (d, e, .3)\}$. En appliquant le lemme 5.2 aux formules $(\neg b \vee c \vee \neg d, .5)$ et $(\neg b \vee c \vee d, .8)$, nous obtenons une base possibiliste quantitative Σ' équivalente à Σ : $\Sigma' = \{(\neg a \vee b, .2), (\neg b \vee c, .5), (\neg b \vee c \vee d, .6), (d, e, .3)\}$.

L'égalité entre les deux distributions de possibilités correspondant à Σ et Σ' est vérifiée : Soit $\omega_1 = a \wedge b \wedge \neg c \wedge \neg d \wedge \neg e$. En utilisant l'équation (5.1), nous obtenons :

$$\begin{aligned}
 \pi_{\Sigma}(\omega_1) &= (1 - .8) * (1 - .3) = .2 * .7 = .14 \text{ et} \\
 \pi_{\Sigma'}(\omega_1) &= (1 - .5) * (1 - .6) * (1 - .3) = .5 * .4 * .7 = .14
 \end{aligned}$$

Lemme 5.3 (Réduction de formules ayant des poids identiques)

Soit Σ une base possibiliste quantitative. Soient $(x \vee p, \alpha)$ et $(x \vee \neg p, \alpha)$ deux formules pondérées de Σ . Soit Σ' une base de connaissances possibiliste quantitative obtenue à partir de Σ en remplaçant $(x \vee p, \alpha)$ et $(x \vee \neg p, \alpha)$ par (x, α) , à savoir $\Sigma' = (\Sigma - \{(x \vee p, \alpha), (x \vee \neg p, \alpha)\}) \cup \{(x, \alpha)\}$.

Alors, Σ' et Σ sont équivalentes.

La preuve est immédiate vu que ce lemme peut être vu comme un cas particulier du lemme 5.2 (lorsque $\alpha = \beta$).

Le lemme suivant est le lemme inverse du lemme 5.3. Il montre que le remplacement de la formule (x, α) par $(x \vee p, \alpha), (x \vee \neg p, \alpha)$ n'affecte pas la distribution de possibilités jointe induite.

Lemme 5.4 (Extension des formules)

Soit Σ une base possibiliste. Soit (x, α) une formule de Σ . Soit Σ' une base de connaissances possibiliste quantitative obtenue à partir de Σ en remplaçant (x, α) par $(x \vee p, \alpha)$ et $(x \vee \neg p, \alpha)$, à savoir $\Sigma' = \Sigma - \{(x, \alpha)\} \cup \{(x \vee p, \alpha), (x \vee \neg p, \alpha)\}$.

Les bases Σ' et Σ sont alors équivalentes.

La preuve est identique à celle du lemme 5.2 (avec $\alpha = \beta$) décrite auparavant.

Le lemme suivant montre comment manipuler les redondances dans une base possibiliste quantitative.

Lemme 5.5 (Redondances)

Soient $(x, \alpha) \in \Sigma$ et $(x, \beta) \in \Sigma$. Soit Σ' la base de connaissances possibiliste quantitative obtenue à partir de Σ en remplaçant $\{(x, \alpha), (x, \beta)\}$ par $(x, \alpha + \beta - \alpha * \beta)$. Ainsi, $\Sigma' = \Sigma - \{(x, \alpha), (x, \beta)\} \cup \{(x, \alpha + \beta - \alpha * \beta)\}$. Alors Σ et Σ' sont équivalentes.

Preuve 5.4 *La preuve se base sur le fait que si une interprétation falsifie (x, α) et (x, β) alors, sa distribution de possibilité va contenir $(1 - \alpha) * (1 - \beta) = 1 - (\alpha + \beta - \alpha * \beta)$. Il en résulte que le résultat est le même si nous considérons uniquement une seule formule $(x, \alpha + \beta - \alpha * \beta)$. D'une manière formelle, considérons les deux situations possibles :*

- $\omega \models x$. Nous notons à ce niveau que si ω est de plus un modèle de Σ , alors c'est également un modèle de Σ' , d'où $\pi_\Sigma(\omega) = \pi_{\Sigma'}(\omega) = 1$.

Si ω n'est pas un modèle de Σ . Alors, en utilisant l'équation (5.1), nous obtenons :

$$\pi_\Sigma(\omega) = * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma \text{ et } \omega \not\models \phi_i\}$$

$$= * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma - \{(x, \alpha), (x, \beta)\} \text{ et } \omega \not\models \phi_i\}$$

(vu que ω satisfait x)

$$= * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma - \{(x, \alpha), (x, \beta)\} \cup \{(x, \alpha + \beta - \alpha * \beta)\} \text{ et } \omega \not\models \phi_i\}$$

(vu que ω satisfait x)

$$= * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma' \text{ et } \omega \not\models \phi_i\}$$

$$= \pi'_{\Sigma'}(\omega)$$

- ω n'est pas un modèle de x . A savoir, $\omega \not\models x$.

En utilisant l'équation (5.2), nous obtenons :

$$\pi_\Sigma(\omega) = * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma \text{ et } \omega \not\models \phi_i\}$$

$$= (1 - \alpha) * (1 - \beta) * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma - \{(x, \alpha), (x, \beta)\} \text{ et } \omega \not\models \phi_i\}$$

(vu que ω falsifie x)

$$= (1 - (\alpha + \beta - \alpha * \beta)) * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma - \{(x, \alpha), (x, \beta)\} \text{ et } \omega \not\models \phi_i\}$$

$$= * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma - \{(x, \alpha), (x, \beta)\} \cup \{(x, \alpha + \beta - \alpha * \beta)\} \text{ et } \omega \not\models \phi_i\}$$

(vu que ω falsifie x)

$$= * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma' \text{ et } \omega \not\models \phi_i\}$$

$$= \pi'_{\Sigma'}(\omega).$$

Exemple 5.7 *Soit Σ la base possibiliste quantitative suivante : $\Sigma = \{(a \vee c, .1), (\neg b \vee \neg c, .4), (\neg c \vee d \vee e, .3), (\neg c \vee d \vee e, .6)\}$. En appliquant le lemme 5.5 aux formules $(\neg c \vee d \vee e, .3)$ et $(\neg c \vee d \vee e, .6)$, nous obtenons une base possibiliste quantitative Σ' équivalente à Σ : $\Sigma' = \{(a \vee c, .1), (\neg b \vee \neg c, .72), (\neg c \vee d \vee e, .6)\}$.*

L'égalité entre les deux distributions de possibilités correspondantes à Σ et à Σ' est vérifiée : Soit $\omega_1 = a \wedge b \wedge c \wedge \neg d \wedge \neg e$. En utilisant l'équation (5.1), nous obtenons :

$$\pi_\Sigma(\omega_1) = (1 - .4) * (1 - .3) * (1 - .6) = .6 * .7 * .4 = .168 \text{ et}$$

$$\pi_{\Sigma'}(\omega_1) = (1 - .4) * (1 - .72) = .6 * .28 = .168$$

5.4.2 Construction du DAG

En exploitant l'ensemble des lemmes décrits précédemment, nous proposons dans cette section de construire le DAG associé à une base de connaissances possibiliste quantitative. Nous supposons au préalable que la base de connaissances ne contient pas de tautologies, vu que ces dernières peuvent être supprimées sans altérer la distribution possibiliste associée.

L'idée pour construire un DAG associé à une base de connaissances possibiliste quantitative est de commencer d'ordonner les variables X_1, X_2, \dots, X_n arbitrairement.

Intuitivement, l'ordre doit refléter le fait que les parents de la variable X_i doivent figurer

parmi $X_{i+1} \dots X_n$. Ainsi, si une variable X_i , n'a pas de parents, alors cela correspondrait à l'ensemble vide.

L'étape suivante consiste à opérer un certain nombre de transformations sur la base de connaissances possibiliste quantitative Σ jusqu'à obtention d'une base de connaissances possibiliste quantitative de la forme : $\Sigma_{X_1} \cup \Sigma_{X_2} \dots \cup \Sigma_{X_n}$. Chaque Σ_{X_i} doit contenir les éléments d'information relatifs à la variable X_i ainsi que ses parents.

Nous commençons d'abord par définir $Par(X_i)$ comme un ensemble de variables X_j (différent de X_i) pour lesquelles il existe au moins une instance de X_j dans Σ_{X_i} . Σ_{X_i} doit alors satisfaire des conditions suivantes :

- Si $X_j \in Par(X_i)$ alors $j > i$.
- Chaque formule dans Σ_{X_i} est de la forme $(x_i \vee u_i, \alpha_i)$ où x_i est une instance de X_i et u_i est une instance de $Par(X_i)$.

La construction de $\Sigma_{X_1} \cup \Sigma_{X_2} \dots \cup \Sigma_{X_n}$ est obtenue en utilisant la procédure suivante :

Procédure 1 : Prétraitement de Σ

Entrée : une base de connaissances possibiliste quantitative Σ

Un ordre total sur les variables

Sortie : Une base de connaissances possibiliste quantitative équivalente à Σ de la forme $\Sigma_{X_1} \cup \Sigma_{X_2} \dots \cup \Sigma_{X_n}$

Pour $i=1$ jusqu'à n faire

Begin

- a1** Application du Lemme 5.3 (réduction de formules ayant le même poids) à toutes les paires de formules contenant une instance de X_i .
- a2** $\Sigma_{X_i} := \{(\phi_i, \alpha_i) : (\phi_i, \alpha_i) \in \Sigma \text{ et } \phi_i \text{ contient une instance de } X_i \text{ et des instances de } X_j \text{ avec } j > i\}$.
 $Par(X_i) := \{X_j : \exists(\phi_i, \alpha_i) \in \Sigma_{X_i} \text{ tel que } \phi_i \text{ contient une instance de } X_j\}$.
- a3** Application du lemme 5.4 qui consiste à étendre chaque formule $(x, \alpha) \in \Sigma_{X_i}$ à toutes les instances de $Par(X_i)$.
- a4** Application du lemme 5.5 à Σ_{X_i} qui consiste à remplacer les redondances.
- a5** Application du lemme 5.2 à Σ_{X_i} qui, dans le cas où $(x \vee p, \alpha) \in \Sigma_{X_i}$ et $(\neg x \vee p, \beta) \in \Sigma_{X_i}$, avec $\alpha < \beta$, consiste à :
 - a5.1** supprimer $(x \vee p, \alpha)$ et $(\neg x \vee p, \beta)$ à partir de Σ_{X_i} ,
 - a5.2** rajouter $(\neg x \vee p, \frac{(1-\alpha)-(1-\beta)}{(1-\alpha)})$ à Σ_{X_i} , et
 - a5.3** rajouter (p, α) à Σ .

Fin

Proposition 3 Soit Σ une base de connaissances possibiliste quantitative. Soit $\Sigma' = \Sigma_{X_1} \cup \Sigma_{X_2} \dots \cup \Sigma_{X_n}$ une base possibiliste obtenue par la procédure décrite ci-dessus. Alors, Σ et Σ' sont équivalentes.

La preuve est immédiate vu que chaque étape de la procédure préserve l'équivalence entre les bases de connaissances.

Définition 5.2 (Construction du DAG) Soit Σ une base de connaissances possibiliste quantitative, et $\Sigma_{X_1} \cup \Sigma_{X_2} \dots \cup \Sigma_{X_n}$ est la base possibiliste obtenue par la procédure décrite précédemment. Le graphe G_Σ associé à Σ est défini comme suit :

- les nœuds sont les variables X_i de Σ ,
- un arc est tracé de X_j à X_i si $X_j \in \text{Par}(X_i)$ (où $\text{Par}(X_i)$ est défini dans l'étape **a2** de la procédure précédente).

Proposition 4 Soit Σ une base de connaissances possibiliste quantitative. Alors le graphe G_Σ associé à Σ est un DAG.

La structure du DAG est assurée par le fait que l'ordre de priorité a priori X_1, X_2, \dots, X_n est fixé, et l'étape **a2** garantie que les parents d'une variable X_i , s'ils existent, sont nécessairement parmi $X_{i+1} \dots X_n$.

Exemple 5.8 Nous allons illustrer la construction du DAG, en utilisant la base possibiliste quantitative suivante :

$$\Sigma = \{(\neg a \vee c \vee d, .6), (\neg a \vee c \vee \neg d, .6), (\neg a \vee h, .8), (a \vee f \vee \neg h, .3), (b \vee \neg e, .7), (c \vee d, .1), (\neg c \vee d \vee \neg h, .7), (d \vee f, .8), (\neg d \vee g, .5), (e \vee \neg g, .2), (f \vee g, .4)\}.$$

Σ contient huit variables ordonnées d'une manière arbitraire comme suit : $X_1 = A, X_2 = B, X_3 = C, X_4 = D, X_5 = E, X_6 = F, X_7 = G$ et $X_8 = H$.

Nous allons appliquer la procédure précédente afin de construire l'ensemble des Σ_{X_i} .

Traitement du nœud A :

La première étape consiste à appliquer le lemme 5.3 (étape **a1**) à $(\neg a \vee c \vee d, .6), (\neg a \vee c \vee \neg d, .6)$. Ceci va entraîner des changements sur Σ :

$$\Sigma = \{(\neg a \vee c, .6), (\neg a \vee h, .8), (a \vee f \vee \neg h, .3), (b \vee \neg e, .7), (c \vee d, .1), (\neg c \vee d \vee \neg h, .7), (d \vee f, .8), (\neg d \vee g, .5), (e \vee \neg g, .2), (f \vee g, .4)\}.$$

L'étape **a2** initialise Σ_A et définit ainsi l'ensemble de parents de A. Nous obtenons :

$$\Sigma_A = \{\neg a \vee c, .6), (\neg a \vee h, .8), (a \vee f \vee \neg h, .3)\}, \text{ et } \text{Par}(A) = \{C, F, H\}.$$

L'étape **a3** consiste à appliquer le lemme 5.4. Les formules $(\neg a \vee c, .6), (\neg a \vee h, .8)$ et $(a \vee f \vee \neg h, .3)$ doivent être étendues aux différentes instances des parents de A.

- L'extension de la formule $(\neg a \vee c, .6)$ donne $\{(\neg a \vee c \vee f \vee h, .6), (\neg a \vee c \vee f \vee \neg h, .6), (\neg a \vee c \vee \neg f \vee h, .6), (\neg a \vee c \vee \neg f \vee \neg h, .6)\}$.
- L'extension de la formule $(\neg a \vee h, .8)$ donne : $\{(\neg a \vee c \vee f \vee h, .8), (\neg a \vee c \vee \neg f \vee h, .8), (\neg a \vee \neg c \vee f \vee h, .8), (\neg a \vee \neg c \vee \neg f \vee h, .8)\}$.

- L'extension de la formule $(a \vee f \vee \neg h, .3)$ donne :
 $\{(a \vee c \vee f \vee \neg h, .3), (a \vee \neg c \vee f \vee \neg h, .3)\}$.

Après avoir appliqué ces étapes, nous obtenons :

$$\Sigma_A = \{(\neg a \vee c \vee f \vee h, .6), (\neg a \vee c \vee f \vee \neg h, .6), (\neg a \vee c \vee \neg f \vee h, .6), (\neg a \vee c \vee \neg f \vee \neg h, .6), (\neg a \vee c \vee f \vee h, .8), (\neg a \vee c \vee \neg f \vee h, .8), (\neg a \vee \neg c \vee f \vee h, .8), (\neg a \vee \neg c \vee \neg f \vee h, .8), (a \vee c \vee f \vee \neg h, .3), (a \vee \neg c \vee f \vee \neg h, .3)\}$$

A ce niveau, il faudrait appliquer le lemme 5.5 (étape **a4**) dans le but d'éliminer les redondances suivantes :

- $(\neg a \vee c \vee f \vee h, .6)$ et $(\neg a \vee c \vee f \vee h, .8)$,
- $(\neg a \vee c \vee \neg f \vee h, .6)$ et $(\neg a \vee c \vee \neg f \vee h, .8)$.

Ainsi, nous obtenons :

$$\Sigma_A = \{(\neg a \vee c \vee f \vee h, .92), (\neg a \vee c \vee f \vee \neg h, .6), (\neg a \vee c \vee \neg f \vee h, .92), (\neg a \vee c \vee \neg f \vee \neg h, .6), (\neg a \vee \neg c \vee f \vee h, .8), (\neg a \vee \neg c \vee \neg f \vee h, .8), (a \vee c \vee f \vee \neg h, .3), (a \vee \neg c \vee f \vee \neg h, .3)\}.$$

Enfin, nous appliquons la dernière étape **a5**, à savoir, la réduction entre $(a \vee c \vee f \vee \neg h, .3)$ et $(\neg a \vee c \vee f \vee \neg h, .6)$ entraîne à remplacer ces deux formules par $(\neg a \vee c \vee f \vee \neg h, .42)$ et à rajouter $(c \vee f \vee \neg h, .3)$ à Σ .

Après avoir appliqué l'ensemble de ces étapes, nous obtenons la base finale Σ_A suivante :

$$\Sigma_A = \{(\neg a \vee c \vee f \vee h, .92), (\neg a \vee c \vee f \vee \neg h, .42), (\neg a \vee c \vee \neg f \vee h, .92), (\neg a \vee c \vee \neg f \vee \neg h, .6), (\neg a \vee \neg c \vee f \vee h, .8), (\neg a \vee \neg c \vee \neg f \vee h, .8), (a \vee \neg c \vee f \vee \neg h, .3)\} \text{ et } \Sigma = \{(b \vee \neg e, .7), (c \vee d, .1), (\neg c \vee d \vee \neg h, .7), (c \vee f \vee \neg h, .3), (d \vee f, .8), (\neg d \vee g, .5), (e \vee \neg g, .2), (f \vee g, .4)\}$$

Traitement du nœud B :

L'application de l'étape 1 de la procédure nous permet d'obtenir Σ_B ainsi que l'ensemble des parents de B comme suit :

$$\Sigma_B = \{(b \vee \neg e, .7)\}$$

d'où : $Par(B) = \{E\}$

Traitement du nœud C :

Pour le nœud C, l'initialisation de Σ_C donne : $\Sigma_C = \{(c \vee d, .1), (\neg c \vee d \vee \neg h, .7), (c \vee f \vee \neg h, .3)\}$.

D'où $Par(C) = \{D, F, H\}$.

En appliquant le lemme 5.4, l'extension de Σ_C aux différentes instances de ses parents produit :

$$\Sigma_C = \{(c \vee d \vee f \vee h, .1), (c \vee d \vee f \vee \neg h, .1), (c \vee d \vee \neg f \vee h, .1), (c \vee d \vee \neg f \vee \neg h, .1), (\neg c \vee d \vee f \vee \neg h, .7), (\neg c \vee d \vee \neg f \vee \neg h, .7), (c \vee d \vee f \vee \neg h, .3), (c \vee \neg d \vee f \vee \neg h, .3)\}$$

L'étape suivante consiste à appliquer le lemme 5.5 sur la redondance aux formules

$(c \vee d \vee f \vee \neg h, .1)$ et $(c \vee d \vee f \vee \neg h, .3)$. Nous obtenons ainsi :

$$\Sigma_C = \{(c \vee d \vee f \vee h, .1), (c \vee d \vee f \vee \neg h, .37), (c \vee d \vee \neg f \vee h, .1), (c \vee d \vee \neg f \vee \neg h, .1), (\neg c \vee d \vee f \vee \neg h, .7), (\neg c \vee d \vee \neg f \vee \neg h, .7), (c \vee \neg d \vee f \vee \neg h, .3)\}$$

Enfinement, la dernière étape consiste à appliquer le lemme 5.2 afin de réduire les formules :

- $(c \vee d \vee f \vee \neg h, .37)$ et $(\neg c \vee d \vee f \vee \neg h, .7)$,
- $(c \vee d \vee \neg f \vee \neg h, .1)$ et $(\neg c \vee d \vee \neg f \vee \neg h, .7)$.

Nous obtenons ainsi, $\Sigma_C = \{(c \vee d \vee f \vee h, .1), (\neg c \vee d \vee f \vee \neg h, .52), (c \vee d \vee \neg f \vee h, .1), (\neg c \vee d \vee \neg f \vee \neg h, .66), (c \vee \neg d \vee f \vee \neg h, .3)\}$

Et les formules $(d \vee f \vee \neg h, .37)$ et $(d \vee \neg f \vee \neg h, .1)$ seront alors rajoutées à Σ .

$$\Sigma = \{(d \vee f, .8), (\neg d \vee g, .52), (d \vee f \vee \neg h, .37), (d \vee \neg f \vee \neg h, .1), (e \vee \neg g, .2), (f \vee g, .4)\}$$

Traitement du nœud D :

Pour le nœud D, nous avons :

$$\Sigma_D = \{(d \vee f, .8), (\neg d \vee g, .52), (d \vee f \vee \neg h, .37), (d \vee \neg f \vee \neg h, .1)\}.$$

Ainsi, $Par(D) = \{F, G, H\}$.

L'extension des formules de Σ_D aux différentes instances des ses parents produit :

$$\Sigma_D = \{(d \vee f \vee g \vee h, .8), (d \vee f \vee g \vee \neg h, .8), (d \vee f \vee \neg g \vee h, .8), (d \vee f \vee \neg g \vee \neg h, .8), (\neg d \vee f \vee g \vee h, .52), (\neg d \vee f \vee g \vee \neg h, .52), (\neg d \vee \neg f \vee g \vee h, .52), (\neg d \vee \neg f \vee g \vee \neg h, .52), (d \vee f \vee g \vee \neg h, .37), (d \vee f \vee \neg g \vee \neg h, .37), (d \vee \neg f \vee g \vee \neg h, .1), (d \vee \neg f \vee \neg g \vee \neg h, .1)\}.$$

A ce stade, nous appliquons le lemme 5.5 (étape a4) afin de traiter les redondances entre les formules suivantes :

- $(d \vee f \vee g \vee \neg h, .8)$ et $(d \vee f \vee g \vee \neg h, .37)$
- $(d \vee f \vee \neg g \vee \neg h, .8)$ et $(d \vee f \vee \neg g \vee \neg h, .37)$

Ainsi, $\Sigma_D = \{(d \vee f \vee g \vee h, .8), (d \vee f \vee g \vee \neg h, .874), (d \vee f \vee \neg g \vee h, .8), (d \vee f \vee \neg g \vee \neg h, .874), (\neg d \vee f \vee g \vee h, .52), (\neg d \vee f \vee g \vee \neg h, .52), (\neg d \vee \neg f \vee g \vee h, .52), (\neg d \vee \neg f \vee g \vee \neg h, .52), (d \vee \neg f \vee g \vee \neg h, .1), (d \vee \neg f \vee \neg g \vee \neg h, .1)\}.$

La dernière étape consiste à appliquer le lemme 5.2 afin de réduire les formules suivantes :

- $(d \vee f \vee g \vee h, .8)$ et $(\neg d \vee f \vee g \vee \neg h, .52)$
- $(d \vee f \vee g \vee \neg h, .874)$ et $(\neg d \vee f \vee g \vee \neg h, .52)$
- $(d \vee \neg f \vee g \vee \neg h, .1)$ et $(\neg d \vee \neg f \vee g \vee \neg h, .52)$

Cette opération induit la base Σ_D suivante :

$$\Sigma_D = \{(d \vee f \vee g \vee h, .58), (d \vee f \vee g \vee \neg h, .737), (d \vee f \vee \neg g \vee h, .8), (d \vee f \vee \neg g \vee \neg h, .874), (\neg d \vee \neg f \vee g \vee h, .52), (\neg d \vee \neg f \vee g \vee \neg h, .466), (d \vee \neg f \vee \neg g \vee \neg h, .1)\}.$$

et le rajout des formules $(f \vee g \vee h, .52)$, $(f \vee g \vee \neg h, .52)$ et $(\neg f \vee g \vee \neg h, .1)$ à Σ .

$$\Sigma = \{(e \vee \neg g, .2), (f \vee g, .4), (f \vee g \vee h, .52), (f \vee g \vee \neg h, .52) \text{ et } (\neg f \vee g \vee \neg h, .1)\}$$

Traitement du nœud E :

$$\Sigma_E = \{(e \vee \neg g, .2)\}.$$

D'où $Par(E) = \{G\}$.

Traitement du nœud F :

Pour le nœud F, la base correspondante est :

$$\Sigma_F = \{(f \vee g, .4), (f \vee g \vee h, .52), (f \vee g \vee \neg h, .52), (\neg f \vee g \vee \neg h, .1)\}.$$

D'où $Par(F) = \{G, H\}$.

L'extension de l'ensemble des formules aux différentes instances des parents de F donne :

$$\Sigma_F = \{(f \vee g \vee h, .4), (f \vee g \vee \neg h, .4), (f \vee g \vee h, .52), (f \vee g \vee \neg h, .52), (\neg f \vee g \vee \neg h, .1)\}$$

L'application du lemme 5.5 permet de réduire :

- les formules $(f \vee g \vee h, .4)$ et $(f \vee g \vee h, .52)$,
- les formules $(f \vee g \vee \neg h, .4)$ et $(f \vee g \vee \neg h, .52)$.

Nous obtenons ainsi : $\Sigma_F = \{(f \vee g \vee h, .712), (f \vee g \vee \neg h, .712), (\neg f \vee g \vee \neg h, .1)\}$

Finalement, nous appliquons le lemme 5.2 aux formules $(f \vee g \vee \neg h, .712)$ et $(\neg f \vee g \vee \neg h, .1)$. Ainsi nous obtenons la sous base Σ_F suivante : $\Sigma_F = \{(f \vee g \vee h, .712), (f \vee g \vee \neg h, .68)\}$

La formule $(g \vee \neg h, .1)$, qui en résulte, est rajoutée à Σ . Ainsi, $\Sigma = \{(g \vee \neg h, .1)\}$

Traitement du nœud G :

Il découle de l'étape précédente que la base correspondante au nœud H est $\Sigma_G = \{(g \vee \neg h, .1)\}$. D'où $Par(G) = \{H\}$

Le nœud H est une racine, ce qui signifie que sa base correspondante est vide.

$$\Sigma_H = \emptyset.$$

Après avoir déterminé les parents de chaque variable, nous obtenons le DAG représenté par la figure 5.3.

5.4.3 Calcul des degrés de possibilités conditionnelles

Soit $\Sigma = \Sigma_{X_1} \cup \Sigma_{X_2} \cup \dots \cup \Sigma_{X_n}$ une base de connaissances possibiliste obtenue à partir de la procédure précédente. Le calcul des degrés de possibilités conditionnelles, pour chaque instance de X_i et chaque instance de parents de X_i , est immédiate.

Plus précisément, soit X_i une variable et $Par(X_i) = \{Y_1, Y_2, \dots, Y_n\}$ l'ensemble de ses parents. Soit x une instance de la variable X_i et $u = y_1 \wedge y_2 \wedge \dots \wedge y_n$ une instance de $Par(X_i)$. Les distributions conditionnelles locales sont définies comme suit :

$$\Pi(x | u) = \begin{cases} 1 - \alpha_i & \text{si } (\neg x \vee \neg u, \alpha_i) \in \Sigma_{X_i} \\ 1 & \text{sinon} \end{cases} \quad (5.5)$$

Proposition 5 Soient Σ une base de connaissances possibiliste quantitative, et $\Sigma_{X_1} \cup \Sigma_{X_2} \dots \cup \Sigma_{X_n}$ la base possibiliste obtenue à partir de la procédure ci-dessus. Soit $\Pi(X_i | Par(X_i))$ les distributions de possibilités conditionnelles données par l'équation (5.5). Alors $\Pi(X_i | Par(X_i))$ sont normalisées.

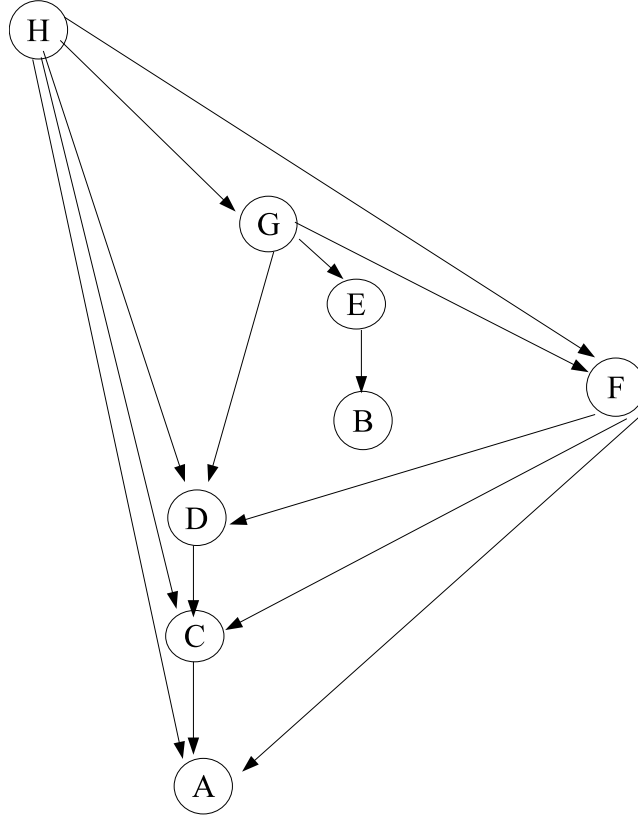


FIG. 5.3 – DAG associé à la base possibiliste de l'exemple 5.3

Preuve 5.5 La preuve est immédiate, vu qu'après application de l'étape **a5** de la procédure précédente, il n'existe aucun x et u (instances de X_i et $\text{Par}(X_i)$) tels que aussi bien $(\neg x \vee u, \alpha)$ que $(x \vee u, \beta)$ appartiennent à Σ_{X_i} .

Par conséquent, au moins un de $(\neg x \vee u, \alpha)$ et $(x \vee u, \beta)$ n'appartient pas à Σ_{X_i} .

D'où, en utilisant l'équation (5.4) au moins un de $\Pi(x \mid \neg u)$ ou $\Pi(\neg x \mid \neg u)$ est égal à 1.

Exemple 5.9 Soit à compléter l'exemple précédent où les Σ_{X_i} ' sont :

$$\Sigma_A = \{(\neg a \vee c \vee f \vee h, .92), (\neg a \vee c \vee f \vee \neg h, .42), (\neg a \vee c \vee \neg f \vee h, .92), (\neg a \vee c \vee \neg f \vee \neg h, .6), (\neg a \vee \neg c \vee f \vee h, .8), (\neg a \vee \neg c \vee \neg f \vee h, .8), (a \vee \neg c \vee f \vee \neg h, .3)\}$$

$$\Sigma_B = \{(b \vee \neg e, .7)\}$$

$$\Sigma_C = \{(c \vee d \vee f \vee h, .1), (\neg c \vee d \vee f \vee \neg h, .52), (c \vee d \vee \neg f \vee h, .1), (\neg c \vee d \vee \neg f \vee \neg h, .66), (c \vee \neg d \vee f \vee \neg h, .3)\}$$

$$\Sigma_D = \{(d \vee f \vee g \vee h, .58), (d \vee f \vee g \vee \neg h, .737), (d \vee f \vee \neg g \vee h, .8), (d \vee f \vee \neg g \vee \neg h, .874), (\neg d \vee \neg f \vee g \vee h, .52), (\neg d \vee \neg f \vee g \vee \neg h, .466), (d \vee \neg f \vee \neg g \vee \neg h, .1)\}$$

$$\Sigma_E = \{(e \vee \neg g, .2)\}$$

$$\begin{aligned} \Sigma_F &= \{(f \vee g \vee h, .712), (f \vee g \vee \neg h, .68)\} \\ \Sigma_G &= \{(g \vee \neg h, .1)\} \\ \Sigma_H &= \emptyset \end{aligned}$$

L'application de l'équation (5.5) permet d'obtenir les distributions de possibilités conditionnelles locales. Elles sont données par les tables (5.9) à (5.15).

TAB. 5.9 – Distributions de possibilités locales conditionnelles finales de $\Pi(A \mid CFH)$

$A \mid CFH$	cfh	$cf\neg h$	$c\neg fh$	$c\neg f\neg h$
a	1	.2	1	.2
$\neg a$	1	1	.7	1

$A \mid CFH$	$\neg cfh$	$\neg cf\neg h$	$\neg c\neg fh$	$\neg c\neg f\neg h$
a	.4	.08	.58	.08
$\neg a$	1	1	1	1

TAB. 5.10 – Distributions de possibilités conditionnelles finales de $\Pi(B \mid E)$

$B \mid E$	e	$\neg e$
b	1	1
$\neg b$.3	1

TAB. 5.11 – Distributions de possibilités locales conditionnelles finales de $\Pi(C \mid DFH)$

$C \mid DFH$	dfh	$df\neg h$	$d\neg fh$	$d\neg f\neg h$
c	1	1	1	1
$\neg c$	1	1	.7	1

$C \mid DFH$	$\neg dfh$	$\neg df\neg h$	$\neg d\neg fh$	$\neg d\neg f\neg h$
c	.34	1	.48	1
$\neg c$	1	.9	1	.9

Proposition 6 Soit Σ une base possibiliste quantitative. Soit Π_G un graphe causal possibiliste associé à Σ , où le DAG associé est obtenu en utilisant la définition (5.2), et les distributions conditionnelles locales sont obtenues en utilisant l'équation (5.5). Alors,

$$\forall \omega, \pi_\Sigma(\omega) = \pi_G(\omega).$$

Preuve 5.6 La preuve se base sur le fait que Σ et $\Sigma_{X_1} \cup \Sigma_{X_2} \dots \cup \Sigma_{X_n}$, obtenue à partir de la procédure précédente, sont équivalentes, et si nous calculons la base de connaissances associée à Π_G en utilisant l'équation (5.4) nous obtenons exactement $\Sigma_{X_1} \cup \Sigma_{X_2} \dots \cup \Sigma_{X_n}$.

TAB. 5.12 – Distributions de possibilités locales conditionnelles finales de $\Pi(D | FGH)$

$D FGH$	fgh	$fg\neg h$	$f\neg gh$	$f\neg g\neg h$
d	1	1	.534	.48
$\neg d$.9	1	1	1
$A FGH$	$\neg fgh$	$\neg fg\neg h$	$\neg f\neg gh$	$\neg f\neg g\neg h$
d	1	1	1	1
$\neg d$.126	.2	.263	.42

TAB. 5.13 – Distributions de possibilités locales conditionnelles finales de $\Pi(E | G)$

$E G$	g	$\neg g$
e	1	1
$\neg e$.8	1

Exemple 5.10 Reconsidérons l'exemple précédent. Soit $\omega_1 = \neg a \vee b \vee c \vee \neg d \vee e \vee \neg f \vee \neg g \vee h$.

En appliquant l'équation (5.1), la distribution de possibilités jointe associée à ω_1 est : $\pi_\Sigma = (1 - .3) * (1 - .7) * (1 - .8) * (1 - .4) = .02$ car l'interprétation ω_1 falsifie les formules $(a \vee f \vee \neg h, .3), (\neg c \vee d \vee \neg h, .7), (d \vee f, .8)$ et $(f \vee g, .4)$.

D'autres part, en appliquant l'équation (5.3), la distribution de possibilités jointe associée à ω_1 est : $\pi_G = \Pi(\neg a | c \wedge \neg f \wedge h) * \Pi(b | e) * \Pi(c | \neg d \wedge \neg f \wedge h) * \Pi(\neg d | \neg f \wedge \neg g \wedge h) * \Pi(e | \neg g) * \Pi(\neg f | \neg g \wedge h) * \Pi(\neg g | h) * \Pi(h) = .7 * 1 * .48 * .263 * 1 * .34 * .9 * 1 = .02$

D'où $\pi_\Sigma(\omega_1) = \pi_G(\omega_1)$

5.5 Conclusion

Dans ce chapitre, nous avons présenté les transformations proposées entre différents formats de représentation des connaissances incertaines dans le cadre de la théorie des possibilités.

la procédure de passage d'un graphe causal possibiliste basé sur le produit vers une base de connaissances possibiliste quantitative est de complexité linéaire.

Sur le plan de la complexité, cette procédure est similaire à la procédure de transformation d'un graphe causal possibiliste basé sur le minimum vers une base possibiliste standard [13].

La transformation inverse d'une base possibiliste quantitative vers un graphe causal possibiliste basé sur le produit est moins directe et nécessite des traitements préliminaires sur la bases de connaissances. De plus, à une base possibiliste quantitative peut correspondre plusieurs réseaux causaux possibilistes basés sur le produit équivalents à une base possibiliste quantitative tout dépend de l'ordre arbitraire fixé sur les variables. D'autres part, le calcul de la complexité de cette transformation dépend du nombre maximal de parents associé aux variables. D'une manière plus précise, cette complexité est égale à $O(2^{nb - \text{maxparents}})$ où $nb - \text{maxparents}$ représente le nombre maximal de

TAB. 5.14 – Distributions de possibilités locales conditionnelles finales de $\Pi(F | GH)$

$F GH$	gh	$g\neg h$	$\neg gh$	$\neg g\neg h$
f	1		1	1
$\neg f$	1	1	.34	.288

TAB. 5.15 – Distributions de possibilités locales conditionnelles finales de $\Pi(G | H)$ et $\Pi(H)$

$\Pi(G H)$	h	$\neg h$	H	$\Pi(H)$
g	1	1	h	1
$\neg g$.9	1	$\neg h$	1

parents associé aux variables.

Ceci implique que le nombre maximal de parents peut être utilisé comme un critère afin de définir l'ordre entre les variables. La cardinalité $\text{Par}(X_i)$ peut alors être utilisée comme un critère afin d'attribuer un ordre entre les variables. A chaque étape, nous pouvons sélectionner la variable ayant le nombre minimal de parents.

En termes de complexité calculatoire, la transformation d'une base possibiliste quantitative vers un graphe causal possibiliste basé sur le produit est plus intéressante si nous la comparons à la procédure de transformation d'une base possibiliste standard (qualitative) vers un graphe causal possibiliste basé sur le minimum décrite dans [13]. En effet, cette dernière transformation nécessite des étapes supplémentaires dans le but d'assurer la cohérence du réseau causal possibiliste basé sur le minimum.

Sur ce même élan des transformations, nous allons exhiber dans le chapitre suivant quelques relations entre :

- une base possibiliste quantitative et une base à pénalités,
- un réseau causal basé sur les fonctions conditionnelles ordinales de Spohn et les réseaux causaux possibilistes basés sur le produit.

Chapitre 6

La théorie des possibilités quantitative, les OCF et la logique des pénalités

6.1 Introduction

Toujours dans l'esprit de faire le parallèle entre les modes de représentations des connaissances incertaines, nous nous intéressons dans ce chapitre aux relations qui existent entre les modes de représentation qu'offre la théorie des possibilités quantitative et d'autres modes de représentation.

Plus précisément, il s'agit de :

- faire le parallèle entre la logique possibiliste quantitative et la logique des pénalités [22, 23],
- d'établir un rapprochement entre les réseaux causaux basés sur le produit et les réseaux causaux basés sur les OCF (Ordinal Conditional Functions) [25, 26].

Nous énonçons un corollaire qui découle de ces équivalences et qui concerne la codification d'une base de pénalités en un réseau causal basé sur les OCF et inversement. Cette étude a fait l'objet de notre travail dans [19].

Les fonctions conditionnelles ordinales de Spohn, offrent des modes de représentations riches. Spohn [25, 26] propose de modéliser l'état épistémique d'un agent par une fonction ordinale conditionnelle, qui peut être considérée comme un classement des interprétations. Une fonction conditionnelle ordinale est une fonction de l'ensemble des mondes possibles vers la classe des ordinaux telle que au moins un monde est associé à 0. Dans [106], Williams propose d'utiliser les fonctions ordinales conditionnelles pour la révision des croyances. Dans [107], les auteurs définissent les réseaux causaux basés sur les OCF. Le chapitre est organisé comme suit. La section 2 fait le parallèle entre la logique possibiliste et la logique des pénalités. La section suivante donne un bref aperçu sur les fonctions conditionnelles ordinales de Spohn. Les sections 4 et 5 mettent en évidence les relations qui existent entre un réseau causal possibiliste basé sur le produit et un réseau causal basé sur les OCF. La section 6 présente le corollaire, qui découle des transfor-

mations présentées dans le chapitre 5 ainsi que les transformations présentées dans ce chapitre et qui met en exergue les relations équivalentes entre un réseau causal basé sur les OCF et une base de pénalités.

6.2 La logique possibiliste quantitative et la logique des pénalités

Comme nous l'avons vu dans le chapitre 2, la logique des pénalités, introduite par Pinkas [22] et développée par Dupin et al. [23], associe à chaque formule de la base de connaissances un poids qui représente le prix à payer si cette formule est violée. Ainsi, une base de pénalités $PK = \{(\phi_i, k_i) : i = 1, \dots, n\}$ est un ensemble fini de formules pondérées où ϕ_i est une formule propositionnelle et $k_i \in \mathbb{N} \cup \{+\infty\}$.

Chaque base de pénalités induit un ordre sur l'ensemble des interprétations, où le coût d'une interprétation $\omega \in \Omega$, noté par $\kappa_{PK}(\omega)$, est égal à la somme des pénalités des formules dans PK violées par ω . D'une manière formelle, :

$$\kappa_{PK}(\omega) = \begin{cases} 0 & \text{si } \forall (\phi_i, k_i) \in PK, \omega \models \phi_i \\ \sum \kappa_i \{\omega \models \neg \phi_i, (\phi_i, k_i) \in PK\} & \text{sinon} \end{cases} \quad (6.1)$$

Exemple 6.1 Soit PK la base de pénalités suivante : $PK = \{(a \vee \neg b, 7), (b \vee c, 3), (\neg a \vee b \vee c, 12), (\neg c, 14)\}$. La table 6.1 représente les valeurs de la fonction coûts associées aux différentes interprétations de la base PK.

TAB. 6.1 – Fonction coût associée à la base de pénalités PK de l'exemple 6.1

ω	$\kappa_{PK}(\omega)$
$abcd$	14
$ab\neg cd$	0
$a\neg bc$	14
$a\neg b\neg c$	15
$\neg abc$	21
$\neg ab\neg c$	7
$\neg a\neg bc$	14
$\neg a\neg b\neg c$	3

Il est clair, qu'il existe des relations étroites entre la logique possibiliste quantitative et la logique des pénalités. Dans [18], l'auteur a mis en exergue les liens sémantiques qui existent entre la logique des pénalités et la logique des possibilités et plus précisément, la relation entre la distribution de possibilités et la fonction coût.

Dans le cadre de notre recherche, nous avons approfondi les liens syntaxiques qu'entretiennent la logique des pénalités et la logique des possibilités quantitative.

Nous allons dans ce qui suit définir la base possibiliste quantitative associée à une base de pénalités puis nous allons proposer une codification possible d'une base possibiliste quantitative en une base de pénalités.

6.2.1 D'une base de pénalités vers une base possibiliste quantitative

Soit PK une base de connaissances à pénalités définie par :

$$PK = \{(\phi_i, k_i) : i = 1, \dots, n\}.$$

La base possibiliste basée sur le produit associée à la base de pénalités PK est définie par : $\Sigma_{PK} = \{(\phi_i, \alpha_i) : (\phi_i, k_i) \in PK \text{ et } \alpha_i = 1 - e^{-k_i}\}$ L'utilisation de la fonction exponentielle permet de transformer l'échelle $[0, +\infty]$ en $[0, 1]$.

La relation qui existe entre les distributions associées à une base possibiliste quantitative Σ_{PK} et à la base de pénalités PK correspondante est donnée par la proposition suivante :

Proposition 7

$$\forall \omega, \pi_{\Sigma_{PK}}(\omega) = e^{-\kappa_{PK}(\omega)}$$

Preuve 6.1 En effet, en partant de la définition de la distribution de possibilité associée à une base possibiliste quantitative $\forall \omega \in \Omega$:

$$\pi_{\Sigma}(\omega) = \begin{cases} 1 & \text{si } \forall (\phi_i, \alpha_i) \in \Sigma, \omega \models \phi_i \\ * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma, \omega \not\models \phi_i\} & \text{sinon} \end{cases} \quad (6.2)$$

Deux cas sont à considérer :

- ω n'est pas un modèle de Σ_{PK} . Alors,

$$\begin{aligned} \pi_{\Sigma}(\omega) &= * \{1 - \alpha_i : (\phi_i, \alpha_i) \in \Sigma_{PK} \text{ et } \omega \not\models \phi_i\} \\ &= * \{1 - (1 - e^{-k_i}) : (\phi_i, k_i) \in PK \text{ et } \omega \not\models \phi_i\} \\ &= * \{e^{-k_i} : (\phi_i, k_i) \in PK \text{ et } \omega \not\models \phi_i\} \\ &= e^{-\sum \{k_i : (\phi_i, k_i) \in PK \text{ et } \omega \not\models \phi_i\}} \\ &= e^{-\kappa_{PK}(\omega)}. \end{aligned}$$

- ω est un modèle de Σ_{PK} .

Dans ce cas, nous avons : $\pi_{\Sigma}(\omega) = e^{-\kappa_{PK}(\omega)}$ vu que $\kappa_{PK}(\omega) = 0$, et $\pi_{\Sigma}(\omega) = 1$.

Exemple 6.2 soit la base de pénalités suivante :

$$PK = \{(\neg a \vee b, 1), (a \vee \neg b \vee c, 4), (\neg b \vee \neg c, 2), (\neg c, 3)\}$$

La base de possibilités quantitative correspondante est :

$$\Sigma_{PK} = \{(\neg a \vee b, .632), (a \vee \neg b \vee c, .981), (\neg b \vee \neg c, .864), (\neg c, .950)\}$$

En utilisant les équations (6.1) et (6.2), nous obtenons les distributions jointes associées à la base de pénalités et à la base de possibilités correspondante, représentées par la table 6.2.

Ainsi, la relation entre les distributions jointes est vérifiée : $\forall \omega, \pi_{\Sigma}(\omega) = e^{-\kappa_{PK}(\omega)}$

TAB. 6.2 – Les distributions jointes associées à PK et à Σ de l'exemple 6.2

ω	$\kappa_{PK}(\omega)$	$\pi_{\Sigma}(\omega) = e^{-\kappa(\omega)}$
$abcd$	5	.0068
$ab\neg cd$	0	1
$a\neg bc$	4	.018
$a\neg b\neg c$	1	.36
$\neg abc$	5	.0068
$\neg ab\neg c$	4	.019
$\neg a\neg bc$	3	.05
$\neg a\neg b\neg c$	0	1

6.2.2 D'une base possibiliste quantitative vers une base à pénalités

La transformation inverse d'une base possibiliste quantitative vers une base de pénalités est également possible. Il suffit pour cela de définir une échelle d'incertitude spécifique quantifiant les formules de la base possibiliste quantitative.

Une échelle possible peut être de la forme : $\{e^{-\infty}, \dots, e^{-2}, e^{-1}, e^0\}$.

6.3 Les fonctions conditionnelles ordinales

6.3.1 concepts de base

Une fonction conditionnelle ordinaire (OCF) ou fonction Kappa définie par Spohn[25, 26] permet de représenter les connaissances incertaines du monde par une fonction notée κ , de \mathcal{L} dans une classe d'ordinaux représentée par l'ensemble des entiers naturels $\mathcal{N} \cup \{+\infty\}$.

$\kappa(\omega)$ peut être vu comme étant le degré d'impossibilité de ω .

Une fonction kappa de Spohn présente les propriétés suivantes :

- $\kappa(\omega)=0$ signifie que rien n'empêche ω d'être le monde réel (avec la convention que 0 correspond à l'impossibilité minimale, c.a.d à la plausibilité maximale).
- $\kappa(\omega)=+\infty$ signifie que ω n'est certainement pas le monde réel.
- Plus $\kappa(\omega)$ est petit plus l'interprétation est préférée. Ainsi, si $\kappa(\omega) < \kappa(\omega')$ alors ω est plus plausible que ω' . C'est la convention inverse de la théorie des possibilités.
- Une OCF est dite admissible par Spohn si $\kappa^{-1}(0) \neq \emptyset$.

A partir d'une distribution kappa κ , nous définissons :

$$\forall \phi, \kappa(\phi) = \min\{\kappa(\omega) : \omega \models \phi\} \quad (6.3)$$

avec

$$\kappa(\perp) = +\infty$$

Nous avons : $\kappa(\phi \vee \psi) = \min\{\kappa(\phi), \kappa(\psi)\}$.

La notion de conditionnement a été également définie dans les OCF. Elle transforme une fonction conditionnelle ordinaire en une autre. Elle est définie comme suit :

$$\kappa(\omega \mid \phi) = \begin{cases} \kappa(\omega) - \kappa(\phi) & \text{si } \omega \models \phi \\ +\infty & \text{sinon} \end{cases} \quad (6.4)$$

En outre, si $\phi \cup \psi \neq \emptyset$ alors

$$\kappa(\psi \mid \phi) = \min\{\kappa(\omega \mid \phi) : \omega \in \phi \wedge \psi\} = \kappa(\phi \wedge \psi) - \kappa(\phi). \quad (6.5)$$

De la définition précédente, découle l'axiome de la conjonction :

$$\kappa(\phi \wedge \psi) = \kappa(\phi) + \kappa(\psi \mid \phi) \quad (6.6)$$

6.3.2 Les OCF, la théorie des possibilités et les probabilités infinitésimales

- Les fonctions conditionnelles ordinales de Spohn présentent des similarités avec la théorie des possibilités et la logique possibiliste [51]. En effet, La fonction $N_\kappa = 1 - e^{-\kappa(\neg(\phi))}$ est une mesure de nécessité satisfaisant la propriété caractéristique de la mesure de nécessité, à savoir :

$$N_\kappa(\phi \wedge \psi) = \min(N_\kappa(\phi), N_\kappa(\psi))$$

L'ensemble $\{\kappa(\omega) : \omega \in \Omega\}$ est la contrepartie de la distribution possibiliste π sur ω telle que $\Pi(\phi) = \{\max\pi(\omega) : \omega \models \phi\}$.

Il en découle la relation suivante :

$$\Pi_\kappa(\phi) = 1 - N_\kappa(\neg\phi) = e^{-\kappa(\neg\phi)}$$

$\kappa(\omega)$ est vue comme le degré d'impossibilité de ω et $\kappa(\phi) = 0$ signifie que ϕ est complètement possible.

Dans l'approche de Spohn, vu que $\kappa(\omega) \in \mathcal{N}$, $\forall \omega \pi_\kappa(\omega) > 0$, ce qui signifie que rien ne peut être considéré comme totalement impossible.

- Les OCF de Spohn ont également des similarités avec les probabilités infinitésimales, comme cité par exemple dans [108].

6.3.3 Les réseaux causaux basés sur les OCF

En se basant sur les fonctions conditionnelles ordinales, des réseaux causaux ont été proposés dans [24].

Un réseau causal basé sur les OCF possède exactement la même structure qu'un réseau causal bayésien, excepté le fait qu'au lieu de fournir des distributions de probabilités conditionnelles, la composante numérique est donnée sous forme d'OCF conditionnelles. La règle de chaînage basée sur les OCF est similaire à celle décrite par la théorie des probabilités. Elle est définie comme suit :

Soit $\omega = (x_1, x_2, \dots, x_n)$

$$\kappa_G(\omega) = \sum \{\kappa(x_i | u_i) : \omega \models x_i \wedge u_i, i = 1, \dots, n\} \quad (6.7)$$

où x_i est une instance de A_i et u_i est une instance de parents de A_i .

Exemple 6.3 *Considérons le réseau causal basé sur les OCF illustré par la figure 6.1. Les distributions d'OCF associées au DAG sont représentées par les tables 6.3 à 6.6. Les distributions de possibilités conditionnelles correspondantes sont données par les tables 6.7 à 6.10.*

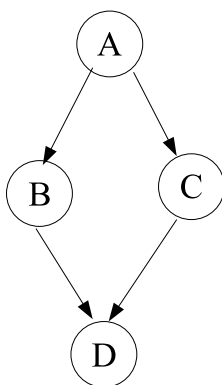


FIG. 6.1 – Exemple d'un DAG

TAB. 6.3 – Distributions d'OCF $\kappa(A)$

A	$\kappa(A)$
a	0
$\neg a$	4

TAB. 6.4 – Distributions d'OCF $\kappa(B | A)$

$\kappa(B A)$	a	$\neg a$
b	0	$+\infty$
$\neg b$	3	0

6.4 D'un réseau causal basé sur les OCF vers un réseau causal possibiliste basés sur le produit

En terme de représentation graphique des connaissances incertaines, il existe des relations étroites entre un réseau causal basé sur les OCF et un réseau causal possibiliste

TAB. 6.5 – Distributions d'OCF $\kappa(C | A)$

$\kappa(C A)$	a	$\neg a$
c	0	2
$\neg c$	$+\infty$	0

TAB. 6.6 – Distributions d'OCF $\kappa(D | BC)$

$\kappa(D BC)$	bc	$b\neg c$	$\neg bc$	$\neg b\neg c$
d	3	0	5	0
$\neg d$	0	$+\infty$	0	6

basé sur le produit.

Plus précisément, chaque réseau causal basé sur les OCF peut être codifié en un réseau causal possibiliste basé sur le produit. Pour cela, il suffit de :

- considérer le même DAG que le réseau causal basé sur les OCF
- définir les distributions de possibilités conditionnelles $\Pi(x_i | u_i)$ à partir des distributions d'OCF $\kappa(x_i | u_i)$ comme suit :

$$\Pi(x_i | u_i) = e^{-\kappa(x_i | u_i)} \quad (6.8)$$

Nous pouvons vérifier l'équivalence entre les distributions de possibilités et les distributions basées sur les OCF par la relation suivante :

Proposition 8

$$\forall \omega \in \Omega, \pi_G(\omega) = e^{-\kappa_G(\omega)} \quad (6.9)$$

Preuve 6.2 La preuve est immédiate :

Par définition : $\pi_G(\omega) = * \{ \Pi(x_i | u_i) : \omega \models x_i \wedge u_i \}$

En utilisant l'équation (6.8), nous obtenons : $\pi_G(\omega) = * \{ e^{-\kappa(x_i | u_i)} : \omega \models x_i \wedge u_i \}$

D'où : $\pi_G(\omega) = e^{-\sum \{ \kappa(x_i | u_i) : \omega \models x_i \wedge u_i \}}$

En utilisant la définition de la fonction coût attribuée à une interprétation : $\kappa_G(\omega) = \sum \{ \kappa(x_i | u_i) : \omega \models x_i \wedge u_i, i = 1, \dots, n \}$, nous obtenons alors :

$$\pi_G(\omega) = e^{-\kappa_G(\omega)} \quad \kappa_G(\omega) = \sum \{ \kappa(x_i | u_i) : \omega \models x_i \wedge u_i \}$$

Exemple 6.4 Reconsidérons le graphe causal basé sur les OCF de l'exemple 6.3. Soit ω_1 l'interprétation suivante : $\omega_1 = (\neg a, \neg b, c, \neg d)$.

En appliquant l'équation (5.3), nous obtenons la distribution de possibilités associée à ω_1 suivante : $\pi(\omega_1) = \Pi(\neg a) * \Pi(\neg b | \neg a) * \Pi(c | \neg a) * \Pi(\neg d | \neg bc) = .0183 * 1 * .1353 * 1 = .0025$

En appliquant l'équation (6.7), nous obtenons la distribution OCF associée à ω_1 suivante :

$$\kappa(\omega) = \kappa(\neg a) * \kappa(\neg b | \neg a) * \kappa(c | \neg a) * \kappa(\neg d | \neg bc) = 4 + 0 + 2 + 0 = 6$$

Ainsi, l'équation (6.9) est vérifiée. En effet, $\pi_G(\omega_1) = e^{-\kappa_G(\omega_1)}$.

TAB. 6.7 – Distributions de possibilités conditionnelles $\Pi(A)$

A	$\Pi(A)$
a	1
$-a$.0183

TAB. 6.8 – Distributions de possibilités conditionnelles $\Pi(B | A)$

$\Pi(B A)$	a	$-a$
b	1	0
$-b$.0498	1

6.5 D'un réseau causal possibiliste basé sur le produit vers un réseau causal basé sur les OCF

Le passage inverse d'un réseau causal possibiliste basé sur le produit vers un réseau causal possibiliste basé sur les OCF est également possible.

Nous remarquons qu'il faudrait pour cela utiliser des échelles spécifiques, comme pour le cas de la transformation d'une base possibiliste quantitative vers une base de pénalités.

En termes de distributions, le passage inverse n'est possible que si :

$\kappa(\omega) = -\log_2(\pi(\omega))$ est une valeur entière. Néanmoins, comme il avait été souligné dans [46], l'avantage de l'utilisation de l'échelle $[0,1]$ est la possibilité d'introduire autant de niveau d'incertitude qu'il est nécessaire.

6.6 Relations entre les réseaux causaux basés sur les OCF et les bases de pénalités

Nous proposons dans cette section de revenir sur l'ensemble des transformations, réalisées lors de cette étude, entre différents modes de représentation des connaissances incertaines autour de la théorie des possibilités quantitative. En effet, nous avons établi :

– d'une part des passages :

1. d'une base de pénalités vers une base de possibilités quantitative (section 6.2.1),
2. d'une base de possibilités quantitative vers un réseau causal possibiliste basé sur le produit (section 5.4),
3. d'un réseau causal possibiliste basé sur le produit vers un réseau causal basé sur les OCF (section 6.5).

– et d'autre part des passages :

1. d'un réseau causal basé sur les OCF vers un réseau causal possibiliste basé sur le produit (section 6.4),
2. d'un réseau causal possibiliste basé sur le produit vers une base de possibilités quantitative (section 5.3),

TAB. 6.9 – Distributions de possibilités conditionnelles $\Pi(C | A)$

$\Pi(C A)$	a	$\neg a$
c	1	.1353
$\neg c$	0	1

 TAB. 6.10 – Distributions de possibilités conditionnelles $\Pi(D | BC)$

$D BC$	bc	$b\neg c$	$\neg bc$	$\neg b\neg c$
d	.0498	1	.0067	1
$\neg d$	1	0	1	.0025

3. d'une base possibiliste quantitative vers une base de pénalités (section 6.2.2).

Il en résulte de l'ensemble de ces procédures de transformation les deux corollaires suivants montrant l'équivalence entre une base de pénalités PK et un graphe causal basé sur les OCF κ_G :

Corollaire 6.1 Soit $PK = \{(\phi_i, k_i), i = 1..n\}$ une base de pénalités. Le graphe causal basé sur les OCF κ_G équivalent à PK est obtenu comme suit :

1. la base PK est transformée en une base possibiliste quantitative équivalente $\Sigma^* = \{(\phi_i, \alpha_i), i = 1..n\}$ avec $(\phi_i, k_i) \in PK$ et $\alpha_i = 1 - e^{-k_i}$
2. la base de possibilités quantitative Σ^* est transformée en un graphe causal possibiliste basé sur le produit Π_G telles que :
 - sa composante graphique est obtenue en utilisant la définition 5.2,
 - sa composante numérique est obtenue en utilisant l'équation 5.5.
3. le graphe possibiliste basé sur le produit Π_G est transformé en un graphe causal basé sur les OCF κ_G comme suit :
 - sa composante graphique est identique à celle du graphe possibiliste Π_G ,
 - sa composante numérique est telle que : à $\kappa(\omega) = -\log_2(\pi(\omega))$ correspond une valeur entière où $\kappa(\omega)$ représente la distribution associée au graphe causal basé sur les OCF et $\pi(\omega)$ représente la distribution associée au graphe possibiliste basé sur le produit.

Corollaire 6.2 Soit κ_G un graphe causal basé sur les OCF représenté par un DAG et des distributions d'OCF $\kappa(x_i | u_i)$ où x_i est une instance d'une variable A_i et u_i une instance des parents de A_i .

La base de pénalités $PK = \{(\phi_i, k_i), i = 1..n\}$ équivalente au graphe causal basé sur les OCF κ_G est obtenue :

1. en transformant le graphe causal basé sur les OCF κ_G en un graphe possibiliste basé sur le produit équivalent Π_G ayant la même composante graphique, représentée par le DAG, et la composante numérique est telle que : $\Pi(x_i | u_i) = e^{-\kappa(x_i|u_i)}$ où $\kappa(x_i | u_i)$ représentent les distributions d'OCF,
2. en transformant le graphe causal Π_G obtenu en une base possibiliste quantitative équivalente $\Sigma^* = \{(\neg x_i \vee \neg u_i, 1 - \alpha_i)\}$ avec $\Pi(x_i | u_i) = \alpha_i \neq 1 \in \Pi_G$,

3. en déduisant la base de pénalités $PK = \{(\neg x_i \vee \neg u_i, k_i)\}$ équivalente à la base de possibilités quantitative Σ^* avec $k_i = -\log(1 - \alpha_i)$

Exemple 6.5 Reconsidérons le graphe causal, noté κ_G , basé sur les OCF de l'exemple 6.4 où la structure graphique est illustrée par la figure 6.1 et les distributions d'OCF sont données par les tables 6.3 à 6.6. Le graphe causal possibiliste basé sur le produit équivalent à κ_G , noté Π_{G^*} , est représenté par la même structure graphique et la composante numérique représentant les distributions de possibilités conditionnelles est donnée par les tables 6.7 à 6.10, en utilisant la procédure décrite dans la section 6.4.

La base possibiliste quantitative équivalente au graphe causal possibiliste basé sur le produit Π_{G^*} , notée Σ_* , est obtenue en suivant la procédure donnée dans la section 5.3 :

$$\Sigma_* = \{(a, .9817), (b \vee \neg a, .9502), (\neg b \vee a, 1), (\neg c \vee a, .8647), (c \vee \neg a, 1), (\neg d \vee \neg b \vee \neg c, .9502), (\neg d \vee b \vee \neg c, .9933), (d \vee \neg b \vee c, 1), (d \vee b \vee c, .9975)\}$$

La base de pénalité PK équivalente à la base de possibilités quantitative Σ_* est obtenue selon la procédure 6.2.2 :

$$PK = \{(a, 4), (b \vee \neg a, 3), (\neg b \vee a, +\infty), (\neg c \vee a, 2), (c \vee \neg a, +\infty), (\neg d \vee \neg b \vee \neg c, 3), (\neg d \vee b \vee \neg c, 5), (d \vee \neg b \vee c, +\infty), (d \vee b \vee c, 6)\}$$

6.7 Conclusion

Les résultats obtenus dans ce chapitre sont importants à plus d'un titre.

- Ils contribuent à clarifier les relations qui existent entre la théorie des probabilités et la théorie des possibilités.

En effet, les réseaux causaux possibilistes basés sur le produit ont des relations étroites avec les probabilités infinitésimales vu que ces dernières permettent de codifier les réseaux causaux basés sur les OCF.

De plus, ils peuvent être vus comme des représentations intermédiaires entre les réseaux causaux numériques représentés par les réseaux bayésiens et les réseaux causaux qualitatifs représentés par les réseaux causaux possibilistes basés sur le minimum.

- Ils font le point entre différentes équivalences. A savoir :
 - entre la logique possibiliste quantitative et la logique des pénalités d'une part,
 - entre les réseaux causaux possibilistes basés sur le produit et les réseaux causaux basés sur les OCF d'autres part.

De ces transformations, il en résulte des transformations équivalentes entre la logique des pénalités et les réseaux causaux basés sur les OCF et ce en considérant certaines échelles d'incertitude.

Après avoir parcouru les différents modes de représentation des connaissances incertaines autour de la théorie des possibilités, nous nous intéressons dans la troisième partie de ce manuscrit aux différents modes de raisonnement relatifs aux modes de représentations de connaissances qui gravitent autour de la théorie de possibilités quantitative, en exploitant les procédures de passages entre les différents types de représentation.

Troisième partie

Algorithmes et Expérimentations

Introduction de la partie III

Dans la deuxième partie, nous avons présenté les premières contributions de notre étude qui touchent l'aspect de la représentation des connaissances incertaines en théorie des possibilités quantitative (entre base possibilistes quantitatives et réseaux causaux possibilistes basés sur le produit) et autour de la théorie des possibilités quantitative (entre les bases possibilistes quantitatives et les bases de pénalités d'une part et entre les réseaux causaux possibilistes basés sur le produit et les réseaux causaux basés sur les OCF).

Les précédents travaux réalisés autour des relations entre la logique des pénalités et la logique possibiliste quantitative ont touché, essentiellement, l'aspect sémantique. Pour notre part, nous avons montré l'équivalence syntaxique entre une base de pénalités et une base de possibilités quantitative.

Dans la troisième partie de cette thèse, nous nous intéressons à l'aspect du raisonnement à partir des connaissances incertaines en théorie des possibilités quantitative. Dans le chapitre 7, nous présentons les algorithmes de propagation développés dans [20] pour les polyarbres et les réseaux et les réseaux à connexions multiples, qui sont des adaptations des algorithmes développés dans le cadre probabiliste pour les polyarbres [10] et pour les réseaux à connexions multiples [11]. Ces algorithmes ont été utilisés pour simuler le raisonnement dans les réseaux causaux possibilistes basés sur le produit.

Dans le chapitre 8, nous proposons un nouvel algorithme d'inférence pour la logique possibiliste quantitative basé sur WMAXSAT (Weighted MAXSAT), où une base possibiliste quantitative est transformée en une base de pénalités correspondante en utilisant la procédure définie dans le chapitre 6, puis, l'algorithme d'inférence sera exécuté sur cette dernière. Cet algorithme contribue à développer les liens qui existent entre la logique possibiliste quantitative et la logique des pénalités. Nous rappelons que les précédents travaux ont seulement montré l'équivalence sémantique. Pour notre part, nous avons déployé d'une part l'équivalence syntaxique et d'autre part, nous avons mis en place un nouvel algorithme pour raisonner sur une base de pénalités obtenue à partir d'une base de possibilités quantitative équivalente.

Le chapitre 9 est consacré à présenter une étude expérimentale en apportant ainsi une évaluation préliminaire, d'une part, sur les différents algorithmes de transformations entre les modes graphiques et les modes logiques et d'autre part sur les algorithmes de raisonnement en théorie des possibilités quantitative.

Chapitre 7

Algorithmes de propagation pour les réseaux causaux possibilistes basés sur le produit

7.1 Introduction

Dans le chapitre 3, nous avons abordé les algorithmes de propagation dans les réseaux causaux possibilistes basés sur le minimum. Nous allons invoquer ici les algorithmes de propagation développés pour les réseaux causaux possibilistes dans [20], qui sont une adaptation directe des algorithmes de propagation dans les réseaux causaux bayésiens. Selon la structure du DAG, deux algorithmes ont été développés l'un pour les polyarbres et l'autre pour les graphes à connexions multiples. Nous allons présenter dans ce chapitre les réseaux causaux possibilistes basés sur le produit, puis nous allons exposer dans les deux sections suivantes les algorithmes de propagation relatifs aux deux types de DAG.

7.2 Réseaux causaux basés sur le produit

Un réseau causal possibiliste basé sur le produit, noté ΠG_p , est un graphe possibiliste où le conditionnement est basé sur le produit :

$$\pi(\omega \mid_p \phi) = \begin{cases} \frac{\pi(\omega)}{\Pi(\phi)} & \text{si } \omega \models \phi \\ 0 & \text{sinon} \end{cases} \quad (7.1)$$

La distribution de possibilités jointe associée à un graphe causal possibiliste basé sur le produit est définie par le biais de la règle de chaînage basée sur le produit :

Définition 7.1 (Règle de chaînage basée sur le produit) *La distribution de possibilités globale jointe d'un réseau causal possibiliste basé sur le produit ΠG_p , à travers un ensemble de variables $V=\{A_1, A_2, \dots, A_N\}$ est exprimée comme le produit des N distri-*

distributions a priori et distributions conditionnelles à travers la règle de chaînage suivante :

$$\pi_p(A_1, \dots, A_N) = \prod_{i=1..N} \Pi(A_i | U_{A_i}) \tag{7.2}$$

Exemple 7.1 Considérons le graphe possibiliste basé sur le produit de la figure 7.1. Les distributions de possibilités conditionnelles sont représentées par la table 7.1. Les distributions de possibilités jointes, obtenues en utilisant la règle de chaînage, sont données par la table 7.2.

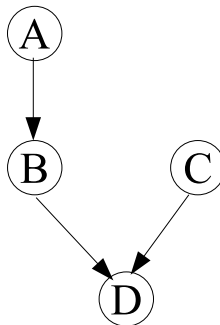


FIG. 7.1 – Exemple d’un DAG

TAB. 7.1 – Distributions initiales

A	$\Pi(A)$	C	$\Pi(C)$	B	A	$\Pi(B A)$
a	.2	c	1	b	a	.6
$\neg a$	1	$\neg c$.7	b	$\neg a$	1
				$\neg b$	a	1
				$\neg b$	$\neg a$.1

D	B	C	$\Pi(D B \wedge C)$	D	B	C	$\Pi(D B \wedge C)$
d	b	c	.9	$\neg d$	b	c	1
d	b	$\neg c$.8	$\neg d$	b	$\neg c$	1
d	$\neg b$	c	1	$\neg d$	$\neg b$	c	.5
d	$\neg b$	$\neg c$	1	$\neg d$	$\neg b$	$\neg c$.3

Les réseaux causaux possibilistes basés sur le produit vérifient deux propriétés primordiales :

1. **Critère de Normalisation** : La marginalisation produit toujours des distributions normalisées. D’une manière formelle, soit X un sous ensemble de V et x une instanciation de X. Soit $Z=V/X$. alors,

$$\max_z \left\{ \prod \{ \Pi(a | u_A) : a \in z, u_A \subseteq z \wedge x \} \right\} = 1 \tag{7.3}$$

TAB. 7.2 – Distributions de possibilités jointes utilisant la règle de chaînage basée sur le produit

A	B	C	D	$\pi_p(A \wedge B \wedge C \wedge D)$	A	B	C	D	$\pi_p(A \wedge B \wedge C \wedge D)$
a	b	c	d	.108	$\neg a$	b	c	d	.90
a	b	c	$\neg d$.012	$\neg a$	b	c	$\neg d$	1
a	b	$\neg c$	d	.0672	$\neg a$	b	$\neg c$	d	.56
a	b	$\neg c$	$\neg d$.084	$\neg a$	b	$\neg c$	$\neg d$.70
a	$\neg b$	c	d	.02	$\neg a$	$\neg b$	c	d	.1
a	$\neg b$	c	$\neg d$.1	$\neg a$	$\neg b$	c	$\neg d$.05
a	$\neg b$	$\neg c$	d	.14	$\neg a$	$\neg b$	$\neg c$	d	.07
a	$\neg b$	$\neg c$	$\neg d$.42	$\neg a$	$\neg b$	$\neg c$	$\neg d$.21

2. **Recouvrement des données initiales.** Il signifie que les distributions de possibilités initiales prodiguées par l'expert sont retrouvées à partir de la distribution de possibilités jointe donnée par l'équation (7.2).

D'une manière formelle, pour un nœud A quelconque d'un graphe possibiliste basé sur le produit noté ΠG_p , la distribution de possibilités conditionnelle $\Pi_p(a | u_A)$ calculée à partir de la distribution de possibilités globale jointe π_p , est égale à la distribution conditionnelle donnée par l'expert concernant le nœud A, notée $\Pi(a | u_A)$. Ainsi,

$$\Pi_p(a | u_a) = \Pi(a | u_A) \quad (7.4)$$

Exemple 7.2 Soit le DAG précédent. A titre d'exemple, l'égalité suivante est vérifiée : $\Pi_p(d | b \wedge \neg c) = \Pi(d | b \wedge \neg c) = .8$, vu que $\Pi(d | b \wedge \neg c) = \frac{\Pi_p(d \wedge b \wedge \neg c)}{\Pi_p(b \wedge \neg c)} = \frac{.56}{.7} = .8$.

7.3 Propagation basée sur le produit dans les polyarbres

L'algorithme présenté dans cette section est une adaptation *possibiliste* de l'algorithme de propagation présenté dans la section 3.2.2.2. Il est basé sur les communications locales via deux types de messages, λ -messages et μ -messages, circulant, respectivement des nœuds enfants vers les nœuds parents et des nœuds parents vers les nœuds enfants.

Remarque 7.1 La formulation mathématique de l'algorithme de propagation pour les réseaux causaux possibilistes dans le cas d'un polyarbre est identique à celle correspondant aux réseaux causaux bayésiens, avec une différence qui consiste à utiliser l'opérateur maximum au lieu de l'opérateur addition.

Ainsi,

- $\forall a \in D_A$, la mesure de probabilité courante de a basée sur l'évidence totale e est définie par :

$$Bel(a) = \Pi_p(a | e) = \alpha * \lambda(a) * \mu(a) \quad (7.5)$$

où $\alpha = \frac{1}{\max_a Bel(a)}$ est le facteur de normalisation

- La valeur λ associée à chaque instance $a \in D_a$ est définie par :

$$\lambda(a) = \Pi_p(a | e_A^-) = \lambda_A(a) * \prod_{j=1}^m \lambda_{Y_j}(a) \quad (7.6)$$

où $\lambda_A(a)$ dénote l'évidence totale relative au nœud A telle que :

$$\lambda_A(a) = \begin{cases} 0 & \text{si } e_A \neq a \text{ (A est instanciée à } a \text{ avec } e_A \neq a) \\ 1 & \text{sinon (A est instanciée à } a \text{ avec } e_A = a \text{ ou A n'est pas instanciée)} \end{cases} \quad (7.7)$$

- La valeur μ associée à chaque instance $a \in D_A$ est définie par :

$$\mu(a) = \Pi_p(a | e_A^+) = \max_u \Pi(a | u) * \prod_{i=1}^n \mu_A(u_i) \quad (7.8)$$

- Le message λ de A vers ses parents U_i , ($i \in \{1, \dots, n\}$) où $U_i = u_i$ est défini par :

$$\lambda_A(u_i) = \Pi_p(e_{U_i A}^- | u_i) = \beta * \max_a \lambda(a) \left[\max_{u_k: k \neq i} \Pi(a | u) * \prod_{k \neq i} \mu_A(u_k) \right] \quad (7.9)$$

où β est une constante de normalisation.

Remarque 7.2 Dans le cas où chaque nœud dans le graphe a un parent unique alors la valeur λ est simplifiée comme suit :

$$\lambda_A(u_i) = \beta * \max_a \lambda(a) * \Pi(a | u)$$

- Le message μ de A à ses enfants Y_j , ($j \in \{1, \dots, m\}$) où $A = a$ est défini par :

$$\mu_{Y_j}(a) = \Pi_p(a | e_A^+ Y_j) = \alpha * \lambda_A(a) * \prod_{i=1, \dots, m, i \neq j} \lambda_{Y_i}(a) * \mu(a) \quad (7.10)$$

7.3.1 Algorithme de propagation :

L'algorithme de propagation s'effectue essentiellement en deux grandes étapes :

- La première phase consiste à sélectionner arbitrairement un nœud pivot dans le plus petit ensemble des nœuds connectés contenant l'ensemble des nœuds modifiés (noté \mathcal{S})
- La deuxième phase consiste à faire passer les messages en deux étapes :
 - *Collecte de l'évidence* où chaque nœud dans \mathcal{S} fait passer un message à ses nœuds adjacents dans la direction du pivot, en commençant par le nœud le plus éloigné du nœud pivot.
 - *Distribution de l'évidence* où chaque nœud du graphe fait passer un message à ses nœuds adjacents, en commençant par le pivot lui même jusqu'à atteindre les feuilles du DAG.

Les grandes lignes de cet algorithme sont les suivantes :

Algorithme 3.8 : Propagation basée sur le produit dans un polyarbre

Début

Choix de la racine de propagation

- Soit \mathcal{S} le plus petit ensemble connecté incluant les nœuds observés ;
- Soit pivot un nœud arbitraire dans \mathcal{S} représentant la racine de la propagation ;
- Soit Postorder un vecteur contenant l'ordre dans lequel les messages sont envoyés dans la phase de collecte de l'évidence (le dernier nœud étant le pivot) ;
- Soit Preorder un vecteur contenant l'ordre dans lequel les messages sont envoyés dans la phase de collecte de l'évidence (le premier nœud est le pivot)

Initialisation

- Initialiser l'ensembles des valeurs λ , μ et messages à 1 ;
- Pour chaque nœud racine A , la valeur $\mu(a)$ est initialisée à $\Pi(a)$, $\forall a \in D_A$.
- Pour chaque nœud A observé, il faut assigner à $\lambda(a)$ la valeur 1 si A est instanciée à a , et lui assigner la valeur 0 dans le cas contraire.

Collecte de l'évidence

Pour $i := 1$ à $\text{length}(\text{Postorder}) - 1$ **Faire**

- $A \leftarrow \text{Postorder}[i]$;
- $B \leftarrow$ un nœud adjacent de A dans Postorder ;
- Calculer $\lambda(A)$ en utilisant l'équation (7.6) ;
- Calculer $\mu(A)$ en utilisant l'équation (7.8) ; **Si** B est un nœud parent de A **alors**
Envoyer un message λ de A vers B en utilisant l'équation (7.9)

Sinon

Envoyer un message μ de A à B en utilisant l'équation (7.10)

Distribution de l'évidence

Pour $i := 1$ à $\text{length}(\text{Preorder}) - 1$ **Faire**

- $A \leftarrow \text{Preorder}[i]$;
- Below \leftarrow un nœud adjacent de A dans Preorder ;
- Calculer $\lambda(A)$ en utilisant l'équation (7.6) ;
- Calculer $\mu(A)$ en utilisant l'équation (7.8) ;

Pour $j := 1$ à $\text{length}(\text{Below})$ **Faire**

- $B \leftarrow \text{Below}[j]$;
- **Si** B est un nœud parent de A **alors**
Envoyer un message λ de A vers B en utilisant l'équation (7.9) **Sinon**
Envoyer un message μ de A à B en utilisant l'équation (7.10)

Marginalisation Pour chaque nœud A , $\text{Bel}(A) = \Pi_p(a \mid e)$ est calculé en utilisant l'équation(7.5) ;

Fin

7.4 Propagation basée sur le produit dans les réseaux causaux possibilistes à connexions multiples

Pour ce type de DAG, le principe de la propagation est similaire à celui de la propagation dans le cadre des réseaux bayésiens. Il consiste à transformer le DAG initial en un arbre de jonction, puis d'effectuer la propagation sur ce dernier.

L'arbre de jonction, noté \mathcal{JT} , est défini et construit de manière similaire que dans le cas probabiliste (voir section 3.2.2.3).

Pour chaque cluster C_i (respectivement séparateur S_{ij}) de \mathcal{JT} , des distributions jointes locales (ou des potentiels) sont assignées, notées π_{C_i} (respectivement $\pi_{S_{ij}}$).

A partir de \mathcal{JT} , une distribution de possibilités jointe globale unique, notée $\pi_{\mathcal{JT}}$, est assignée. Elle est définie par :

$$\pi_{\mathcal{JT}}(A_1, \dots, A_N) = \frac{\prod_{i=1}^m \pi_{C_i}}{\prod_{j=1}^{m-1} \pi_{S_{ij}}} \quad (7.11)$$

où m est le nombre de clusters dans le \mathcal{JT} .

Un arbre de jonction est caractérisé par la propriété de consistance suivante :

Définition 7.2 Soient C_i et C_j deux clusters adjacents dans un arbre de jonction \mathcal{JT} et soit S_{ij} leur séparateur. Un lien entre C_i et C_j est dit stable ou consistant si

$$\max_{C_i \setminus S_{ij}} \pi_{C_i} = \pi_{S_{ij}} = \max_{C_j \setminus S_{ij}} \pi_{C_j} \quad (7.12)$$

où $\max_{C_i \setminus S_{ij}} \pi_{C_i}$ est la distribution marginale de S_{ij} définie à partir de π_{C_i} .

Si l'ensemble des liens d'un arbre de jonction est consistant, alors l'arbre de jonction est dit globalement consistant.

Lorsque l'arbre de jonction est globalement consistant, alors le potentiel attaché à chaque cluster correspond à sa distribution jointe locale calculée à partir du graphe initial. Formellement,

$$\pi_{C_i} = \Pi_p(C_i) \quad (7.13)$$

A partir de cette dernière équation, la distribution de possibilités d'une variable $A \in V$, dans un arbre de jonction globalement consistant est calculée en utilisant un cluster quelconque C_i contenant A en marginalisant son potentiel sur A comme suit :

$$\Pi_p(A) = \max_{C_i \setminus A} \pi_{C_i} \quad (7.14)$$

Dans ce qui suit, $\pi_{C_i}^I$ représente le potentiel du cluster C_i lors de l'étape d'initialisation et $\pi_{C_i}^C$ représente le potentiel du cluster C_i lors de l'étape de consistance globale.

7.4.1 Algorithme de la propagation

Après avoir construit l'arbre de jonction optimal associé au graphe causal original, l'étape suivante consiste à calculer la composante numérique associée à l'arbre de jonction. En d'autres termes, pour chaque variable $A \in V$, il faudrait calculer le degré de

possibilité $\Pi(A)$. Le problème de la prise en charge d'une évidence totale e (calculer $\Pi(a | e)$) sera abordé par la suite.

Après la phase de transformation graphique, la propagation s'effectue en trois phases :

1. **Initialisation** : Cette étape consiste à quantifier l'arbre de jonction, en utilisant les probabilités conditionnelles initiales : **Algorithme 3.9 : Initialisation**

Debut

Pour chaque cluster C_i **Faire** $\pi_{C_i}^I \leftarrow 1$;
Pour chaque séparateur S_{ij} **Faire** $\pi_{S_{ij}}^I \leftarrow 1$;
Pour chaque variable A **Faire**

Sélectionner un cluster C_i contenant $\{A\} \cup U_A$;

$\pi_{C_i}^I := \pi_{C_i}^I * P(A | U_A)$;

Fin

L'arbre de jonction ainsi initialisé codifie les mêmes distributions que le graphe initial :

$$\pi_p = \pi_{\mathcal{JT}}^I \quad (7.15)$$

où π_p est la distribution de possibilités jointe du réseau causal initial obtenue en utilisant l'équation (7.2) et $\pi_{\mathcal{JT}}^I$ représente la distribution de probabilités jointe associée à l'arbre de jonction obtenue en utilisant (7.11). Après avoir initialisé les potentiels de l'arbre de jonction, il faudrait s'assurer de la consistance locale en effectuant la propagation globale décrite comme suit :

2. **La propagation globale** : La propagation globale est réalisée par le mécanisme de passages de messages entre chaque cluster C_i et ses clusters adjacents C_j en deux phases, en commençant par un cluster arbitraire dit nœud pivot :

– **Collecte de l'évidence** :

Chaque cluster passe un message à ses clusters adjacents dans la direction du pivot, en commençant par les clusters les plus éloignés de celui-ci. L'ordre dans lequel les messages sont envoyés est matérialisé par *Postorder*.

– **Distribution de l'évidence** :

Chaque cluster envoie des messages à ses clusters adjacents, en commençant par le pivot lui même jusqu'à atteindre les feuilles. L'ordre dans lequel les messages sont envoyés est matérialisé par *Preorder*.

Pour les deux phases, lorsqu'un cluster C_i fait passer un message à ses nœuds adjacents C_j , les potentiels associés à C_i, C_j et leur séparateur S_{ij} sont mis à jour comme suit :

- (a) Preserver le même potentiel pour C_i :

$$\pi_{C_i}^{t+1} \leftarrow \pi_{C_i}^t \quad (7.16)$$

(b) Assigner un nouveau potentiel à S_{ij} :

$$\pi_{S_{ij}}^{t+1} \leftarrow \max_{C_i \setminus S_{ij}} \pi_{C_i}^t \quad (7.17)$$

(c) Assigner un nouveau potentiel à C_j :

$$\pi_{C_j}^{t+1} \leftarrow \pi_{C_j}^t * \frac{\pi_{S_{ij}}^{t+1}}{\pi_{S_{ij}}^t} \quad (7.18)$$

La procédure de propagation globale peut être résumée comme suit [20] : **Algorithme**

3.10 : Propagation globale

Debut

Selection de la racine de la propagation

Soit Pivot un cluster arbitraire C_i représentant la racine de la propagation ;

Soit Postorder un vecteur contenant l'ordre dans lequel les messages vont être envoyés dans la phase de la collecte de l'évidence ;

Soit Preorder un vecteur contenant l'ordre dans lequel les messages vont être envoyés dans la phase de la distribution de l'évidence ;

Collecte de l'évidence

Pour $i := 1$ jusqu'à $\text{lenght}(\text{Postorder})-1$ Faire

$C_i \leftarrow \text{Postorder}[i]$;

$C_j \leftarrow$ cluster adjacent de C_i dans Postorder ;

Envoyer le message de C_i vers C_j en utilisant les équations (7.16), (7.17) et (7.18) ;

Distribution de l'évidence

Pour $i := 1$ jusqu'à $\text{lenght}(\text{Preorder})$ Faire

$C_i \leftarrow \text{Preorder}[i]$;

$\text{Below} \leftarrow$ cluster adjacents de C_i dans Preorder ;

Pour $j := 1$ jusqu'à $\text{lenght}(\text{Below})$ Faire

$C_j \leftarrow \text{Below}[j]$;

Envoyer le message de C_i à C_j en utilisant les équations (7.16), (7.17) et (7.18) ;

Fin

Remarque 7.3 *A partir de ces deux procédures, il est important de relever certains points :*

- *A chaque étape t de la procédure de la propagation globale, l'arbre de jonction codifie la même distribution jointe. Formellement,*

$$\pi_{\mathcal{JT}}^t = \pi_{\mathcal{JT}}^{t+1} \quad (7.19)$$

où $\pi_{\mathcal{JT}}^t$ représente la distribution jointe relative à l'arbre de jonction \mathcal{JT} à l'étape t .

- *A partir des équations (7.15) et (7.19), nous déduisons que de la phase d'initialisation jusqu'à la phase de consistance globale, l'arbre de jonction codifie la même distribution jointe. Formellement,*

$$\pi_p = \pi_{\mathcal{JT}}^C \quad (7.20)$$

où π_p est la distribution jointe de ΠG_p et $\pi_{\mathcal{JT}}^C$ est la distribution jointe associée à \mathcal{JT} après la procédure de propagation globale.

- *La consistance globale est atteinte après les phases de collecte de l'évidence et de distribution de l'évidence.*

3. **Marginalisation** A partir de l'arbre de jonction consistant obtenu lors de la phase précédente, la mesure de possibilités associée à chaque variable d'intérêt A , notée $\Pi_p(A)$ peut être calculée par l'algorithme suivant :

Algorithme 3.11 : Marginalisation

Debut

Identifier un cluster C_i contenant A ;

Calculer $\Pi_p(A)$ en marginalisant $\pi_{C_i}^C$ sur A :

$$\Pi_p(A) \leftarrow \max_{C_i \setminus A} \pi_{C_i}^C ;$$

Fin

Prise en charge de l'évidence : La prise en charge d'une évidence totale e se fait par la généralisations des différents algorithmes précédents. Il s'agit de calculer $\forall A \in V, \Pi_p(A \wedge e)$. L'évidence peut être codifiée en utilisant la fonction *likelihood* définie par :

$$\Lambda_A(a) = \begin{cases} 1 & \text{si } A \text{ n'est pas instanciée} \\ 1 & \text{si } A \text{ est instanciée à } a \\ 0 & \text{si } A \text{ est instanciée mais pas à } A \end{cases} \quad (7.21)$$

La prise en charge de l'évidence e nécessite l'extension de la procédure d'initialisation afin d'incorporer la connaissance certaine. En effet, l'évidence e doit être codifiée comme une fonction *likelihood*, en utilisant l'équation (7.21), puis elle doit être intégrée dans l'arbre de jonction en ajoutant des étapes supplémentaires à la procédure d'initialisation. Elle se résume comme suit :

- Pour chaque instance de la variable A , coder l'observation $A=a$ comme une fonction *likelihood*, en utilisant l'équation (7.21).
- Identifier un cluster C_i contenant A :

$$\pi_{C_i}^I \leftarrow \pi_{C_i}^I * \Lambda_A$$

La marginalisation d'un potentiel d'un cluster quelconque $\pi_{C_i}^C$ dans une variable A , en utilisant l'équation (7.14) permet d'obtenir la mesure de possibilité de A et e :

$$\Pi_p(A, e) = \max_{C_i \setminus A} \pi_{C_i}^C$$

Le calcul de $\Pi_p(A | e)$ est effectué à partir de $\Pi_p(A \wedge e)$ en appliquant la définition du conditionnement basé sur le produit comme suit :

$$\Pi_p(A | e) = \frac{\Pi_p(A \wedge e)}{\Pi_p(e)} = \frac{\Pi_p(A \wedge e)}{\max_A \Pi_p(A \wedge e)} \quad (7.22)$$

Remarque 7.4 *Nous rappelons que la tâche de trouver une triangulation optimale est un problème NP-complet [87].*

7.5 conclusion

Les algorithmes de propagation présentés dans ce chapitre sont des adaptations directes des algorithmes de propagation développés pour les réseaux causaux bayésiens. En effet, l'algorithme pour les polyarbres est une adaptation de la version centralisée de Pearl [85, 90]. Pour les graphes à connexions multiples, l'algorithme est une adaptation de l'algorithme pour les arbres de jonction développé dans [11]. Ces deux algorithmes de propagation vont être utilisés dans notre travail comme mode de raisonnement sur les connaissances représentées graphiquement.

D'autres part, il ressort que les réseaux causaux possibilistes entretiennent des relations étroites avec :

- les réseaux VBS [61, 62],
- les réseaux OCF [24]

Après avoir abordé les modes de raisonnement pour les modèles graphiques en théorie des possibilités, nous nous intéressons dans le chapitre suivant aux modes de raisonnement pour les modèles logiques. Dans le cadre de la logique possibiliste qualitative, il y a eu développement d'algorithme d'inférence. Dans le contexte quantitatif, un algorithme d'inférence a été développé dans le cadre de ce travail, en exploitant l'équivalence qui existe entre une base logique possibiliste quantitative et une base de pénalités.

Chapitre 8

Algorithmes d'inférence pour la logique possibiliste quantitative

8.1 Introduction

Nous rappelons que la logique des possibilités, issue de la logique floue, a été conçue pour raisonner automatiquement sur des connaissances incertaines. Une base de connaissance possibiliste est composée d'un ensemble de couples (ϕ, α) , où ϕ est une formule logique classique $\in \mathcal{L}$ et α représente le seuil minimal du degré de nécessité de ϕ . Autrement dit : $N(\phi) \geq \alpha$.

Cette valuation peut être interprétée différemment selon le contexte. En effet,

- Dans un contexte qualitatif, l'attribution d'une valuation α à une formule ϕ permet d'établir un ordre implicite entre les connaissances. Cette interprétation des degrés de certitude définit *la logique possibiliste standard*.
- Dans un contexte numérique ou quantitatif, la valeur α est interprétée comme une valuation numérique de la connaissance ϕ quantifiant l'incertitude qui lui est affiliée définissant ainsi la notion de *logique possibiliste quantitative* qui est souvent comparée à la logique probabiliste.

Vu cette distinction, les mécanismes d'inférence mis en œuvre sont certainement différents.

Dans le cadre de la logique possibiliste standard, plusieurs algorithmes issus de la déduction possibiliste ont été développés exploitant la relation d'ordre établie entre les formules [47]. Ces approches ordinales sélectionnent des sous-ensembles consistants préférés, parmi tous les sous-ensembles consistants de formules. Les bases de connaissances considérées dans ce type d'approche sont stratifiées. Ainsi, les formules des strates les moins prioritaires, étant les moins certaines, sont rejetées plus facilement en cas d'incohérence. Cependant, est-il toujours souhaitable de rejeter un grand nombre de formules de faible priorité.

Justement, les approches quantitatives (telle que la logique des pénalités) se basent sur d'autres critères en tenant compte des poids affectés aux formules. Il est à souligner que dans le cadre de la logique possibiliste quantitative, il n'y a pas eu de développement

d'algorithme d'inférence. Nous proposons justement dans ce chapitre de présenter un nouvel algorithme d'inférence pour la logique possibiliste quantitative.

Dans la section 2, nous allons exposer les principes de raisonnement dans la logique possibiliste standard développés dans [45]. La section 3 sera consacrée à présenter un nouvel algorithme d'inférence que nous avons développé, pour la logique des possibilités quantitative.

8.2 L'inférence en logique possibiliste standard (SPL)

En logique possibiliste standard, les bases de connaissances sont stratifiées et la strate la plus élevée correspond aux connaissances les plus certaines. Ainsi, le mécanisme de raisonnement consiste, à partir d'une SPL Σ , à dériver des conclusions en utilisant la partie la plus certaine de celle-ci.

Le problème de déduction en logique possibiliste standard est basé sur le principe d'ordre induit par les connaissances. Il peut être résumé comme suit :

Soit Σ une base possibiliste standard et ϕ une formule logique classique que l'on veut déduire de Σ avec un certain degré. Le processus d'inférence revient à calculer la plus grande valeur de α telle que (ϕ, α) est une conséquence logique de Σ . Pour aborder les différents mécanismes de raisonnement en SPL, il est nécessaire de présenter certains concepts.

- **La satisfaction d'une base de connaissances par une distribution de possibilités :** Une distribution de possibilités π satisfait une formule SPL (ϕ, α) (noté $\pi \models (\phi, \alpha)$) ssi $N(\phi) \geq \alpha$, où N est la mesure de nécessité induite par π . Par extension, une distribution de possibilités π satisfait une base de connaissances SPL $\Sigma = \{(\phi_i, \alpha_i), i = 1..n\}$ (noté $\pi \models \Sigma$) ssi $\forall i, \pi \models (\phi_i, \alpha_i)$

- **La conséquence logique :** Une formule (ϕ, α) est une conséquence logique d'une base de connaissances SPL Σ ssi toute distribution de possibilités π qui satisfait Σ , satisfait alors (ϕ, α) .

Le problème de déduction est alors présenté comme suit :

Soit Σ une base de connaissances SPL et ϕ une formule logique classique à déduire à partir de Σ avec un certain degré. Il faudrait alors calculer le plus grand degré α tel que (ϕ, α) est une conséquence logique de Σ . En d'autres termes, il faut calculer :

$$Val(\phi, \Sigma) = \sup\{\alpha \in [0, 1] : \Sigma \models (\phi, \alpha)\}$$

- **Le principe du minimum de spécificité :** Un résultat fondamental qui découle de la déduction à partir d'une base de connaissances SPL est qu'il existe toujours une distribution de possibilités la moins spécifique qui satisfait la base de connaissances Σ .

Formellement, si $\Sigma = \{(\phi_i, \alpha_i), i = 1..n\}$, alors la distribution de possibilités la

moins spécifique π_Σ qui satisfait Σ est définie par :

$$\pi_\Sigma(\omega) = \begin{cases} 1 & \text{si } \omega \models \phi_1 \wedge \dots \wedge \phi_n \\ \min\{1 - \alpha_i : \omega \models \neg\phi_i, (\phi_i, \alpha_i) \in \Sigma\} & \text{sinon} \end{cases} \quad (8.1)$$

Ainsi, pour toute distribution de possibilités [47] π , π satisfait Σ ssi $\pi \geq \pi_\Sigma$ D'où :

$$\Sigma \models (\phi, \alpha) \text{ssi} \pi_\Sigma \models (\phi, \alpha)$$

L'incohérence partielle : Un des objectifs majeurs de la logique possibiliste standard est qu'elle permet de considérer l'incohérence partielle et donc de permettre la déduction à partir d'une base de connaissances partiellement incohérente.

Une base de connaissances possibiliste Σ est partiellement incohérente si sa distribution possibiliste est telle que :

$$0 < \sup \pi_\Sigma < 1$$

Mesurer le degré d'incohérence de Σ revient à évaluer à quel degré il existe au moins une interprétation complètement possible pour Σ . Le degré d'incohérence d'une base de connaissances possibiliste est défini par :

$$Incons(\Sigma) = 1 - \sup_{\omega \in \Omega} \pi_\Sigma(\omega)$$

- **La déduction sous l'incohérence partielle :** Le processus de déduction en SPL est proche du processus en logique classique. Le lien entre le degré d'incohérence en SPL et l'incohérence d'une base classique est défini par :

$$Incons(\Sigma) = \sup\{\alpha : \Sigma_\alpha^* \text{ est incohérente}\} = \inf\{\alpha : \Sigma_\alpha^* \text{ est consistante}\}$$

où $\Sigma = \{(\phi_i, \alpha_i) : i = 1 \dots n\}$ est une base possibiliste standard,
 $\Sigma_\alpha^* = \{\phi : (\phi, \beta) \in \Sigma, \beta \geq \alpha\}$ est la projection classique de la α -cut de Σ ,
 $\Sigma_\alpha^* = \{\phi : (\phi, \beta) \in \Sigma, \beta > \alpha\}$ est la projection classique de la α -cut stricte de Σ .

Les généralisations des théorèmes de déduction et de réfutation de la logique classique en SPL sont représentées par les résultats suivants [47] :

- La déduction :

$$\Sigma \cup \{(\phi, 1)\} \models (\psi, \alpha) \text{ssi} (\phi \rightarrow \psi, \alpha)$$

- La réfutation :

$$\Sigma \models (\phi, \alpha) \text{ssi} \Sigma \cup \{(\neg\phi, 1)\} \models (\perp, \alpha)$$

qui est équivalent à :

$$Val(\phi, \Sigma) = Incons(\Sigma \cup \{(\neg\phi, 1)\})$$

Ce résultat montre que tout problème de déduction en logique possibiliste revient à un calcul d'un degré d'incohérence.

L'équation suivante montre qu'afin de déduire (ϕ, α) , seules les formules dont les poids sont supérieurs ou égaux à α sont utilisées. Ainsi,

$$\Sigma \models (\phi, \alpha) \text{ssi} \Sigma_{\alpha} \models (\phi, \alpha)$$

Ainsi, ces résultats permettent de calculer $\text{Val}(\phi, \Sigma)$ par dichotomie, où Σ est une SPL et ϕ est une formule propositionnelle à déduire et ceci en utilisant n'importe quel prouveur pour le problème de satisfiabilité propositionnelle SAT.

Soit $\alpha_0 = 0$ et soient $\alpha_1, \dots, \alpha_n$ les valuations distinctes qui apparaissent dans Σ , ordonnées d'une manière croissante : $0 < \alpha_1, \dots, \alpha_n \leq 1$.

La procédure de calcul de $\text{Val}(\phi, \Sigma)$ peut être résumée comme suit :

Algorithme 8.1 : Calcul de $\text{Val}(\phi, \Sigma)$

Debut

$l := 0;$

$u := n;$

Tant que $(l < u)$ **Faire**

$r := \lfloor (l+u)/2 \rfloor;$

Si $(\Sigma_{\alpha_r}^* \wedge \neg\phi)$ **est consistante** **Alors**

$u := r-1;$

sinon

$l := r;$

$\text{Val}(\phi, \Sigma) := \alpha_r;$

Fin

Cet algorithme contient exactement $\lceil \log_2 n \rceil$ appels à SAT pour le test de satisfiabilité, où n représente le nombre des différents strates qui existent dans Σ [45].

Le problème de déduction en logique possibiliste standard revient donc à calculer un degré d'incohérence. L'algorithme proposé par Lang utilise la stratification des bases de connaissances. Les formules des strates les moins prioritaires étant les moins certaines sont rejetées plus facilement en cas d'incohérence.

Dans cette approche ordinale, les formules de faible priorité sont systématiquement rejetées dans le cas où elles sont en contradiction avec une formule donnée. Mais cette philosophie peut s'avérer inadéquate pour certains domaines. En effet, il est parfois souhaitable de faire intervenir la notion de compensation qui permet de décider s'il faudrait rejeter une seule formule de forte priorité ou rejeter un très grand nombre de formules de faible priorité en fonction du nombre et de l'importance des formules.

8.3 Un nouvel algorithme pour raisonner en logique des possibilités quantitative

D'une manière générale, en logique, l'objectif du raisonnement consiste à définir des mécanismes permettant de déduire de nouvelles connaissances à partir de celles qui existent. La majorité des algorithmes d'inférence se basent sur des tests de satisfiabilité utilisant des prouveurs de la famille SAT. Le problème SAT fût le premier problème prouvé comme étant NP-complet [109].

8.3.1 Le problème WMAXSAT

Le problème de satisfiabilité, SAT en abrégé, constitue l'épine dorsale de la discipline de la NP-complétude. Malgré le nombre important d'études effectuées dans ce domaine, il demeure pour les chercheurs, une de leurs importantes préoccupations. SAT est défini par les composantes suivantes :

- Soit $X = \{x_1, x_2, \dots, x_n\}$, un ensemble de n variables booléennes.
- Soit $C = \{c_1, c_2, \dots, c_m\}$, un ensemble de m clauses où :
 - chaque clause est une disjonction de littéraux,
 - chaque littéral est une variable (un littéral positif) ou sa négation (un littéral négatif).
- Soit D , la donnée SAT, composée par une conjonction de littéraux.

Le problème SAT consiste à déterminer s'il existe une assignation des variables x_i de X telle que la donnée D soit satisfaite (évaluée à VRAI). SAT est un problème de décision qui admet une réponse 'Oui' s'il existe une assignation de variables qui satisfait toutes les clauses ou 'Non' si aucune assignation n'existe.

Néanmoins, si la réponse est négative, une question importante en découle : quel est le nombre maximal de clauses pouvant être satisfaites simultanément ? A partir de cette question, est défini le problème du "Maximum Satisfiability", MAXSAT en abrégé. Le problème MAXSAT est un problème d'optimisation dont la fonction objective est le nombre de clauses satisfaites. Le problème MAXSAT est un problème NP-difficile.

Si de plus, nous associons à chaque clause c_i un poids $w(c_i)$, l'objectif revient donc à trouver une assignation aux variables x_i de X qui maximise la somme des poids des clauses satisfaites simultanément ou encore qui minimise la somme des poids des clauses insatisfaites, d'où la définition du problème MAX-SAT pondéré ou WMAXSAT (Weighted MAX-SAT).

Remarque 8.1 *Le problème MAX-SAT est un cas particulier du problème WMAX-SAT, et ceci en associant à toutes les clauses le même poids.*

Pour les prouveurs de type WMAXSAT, les pondérations des clauses sont des entiers.

Nous proposons dans cette section un nouvel algorithme d'inférence basé sur une approche numérique où une base de connaissances possibiliste quantitative est transformée en une base de pénalités. Dans une base possibiliste quantitative, à chaque formule est associée le seuil minimal de son degré de nécessité. Pour une base de pénalités, à

chaque formule est associée une pondération qui correspond au prix à payer pour pouvoir la rejeter.

La philosophie attribuée aux pondérations affectées aux clauses dans une base de pénalités adhère au fonctionnement des prouveurs WMAXSAT. Ainsi, l'algorithme repose sur un prouveur WMAXSAT. Et comme les pondérations des clauses WMAXSAT doivent être de type Entier, nous considérons, dans cette proposition, que les coûts des clauses sont des entiers.

8.3.2 Principe de l'algorithme

En logique possibiliste quantitative, le principe de la déduction classique est le suivant :

- Soit $\Sigma = \{(\phi_i, \alpha_i), i = 1..n\}$ une base de possibilités quantitative où ϕ_i est une formule propositionnelle et α_i le seuil minimal du degré de nécessité associé à ϕ_i .
- Soit E l'évidence composée par une conjonction de littéraux portant sur les variables de la base auxquelles sont associées des poids égaux à 1.
- Soit ψ , une variable d'intérêt à déduire.

Il s'agit donc de tester si :

$\Sigma \cup \{(E, 1)\} \models (\psi, \beta)$. Dans les cas positif, il faudrait calculer le degré de possibilité $\Pi(\psi) = \beta$ associé à ψ .

Un algorithme d'inférence possible consiste à :

- exploiter la transformation d'une base possibiliste quantitative vers une base de pénalités, présentée dans la section 6.2.2,
- appliquer un mécanisme d'inférence sur cette dernière, en utilisant un prouveur de satisfiabilité adéquat afin de déduire l'interprétation qui fournit la somme minimale des formules qu'elle falsifie,
- déduire le degré de possibilité β affilié à ψ

Ainsi, le processus d'inférence se déroule en trois étapes :

1. **Étape 1** : Elle consiste à déduire la base de pénalités PK associée à la base possibiliste quantitative Σ donnée. Soit

$$\Sigma = \{(\phi_i, \alpha_i), i = 1..n\} \text{ avec } \alpha_i \in [0, 1]$$

La base de pénalités PK correspondante est définie par :

$$PK = \{(\phi_i, k_i), i = 1..n\} \text{ avec } k_i \in \mathcal{N} + \infty \text{ où } k_i = -\log(1 - \alpha_i).$$

Remarque 8.2 Dans le cas général, les pondérations k_i associées aux clauses dans une base de pénalités $PK = \{(\phi_i, k_i), i = 1..n\}$ sont des réels. Vu que nous allons utiliser un prouveur WMAXSAT pour simuler l'algorithme d'inférence nous considérons que les différentes pondérations k_i dans PK sont des entiers.

Cette restriction n'est pas pénalisante. Dans le cas où l'on voudrait élargir le domaine des pondérations k_i , il faudrait utiliser un autre type de prouveur.

L'équivalence entre deux bases de connaissances est généralement prouvée à partir de la relation d'équivalence entre leurs distributions induites.

En effet, dans la section 6.2, nous avons noté la relation suivante :

$$\forall \omega, \pi_{\Sigma_{PK}}(\omega) = e^{-\kappa_{PK}(\omega)}$$

où $\pi_{\Sigma_{PK}}$ représente la distribution de possibilité associée à la base possibiliste quantitative Σ_{PK} et κ_{PK} représente la distribution associée à la base de pénalités PK. D'autre part, et comme nous l'avons introduit dans la section 2.4, la logique des pénalités définit les deux mesures suivantes :

- **Le coût d'une interprétation ω** : Il est noté par $\kappa_{PK}(\omega)$ et il est égal à la somme des pénalités des formules dans PK violées par ω . D'une manière formelle, il est défini par :

$$\kappa_{PK}(\omega) = \begin{cases} 0 & \text{si } \forall (\phi_i, k_i) \in PK, \omega \models \phi_i \\ \sum \kappa_i \{ \omega \models \neg \phi_i, (\phi_i, k_i) \in PK \} & \text{sinon} \end{cases} \quad (8.2)$$

- **Le coût d'une formule ϕ** : Il est noté par $K_{PK}(\phi)$ et il correspond au coût à payer afin de rendre la formule ϕ consistante avec PK. D'une manière formelle, il est défini par :

$$K_{PK}(\phi) = \min(\kappa_{PK}(\omega) : \omega \models \phi) \quad (8.3)$$

2. **Étape 2** : Il s'agit à ce stade de définir un algorithme d'inférence pour la logique des pénalités. Nous rappelons que la philosophie d'une base de pénalités réside dans le fait que les poids sont des pénalités à payer afin de violer une formule. Ainsi, une formule est d'autant plus importante que son poids est élevé. Une formule inviolable (donc certaine) a un poids infini.

Dans cette optique, la recherche d'un sous ensemble consistant de formules est réalisée par le paramètre qui correspond à la somme minimale des pénalités des formules qui doivent être rejetées afin de restaurer la cohérence de la base. Ceci nous conduit donc à utiliser un prouveur de type WMAXSAT (Weighted MAX-SAT).

Un prouveur WMAXSAT prend comme paramètre d'entrée une base pondérée et en sortie, il fournit :

- l'interprétation ω qui minimise la somme des coûts des formules qu'elle falsifie
- la somme minimale des coûts des formules falsifiées par ω .

Le processus de fonctionnement de WMAXSAT peut être résumé comme suit :

Algorithme 8.2 : Procédure WMAXSAT($PK, \omega, k_{PK}(\omega)$)**Debut****Entrée :** la base de pénalités PK **Sorties :** $\omega, k_{PK}(\omega)$ (somme minimale des coûts des formules falsifiées par ω)**Répéter**choisir une interprétation ω' de Ω $k' := 0$;sommecouts $:= \infty$;**Pour** chaque formule ϕ_i in PK **Faire****Si** $\omega' \not\models \phi_i$ **Alors** $k' := k' + \alpha_i$;**Si** $k' < \text{sommecouts}$ **Alors**sommecouts $:= k'$; $\omega := \omega'$;**jusqu'à** (sommecouts=0) ou (il n'existe plus d'interprétation) $k_{PK}(\omega) := \text{sommecouts}$;retourner($\omega, k_{PK}(\omega)$);**Fin**

- Le prouveur WMAXSAT, fournit donc l'interprétation ω qui minimise la somme des coûts des formules qu'elle falsifie, ainsi que la somme minimale des coûts des formules falsifiées.

Si ω satisfait toutes les formules de PK , alors son coût est nul. Ceci correspond intégralement à la définition du coût d'une interprétation, donnée par l'équation 8.2. C'est ce qui justifie l'utilisation d'un prouveur WMAXSAT.

- De plus, si nous utilisons un prouveur MAXSAT au lieu d'un prouveur WMAXSAT, il serait alors nécessaire de dupliquer pour chaque formule de type (ϕ, α) , les formules $(\phi, 1)$, α fois.

Donc le problème d'inférence est redéfini comme suit :

- soit $PK = \{(\phi_i, k_i) \mid i = 1, \dots, n \text{ et } k_i \in \mathcal{N} \cup \infty\}$,
- soit E une évidence, à laquelle est associée un coût infini, qui illustre le fait qu'elle est inviolable,
- soit ψ la variable d'intérêt à inférer.

Il s'agit alors de tester si $PK \cup (E, \infty) \models (\psi, \beta)$, en déduisant son coût β .

D'une manière formelle, WMAXSAT fournit une interprétation ω' qui minimise la somme des coûts des formules falsifiées dans $PK \cup \{(Evidence, +\infty) \cup (\psi, +\infty)\}$ ainsi que ce coût $k_{PK}(\omega')$. Deux cas peuvent se présenter :

- si $k_{PK}(\omega') = +\infty$, alors $PK \cup \{(E, \infty)\} \not\models \psi$.
- si $k_{PK}(\omega') \neq +\infty$, alors $\omega' \models \{(E, +\infty)\}$ et $\omega' \models \{(\psi, +\infty)\}$.

Dans le deuxième cas, en utilisant l'équation 8.3, nous énonçons la proposition suivante :

Proposition 9

$$K_{PK}(\psi) = \min(k_{PK}(\omega) : \omega \models \psi) = k_{PK}(\omega')$$

Preuve 8.1 Par définition, en utilisant l'équation 8.3,

$$(a) \quad K_{PK}(\psi) = \min(k_{PK}(\omega) : \omega \models \psi)$$

(b) En utilisant l'équation 8.2, nous avons

$$k_{PK}(\omega') = \sum \kappa_i \{ \omega' \models \neg \phi_i, (\phi_i, k_i) \in PK \cup (E, +\infty) \cup (\psi, +\infty) \}$$

avec :

$$- \omega' \models \psi \text{ et } k_{PK}(\omega') = \min \{ \sum(k_i) : (\phi_i, k_i) \in PK \cup (E, +\infty) \cup (\psi, +\infty), \omega' \not\models \phi \}$$

– et ω' est l'interprétation fournie par l'appel à $WMAXSAT(PK \cup (E, +\infty) \cup (\psi, +\infty), \omega', k_{PK}(\omega'))$

ce qui implique que $k_{PK}(\omega') = \min(k_{PK}(\omega))$ et $\omega' \models \psi$

Ainsi, $K_{PK}(\psi) = k_{PK}(\omega')$

- L'attribution d'un coût infini à l'évidence E reflète le fait qu'elle est inviolable. Donc si une interprétation ω viole l'évidence, son coût serait alors infini.
- Pour la variable d'intérêt ψ , nous lui assignons également un coût infini pour signifier qu'elle doit être également satisfaite.

Ainsi, l'interprétation fournie par le prouveur WMAXSAT doit satisfaire $(E, +\infty)$ et $(\psi, +\infty)$.

3. **Étape 3** : Dans le cas où $PK \cup (E, +\infty) \models (\psi, K_{PK}(\psi))$, il faudrait calculer alors le degré de possibilité de ψ .

Le calcul du degré de possibilité d'une formule ψ , noté $\Pi(\psi)$ à partir du coût de pénalité de cette formule, noté $K_{PK}(\psi)$ est donné par la proposition suivante :

Proposition 10 La relation qui lie le degré de possibilité d'une formule ψ au coût de cette formule est la suivante :

$$\Pi(\psi) = e^{-K_{PK}(\psi)} \tag{8.4}$$

Preuve 8.2 En effet,

- comme il a été défini en théorie des possibilités, le degré de possibilité d'une formule ψ , $\Pi(\psi)$ est égale à :

$$\Pi(\psi) = \max_{\omega \in \Omega} \{ \pi(\omega) : \omega \models \psi \}$$

- En utilisant la relation établie dans la section 6.2.1 qui lie la distribution de possibilité, $\pi(\omega)$, associée à la base de possibilité au coût d'une interprétation $K_{PK}(\omega)$ associé à une base de pénalités, nous avons :

$$\pi(\omega) = e^{-K_{PK}(\omega)}$$

- Nous obtenons : $\Pi(\psi) = \max_{\omega \in \Omega} \{e^{-K_{PK}(\omega)} : \omega \models \psi\}$
 $= e^{-\min\{K_{PK}(\omega) : \omega \models \psi\}}$

- En utilisant l'équation 8.3, nous obtenons :

$$\Pi(\psi) = e^{-K_{PK}(\psi)}$$

A partir de ces résultats, nous pouvons décrire l'algorithme pour le problème de déduction pour la logique possibiliste quantitative comme suit :

Algorithme 8.3 : Procédure infquant()

Debut

genere(sigma, evidence, interet);

PK := transformer(sigma);

infini := calculinfini;

inserer(PK, evidence, infini);

inserer(PK, interet, infini);

WMAXSAT(PK, interpretation, coutinter);

Si (*coutinter=infini*) **Alors**

ecrire("sigma n'infère pas la variable d'intérêt");

sinon

coutinteret := coutinter;

degreinteret := e^{-coutinteret};

ecrire ("sigma infère", interet, "avec un degré", degreinteret);

Fin

Cet algorithme consiste à :

1. générer aléatoirement une base possibiliste quantitative, une évidence et une variable d'intérêt,
2. transformer la base possibiliste en une base de pénalités,
3. insérer l'évidence dans la base de pénalités avec un coût infini,
4. insérer la variable d'intérêt dans la base de pénalités avec un coût infini,
5. exécuter le prouver WMAXSAT avec la base de pénalités comme paramètre d'entrée,

6. calculer le degré de possibilité associé à la variable d'intérêt à partir du coût d'interprétation fournit par WMAXSAT,

Exemple 8.1 Soit à dérouler la procédure d'inférence sur la base possibiliste quantitative suivante :

$$\Sigma = \{(\neg a, 0.999088), (b \vee a, 0.632121), (\neg b \vee \neg a, 0.997521), (c \vee a \vee \neg b, 0.864665), (\neg c \vee \neg a \vee b, 0.981684), (\neg c \vee \neg a \vee \neg b, 0.999088), (d \vee c, 0.999088), (\neg d \vee \neg c, 0.981684), (e \vee a \vee b \vee d, 0.981684), (e \vee a \vee \neg b \vee d, 0.999088), (e \vee a \vee b \vee \neg d, 0.999088), (e \vee a \vee \neg b \vee \neg d, 0.997521), (\neg e \vee \neg a \vee b \vee d, 0.999088), (\neg e \vee \neg a \vee \neg b \vee d, 0.950213), (\neg e \vee \neg a \vee b \vee \neg d, 0.999877), (\neg e \vee \neg a \vee \neg b \vee \neg d, 0.999665)\}.$$

Soient d la variable d'évidence et $\neg c$ la variable d'intérêt.

- **Étape 1** : Elle consiste à déduire la base de pénalités PK équivalente à Σ .
 $PK = \{(\neg a, 7), (b \vee a, 1), (\neg b \vee \neg a, 6), (c \vee a \vee \neg b, 2), (\neg c \vee \neg a \vee b, 4), (\neg c \vee \neg a \vee \neg b, 7), (d \vee c, 7), (\neg d \vee \neg c, 4), (e \vee a \vee b \vee d, 4), (e \vee a \vee \neg b \vee d, 7), (e \vee a \vee b \vee \neg d, 7), (e \vee a \vee \neg b \vee \neg d, 6), (\neg e \vee \neg a \vee b \vee d, 7), (\neg e \vee \neg a \vee \neg b \vee d, 3), (\neg e \vee \neg a \vee b \vee \neg d, 9), (\neg e \vee \neg a \vee \neg b \vee \neg d, 8)\}$
- **Étape 2** : Elle consiste à exécuter PEN -WMAXSAT ayant comme arguments d'entrée la base de pénalités PK , l'évidence d et la variable d'intérêt $\neg c$.
 L'appel à la procédure PEN -WMAXSAT nous fournit une interprétation ω' telle que la somme des coûts des formules falsifiées par cette interprétation est minimale :
 $\omega' = \neg a \vee \neg b \vee \neg c \vee d \vee f$ et $k_{PK}(\omega') = K_{PK}(\text{interet}) = 1$.
- **Étape 3** : Elle consiste à déduire le degré de possibilité de la variable d'intérêt à partir de son coût de pénalité en utilisant l'équation 8.4 :
 $\Pi(\text{interet}) = e^{-K_{PK}(\text{interet})} = e^{-1} = .367879$

8.3.3 Complexité de l'algorithme d'inférence en logique des pénalités

Il est très attrayant d'utiliser la logique des pénalités comme mode de représentation des connaissances incertaines, vu qu'elle présente des propriétés intéressantes pour l'inférence. En effet, la complexité du problème de déduction pour une base de connaissances à pénalités est une des plus simples parmi les modes d'inférence. Elle est de l'ordre de (Δ_2^p) [18].

La classe (Δ_2^p) est une classe de problèmes pouvant être résolus en un temps polynomial par une machine rehaussée avec un NP oracle. La classe NP est un sous ensemble de la classe (Δ_2^p) .

Cette complexité est considérée comme la plus simple [110]. Cependant, elle peut être excessive dans le cadre d'applications pratiques [18].

8.4 Conclusion

Le thème principal de cette thèse est de développer la théorie des possibilités quantitative. Dans les chapitres 5 et 6, nous nous sommes intéressé au volet de la représenta-

tions des connaissances incertaines en théorie des possibilités quantitative et autour de cette théorie. Dans ce chapitre, nous nous sommes penché sur le volet du raisonnement. Contrairement à la logique des possibilités standard, il n'y a pas eu de propositions d'algorithmes d'inférence pour la logique des possibilités quantitative. Le nouvel algorithme d'inférence présenté dans ce chapitre constitue une des contributions principales dans notre travail. Il utilise la procédure de transformation d'une base possibiliste quantitative vers une base de pénalités et il se base sur un prouveur WMAXSAT.

Dans le dernier chapitre de cette thèse, nous allons présenter l'étude expérimentale relative aux différents travaux développés dans cette thèse. Nous allons exposer les différents outils et environnements utilisés pour l'implémentation aussi bien des procédures de transformations que des algorithmes de raisonnement.

D'autre part, nous allons simuler plusieurs scénarios de raisonnement pour les deux modes de représentations afin d'apporter des analyses pertinentes.

Chapitre 9

Implémentation et résultats expérimentaux

9.1 Introduction

Nous présentons dans ce chapitre l'implémentation des différents algorithmes de transformation présentés dans les chapitres 5 et 6, les algorithmes de propagation graphique exposés dans le chapitre 7, l'algorithme d'inférence donné dans le chapitre 8, ainsi qu'une étude expérimentale que nous avons menée afin d'estimer les performances de ces différentes procédures.

Ce chapitre est composé de deux parties. La section 2 expose les différents outils et environnements utilisés pour l'implémentation de l'ensemble des algorithmes de transformations et de raisonnement. La section 3 donne un aperçu sur l'étude expérimentale que nous avons réalisée et qui nous mènera à des analyses pertinentes.

9.2 Implémentation

Nous allons aborder dans cette section les détails sur l'implémentation des différentes procédures effectuée lors de notre étude. Elle concerne principalement :

1. la génération aléatoire de graphe possibiliste,
2. les deux algorithmes de propagations décrits dans les sections 7.3 et 7.4 pour les deux types de DAG : polyarbres et les graphes à connexions multiples,
3. l'algorithme de transformation d'un graphe causal possibiliste basé sur le produit vers une base possibiliste quantitative décrit dans la section 5.3,
4. l'algorithme de passage d'une base possibiliste quantitative vers une base de pénalités décrit dans la section 6.2.2,
5. l'algorithme d'inférence appliqué à une base de pénalité décrit dans la section 8.3.

9.2.1 Mise en œuvre du processus de propagation dans le mode graphique

Pour le mode graphique, nous avons essentiellement exploité les implémentations des algorithmes de propagation développés dans [20] sous Matlab pour les réseaux causaux possibilistes basés sur le produit, qui sont à leur tour, des adaptations des algorithmes de propagation développés dans [21] pour les réseaux causaux probabilistes.

L'implémentation démarre par la génération aléatoire d'un graphe causal possibiliste basé sur le produit. Cet algorithme se résume à :

1. définir les paramètres du réseau : le nombre de nœuds et le nombre maximal de parents.
2. générer pour chaque variable la liste de ses parents afin de spécifier la structure du DAG, en créant ainsi la matrice d'adjacence.
3. spécifier pour chaque nœud (représentant des variables discrètes), sa taille correspondant au nombre de valeurs possibles de la variable qui la représente. Nous considérons pour notre implémentation les variables binaires.

A ce stade de spécification, le réseau bayésien est créé.

4. définir les paramètres du graphe représentés par des objets CPD (Conditional Possibility Distribution) qui représentent des distributions de possibilités a priori des nœuds racines ou les distributions de possibilités conditionnelles connaissant leurs parents pour les autres types de nœuds. Ces distributions sont générées aléatoirement en respectant la condition de normalisation pour chaque nœud qui stipule :

Soit A une variable binaire et u une instance de ses parents : alors $\max(\Pi(a | u), \Pi(\neg a | u)) = 1$.

Vu que lors du processus expérimental, nous allons associer au graphe généré la base de possibilités quantitative équivalente puis la base de pénalités correspondant à cette dernière et que l'algorithme d'inférence appliqué à la base de pénalités appartient à la classe WMAXSAT, nous apportons à ce niveau une restriction qui consiste à faire correspondre aux clauses des bases de connaissances des poids de type *entier*.

Pour cela, au niveau de la génération des distributions conditionnelles de type $\Pi(a | u) = \beta$ avec a une instance d'une variable A et u une instance de $\text{Par}(A)$:

- nous générons d'abord une valeur entière aléatoire k qui va correspondre au coût de pénalité associé à la clause $(\neg a \vee \neg u)$ correspondant à cette distribution,
- puis nous associons à cet entier la valeur e^{-k} qui correspond à la valeur de la distribution de possibilité conditionnelle.

Pour plus de clarté, nous rappelons qu'à une distribution conditionnelle de type $\Pi(a | u) = \beta \neq 1$ et $\beta \in [0, 1]$ où a est une instance d'une variable A associée à un nœud du graphe et u une instance de $\text{Par}(A)$, correspond une clause dans la base de possibilités $(\neg a \vee \neg u, \alpha)$ avec $\alpha \in [0, 1]$ et $\alpha = 1 - \beta$. De plus à une clause possibiliste $(\neg a \vee \neg u, \alpha)$ correspond une clause dans la base de pénalités $(\neg a \vee \neg u, k)$ avec k un entier et $k = -\log(1 - \alpha) = -\log(\beta)$. D'où $\beta = e^{(-k)}$.

5. générer aléatoirement une ou (plusieurs) évidence(s) avec son (leurs) instance(s),
6. générer aléatoirement une variable d'intérêt avec son instance,
7. tester la topologie du graphe afin de lui associer le moteur d'inférence adéquat développé sous Matlab. Si le graphe est un polyarbre, nous utilisons le moteur qui correspond à l'adaptation de l'algorithme de Pearl [10], développé dans [20] pour les réseaux causaux possibilistes basés sur le produit et décrit par la section 7.3. Si le graphe est à connexions multiples, nous utilisons le moteur qui correspond à l'adaptation de l'algorithme de l'arbre de jonction de [12, 11], développé dans [20] pour les réseaux causaux possibilistes basés sur le produit et décrit par la section 7.4.
8. rajouter la ou (les) variable(s) d'observation ou évidence(s) au moteur,
9. injecter la variable d'intérêt au moteur,
10. calculer la valeur de $\text{Bel}(\text{interet}) = \Pi(\text{interet} \mid \text{evidence})$ ainsi que le temps de propagation.

Remarque 9.1 *L'ensemble des paramètres du réseau sont sauvegardés dans un fichier afin de permettre le passage à la base possibiliste quantitative correspondante.*

9.2.2 Mise en œuvre du processus de transformation des connaissances

La procédure de transformation des connaissances s'effectue en deux étapes :

La première étape consiste à transformer le graphe possibiliste basé sur le produit en une base possibiliste quantitative en utilisant la transformation immédiate décrite dans la section 5.3. Elle s'effectue de la manière suivante :

Si Π_G est un graphe causal possibiliste basé sur le produit défini par :

$$\mathcal{P}_G = \{(x_i, u_i, \beta_i) : \Pi(x_i \mid u_i) = \beta_i \neq 1 \in \Pi_G\},$$

où x_i est une instance de la variable A_i et u_i est une instance de $Par(A_i)$,

Alors la base possibiliste quantitative Σ_G associée à Π_G est définie par :

$$\Sigma_G = \{(\neg x_i \vee \neg u_i, \alpha_i) : (x_i, u_i, \beta_i) \in \mathcal{P}_G, \alpha_i = 1 - \beta_i\}$$

Remarque 9.2 *Nous ne considérons que les distributions conditionnelles $\Pi(x_i \mid u_i) = \beta_i \neq 1$, car "1" représente un élément neutre pour le produit.*

La seconde étape consiste à associer à la base possibiliste quantitative la base de pénalités correspondante en exploitant la procédure décrite par la section 6.2.2 et qui stipule :

Si Σ est une base de connaissances possibiliste quantitative définie par : $\Sigma = \{(\phi_i, \alpha_i) : i = 1, \dots, n\}$.

Alors, la base de pénalité PK associée à la base de possibilité quantitative Σ est définie par :

$PK = \{(\phi_i, k_i) : (\phi_i, \alpha_i) \in \Sigma \text{ et } k_i = -\log(1 - \alpha_i)\}$.

Vu que cette base de pénalités sera exploitée par un prouveur de type WMAXSAT, La base respecte le format WCNF (Weighted CNF) où les différentes pondération k_i affectées à chaque clauses sont des entiers et les clauses sont données sous le format DIMACS.

9.2.3 Mise en œuvre de l'algorithme d'inférence

Comme nous l'avons vu dans la section 8.3, l'algorithme d'inférence pour la logique des possibilités quantitative exploite la relation qui existe entre une base de possibilités quantitative et une base de pénalités. De plus, il est basé sur un prouveur de type WMAXSAT sur une base pondérée. Pour l'implémentation de cet algorithme, nous avons opté pour la toolbar de Givry, Heras, Larrosa et Schiex qui fût classée le premier solver WMAXSAT lors de la compétition WMAXSAT 2006 permettant d'évaluer les performances de différents prouveurs. Cette toolbar est solveur générique de type *open source*, pour les instances WCSP, MAXSAT et WMAXSAT, développé sous Linux.

Étant donné que la majorité des prouveurs SAT disponibles sur internet sont développées dans des plateformes UNIX, nous avons eu recours à CYGWIN qui est un émulateur de l'environnement LINUX sous WINDOWS de type *opensource*. Les algorithmes ont été programmés avec le compilateur GCC (GNU C Compiler) sous CYGWIN.

Il est également à noter que nous avons opté pour la plateforme windows vu que les algorithmes de propagations que nous avons utilisés ont été développés dans [20] avec MATLAB sous WINDOWS.

9.2.4 Formats des fichiers utilisés

Comme toute implémentation, les données ainsi que les résultats sont représentées sous forme de fichiers répondant aux formats exigés par les outils qui vont les exploiter.

9.2.4.1 Les connaissances

Les connaissances manipulées dans cette étude sont de trois natures : graphe causal possibiliste basé sur le produit, base possibiliste quantitative et base de pénalités. Sur le plan pratique, ces connaissances sont générées comme suit :

1. génération aléatoire d'un graphe causal possibiliste basé sur le produit,
2. passage du graphe causal possibiliste basé sur le produit vers une base possibiliste quantitative,
3. transformation de la base possibiliste quantitative en une base de pénalités

D'une manière pratique, ces différentes connaissances sont représentées en répondant aux exigences des outils qui vont les exploiter.

Exemple 9.1 Afin d'illustrer de quelle manière ces connaissances sont représentées, nous considérons le graphe possibiliste basé sur le produit représenté par la figure 9.1. Les distributions de possibilités conditionnelles sont données par les tables 9.1 et 9.2.

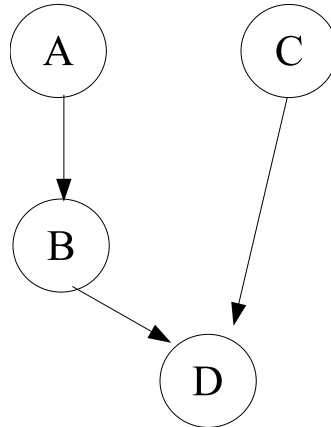


FIG. 9.1 – Exemple d’un DAG

TAB. 9.1 – Distributions a priori associées au graphe 9.1

$\Pi(\neg a)$	$\Pi(a)$	$\Pi(\neg c)$	$\Pi(c)$
1	.006738	1	.049787

La table 9.3 résume de quelle manière les connaissances sont représentées.

La première colonne donne le descriptif du fichier correspondant au graphe de la figure 9.1. Il englobe :

1. les paramètres du réseau : le nombre de variables, le nombre maximal de parents.
2. la matrice d’adjacence qui décrit la structure du graphe. Si $DAG(i,j)=1$ alors $variable_i \in Par(variable_j)$.
3. Les distributions a priori pour les variables racines et les distributions conditionnelles pour les autres variables.
4. les variables d’observations et leurs instances suivies par la variable d’intérêt et son instance. Une instance négative d’une variable est représentée par 1 et une instance positive est représentée par 2. Dans l’exemple, les évidences sont : $c \wedge \neg a$ et la variable d’intérêt est $\neg d$.

La seconde colonne représente, sous le format DIMACS, la base possibiliste quantitative associée au graphe représenté par la colonne 1 de la table 9.3. Elle englobe :

1. une entête "p wcnf nbvar nbclauses" qui indique le type WCNF (Weighted CNF), le nombre de variables ainsi que le nombre exact de clauses contenues dans le fichier.
2. les clauses de la bases. Chaque ligne débute par le poids affilié à la clause qui correspond à son degré de possibilité qui est suivi par la clause. Chaque clause est formée d’une séquence de nombres non nuls distincts compris entre 1 et nbvar, représentant les variables, qui se termine par 0. Les nombres

TAB. 9.2 – Distributions conditionnelles associées au graphe 9.1

A	$\Pi(\neg b)$	$\Pi(b)$	B C	$\Pi(\neg d)$	$\Pi(d)$
$\neg a$.000123	1	$\neg b \neg c$.000123	1
a	1	.049787	b \neg c	1	.135335
			$\neg bc$.367879	1
			b c	1	.000912

positifs dénotent les variables correspondantes. Les nombres négatifs dénotent la négation des variables correspondantes.

La dernière colonne de la table 9.3 décrit la base de pénalités qui correspondent à la base de possibilités quantitative donnée par la deuxième colonne de la table 9.3. Une base de pénalités est composée de la même entête que la base possibiliste, qui est suivie par les clauses de la base. Le coût de pénalité de chaque clause est identifié par le premier entier compris dans l'intervalle $[1, 2^{20}]$. Pour les clauses, elles sont décrites de la même manière que les clauses de la base possibiliste.

Les clauses sont suivies par les evidences et la variable d'intérêt mises sous forme clause.

9.2.4.2 Les fichiers résultats

Les résultats des algorithmes de raisonnement sont récupérés dans un fichier dont le contenu est illustré par la table 9.4. Ce fichier englobe deux parties. La première partie résume les paramètres du réseau ainsi que les résultats fournis. La seconde partie donne des informations sur le processus d'inférence. Ainsi, la consultation de ce fichier donne un aperçu sur le déroulement des deux algorithmes de raisonnement.

Remarque 9.3 *Dans le cas où la valeur du degré de possibilité conditionnelle $\Pi(\text{interet} \mid \text{evidences})$ est nulle, les algorithmes de raisonnement produisent un fichier résultat similaire à celui illustré par la table 9.5.*

L'exécution de la toolbar avec un fichier de type wcnf fournit un fichier de sortie de la forme suivante :

```

c —————
c — Toolbar Version 3.0 —————
c — Local Consistency : EDAC—————
c — Max-SAT technique : HyperResolution (-16)-

```

```

c-----
o 674
o 315
o 291
o 283
o 153
o 118
o 32
s OPTIMUM FOUND
v 1 -2 -3 -4 -5 -6 -7 -8 -9 10 -11 -12 -13 -14 -15 -16 -17 -18 -19 -20 -21 -22 23 -24
c Total time 0 seconds and 24 nodes.
Optimum : 32 in 24 nodes and 0.041 seconds.

```

Les ligne commençant par "c" sont des lignes commentaires. Les lignes qui débutent par "o" indiquent les solutions optimales courantes qui correspondent à la somme minimale des coûts des clauses insatisfaites par la solution courante. La solution optimale finale est donnée par la dernier ligne "o". Elle sera suivie par l'interprétation correspondante.

9.3 Etude expérimentale

Dans cette section, nous exposons quelques expérimentations que nous avons effectuées afin de pouvoir évaluer les performances des algorithmes de raisonnement à savoir :

- l'algorithme de propagation appliqué à un graphe causal possibiliste basé sur le produit dans le cas des polyarbres ou l'algorithme de propagation appliqué à un réseau à connexions multiples,
- l'algorithme d'inférence appliqué à une base de pénalités déduite à partir de la base possibiliste quantitative correspondant au graphe causal possibiliste.

Nous avons considéré comme critère de comparaison le temps d'exécution estimé en secondes. Ces différentes expérimentations ont été élaborées comme suit :

- La première phase consiste à générer d'une manière aléatoire un graphe causal possibiliste basé sur le produit en utilisant l'adaptation réalisée dans ce travail sur l'algorithme de génération aléatoire développé dans [20].

Ce processus prend comme paramètre en entrée le nombre de variables ainsi que le nombre maximal de parents et permet de générer la structure du graphe, les distributions a priori (cas des nœuds racines) ou les distributions conditionnelles associées aux autres variables, les evidences ainsi que la variable d'intérêt. Ces données sont sauvegardées dans un fichier de type Matlab dont la structure est décrite par la colonne 1 de la table 9.3,

- La seconde étape consiste à appliquer l'algorithme de propagation à la structure du graphe généré. Pour cela, nous avons utilisé les noyaux implémentés dans la PNT (Possibilistic Net Toolbox) et développés dans [20], l'un pour les polyarbres (décrit dans la section 7.3) et l'autre pour les graphes à connexions multiples

- (décrit dans la section 7.4),
- La troisième phase consiste à appliquer l’algorithme de transformation linéaire d’un graphe possibiliste basé sur le produit vers une base possibiliste quantitative, en utilisant la procédure décrite dans la section 5.3,
 - La quatrième phase consiste à faire correspondre à la base possibiliste générée lors de l’étape précédente, la base de pénalités équivalente et ceci en utilisant la procédure décrite dans la section 6.2.2
 - La dernière étape consiste à exécuter l’algorithme d’inférence, décrit dans la section 8.3, sur la base de pénalités générée lors de la phase précédente, qui fournit le coût de la variable d’intérêt. Le degré de possibilité de la variable d’intérêt sera alors calculé en appliquant la proposition 9.

9.3.1 Comportements des algorithmes de raisonnement

Nous avons élaboré divers scénarios d’expérimentation. Pour les centaines de tests effectués, nous avons obtenu une égalité exacte entre le degré de possibilité conditionnel $\Pi(\textit{interet} \mid \textit{evidence})$ obtenu par l’algorithme de propagation et le degré de possibilité $\Pi(\textit{interet})$ obtenu par l’algorithme d’inférence.

Afin de diversifier l’étude expérimentale, nous avons considéré trois catégories de graphes : les polyarbres, les graphes de complexité moyenne et les graphes denses. Pour chaque catégorie de graphe, nous avons fait varier le nombre des variables d’évidence, vu qu’en théorie, l’évidence est composée par une conjonction de littéraux.

9.3.1.1 Cas des polyarbres

L’algorithme de propagation utilisé pour cette catégorie de graphes est l’adaptation possibiliste de l’algorithme de Pearl [10] réalisée dans [20]. Cet algorithme est de complexité polynômiale. Pour l’algorithme d’inférence, il appartient à la classe Δ_2^p . Le graphe 9.2 (respectivement 9.3) illustre les comportements des algorithmes de propagation dans les réseaux de croyances et d’inférence pour des bases de pénalités dans le cas d’une évidence (respectivement de deux évidences). Il est clair que l’algorithme de propagation est plus performant que l’algorithme d’inférence. Ce résultat est en diapason avec les classes de complexité auxquelles appartiennent les deux algorithmes de raisonnement.

9.3.1.2 Cas des graphes de complexité moyenne

La densité d’un graphe est estimée par le critère du nombre maximal de parents. Pour ce scénario, nous avons fixé ce paramètre à 3. Dans le cas où le graphe généré présente une complexité moyenne, l’algorithme de propagation adopte divers comportements :

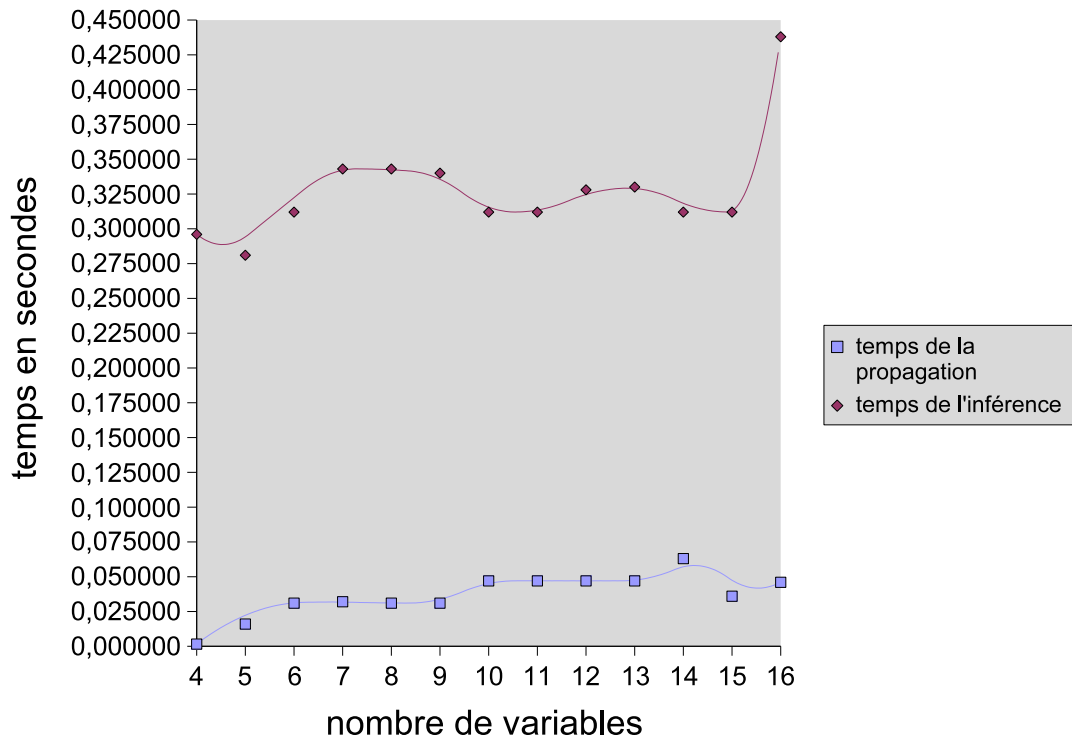


FIG. 9.2 – Temps d'exécution dans le cas des polyarbres avec une évidence

- si le nombre de nœuds du graphe est inférieur à un certain seuil (95 nœuds), il est plus intéressant d'opter pour la structure graphique car l'algorithme correspondant est plus performant.
- si le nombre de nœuds du graphe est assez grand (supérieur à 95), la représentation logique est alors à préconiser car l'algorithme correspondant est plus optimal lors de l'exécution.

Ces différentes constations sont illustrées par le graphe 9.4.

9.3.1.3 Cas des graphes complexes

Dans les graphes à connexions multiples, la complexité du problème de construire un arbre de jonction optimal est un problème NP-complet. Ceci se reflète sur le comportement de l'algorithme de propagation.

La figure 9.5 illustre ce comportement dans le cas de graphes denses (nombre de parents maximal=7). Dans le cas où le nombre de nœuds n'est pas élevé (inférieur à 45), l'algorithme de propagation graphique est plus rapide. Mais dans le cas de graphe de plus en plus complexe, il s'avère que l'algorithme d'inférence devient plus performant.

De plus, lorsque le nombre de nœuds dépasse un certain seuil (supérieur à 60), l'algo-

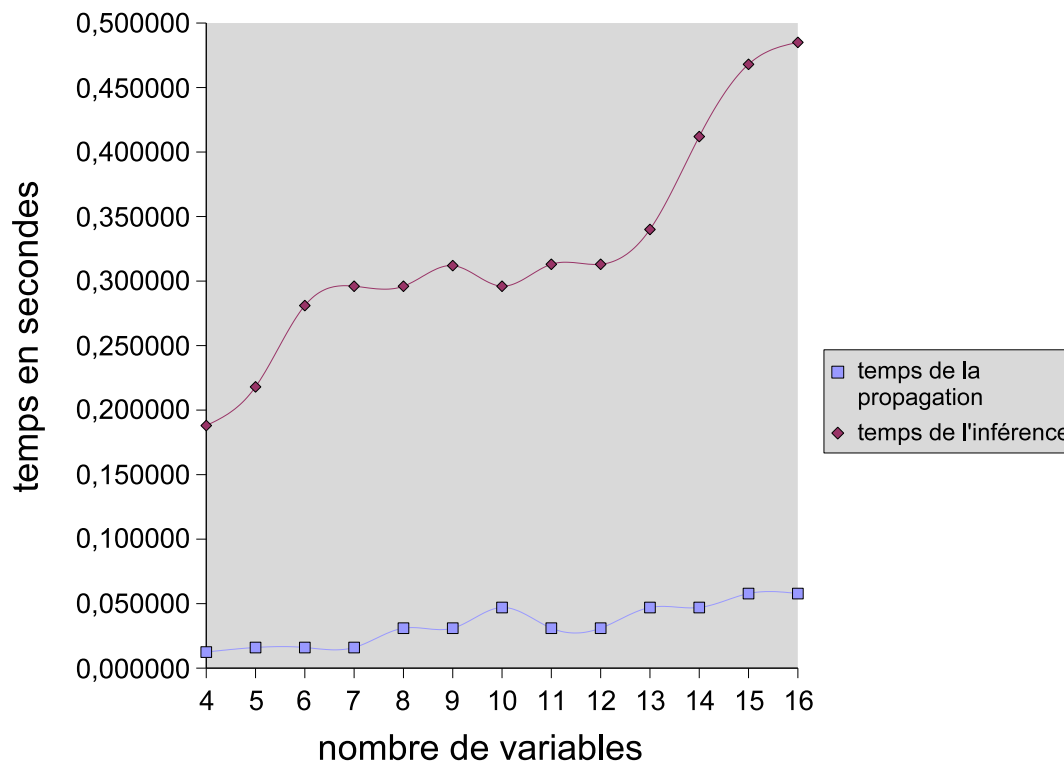


FIG. 9.3 – Temps d'exécution dans le cas des polyarbres avec deux évidences

rithme de l'arbre de jonction bloque car il engendre des grands clusters pour lesquels les calculs locaux deviennent impossibles. La figure 9.6 illustre ce phénomène et montre que l'algorithme d'inférence ne présente pas les limites induites par l'algorithme de l'arbre de jonction, vu qu'il repose essentiellement sur un prouveur de type WMAXSAT.

Les figures 9.7 et 9.8 montrent l'effet du second paramètre de complexité (nombre maximal de parents) sur l'algorithme de propagation dans les cas d'une ou de plusieurs évidences. A titre expérimental, nous avons fixé le nombre de variables à 50. Ces deux graphes ont la même allure, ce qui implique que le nombre de variables d'évidence n'a aucune influence sur les algorithmes.

D'autres part, nous remarquons deux comportements selon le nombre maximal de parents :

- s'il n'est pas trop élevé (inférieur à 5), le raisonnement en mode graphique est plus optimal,
- s'il est élevé, le raisonnement en mode logique devient plus avantageux.

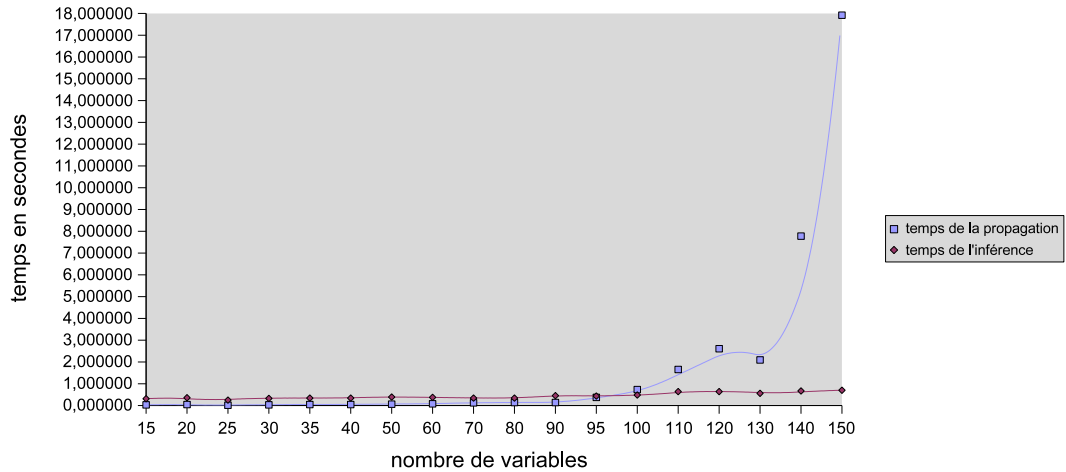


FIG. 9.4 – Temps d'exécution dans le cas des graphes moyennement denses (nbpar-max=3)

9.4 Conclusion

La première partie de ce chapitre était consacrée à présenter les différents outils et environnements utilisés afin de mener l'étude expérimentale.

La seconde partie présente les résultats expérimentaux obtenus. Ces derniers sont conformes avec la complexité des algorithmes d'inférences et les outils déployés pour simuler les mécanismes de raisonnement.

Ces résultats illustrent le fait que les performances des deux algorithmes de raisonnement dépendent de la structure du graphe :

- dans le cas où la structure du graphe généré est un polyarbre, il est clair que l'algorithme de propagation de Pearl est plus performant que l'algorithme d'inférence, et ce même dans le cas d'un DAG ayant un très grand nombre de nœuds,
- dans le cas où la structure du graphe généré est DAG à connexions multiples, les performances des algorithmes dépendent de la complexité du DAG :
 - si le DAG est de complexité moyenne, vu que le problème de construire l'arbre de jonction le plus optimal est un problème NP-complet, l'algorithme de propagation reste toujours plus performant que l'algorithme d'inférence.
 - si le DAG est complexe (plus de 50 nœuds) et dense (avec un nombre de parents maximal > 4), l'algorithme d'inférence logique donne des résultats plus performants. Dans ce cas, il serait plus judicieux :
 - de transformer le graphe en une base de possibilités,
 - puis de transformer la base de possibilités en une base de pénalités correspondantes,
 - puis d'appliquer l'algorithme d'inférence logique.

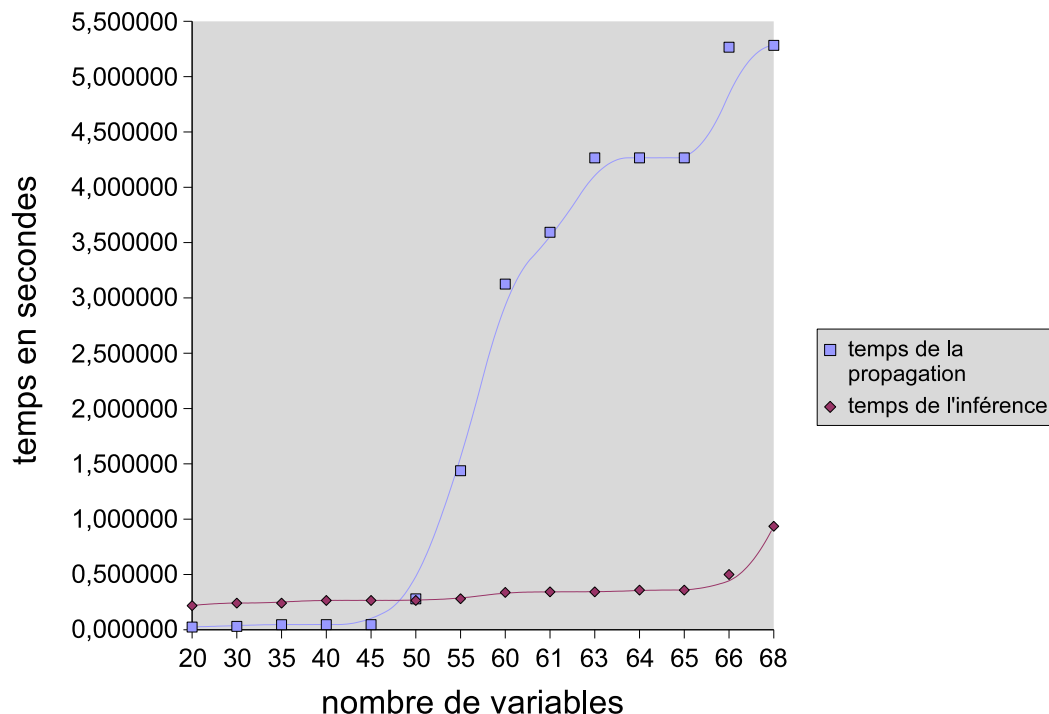


FIG. 9.5 – Temps d'exécution dans le cas de graphes denses avec une évidence

Dans le cas des graphes complexes, l'algorithme de l'arbre de jonction engendre des grands clusters pour lesquels les calculs locaux deviennent impossibles, et l'algorithme de propagation bloque. En effet, les différents scénarios élaborés montrent bien qu'au delà de 60 nœuds, l'algorithme de propagation bloque. Cette limite n'est pas imposée à l'algorithme d'inférence.

Cette étude expérimentale est en corrélation totale avec notre étude théorique en rappelant que :

- la complexité de l'algorithme de propagation dans le cas des polyarbres est de l'ordre de $\mathcal{O}(s * N)$ où s est la taille maximale des tables de distributions de possibilités conditionnelles et N est le nombre de variables,
- dans l'algorithme de propagation dans les réseaux à connexions multiples, la construction de l'arbre de jonction optimal est un problème NP-complet [87]
- la complexité du problème d'inférence appliqué à une base possibiliste quantitative est de l'ordre (Δ_2^p) .

Pour conclure, chacun des modes de raisonnement présente des avantages et des inconvénients. L'algorithme de l'arbre de jonction est limité par l'espace par contre l'algorithme d'inférence basé sur WMAXSAT peut être limité par le temps.

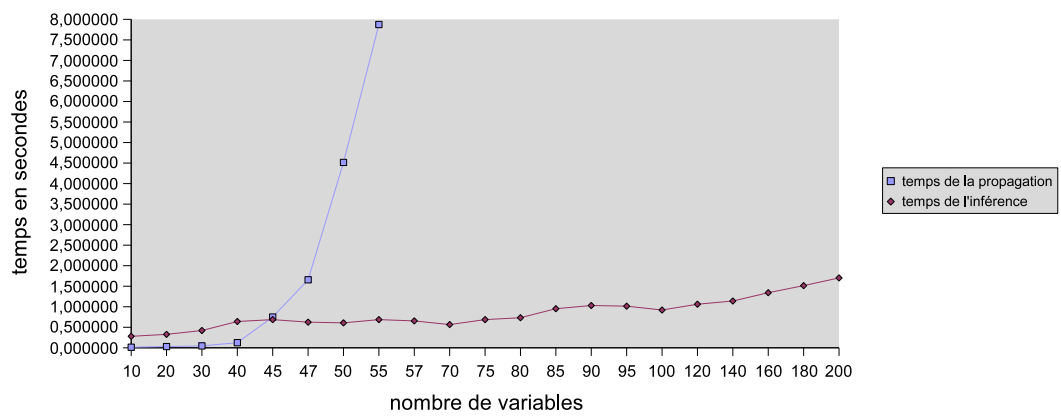


FIG. 9.6 – Comportement des algorithmes de raisonnement (nbparmax=7)

TAB. 9.3 – Description des formats des connaissances

Graphe causal possibiliste basé sur le produit	Base possibiliste quantitative	Base de pénalités sous le format DIMACS
paramètres du réseau	nombre de variables	Entête : nombre de variables
nombre de variables	nombre de clauses	et nombre de clauses
nombre de parents max	4 8	
4 2	p wcnf 4 11	p wcnf 4 11
Composante graphique	Clauses	Clauses
0 1 0 0		5 -1 0
0 0 0 1	0.993262 -1 0	9 2 1 0
0 0 0 1	0.999877 2 1 0	3 -2 -1 0
0 0 0 0	0.950213 -2 -1 0	3 -3 0
Composante numérique	0.950213 -3 0	9 4 2 3 0
1.000000	0.999877 4 2 3 0	1 4 2 -3 0
0.006738	0.632121 4 2 -3 0	2 -4 -2 3 0
0.000123	0.864665 -4 -2 3 0	7 -4 -2 -3 0
1.000000	0.999088 -4 -2 -3 0	Variables d'évidence
1.000000		39 3 0
0.049787		39 -1 0
1.000000		Variable d'intérêt
0.049787		39 -4 0
0.000123		
1.000000		
0.367879		
1.000000		
1.000000		
0.135335		
1.000000		
0.000912		
Variables d'évidence avec leurs instances		
3		
2		
1		
1		
Variable d'intérêt avec son instance		
4		
1		

TAB. 9.4 – prototype du fichier résultat

Résultat obtenu dans le cas où la base n'infère pas la variable d'intérêt
*****Résultats de la propagation graphique*****
Nombre de variables : 8
Nombre de parents max : 3
Évidence 1 : 5
Évidence 2 :-1
Variable d'intérêt : 6
Distribution de possibilité conditionnelle est de : 0.135335
Temps de propagation est de : 0.016000 secondes
*****Résultats de l'inférence logique*****
Le nombre de clauses dans les bases est de : 27
La variable d'intérêt est inférée à partir de la base de pénalités
Le coût de pénalité est de : 2
Le degré de possibilité égale à : 0.135335
La durée de l'inférence est de : 0.234000 secondes

TAB. 9.5 – prototype du fichier résultat dans le cas où la base n'infère pas la variable d'intérêt

*****Résultats de la propagation graphique*****
Nombre de variables : 6
Nombre de parents max : 1
Évidence : 2
Variable d'intérêt : 1
Possibilité conditionnelle de l'intérêt évidence 0.000000
Temps de propagation 0.110000 secondes
*****Résultats de l'inférence logique*****
La variable d'intérêt n'est pas inférée à partir de la base de pénalités
Le coût de pénalité est de : 15
La durée de l'inférence est de : 0.343000 secondes

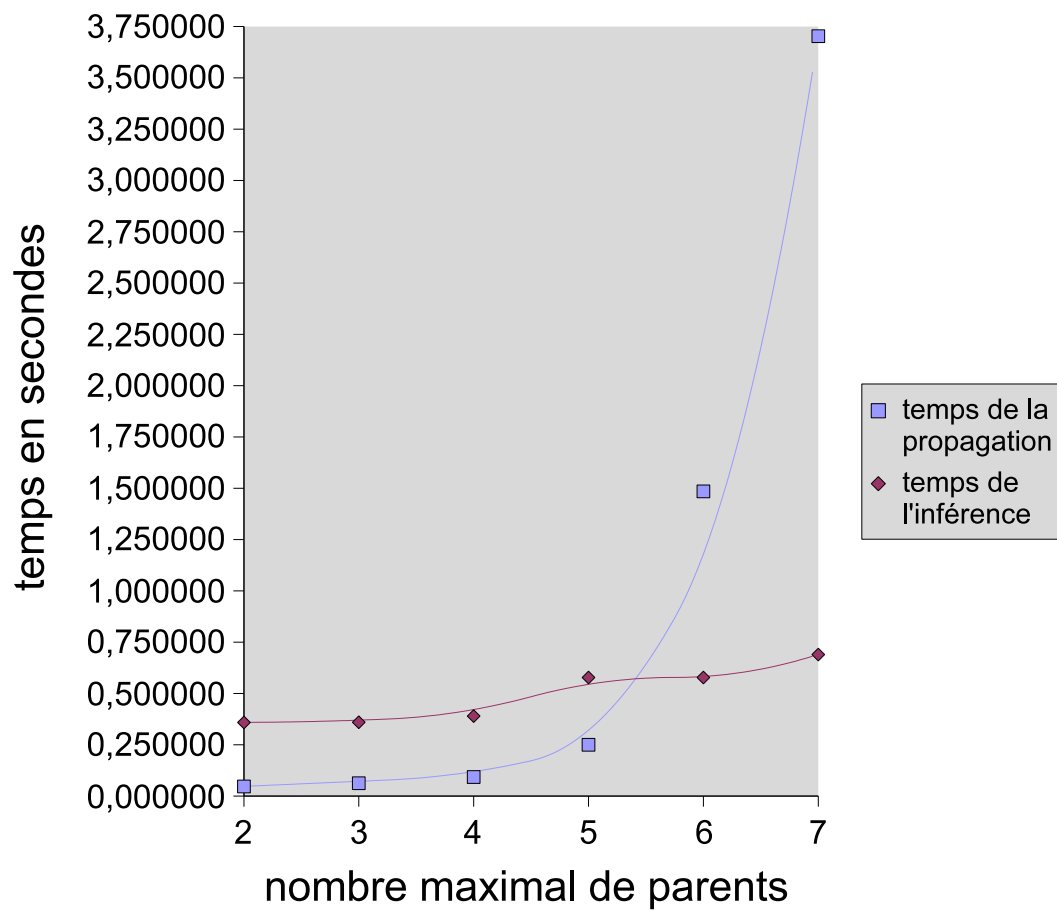


FIG. 9.7 – Temps d'exécution dans le cas des graphes à connexion multiples avec une évidence

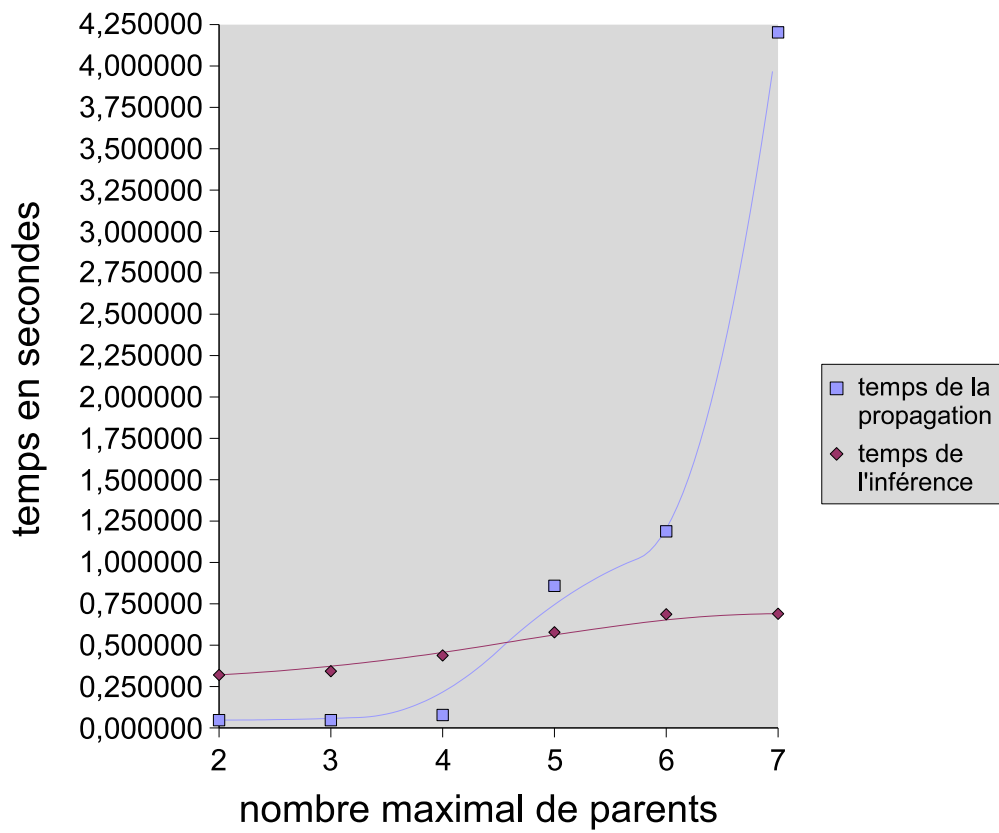


FIG. 9.8 – Temps d'exécution dans le cas des graphes à connexions multiples avec deux évidences

Conclusion Générale

Le thème général de cette thèse tourne autour d'une théorie non probabilisée qui est la théorie des possibilités quantitative pour représenter et raisonner sur des connaissances dans un contexte incertain.

La synthèse de l'état de l'art que nous avons présenté nous a permis de constater la richesse et la panoplie des modes de représentation des connaissances. Ces modes reposent sur des théories de l'incertain telles que la théorie des probabilités, la théorie de l'évidence et la théorie des possibilités. Les chercheurs pionniers en intelligence artificielle ont d'abord été des adeptes pour les modes logiques vu la robustesse des algorithmes d'inférence. Mais ces dernières années, l'avènement des modèles graphiques tels que les réseaux bayésiens a permis de considérer d'autres types de représentation plus expressifs, offrant des algorithmes de raisonnement très puissants.

La théorie des possibilités dans ces deux contextes (qualitatif et quantitatif) offre les deux modes de représentation (à savoir le mode logique et le mode graphique).

Dans le cadre qualitatif, plusieurs chercheurs se sont intéressés à faire des correspondances entre ces deux modes de représentation. Néanmoins, peu de travaux ont concerné la théorie des possibilités quantitative.

Principales contributions

Les travaux réalisés durant cette thèse apportent diverses contributions autour de la théorie des possibilités quantitative tant sur le plan de la représentation des connaissances que sur le plan du raisonnement.

Sur le plan de la représentation des connaissances, ils ont permis d'identifier de manière précise les relations entre le modèle logique et le modèle graphique en théorie de possibilités quantitative. Précisément, ils montrent que toute base possibiliste quantitative peut être codée de façon équivalente par un réseau causal possibiliste basé sur le produit. Ces procédures donnent ainsi une liberté aux experts de pouvoir exprimer leurs connaissances incertaines en utilisant les deux modes de représentation. Ces procédures peuvent être également exploitées dans le cas de fusion de connaissances.

D'autres part, cette étude s'est intéressée à d'autres modes de représentation des connaissances incertaines, en particulier à la logique des pénalités et aux réseaux causaux basés sur les OCF. En effet, elle montre que les réseaux de croyances basés sur les "fonctions de conditionnement ordinal" ainsi que les bases de pénalités peuvent être naturellement représentés avec les bases de connaissances possibilistes quantitatives.

Ces résultats sont importants et peuvent être vus comme une clarification des relations qu'entretiennent la théorie des probabilités et la théorie des possibilités. En effet, les réseaux causaux possibilistes ont des relations étroites avec les probabilités infinitésimales vu leurs relations avec les réseaux causaux basés sur les "fonctions de conditionnement ordinal". De plus, ils peuvent être vu comme des représentations intermédiaires entre les réseaux causaux *numériques*, représentés par les réseaux causaux bayésiens et les réseaux causaux *qualitatifs*, représentés par les réseaux causaux possibilistes basés sur le minimum.

Sur le plan du raisonnement, ce travail a permis :

- d'exploiter les algorithmes de propagation pour les réseaux causaux possibilistes basés sur le produit, développés dans [20], pour les deux types de structure, les polyarbres et les graphes à connexions multiples. Ces algorithmes étant des adaptations directes des algorithmes de propagation dans le cadre probabiliste, de part leurs similitudes par rapport à la nature des connaissances manipulées (numériques).
- de développer un algorithme d'inférence pour une base de pénalités basé sur un prouveur de type WMAXSAT. Cet algorithme peut être appliqué à une base possibiliste quantitative en exploitant la procédure de transformation linéaire d'une base possibiliste quantitative vers une base à pénalités. Ainsi, la théorie des possibilités quantitative dispose d'une implémentation alternative à celle des réseaux causaux quantitatifs.

En termes de complexité, l'algorithme de propagation dans un polyarbre est de complexité polynômial ($\mathcal{O}(s * N)$) où s est la taille maximale des tables de distributions de possibilités conditionnelles et N est le nombre de variables). Par contre dans le cas d'un graphe à connexions multiples, le problème est NP-complet. Pour l'algorithme d'inférence logique, il est de l'ordre de (Δ_2^p) .

En marge de ces développements, nous avons implémenté les algorithmes mis en œuvre afin de pouvoir en estimer les performances. Les algorithmes de propagation sont efficaces dans le cas de graphes non complets.

Seulement, dès que la densité du graphe s'accroît, le temps de propagation devient plus important si nous le comparons à celui de l'algorithme d'inférence.

En pratique, dans une optique de multi-expertise, un agent, pourra opter pour un mode de représentation selon la structure du graphe représentant ces connaissances et utiliser l'algorithme de raisonnement qui correspond au mode sélectionné. Ce travail est très intéressant car il donne accès à plusieurs pistes de recherche.

Perspectives de recherche

Cette étude augure diverses perspectives.

Améliorations

Les procédures développées dans le cadre de ce travail peuvent être optimisées comme suit :

- Les différentes procédures de transformations ainsi que les algorithmes de raisonnement, prennent en compte uniquement les variables binaires. La prise en charge de l’extension du domaine des variables, est immédiate.
- Pour la procédure de passage d’une base possibiliste quantitative vers un graphe causal basé sur le produit, l’affectation de l’ordre arbitraire entre les variables influe sur la structure du graphe. Il serait donc intéressant de définir des heuristiques afin d’obtenir la structure de graphe la moins complexe possible afin d’optimiser le temps de la propagation.
- Le problème d’inférence à partir d’une base possibiliste quantitative est un problème dont la complexité est de l’ordre de Δ_2^p . Il serait transformer la base de connaissances afin de pouvoir utiliser un prouveur de la classe MAXSAT connu pour être un des problèmes NP-difficiles.

Approfondissements

- En marge de ce travail, d’autres pistes de recherches peuvent être envisagées :
- Sur le plan logique, il serait intéressant d’exploiter les techniques de compilation développées dans [111, 112]. Elles consistent à transformer les bases de connaissances en utilisant le format WDDNF dans une phase *offline* et de procéder à l’inférence à partir de la base compilée *online*. La complexité du processus d’inférence sera ainsi réduite.
 - Sur le plan graphique, nous envisageons de nous intéresser aux Valuations Based Systems (VBS) [113, 114, 115, 116, 62]. Les VBS incluent deux opérateurs : la combinaison et la marginalisation qui opèrent sur les valuations. Le processus de raisonnement dans les VBS peut être décrit simplement comme un calcul de la valuation jointe marginale associée à chaque variable du système. Le cadre des VBS est assez général pour représenter plusieurs domaines tels que la théorie des probabilités [116], la théorie des croyances de la théorie de Dempster-Shafer [35, 25, 61], la logique proportionnelle [114, 115], la théorie de la décision [117] et la théorie des possibilités [62].
- Dans cet esprit, il serait judicieux de se pencher et d’exploiter les similitudes qui existent entre les réseaux causaux possibilistes basés sur le produit et les VBS. En effet, les deux types de représentations graphiques reposent sur la même définition du conditionnement, basée sur le produit. Il est à noter que dans [62], Shenoy a proposé une version possibiliste de l’algorithme de propagation dans les VBS utilisant uniquement l’opérateur produit.

Développements

- En vue d’exploiter la relation qui existe entre un graphe causal possibiliste basé sur le produit et un réseau causal basé sur les OCF, nous envisageons de développer des algorithmes de propagation dans les réseaux basés sur les OCF. Pour cela, il va falloir d’abord faire une étude détaillée sur les relations d’indépendances dans les réseaux causaux basés sur les OCF.

- Enfin, pour clôturer ce travail, nous souhaitons souligner le fait que les efforts fournis tout au long de cette thèse débordent de la théorie des possibilités quantitative. En effet, dans le contexte qualitatif de cette théorie, nous avons développé, dans le cadre d'un encadrement de magistère [118], deux algorithmes pour le calcul de la décision optimale dans le cas pessimiste et le cas optimiste. Nous avons translaté ce domaine en un problème de fusion de connaissances de deux distributions possibilistes particulières [15] : l'une représentant les croyances d'un agent et l'autre représentant l'utilité qualitative. Nous avons montré que le calcul de la décision optimale est équivalent au calcul d'un degré d'inconsistance d'une base possibiliste qui résulte de la fusion entre les croyances et les préférences d'un agent. Nous avons ainsi remédié aux limites présentées par l'approche basée sur les ATMS.

Bibliographie

- [1] D. Dubois and H. Prade. Uncertainty issues in artificial intelligence. In *Workshop on Philosophy of Probability*, 1990.
- [2] D.Kaysers. *La représentation des connaissances*. Editions Hermès, 1997.
- [3] P.Smets. Varieties of ignorance and the need for well-founded theories. 1999.
- [4] P.Smets. Imperfect information : Imprecision- uncertainty. *UMIS*, 1999.
- [5] B.Bouchon-Meunier. *La logique floue et ses applications*. Addison-Wesley France, SA, 1995.
- [6] G. Shafer. A mathematical theory of evidence. In *New Jersey : Princeton University Press*, 1976.
- [7] L. Zadeh. Fuzzy sets. *Information and Control*, 8 :338–353, 1965.
- [8] L. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1 :3–28, 1978.
- [9] D. Dubois and H. Prade. Necessity measures and the resolution principle. *IEEE Trans. on Systems, Man and Cybernetics*, 17 :474–478, 1987.
- [10] J. Pearl. Probabilistic reasoning in intelligent systems : Networks of plausible inference. Morgan Kaufmann Publ. Inc., San Mateo, Ca, 1988.
- [11] F. V. Jensen. An introduction to Bayesian Networks. University College London : UCL Press, 1996.
- [12] S.L. Lauritzen and D.J. Spiegelhalter. Local computations with probabilities on graphical structures and their application to expert systems (with discussion). In *J Royal Statist Soc B*, volume 50, pages 157–224, 1988.
- [13] S. Benferhat, D. Dubois, L. Garcia, and H. Prade. On the transformation between possibilistic logic bases and possibilistic causal networks. *International Journal of Approximate Reasoning*, 29 :135–173, 2002.
- [14] S. Benferhat, D. Dubois, S. Kaci, and H. Prade. Graphical readings of possibilistic logic bases. In *Seventeenth International Conference on Uncertainty in Artificial Intelligence (UAI'01)*, pages 24–31, 2001.
- [15] S. Benferhat, D. Dubois, and H. Prade. From semantic to syntactic approaches to information combination in possibilistic logic. *Aggregation and Fusion of Imperfect Information*, Physica Verlag, pages 141–151, 1997.

- [16] S.Benferhat, F.Khellaf-Haned, and A.Mokhtari. Réseau causal possibiliste basé sur le produit et logique possibiliste quantitative. In *14 ème Congrès Francophone AFRIF ARIA de reconnaissances des formes et Intelligence artificielle*, Janvier 2004.
- [17] S.Benferhat, F.Khellaf-Haned, and A.Mokhtari. Product-based causal networks and quantitative possibilistic base. In *The Florida Artificial Intelligence Research Society FLAIRS04 Floride*, May 2004.
- [18] F. Bannay-Dupin de Saint-Cyr. Gestion de l'évolutif et de l'incertain en logiques pondérées. In *Thèse de doctorat. Université Paul Sabatier. Toulouse.*, 1996.
- [19] S.Benferhat, F.Khellaf-Haned, and A.Mokhtari. Product-based causal networks and quantitative possibilistic base. *International journal of Uncertainty, Fuzziness and Knowledge-based Systems IJUFKS*, 13(5) :469–493, October 2005.
- [20] N.Ben Amor. *Qualitative possibilistic graphical models :from independance to propagation algorithms*. PhD thesis, ISG Université de Tunis, 2002.
- [21] K.P.Murphy. The bayes net toolbox for matlab. In Computing Science and Statistics, editors, *Proceeding of interface*,33, 2001.
- [22] G. Pinkas. Propositional nonmonotonic reasoning and inconsistency in symmetric neural networks. In *12th IJCAI, Sydney, Australia, Morgan-Kaufmann*, pages 525–530, 1991.
- [23] F.Dupin de Saint-Cyr, J.Lang, and T.Schiex. Gestion de l'inconsistance dans les bases de connaissances ; une approche syntaxique basée sur la logique des pénalités. In *9ème congrés Reconnaissance des Formes et Intelligence Artificielle*. Pris, France, 1994.
- [24] M. Goldszmidt. Belief-based irrelevance and networks : Towards faster algorithms for prediction. In *AAAI'94 Fall Symposium Series : Relevance*. <http://www.pitt.edu/~druzdzel/papers/moises.html>, pages 89–94, 1994.
- [25] W. Spohn. Ordinal conditional functions : a dynamic theory of epistemic states. In *Causation in Decision, Belief Change, and Statistics*, Vol. 2 :pages 105–134, 1988.
- [26] W.Sphon. A general non probabilistic theory of inductive reasoning. In *In proceeding of the fourth conference of uncertainty in Artificial Intelligence (UAI88)*, pages 149–158, 1988.
- [27] L.A.Zadeh. The concept of a linguistic variable and its application to approximate reasoning. *Information science*, 9 :43–80, 1975.
- [28] P.Fabiani. *Représentation dynamique de l'incertain et stratégies de prise d'information pour un système autonome en environnement évolutif*. PhD thesis, Ecole Nationale Supérieure de l'Aéronotique et de l'Espace, 1996.
- [29] D.Dubois and H.Prade. Representation and combinaison of uncertainty with belief functions and possibility mesures. In *Computer Intelligence 4*, pages 244–264, 1988.

- [30] D.Dubois and H.Prade. *Introduction aux systèmes experts de gestion*, chapter Le traitement de l'imprécision et de l'incertitude dans les modèles de raisonnement des experts. Eyrolles, 1985.
- [31] P.Smets. What is dempster-shafer model? In J.Kacprzyk R.Yager and M.Fedrizzi, editors, *Advances in Dempster-Shafer theory of evidence*, pages 5–34, 1994.
- [32] P.Smets and P.Magrez. La représentation de la croyance. *Revue de l'intelligence artificielle*, pages 31–46, 1987.
- [33] P.Smets, E.H.Mamdani, D.Dubois, and H.prade. *Non-Standard Logics for Automated Reasoning*. 1988.
- [34] E.Ouzounova. *Incertainitude et mesure de performance*. PhD thesis, Université de Fribourg Suisse, 2005.
- [35] A.P.Dempster. Upper and lower probabilities induced by a multivalued mapping. In *Ann Math Stat 38*, pages 325–339, 1967.
- [36] S.Parsons and E.H.Mamdani. Qualitative dempster-shafer theory. In *Decision Technologies*. Barcelone, June 1993.
- [37] N.Wilson. *Some theoretical aspects of the Dempster-Shafer theory*. PhD thesis, Departement of computing and mathematical sciences Oxford Polytechnic, 1992.
- [38] P. Smets. Belief Functions. In *Non-Standard Logics for Approximative Reasoning (P. Smets, A. Mamdani, D. Dubois and H. Prade eds)*, pages 253–286. New York : Academic Press, 1988.
- [39] P. Smets. The transferable belief model for quantified belief representation. In *Handbooks of Defeasible Reasoning and Uncertainty Management Systems, Vol. 1*, pages 267–301. (D. Gabbay and P. Smets, eds.) Kluwer, Dordrecht, 1998.
- [40] D.Dubois and H.Prade. Possibility theory and its applications : A retrospective and prospective view. In *IEEE International Conference on Fuzzy Systems*, pages 3–11, 2003.
- [41] D. Dubois and H. Prade. (with the collaboration of H. Farreny, R. Martin-Clouaire and C. Testemale). Possibility Theory - an Approach to Computerized Processing of Uncertainty. 1988.
- [42] D.Dubois and H.Prade. An introduction to possibilistic and fuzzy logics. In *Non-Standard Logics for Approximative Reasoning (P. Smets, A. Mamdani, D. Dubois and H. Prade eds)*, pages 287–326. New York : Academic Press, 1988.
- [43] E. Hisdal. Conditional possibilities independence and non interaction. *Fuzzy Sets and Systems*, 1, 1978.
- [44] Allel Hadj Ali. *Raisonnement qualitatif sur les ordres de grandeur relatif : Formalisation et traitement en logique floue*. Thèse de doctorat d'état, Université Mouloud Mameri Tizi Ouzou, 2004.
- [45] J. Lang. Possibilistic logic : Complexity and algorithms. In *Handbook of Defeasible Reasoning and Uncertainty Management Systems*, 5 :179–220, 2000.

- [46] S.Benferhat. *Evolution des croyances et raisonnement plausible - Modèles qualitatifs*. Thèse d'habilitation à diriger les recherches, Université Paul Sabatier Toulouse, 2000.
- [47] D. Dubois, J. Lang, and H. Prade. Possibilistic logic. In *Handbook of Logic in Artificial Intelligence and Logic Programming*, (D. Gabbay et al., eds, 3, Oxford University Press :pages 439–513, 1994.
- [48] D. Dubois and H. Prade. Possibility theory as a basis for qualitative decision theory. In *14th International Joint Conference on Artificial Intelligence (IJCAI'95), MontréAlal*, pages 1924–1930, 1995.
- [49] S.Benferhat and F. Bannay-Dupin de Saint-Cyr. Contextual handling of conditional knowledge. In *5th Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, IPMU96*, pages 1369–1374, 1996.
- [50] M.Freund. Default extension : an alternative to the probabilistic approach. 1993.
- [51] D. Dubois and H. Prade. Epistemic entrenchment and possibilistic logic. *Artificial Intelligence*, 50 :223–239, 1991.
- [52] R.Carnap. Logical foundations of probability. In *Routledge and Kegan Paul, London*, 1950.
- [53] J.Y.Halpern. An analysis of first-order logics of probability. In *In proceeding of the 11th International Joint Conference on Artificial Intelligence (IJCAI-89)*, pages 1375–1381, August 20-25 1989.
- [54] L.Farinas del Cerro, A.Hezig, and J.Lang. Ordering-based non monotonic reasoning. *Artificial Intelligence*, (66) :375–393, 1994.
- [55] E.Rich. Default resoning as likelihood reasoning. In *American Association for Artificial Intelligence Conference*, pages 348–351. Washington, DC, 1983.
- [56] H.Farreny and H.Prade. Default and inexact reasoning with possibility degrees. In *IEEE Trans System Man Cyber 16*, pages 270–276, 1986.
- [57] H.Raiffa. *Decision Analysis*. Addison-Welsley Publishing Compagny Toronto, 1968.
- [58] R.A.Howard and J.E.Matheson. Influence diagramms. *The principles and applications of decision Analysis*, 2 :720–761, 1984.
- [59] R.D.Shachter. Evaluating influence diagramms. *Operations Research*, 34 :871–882, 1986.
- [60] R.D.Shachter. A graph-based inferences method for conditional independance. In *Seventh Conference on Uncertainty in Artificial Intelligence UAI91*, pages 353–360. Morgan Kaufmann Publishers, 1991.
- [61] P.P.Shenoy. A fusion algorithm for solving bayesian decision problem. In *Seventh Conference on Uncertainty in Artificial Intelligence UAI91*, pages 169–198. Morgan Kaufmann Publishers, 1991.
- [62] P.P.Shenoy. Using possibility theory in expert systems. *Fuzzy Sets and Systems*, 52 :129–142, 1992.

- [63] P.P.Shenoy. Valuation-based systems for bayesian decision analysis. *Operations research*, 40 :463–484, 1992.
- [64] P.P.Shenoy. A comparaison of graphical technical techniques for decision analysis. *European Journal of Operational Reasearch*, 78 :1–21, 1994.
- [65] P.P.Shenoy. Conditional independance in valuations-based systems. *International Journal of Approximate Reasoning*, 10 :203–234, 1994.
- [66] D. Dubois and H. Prade. Inference in probabilstic hypergraph. In *International Conference on Information Processing of Uncertainty in Knowledge Bases Systems (IPMU'90)*, pages 228–230. Paris France, 1990.
- [67] P.Fonck. *Réseaux d'inférence pour le raisonnement possibiliste*. PhD thesis, Université de Liège, Faculté des sciences, 1994.
- [68] J. Gebhardt C.Borgelt and R. Kruse. Inference methods. In *Handbook of Fuzzy Computation*, pages 1–13. Institute of Physics Publishing,Bristol,United Kingdom, 1998.
- [69] J. Gebhardt and R. Kruse. Background and perspectives of possibilistic graphical models. In *4th European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty (ECSQARU'97)*, LNAI 2143, pages 108–121, 1997.
- [70] J. Gebhardt and R. Kruse. Possinfer- a software tool for possibilistic inference. In *Fuzzy Set Methods in Information Engineering : A guide tour of Applications*. D.Dubois, H.Prade, R.Yager,J.Wiley and sons, New York, 1995.
- [71] J.Cano, M.Delgado, and S.Moral. An axiomatic framework for propagating uncertainty in directed acyclic networks. *International Journal of Approximate Reasoning*, 8 :253–380, 1993.
- [72] F. V. Jensen. *Bayesian networks and Decision graphs*. Springer-Verlag, New York, 2001.
- [73] Mc Carthy and P.J.Hayes.
- [74] V.Isham. An introduction to spatial point processes and markov random fields. *International Statistical Review*, 49 :21–43, 1981.
- [75] D.R.Cox and N.Wermuth. *Multivariate Dependencies-Models, Analysis and Interpretation*. Chapman Hall, London, 1996.
- [76] S.L.Lauritzen. *Graphical Models*. Clarendon, Oxford UK, 1996.
- [77] S.L.Lauritzen. *Lectures on contengency tables*. University of Aalborg Press,Aalborg, Denmark, 1982.
- [78] S.L.Lauritzen and N.Wermuth. Graphical and recursive models for contingency tables. *biometrika*, 70 :37–52, 1983.
- [79] J.B.Carlin, H.Kiiveri, and T.P.Speed. Recursive causal models. *Journal of the Australian Mathematical Society*, 36 :30–52, 1984.
- [80] T.Verma. Causal networks : Semantics and expressiveness. In Inc Elsevier Publishing Compagny, editor, *Fourth Conference on Uncertainty in Artificial Intelligence (UAI'88)*, pages 69–76, New York, 1988.

- [81] D.Geiger, J.Pearl, and T.Verma. D-separation : from theorems to algorithms. In Inc Elsevier Publishing Compagny, editor, *Fifth Conference on Uncertainty in Artificial Intelligence (UAI'89)*, pages 139–148, New York, 1989.
- [82] D.Geiger and A.Shaffer. *Fourth Conference on Uncertainty in Artificial Intelligence (UAI'88)*. Number 48. 1998.
- [83] D.Geiger, J.Pearl, and T.S.Verma. Identifying independence in bayesian networks. *Networks*, 20 :507–534, 1988.
- [84] J.Pearl and T.S.Verma. Causal networks : Semantics and expressiveness. In *Fourth Workshop on Uncertainty in Artificial Intelligence*, pages 352–359. Mountain viex,CA, 1988.
- [85] J. Kim and J. Pearl. A computational model for causal and diagnostic reasoning in inference systems. In *Proceedings of Seventeenth International Joint Conference on Artificial Intelligence (IJCAI'83)*, pages 190–193, 1983.
- [86] J. Pearl. Fusion propagation and structuring in belief networks. In *Artificial Intelligence*, pages 241–288, 1986.
- [87] G.F.Cooper. Computational complexity of probabilistic inference using bayesian belief networks. *Artificial Intelligence*, 20 :661–685, 1990.
- [88] E.Castillo, J.M.Gutiérrez, and A.S.Hadi. *Expert systems and probabilistic network models*. Springer, 1997.
- [89] M.I.Jordan, K.P.Murphy, and Y.Weiss. Loopy belief propagation for approximate inference : An empirical study. In *Fifteenth Conference on Uncertainty in Artificial Intelligence UAI99*, pages 467–475, 1999.
- [90] M.A.Peot and R.D.Shachter. Fusion and propagation with multiple observations in belief networks. *Artificial Intelligence*, 48 :299–318, 1991.
- [91] A.P.Dawid, S.L.Lauritzen, B.N.Larsen, and H.G.Leimer. Independence proprieties of directed markov fields. *Networks*, 20 :491–505, 1990.
- [92] M.Yannakakis. Computing the minimum fill-in is np-complete. *SIAM Journal on Algebrics and Discrete Methods*, 2 :77–79, 1981.
- [93] S.Arnborg, D.G.Corneil, and A.Prokurowski. Complexity of finding embeddings in a k-tree. *SIAM Journal on Algebrics and Discrete Methods*, 8 :277–284, 1987.
- [94] U.Kjærulff. Triangulation of graphs- algorithms giving small total state space. Technical report, Departement of Mathematics and computer Science, Aalborg University, Denmark, 1990.
- [95] C.Golumbic. *Triangulated graphs*, chapter Algorithms graph theory and perfect graphs, pages 98–100. Academic Press, New York University, 1980.
- [96] F.V.Jensen and F.Jensen. Optimal junction trees. In *Tenth Conference on Uncertainty in Artificial Intelligence*, pages 360–366. Seatle, Wash, 1994.
- [97] A.Darwiche and C.Huang. Inference in belief networks : A procedural guide. *International Journal of Approximate Reasoning*, 15 :225–263, 1996.

- [98] F.V.Jensen, S.L.Lauritzen, and K.G.Olesen. Bayesian updating in causal probabilistic networks by local computations. *Comput.Statist.Quart*, 4 :269–292, 1990.
- [99] C.Borgelt and R.Kruse. Some experimental results on learning probabilistic and possibilistic networks with different evaluation measures. In *European Conference of Symbolic and Quantitative Approaches to Reasoning and Uncertainty (ECSQARU97)*, pages 71–85. Bad Honnef (Germany), 1997.
- [100] S. Benferhat, D. Dubois, L. Garcia, and H. Prade. Directed possibilistic graphs and possibilistic logic. In *15th International Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU'98), Paris*, pages 1470–1477, 1998.
- [101] S. Benferhat, D. Dubois, L. Garcia, and H. Prade. Possibilistic logic bases and possibilistic graphs. In *15th Conference on Uncertainty in Artificial Intelligence (UAI'99), Stockholm*, pages 57–64, 1999.
- [102] S. Benferhat, D. Dubois, S. Kaci, and H. Prade. Bridging logical, comparative and graphical possibilistic representation frameworks. In *Proceedings of the 6th Eur. Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty (ECSQARU'01)*, pages 422–431, 2001.
- [103] S. Benferhat, D. Dubois, and H. Prade. Representing default rules in possibilistic logic. In *Proceedings of 3rd International Conference of Principles of Knowledge Representation and Reasoning (KR'92)*, pages 673–684, 1992.
- [104] C.Boutilier, T.Deans, and S.Hanks. Decision theoretic planning : Structural assumptions and computational leverage. *Journal of Artificial Intelligence Research*, 11 :1–94, 1999.
- [105] D. Poole. Logic, knowledge representation, and Bayesian decision theory. In *Computational Logic - CL 2000*, pages 70–86. LNAI 1861, Springer, 2000.
- [106] M.A. Williams. Iterated Theory Base Change :A Computational Model. In *Proceedings of 14th International Joint Conference on Artificial Intelligence (IJCAI'95)*, pages 1541–1547, 1995.
- [107] Adnan Darwiche and Moises Goldszmidt. On the relation between kappa calculus and probabilistic reasoning. In *The 10th Annual Conference on Uncertainty in Artificial Intelligence UAI 94*, San Francisco, CA, 1994. Morgan Kaufmann.
- [108] M. Goldszmidt and J. Pearl. Qualitative probabilities for default reasoning, belief revision, and causal modeling. *Artificial Intelligence*, 1996.
- [109] S.A.Cook. The complexity of theory -proving procedures. In *Third Annual ACM Symposium on Theory of Computing*, pages 151–158, ACM New york, 1971.
- [110] C. Cayrol and M. Lagasquie-Schiex. On the complexity of non-monotonic entailment in syntax-based approaches. In *ECAI94 Workshop on Algorithms, Complexity and Common-sens reasoning Amsterdam*, 1994.
- [111] A. Darwiche. Compiling knowledge into decomposable negation normal form. In *Proceedings of 16th International Joint Conference on Artificial Intelligence (IJCAI'99), Stockholm, Sweden*, pages 284–289, 1999.

- [112] A. Darwiche and P. Marquis. Compilation of Propositional Weighted Bases. In *International Workshop on Nonmonotonic Reasoning (NMR' 2002)*, pages 6–14, 2002.
- [113] P.P.Shenoy. A valuation-based language for expert systems. *International Journal of Approximate Reasoning*, 3 :383–411, 1989.
- [114] P.P.Shenoy. Valuations-based systems for propositional logic. *Methodologies for Intelligent Systems*, 5 :305–312, 1990.
- [115] P.P.Shenoy. Consistency in valuation-based systems. Working Paper 216, School of Business, 1990.
- [116] P.P.Shenoy. *Fuzzy logic for the management of uncertainty in expert systems*, chapter Valuation-based systems : A framework for managing uncertainty in expert systems. Jhon Wiley and Sons New York, 1991.
- [117] P.P.Shenoy. Using dempster-shafer's belief-functions theory in expert systems. Working paper 234, School of Business, University of Kansas, 1991.
- [118] I.Zeddigha. La decision possibiliste qualitative basée sur la technique de fusion syntaxique. Thèse de Magister USTHB - Faculté d'Electronique et d'Informatique - Département d'Informatique, 2006.

Résumé

Les travaux réalisés dans cette thèse se situent dans le cadre de la représentation et du raisonnement sur des connaissances incertaines en se basant sur une approche non-probabilisée : la théorie des possibilités quantitative. Deux formats de représentation ont été étudiés : un format logique basé sur les bases possibilistes quantitatives et un format graphique basé sur les réseaux de croyances possibilistes.

Sur le plan de la représentation, cette étude a contribué à développer des équivalences entre une base possibiliste quantitative et un réseau causal possibiliste basé sur le produit. De plus, elle montre que les réseaux de croyances basés sur les fonctions du conditionnement ordinal de Spohn ainsi que les bases de pénalités peuvent être naturellement représentés avec les bases de connaissances possibilistes quantitatives.

Sur le plan du raisonnement, il a été développé un algorithme original pour l'inférence à partir des bases de connaissances possibilistes quantitatives. Une étude expérimentale a été également effectuée confirmant ainsi l'intérêt de l'outil d'inférence basé sur la logique possibiliste quantitative.

Grâce aux transformations équivalentes développées, la théorie des possibilités dispose maintenant d'une implémentation alternative à celle des réseaux possibilistes quantitatifs.

Mots-clés : Intelligence artificielle, Théories de l'incertain, Théorie des possibilités, Réseaux de croyances possibilistes, Bases possibilistes quantitatives, Logique des pénalités, Réseaux causaux basés sur les OCF, Inférence en logique des possibilités quantitative.

Abstract

This work is intended to provide some contributions in the framework of the representation and the reasoning on uncertain knowledge based on a non probabilistic approach : the quantitative possibilistic theory. Two formats of representation were studied : a logical format based on quantitative possibilistic bases and a graphical format based on the beliefs possibilistic networks.

In the field of the representation, this study has contributed to develop equivalences between a quantitative possibilistic base and a possibilistic causal network based on the product. Moreover, it shows that the beliefs networks based on the ordinal conditioning functions of Spohn as well as the penalties bases can be naturally represented with the quantitative possibilistic bases.

In the field of the reasoning, an original algorithm was developed for the inference starting from the quantitative knowledge bases. An experimental study was also carried out thus confirming the advantage of inference tool based on possibilistic quantitative logic. Thanks to the equivalents transformations developed, now the possibility theory has an alternative implementation to that of the quantitative possibilistic networks.

Keywords : Artificial intelligence, Uncertain theories, Possibilistic theory, Possibilistic belief networks, Quantitative possibilistic bases, Penalty logic, OCF based causal networks, Quantitative possibilistic logic inference.