

MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR
ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITÉ DE LA MANOUBA
ÉCOLE NATIONALE DES SCIENCES DE L'INFORMATIQUE



THÈSE

Présentée en vue de l'obtention du diplôme de
DOCTEUR EN INFORMATIQUE
par

AMEN AJROUD

LES RÉSEAUX POSSIBILISTES GARANTIS ET ALGORITHMES DE PROPAGATION APPROXIMATIFS

Sous la direction de :

Prof. Habib YOUSSEF et Prof. Salem BENFERHAT

Réalisé au sein de l'unité de Recherche PRINCE (ISITCom - H. Sousse)



Soutenue publiquement le jeudi 24 juin 2010, devant le jury composé de :

| | | |
|-----------------------|------------------------|--|
| Président : | Faouzi GHORBEL | Professeur à l'ENSI |
| Rapporteurs : | Mohamed BEN AHMED | Professeur Émérite à l'ENSI |
| | Engelbert Mephu NGUIFO | Professeur à l'Université de Clermont Ferrand II |
| Directeurs de thèse : | Habib YOUSSEF | Professeur à l'ISITCom de Hammam Sousse |
| | Salem BENFERHAT | Professeur des Universités à l'Université d'Artois |

Résumé

Cette thèse a pour objectif le traitement de deux problèmes importants : le raisonnement sous incertitude avec des algorithmes de propagation d'évidence et la représentation des préférences avec les réseaux possibilistes garantis. Dans un premier temps, nous avons passé en revue les différentes théories de l'incertain qui ont été proposées et mises en place afin de représenter et manipuler les informations incertaines et imprécises. La résolution de problèmes et la prise de décision dans un système d'aide à la décision sont souvent entachées par cette incertitude, imprécision et ce manque des données à traiter. Pour une meilleure modélisation de ces données, nous avons montré que les modèles graphiques (probabilistes et possibilistes) apparaissent comme des outils efficaces pour une telle représentation. Ensuite, nous avons abordé le problème crucial de propagation (exacte et approximative) d'évidence dans les réseaux probabilistes et possibilistes. Dans le cas d'un réseau à connexions multiples, les algorithmes exacts de propagation souffrent de la transformation du graphe initial vers une structure secondaire ; la complexité de cette transformation appartient à la classe des problèmes NP-difficiles. Les algorithmes approximatifs de propagation apparaissent alors comme des alternatives aux algorithmes exacts. Dans cette optique, nos travaux se sont focalisés sur l'adaptation d'un algorithme approximatif de propagation au cadre possibiliste quantitatif et qualitatif. Les résultats expérimentaux de l'implémentation de cette adaptation ont montré les bonnes performances de l'algorithme adapté. Dans un second temps, nous nous sommes intéressés à l'utilisation de la mesure de possibilité garantie Δ pour définir un réseau possibiliste garanti basé sur cette mesure. La définition de ce réseau garanti engendre la définition d'autres aspects liés comme : le conditionnement possibiliste, la règle de chaînage et la distribution de possibilités garanties. Nous avons confirmé que les réseaux possibilistes garantis sont capables de modéliser et d'ordonner les préférences utilisateur.

Mots clés : Théorie des possibilités, réseaux possibilistes, algorithmes de propagation exacts et approximatifs, mesure de possibilité garantie, réseaux possibilistes garantis, représentation des préférences.

Remerciements

Je tiens à exprimer ici tous mes remerciements ainsi que toute ma gratitude à toutes les personnes sans lesquelles ce travail n'aurait pas été mené à terme. Sachant à l'avance que je ne saurai les remercier assez, j'aimerai tout de même leur signifier ma reconnaissance.

Mes remerciements vont tout d'abord à mes deux directeurs de thèse M. Habib Youssef, Professeur à l'ISITCom de Hammam Sousse et M. Salem Benferhat, Professeur des Universités à l'Université d'Artois. Je tiens à vous exprimer toute ma reconnaissance pour ces quelques années que j'ai passé à vos côtés. Au cours de ces années, votre grande disponibilité, votre rigueur scientifique, votre enthousiasme et vos précieux conseils m'ont permis de travailler dans les meilleures conditions. La confiance que vous m'avez accordée ainsi que nos diverses discussions m'ont permis de progresser et de mieux appréhender le métier d'enseignant-chercheur. Soyez assurés, Messieurs, de toute mon estime et de mon profond respect. Mes remerciements vont également à M. Mohamed Nazih Omri, Maître de Conférences à la Faculté des Sciences de Monastir et co-encadreur Tunisien de cette thèse, pour tous ses conseils et les efforts qu'il a consentis pour que les travaux de cette thèse avancent et soient menés dans les meilleures conditions.

Je suis très reconnaissant à M. Mohamed Ben Ahmed, Professeur Émérite à l'ENSI et à M. Engelbert Mephu Nguifo, Professeur à l'Université de Clermont Ferrand II, d'avoir accepté, tous les deux, de rapporter cette thèse. Les commentaires et les suggestions de ces personnalités scientifiques, tant sur la forme du mémoire que sur son fond, ont contribué de manière significative à son amélioration.

J'exprime toute ma gratitude à M. Faouzi Ghorbel, Professeur à l'ENSI, qui m'a fait l'honneur d'accepter de présider ce jury de thèse. Qu'il trouve ici l'expression de mon plus profond respect.

Je voudrais adresser un remerciement particulier à tous les membres de l'unité de recherche PRINCE pour l'ambiance de travail et pour leur collaboration. J'aimerai aussi remercier l'ensemble des membres et personnels du CRIL pour m'avoir accueilli 3 fois au cours de cette thèse.

Un grand merci à Jalel Boughizane, Lotfi Kedhiri, Hédi Magroun, Mourad Saadana et Khaled Boufares, mes collègues à l'ISET de Sousse qui m'ont soutenu et encouragé pendant les moments difficiles.

Je ne pourrai rater l'occasion qui m'est offerte ici pour dire un grand merci à mes parents, à mes sœurs et leurs époux, à ma belle-famille et à mes amis pour leur soutien.

Mes derniers remerciements vont à ceux qui m'ont soutenu de très près : ma femme Haifa et mes enfants Mohamed Ali, Melek et Maram.

Table des matières

| | |
|--|-----------|
| Résumé | 1 |
| Remerciements | 2 |
| Table des matières | 4 |
| Table des figures | 9 |
| Liste des tableaux | 11 |
| Introduction générale | 13 |
| 1 Théories de l'incertain | 20 |
| 1.1 Introduction | 20 |
| 1.2 La théorie des probabilités | 21 |
| 1.2.1 Espace d'état et événements | 21 |
| 1.2.2 Mesure de probabilités | 22 |
| 1.2.3 Probabilité conditionnelle et indépendance | 23 |
| 1.3 La théorie des ensembles flous | 25 |
| 1.3.1 Définition | 25 |
| 1.3.2 Propriétés | 25 |
| 1.3.3 Opérations sur les ensembles flous | 27 |
| 1.4 Les fonctions conditionnelles ordinales de Spohn | 28 |
| 1.5 Le Modèle des Croyances Transférables | 31 |
| 1.5.1 Représentation des connaissances | 31 |
| 1.5.2 Niveau pignistique | 32 |
| 1.6 La théorie des possibilités | 33 |
| 1.6.1 Distribution de possibilités | 33 |
| 1.6.2 Mesure de possibilité et de nécessité | 34 |
| 1.6.2.1 Mesure de possibilité Π | 35 |
| 1.6.2.2 Mesure de nécessité N | 35 |
| 1.6.3 Conditionnement possibiliste | 36 |
| 1.6.3.1 Conditionnement possibiliste quantitatif | 36 |
| 1.6.3.2 Conditionnement possibiliste qualitatif | 37 |
| 1.6.4 Indépendance possibiliste | 38 |

| | | |
|----------|--|-----------|
| 1.6.4.1 | Indépendance possibiliste quantitative | 39 |
| 1.6.4.2 | Indépendance possibiliste qualitative | 39 |
| 1.7 | Logique possibiliste | 40 |
| 1.8 | Conclusion | 41 |
| 2 | Les modèles graphiques probabilistes et possibilistes | 42 |
| 2.1 | Introduction | 42 |
| 2.2 | Les modèles graphiques | 43 |
| 2.2.1 | Exemples d'utilisation des modèles graphiques | 44 |
| 2.2.2 | Exemples de modèles graphiques | 48 |
| 2.2.2.1 | Les réseaux Bayésiens | 48 |
| 2.2.2.2 | Les réseaux possibilistes | 49 |
| 2.2.2.3 | Les Chaînes de Markov | 49 |
| 2.2.2.4 | Les modèles de Markov cachés | 51 |
| 2.3 | Les réseaux Bayésiens | 51 |
| 2.3.1 | Les relations de dépendance | 51 |
| 2.3.2 | Les réseaux Bayésiens probabilistes | 54 |
| 2.4 | Les réseaux de croyances possibilistes | 56 |
| 2.4.1 | Définition | 56 |
| 2.4.2 | Réseaux possibilistes quantitatifs | 57 |
| 2.4.3 | Réseaux possibilistes qualitatifs | 59 |
| 2.5 | Conclusion | 60 |
| 3 | Les algorithmes de propagation probabilistes et possibilistes | 61 |
| 3.1 | Introduction | 61 |
| 3.2 | Propagation dans les réseaux probabilistes | 62 |
| 3.2.1 | Les algorithmes exacts de propagation | 62 |
| 3.2.1.1 | Propagation dans les polyarbres | 62 |
| 3.2.1.2 | Propagation dans les DAGs à connexions multiples | 66 |
| 3.2.1.3 | Autres méthodes d'inférence exactes | 69 |
| 3.2.2 | Les algorithmes approximatifs de propagation | 70 |
| 3.2.2.1 | Les algorithmes d'échantillonnage par importance | 71 |
| 3.2.2.2 | Les méthodes de Monte-carlo par chaînes de Markov | 74 |
| 3.2.2.3 | L'algorithme <i>Loopy Belief Propagation</i> | 74 |
| 3.2.2.4 | Les méthodes de simplification de modèle | 75 |
| 3.2.2.5 | Autres méthodes d'inférence approchées | 76 |
| 3.3 | Propagation dans les réseaux possibilistes | 76 |
| 3.3.1 | Propagation exacte dans les réseaux possibilistes quantitatifs | 77 |

| | | |
|----------|---|------------|
| 3.3.1.1 | Adaptation de l'arbre de jonction aux réseaux quantitatifs | 77 |
| 3.3.1.2 | Algorithme de propagation dans les réseaux possibilistes hybrides | 85 |
| 3.3.2 | Algorithme de propagation <i>anytime</i> pour les réseaux qualitatifs | 87 |
| 3.4 | Conclusion | 89 |
| 4 | Un nouvel algorithme approximatif pour les réseaux possibilistes | 91 |
| 4.1 | Introduction | 91 |
| 4.2 | L'algorithme approximatif ' <i>Loopy Belief Propagation</i> ' | 92 |
| 4.3 | Propagation dans les réseaux quantitatifs simplement connectés . . . | 96 |
| 4.4 | Adaptation de l'algorithme <i>LBP</i> aux réseaux quantitatifs | 110 |
| 4.5 | L'algorithme <i>LBP</i> possibiliste appliqué sur des réseaux quantitatifs certains | 117 |
| 4.5.1 | Les réseaux certains | 118 |
| 4.5.2 | Application de <i>LBP</i> possibiliste sur des réseaux quantitatifs certains | 119 |
| 4.6 | Adaptation de l'algorithme <i>LBP</i> aux réseaux qualitatifs | 121 |
| 4.6.1 | Propagation dans les réseaux qualitatifs simplement connectés | 121 |
| 4.6.2 | <i>LBP</i> possibiliste pour les réseaux qualitatifs | 123 |
| 4.7 | Application de <i>LBP</i> possibiliste sur des réseaux qualitatifs certains . | 129 |
| 4.8 | Conclusion | 132 |
| 5 | Implémentation et expérimentations | 133 |
| 5.1 | Introduction | 133 |
| 5.2 | L'environnement expérimental | 133 |
| 5.3 | Les réseaux expérimentaux | 134 |
| 5.3.1 | Exemple 1 : réseau ALARM | 134 |
| 5.3.2 | Exemple 2 : réseau aléatoire | 135 |
| 5.3.3 | Exemple 3 : réseau QMR-DT | 135 |
| 5.4 | Expérimentations sur les réseaux quantitatifs | 137 |
| 5.4.1 | Évaluation des performances de l'algorithme <i>LBP</i> possibiliste . | 137 |
| 5.4.2 | Étude de la convergence | 138 |
| 5.4.2.1 | Exemple 1 : réseau ALARM | 139 |
| 5.4.2.2 | Exemple 2 : réseau aléatoire | 139 |
| 5.4.2.3 | Exemple 3 : réseau QMR-DT | 139 |
| 5.4.3 | Étude de l'exactitude | 141 |
| 5.4.3.1 | Exemple 1 : réseau ALARM | 141 |
| 5.4.3.2 | Exemple 2 : réseau aléatoire | 141 |

| | | |
|----------|---|------------|
| 5.4.3.3 | Exemple 3 : réseau QMR-DT | 143 |
| 5.4.4 | Détermination du score MAR | 143 |
| 5.4.5 | Expérimentations sur les réseaux quantitatifs certains | 144 |
| 5.5 | Expérimentations sur les réseaux qualitatifs | 145 |
| 5.5.1 | Étude de la convergence | 146 |
| 5.5.1.1 | Exemple 1 : réseau ALARM | 146 |
| 5.5.1.2 | Exemple 2 : réseau aléatoire | 146 |
| 5.5.1.3 | Exemple 3 : réseau QMR-DT | 147 |
| 5.5.2 | Étude de l'exactitude | 147 |
| 5.5.2.1 | Exemple 1 : réseau ALARM | 148 |
| 5.5.2.2 | Exemple 2 : réseau aléatoire | 149 |
| 5.5.2.3 | Exemple 3 : réseau QMR-DT | 149 |
| 5.5.3 | Détermination du score MAR | 150 |
| 5.5.4 | Expérimentations sur les réseaux qualitatifs certains | 150 |
| 5.6 | Conclusion | 152 |
| 6 | Un nouveau format de représentation des préférences : les réseaux possibilistes garantis | 154 |
| 6.1 | Introduction | 154 |
| 6.2 | Notion de préférences | 155 |
| 6.3 | Les CP-Nets | 156 |
| 6.4 | La logique QCL | 158 |
| 6.5 | Révisions et modifications de la logique QCL | 159 |
| 6.5.1 | Les variantes de la logique QCL | 159 |
| 6.5.2 | La Logique du Choix Qualitatif Minimale $MQCL$ | 160 |
| 6.5.3 | La Logique du Choix Qualitatif Prioritaire $PQCL$ | 161 |
| 6.5.4 | La Logique du Choix Qualitatif Positive $QCL+$ | 162 |
| 6.6 | Mesure de possibilité garantie et représentation logique | 163 |
| 6.6.1 | Mesure de possibilité garantie Δ | 163 |
| 6.6.2 | Bases de connaissances possibilistes garanties | 164 |
| 6.7 | Représentation logique des préférences | 166 |
| 6.8 | Réseaux possibilistes garantis | 169 |
| 6.8.1 | Conditionnement possibiliste basé sur la mesure Δ | 169 |
| 6.8.2 | Réseaux possibilistes garantis | 170 |
| 6.9 | Réseaux possibilistes garantis et préférences | 173 |
| 6.9.1 | Préférences basiques inconditionnelles : | 173 |
| 6.9.2 | Préférences conditionnelles | 176 |
| 6.9.3 | Un exemple illustratif | 178 |
| 6.10 | Représentation logique des réseaux garantis Δ | 180 |

| | |
|---|------------|
| 6.11 Équivalence syntactique entre réseaux possibilistes garantis et réseaux qualitatifs | 182 |
| 6.12 Conclusion | 184 |
| Conclusion générale | 186 |
| Bibliographie | 189 |
| Annexe A | 206 |
| Annexe B | 211 |

Table des figures

| | | |
|-----|---|-----|
| 1.1 | Représentation de la fonction d'appartenance μ_B en logique floue . . . | 26 |
| 1.2 | Représentation de la température de l'eau par la logique classique . . . | 28 |
| 1.3 | Représentation de la température de l'eau par la logique floue | 29 |
| 2.1 | Exemple de modèle graphique représentant des relations d'influences . | 44 |
| 2.2 | Exemple d'une chaîne de Markov à trois états | 50 |
| 2.3 | Connexion en série | 52 |
| 2.4 | Connexion divergente | 52 |
| 2.5 | Connexion convergente | 53 |
| 2.6 | Exemple d'un réseau bayésien | 55 |
| 2.7 | Exemple de réseau possibiliste Π_G | 57 |
| 3.1 | Partition de l'ensemble E en deux sous-ensembles E_i^+ et E_i^- | 64 |
| 3.2 | Messages échangés dans le polyarbre lors de la propagation d'évidence | 66 |
| 3.3 | Exemple d'un arbre de jonction formé par trois cliques | 67 |
| 3.4 | Etape de moralisation d'un graphe | 67 |
| 3.5 | Etape de triangulation d'un graphe | 68 |
| 3.6 | Etape de construction de l'arbre de jonction | 68 |
| 3.7 | Exemple de réseau possibiliste quantitatif à connexions multiples . . . | 77 |
| 3.8 | Étapes de transformation du réseau ΠG_p en un arbre de jonction . . . | 79 |
| 3.9 | L'ordre des messages échangés pendant la phase de propagation globale | 82 |
| 4.1 | Réseau probabiliste à connexions multiples | 94 |
| 4.2 | Messages sortants de chaque nœud pendant l'itération $t=0$ et $t=1$. . | 94 |
| 4.3 | Partition de l'ensemble E en deux sous-ensembles E_i^+ et E_i^- | 97 |
| 4.4 | Les différents messages reçus et envoyés par le nœud A_i | 100 |
| 4.5 | Réseau possibiliste quantitatif en polyarbre | 102 |
| 4.6 | Messages envoyés lors de l'étape de collecte de l'évidence | 106 |
| 4.7 | Messages envoyés lors de l'étape de distribution de l'évidence | 109 |
| 4.8 | Messages entrants au nœud A_i à l'itération t | 111 |
| 4.9 | Messages sortants du nœud A_i pour l'itération $t + 1$ | 111 |

| | | |
|------|--|-----|
| 4.10 | Messages sortants de chaque nœud à la fin de l'itération $t=0$ | 115 |
| 4.11 | Messages sortants de chaque nœud à la fin de l'itération $t=1$ | 117 |
| 5.1 | Structure du réseau ALARM | 136 |
| 5.2 | Structure d'un réseau aléatoire | 137 |
| 5.3 | Structure d'un réseau QMR-DT | 137 |
| 5.4 | Nombre d'itérations effectuées pour la convergence dans le cas d'un réseau aléatoire quantitatif | 140 |
| 5.5 | Nombre d'itérations effectuées pour la convergence dans le cas du réseau QMR-DT | 140 |
| 5.6 | Mesure de l'exactitude des valeurs issues de l'arbre de jonction exacte et de LBP possibiliste approximatif pour le cas du réseau ALARM quantitatif | 142 |
| 5.7 | Mesure de l'exactitude des valeurs issues de l'arbre de jonction exacte et de LBP possibiliste approximatif pour le cas des réseaux aléatoires à 30, 40 et 50 nœuds | 142 |
| 5.8 | Mesure de l'exactitude des valeurs issues de l'arbre de jonction exact et de LBP possibiliste approximatif pour le cas de réseaux QMR-DT quantitatif formés par 30, 40 et 50 nœuds | 143 |
| 5.9 | Nombre d'itérations effectuées pour la convergence dans le cas d'un réseau aléatoire qualitatif | 147 |
| 5.10 | Nombre d'itérations effectuées pour la convergence dans le cas de réseau QMR-DT qualitatif | 148 |
| 5.11 | Mesure de l'exactitude des valeurs issues de l'arbre de jonction exacte et de LBP possibiliste approximatif pour le cas du réseau ALARM qualitatif | 148 |
| 5.12 | Mesure de l'exactitude des valeurs issues de l'arbre de jonction exacte et de LBP possibiliste approximatif pour le cas des réseaux aléatoires à 30, 40 et 50 nœuds | 149 |
| 5.13 | Mesure de l'exactitude des valeurs issues de l'arbre de jonction exact et de LBP possibiliste approximatif pour le cas de réseaux QMR-DT qualitatif formés par 30, 40 et 50 nœuds | 150 |
| 6.1 | Exemple d'un CP-Net | 158 |
| 6.2 | Exemple d'un réseau possibiliste garanti \mathcal{T} basé sur Δ | 171 |
| 6.3 | Réseau possibiliste garanti | 177 |
| 6.4 | Exemple de réseau possibiliste garanti représentant des préférences . . | 179 |

Liste des tableaux

| | | |
|-----|--|-----|
| 1.1 | Correspondance entre théorie des ensembles et Probabilités | 22 |
| 1.2 | Probabilités jointes sur A et B | 24 |
| 1.3 | Probabilités marginales sur A et B | 25 |
| 1.4 | Distribution κ associée à Ω | 30 |
| 1.5 | Exemple de fonction de masse $m(A)$, de croyance $bel(A)$ et combinaison conjonctive entre m et m_s | 33 |
| 1.6 | Distribution de possibilités jointe sur A et B | 36 |
| 1.7 | Distribution de possibilités sur A et B après conditionnement quantitatif par $B = \neg b$ | 37 |
| 1.8 | Distribution de possibilités sur A et B après conditionnement qualitatif par $B = \neg b$ | 38 |
| 2.1 | Distributions de probabilités conditionnelles initiales | 55 |
| 2.2 | Distributions de possibilités conditionnelles associées à Π_G | 57 |
| 2.3 | Distribution de possibilités jointe associée à Π_{G_p} | 58 |
| 2.4 | Distribution de possibilités jointe associée à Π_{G_m} | 59 |
| 3.1 | Distributions initiales de possibilités conditionnelles associées à Π_{G_p} . | 78 |
| 3.2 | Distributions initiales de possibilités conditionnelles associées à Π_{G_p} . | 78 |
| 3.3 | Distribution de possibilités jointe $\pi_{\Pi_{G_p}}$ | 79 |
| 3.4 | Distributions initiales de possibilités associées aux cliques | 80 |
| 3.5 | Distributions de possibilités associées aux cliques après propagation . | 84 |
| 3.6 | Distributions de possibilités associées aux séparateurs après propagation | 84 |
| 4.1 | Distributions de probabilités conditionnelles initiales | 95 |
| 4.2 | Distributions de possibilités conditionnelles initiales associées à Π_{G_p} . | 104 |
| 4.3 | Distributions de possibilités conditionnelles initiales | 113 |
| 4.4 | Distributions de possibilités conditionnelles initiales certaines | 118 |
| 5.1 | Nombre d'itérations nécessaire pour atteindre la convergence sur un réseau ALARM quantitatif | 139 |

| | | |
|------|--|-----|
| 5.2 | Répartition du score MAR sur différents intervalles de valeurs pour 100 exécutions de LBP possibiliste sur des réseaux aléatoires et de type QMR | 144 |
| 5.3 | Expérimentations de LBP possibiliste sur les réseaux quantitatifs certains aléatoires | 145 |
| 5.4 | Expérimentations de LBP possibiliste sur les réseaux quantitatifs certains de type QMR-DT | 145 |
| 5.5 | Nombre d'itérations nécessaire pour atteindre la convergence sur un réseau ALARM qualitatif | 146 |
| 5.6 | Répartition du score MAR sur différents intervalles de valeurs pour 100 exécutions de LBP possibiliste sur des réseaux aléatoires et de type QMR | 151 |
| 5.7 | Expérimentations de LBP possibiliste sur les réseaux qualitatifs certains aléatoires | 151 |
| 5.8 | Expérimentations de LBP possibiliste sur les réseaux qualitatifs certains de type QMR-DT | 152 |
| 6.1 | Distribution de possibilités $\pi_{\mathfrak{R}}$ associée à la base \mathfrak{R} | 165 |
| 6.2 | Distribution de possibilités π_R associée à l'ensemble des rejets R . . . | 167 |
| 6.3 | Distribution de possibilités π_B associée à l'ensemble des buts B . . . | 168 |
| 6.4 | Distributions de possibilités conditionnelles $\pi_m(\omega \varphi)$ et $\Delta(\omega \varphi)$. . . | 170 |
| 6.5 | Distributions initiales de possibilités conditionnelles associées à \mathcal{T} . . | 172 |
| 6.6 | Distributions initiales de possibilités conditionnelles associées à \mathcal{T} . . | 172 |
| 6.7 | Distribution de possibilités jointe de $\pi_{\mathcal{T}}$ | 174 |
| 6.8 | Distribution initiale de possibilités garantie associée à \mathcal{T} | 175 |
| 6.9 | Distributions initiales conditionnelles de possibilités garanties | 177 |
| 6.10 | Distributions de possibilités garantis conditionnelles initiales | 180 |
| 6.11 | Distributions initiales de possibilités conditionnelles associées à ΠG_m | 184 |
| 6.12 | Distributions initiales de possibilités conditionnelles associées à ΠG_m | 184 |
| 6.13 | Les degrés de possibilités a posteriori des variables non observés . . . | 185 |
| 14 | Distributions de possibilités associées aux bases \mathcal{B}_1 , \mathcal{B}_2 et \mathcal{B}_3 | 207 |
| 15 | Distribution de possibilités π associée à Ω | 207 |
| 16 | Distribution de possibilités $\pi_{\mathcal{B}}$ obtenue à partir de la base \mathcal{B} | 208 |
| 17 | Distributions de possibilités π_1 et π_2 associées à \mathcal{B}_1 et \mathcal{B}_2 | 209 |
| 18 | Distribution de possibilités résultante $\pi^* = \pi_1 * \pi_2$ associée à \mathcal{B}^* . . . | 210 |

Introduction générale

Depuis plusieurs années, les travaux en Intelligence Artificielle (IA), dont l'objectif général était d'essayer d'imiter l'intelligence humaine, ont permis de constater que la modélisation du raisonnement humain est une tâche ardue et très complexe. La logique classique a depuis longtemps été considérée comme un outil privilégié des chercheurs en Intelligence Artificielle pour la formalisation et l'exploitation des connaissances. Toutefois, elle reste inadéquate pour la représentation et le raisonnement avec des connaissances imparfaites, comme par exemple des connaissances incertaines, imprécises ou encore incomplètes.

Les systèmes d'aide à la décision représentent une des applications les plus répandues dans le domaine de l'Intelligence Artificielle. Ces systèmes sont mis en œuvre dans divers domaines où la prise de décision est vitale tels que la médecine, la gestion des entreprises, l'économie, l'aéronautique... Différents formalismes existent pour exprimer les problèmes de décision. Malheureusement, certains facteurs entachent ces problèmes : il s'agit de l'incertitude et de l'imprécision qui affectent aussi bien les connaissances brutes que la méthode de résolution (de raisonnement). En effet, la description du monde réel par des connaissances est souvent altérée par l'incertitude. Par ailleurs, toute résolution automatisée de problèmes, "aussi fidèle qu'elle soit" au comportement humain, est ternie d'imprécision.

Plusieurs théories (de l'incertain) ont été proposées et mises en place afin de représenter et manipuler les informations incertaines et imprécises. Certaines (en plus de la théorie standard des probabilités [She97, Dur96]) se basent sur la logique non-monotone comme la théorie des ensembles flous [Zad65], les fonctions conditionnelles ordinales de Spohn [Spo88b, Spo88a], le Modèle des Croyances Transférables [SK94] et la théorie des possibilités [DP04b, DP98, DP88].

Pour que le contenu de la base de connaissances et la base de règles d'un système d'aide à la décision soit compréhensible aussi bien par l'expert que par la machine, les modèles graphiques apparaissent comme des outils efficaces pour une telle modélisation.

Les modèles graphiques [Pea88, Lau96, Bis06, CGH96b], issus d'une combinaison de la théorie de graphes [CGH96b, Bol02] et de la théorie des probabilités [She97, Dur96], permettent l'acquisition, la représentation et la manipulation des connaissances incertaines, imprécises et incomplètes dans le contexte d'Intelligence Artificielle. Ces modèles peuvent représenter intuitivement un domaine de connaissance : il est souvent plus clair pour un expert de formaliser les connaissances d'un système d'aide à la décision sous la forme d'un graphe plutôt que sous la forme d'une base de règles.

Le succès de ces modèles graphiques est dû à leurs caractéristiques intrinsèques. Premièrement, ces modèles graphiques représentent visuellement le contenu du système à modéliser, ce qui offre aux différents experts du domaine une lisibilité de la structure du système. Deuxièmement, les modèles graphiques expriment naturellement les relations de dépendance et/ou d'indépendance, à travers des «liens» entre les variables constituant le système à modéliser. Troisièmement, ces modèles sont modulaires c'est-à-dire que certains problèmes et systèmes complexes peuvent être décrits et manipulés en les divisant en composants élémentaires.

Il existe deux familles de modèles graphiques : les graphes orientés et les graphes non orientés. Dans cette thèse, nous nous intéressons aux modèles graphiques orientés. Parmi les modèles graphiques les plus connus, citons les réseaux bayésiens probabilistes [Pea88, LS90, Jen01, NWL⁺07], les réseaux possibilistes [GK97, Fon94, BDGP99, BGK98], les diagrammes d'influence [HJ84, Pea88, Sha86b, Sha88], les modèles de Markov cachés [CMT05, Rab90, EM02] et les arbres de décision [ZR00].

La plupart des modèles existants s'appuient sur la théorie des probabilités pour quantifier l'incertitude au niveau des connaissances traitées. En particulier, pour les réseaux bayésiens, des applications réelles, de plus en plus nombreuses, mettent en œuvre certaines techniques (inférence, apprentissage) liés à ces réseaux. Les systèmes de détection de fraude dans le domaine des télécommunications [ES95], les modèles de détection d'intrusions [JZA07, BT08a, BT08b, JZA09] et les systèmes d'aide à la décision en situation critique [HB95] sont des exemples de systèmes fondés sur les réseaux bayésiens.

Malgré le succès de ces applications (et bien d'autres) qui prouvent à quel point les réseaux bayésiens sont des outils privilégiés pour les problèmes de diagnostic et prise de décision, la théorie des probabilités n'offre pas un cadre toujours adéquat pour le raisonnement en présence de données numériques incomplètes. En effet, une distribution de probabilités permet uniquement de représenter la notion d'indécision mais reste inadéquate lorsqu'il s'agit de représenter la notion d'ignorance totale. Effectivement, étant donné un événement quelconque, l'interprétation de la valeur

du degré de probabilité de cet événement n'arrive pas à distinguer notre indifférence de l'ignorance totale concernant la réalisation de cet événement.

Issue de la théorie des ensembles flous, la théorie des possibilités [Zad78, DP04b, Kli99, DP98] permet d'évaluer l'incertitude sur l'intervalle unitaire à l'aide du degré de possibilité et de son dual, le degré de nécessité (ou de certitude) [Fab96]. Il existe deux cadres de raisonnement : quantitatif et qualitatif [DP98], selon la nature des connaissances initiales manipulées. Dans le cadre quantitatif, les mesures de possibilité prennent des valeurs significatives et précises dans l'intervalle unitaire ; ces valeurs sont considérées comme une borne supérieure des degrés de probabilité. Tandis que dans le cadre qualitatif, ces valeurs établissent un pré-ordre total sur l'ensemble des interprétations du monde à modéliser et permettant ainsi d'exprimer les préférences quant aux différentes interprétations de cet ensemble.

L'arrivée d'une nouvelle observation permet de changer les connaissances d'un expert qui modifie (révise), en conséquence, la mesure de l'incertitude attribuée à chaque interprétation : c'est le conditionnement. Plusieurs définitions du conditionnement ont été proposées dans la littérature [BMCM02, DCK94, Fon97, His78, DP88, Van03]. Dans le cadre possibiliste, le conditionnement diffère selon la nature de l'information disponible : quantitative ou qualitative. La logique possibiliste [BP06, DP04b, BDGP02] est une extension de la logique classique qui permet de traiter des informations totalement ordonnées. Les priorités entre ces informations sont représentées à l'aide de formules pondérées, où chaque poids représente la borne inférieure d'une mesure de nécessité.

Les contreparties des réseaux bayésiens probabilistes en théorie des possibilités sont les réseaux possibilistes [GK97, Fon94, BDGP99, BGK98]. Ces derniers représentent et raisonnent avec des informations incertaines et imprécises et même manquantes et s'adaptent aux deux cadres de raisonnement qualitatif et quantitatif. Comme exemple d'application récente basée sur les réseaux possibilistes, citons le système de recherche d'information des documents Web : SARIPOD [EEZA07, EEZA09].

La plupart des modèles probabilistes et possibilistes ont été développés pour résoudre certains problèmes de prédiction et rétrodiction en Intelligence Artificielle. Ceci est rendu possible grâce à un mécanisme très puissant, qui est l'inférence.

D'une manière générale, l'inférence probabiliste [Sha88, Mac03, DLN⁺90] consiste à propager une ou plusieurs informations certaines (appelées "évidence") dans un modèle graphique afin de réviser les croyances concernant l'incertitude exprimée dans chaque variable du modèle. Dans un cadre probabiliste, l'inférence dans les réseaux

admettant des connexions multiples est un problème de décision, de la forme "Est-ce que $P(A = a) = x$?" qui appartient à la classe des problèmes NP-difficile [Coo90]. Ce problème de propagation devient polynomial dans les réseaux à connexions simples. Le mécanisme d'inférence cherche à conclure dans divers sens : il permet de prédire des effets (conséquences) sachant que certaines causes (conditions) ont été observées. Inversement, il permet à partir des faits observés de remonter jusqu'aux causes supposées. Deux classes principales d'algorithmes d'inférence ont été développées [GH02] : l'approche exacte et l'approche approximative.

1. Les algorithmes exacts d'inférence [LS90, Jen96, Pea88, CGH96b] exploitent les indépendances conditionnelles contenues dans les réseaux et donnent les valeurs exactes des probabilités a posteriori, à chaque fois que l'inférence est mise en œuvre. Selon la topologie des réseaux (à connexions simples ou multiples), une technique de propagation de messages est appliquée.

Dans le cas d'un réseau à connexions simples (ou polyarbre), cette technique associe à chaque nœud un mécanisme qui lui permet d'envoyer des messages de façon asynchrone à ses voisins jusqu'à ce qu'un équilibre soit atteint, en un nombre fini d'étapes. Cette méthode a été depuis étendue à n'importe quel topologie de réseaux (à connexions multiples) pour donner l'algorithme de l'Arbre de jonction [LS90, Jen96]. Cet algorithme transforme, d'abord, le réseau initial en supprimant les boucles afin d'obtenir une structure secondaire appelée arbre de jonction et formée par des cliques de nœuds. La construction de l'arbre de jonction est un problème NP-difficile. Ensuite, il procède à un échange de messages entre les cliques, identique à celui de propagation des messages dans les polyarbres. Le coût de l'algorithme arbre de jonction est déterminé par la taille de la plus grande clique ainsi formée et est exponentiel en espace par rapport à la taille de cette plus grande clique. Malheureusement, dans certains cas de graphes à connexions multiples, l'inférence exacte peut s'avérer impraticable [BA02]. Ceci est du essentiellement à la grande taille des cliques qui empêche souvent l'algorithme arbre de jonction de s'initialiser.

2. L'inférence approximative [GG84, GRS96, MWJ99, Bea03] est une alternative de l'approche exacte lorsque celle-ci devient inapplicable. Les algorithmes approximatifs d'inférence fournissent des valeurs approchées des probabilités calculées. Plusieurs méthodes approximatives d'inférence ont été développées dans le cadre probabiliste et sont regroupées en deux familles : celles qui appliquent les algorithmes exacts sur des topologies approchées (ou sur une partie) du réseau original, et celles qui appliquent des algorithmes approximatifs directement sur le réseau original. Dans le cadre de cette thèse, nous nous

intéressons à cette dernière famille des algorithmes approximatifs. Comme exemples de ces algorithmes nous citons : les méthodes de *Markov chain Monte Carlo* [GG84, GRS96] qui utilisent de méthodes d'échantillonnage à partir de distributions de probabilités et la méthode *Loopy Belief Propagation* (LPB) [MWJ99, Wei00] qui itère l'algorithme de propagation de messages sur un graphe à connexions multiples. Bien qu'il soit approximatif, plusieurs applications réelles ont prouvé les performances de l'algorithme LBP dans la résolution des problèmes d'inférence et d'optimisation [TIT03, BGT93]. Ceci est dû aux valeurs de probabilités fournies par l'algorithme LBP qui se rapprochent des valeurs exactes.

La plupart des algorithmes de propagation possibilistes exacts sont des adaptations directes des algorithmes probabilistes tenant compte du cadre de raisonnement qualitatif ou quantitatif. Essentiellement, ce sont des adaptations des algorithmes exacts *propagation de messages* et *arbre de jonction*. Tout comme dans le cadre probabiliste, le problème d'inférence dans le cadre possibiliste est un problème de décision "Est-ce que $\Pi(A = a) = x$?" NP-difficile [Fon94]. À notre connaissance, aucun algorithme approximatif d'inférence probabiliste n'a été adapté aux réseaux possibilistes quantitatifs.

Le premier objectif principal de cette thèse sera d'adapter l'algorithme Loopy Belief Propagation probabiliste aux réseaux possibilistes quantitatifs puis aux réseaux qualitatifs. Essentiellement, l'algorithme adapté propage l'évidence dans le réseau possibiliste par un échange itératif de messages entre les nœuds voisins jusqu'à atteindre un état de convergence. En effet, pendant chaque itération, chacun des nœuds du réseau reçoit des messages entrants provenant de tous ses voisins (parents et fils). Ces messages serviront d'une part, à évaluer la condition de convergence qui exprime l'état d'équilibre dans le réseau possibiliste et d'autre part, à calculer, éventuellement, l'expression des messages sortants à envoyer aux parents et fils de chaque nœud si l'état de convergence n'est pas atteint. L'implémentation et l'expérimentation de cet algorithme adapté sur différentes structures de réseaux possibilistes et la comparaison des valeurs approchées calculées avec les valeurs exactes ont montré des résultats satisfaisants.

En théorie des possibilités, il existe deux grandeurs standards (le degré de possibilité Π et le degré de nécessité N) permettant de mesurer l'incertitude associée à une information exprimée sous la forme d'une formule logique. Il existe aussi une autre mesure proposée récemment [DHP00, BK03], permettant l'évaluation de l'incertitude au niveau d'une information, basée sur la notion de mesure de possibilité garantie notée par Δ . Cela signifie que, pour une formule logique pondérée (φ, a) ,

le degré de possibilité garanti de φ est au moins égal au poids a ; c-à-d $\Delta(\varphi) \geq a$. Intuitivement, si la formule φ exprime un objectif visé (ou une préférence souhaitée) par un agent, alors $\Delta(\varphi) \geq a$ signifie que toute solution qui satisfait l'objectif (ou la préférence) φ lui est satisfaisante avec un degré au moins égal à a .

La notion de préférence est sollicitée lorsqu'un agent se trouve confronté à un ensemble d'alternatives à partir desquelles il doit faire un choix. Pour que le choix effectué réponde au mieux à la préférence souhaitée, on doit procéder à une évaluation de ces alternatives en tenant compte des informations dont l'agent dispose. Cette évaluation cherche à comparer, en termes de préférence, les différentes alternatives pour la prise de décision.

Dans la littérature, Il existe deux familles pour la modélisation des préférences [Sed08] :

- La première famille est dite représentation quantitative ou numérique des préférences où à chaque alternative est affectée une ou plusieurs valeurs numériques. Cette représentation se base aussi sur la notion de fonction d'utilité.
- La deuxième famille est appelée représentation qualitative ou logique des préférences dans laquelle une relation binaire R est utilisée pour comparer les alternatives.

Dans cette thèse, nous nous restreindrons à la deuxième famille de modélisation ; nous nous intéresserons, plus précisément, à la logique pondérée qui est une représentation des préférences basée sur la théorie des possibilités.

Le deuxième objectif principal de cette thèse consiste à définir un nouveau format de représentation des préférences par le biais des réseaux possibilistes garantis basé sur la mesure Δ . La structure graphique de ce réseau est identique à celle des réseaux possibilistes standards, c'est-à-dire un DAG. Cependant, la composante numérique de ce réseau est un degré de possibilité garanti α qui est associé à chaque état d'un nœud et signifiant que cet état est garanti au minimum avec un degré de possibilité égal à α . La définition de ce réseau garanti engendre la définition d'autres aspects liés : le conditionnement possibiliste, la règle de chaînage et la distribution de possibilités garantie.

Cette thèse est organisée en six chapitres :

Le premier chapitre propose un état de l'art sur les théories de l'incertain permettant de représenter et manipuler l'incertitude. Nous insisterons, en particulier, sur deux théories : la théorie des probabilités et la théorie des possibilités.

Le second chapitre expose les modèles graphiques probabilistes et possibilistes

les plus connus. Nous présenterons leurs caractéristiques, leurs utilisations et leurs possibilités de traiter les informations incertaines.

Le troisième chapitre aborde, en détails, le mécanisme d'inférence dans les réseaux probabilistes et possibilistes. Selon la topologie des composantes graphiques des ces réseaux, nous exposerons les algorithmes exacts et approximatifs de propagation dans les réseaux bayésiens probabilistes et les réseaux possibilistes qualitatifs et quantitatifs.

Le quatrième chapitre présente notre première contribution à savoir l'algorithme *Loopy Belief Propagation* probabiliste adapté aux réseaux possibilistes quantitatifs puis à ceux qualitatifs. Nous argumentons notre démarche théorique par des exemples détaillés.

Le cinquième chapitre traite l'implémentation et les résultats expérimentaux de l'algorithme LBP possibiliste présenté dans le quatrième chapitre. Nous expérimentons, en particulier, les aspects de convergence et d'exactitude pour l'algorithme adapté sur trois structures de réseaux quantitatifs différentes.

Le sixième chapitre propose notre deuxième contribution dans laquelle nous définissons un réseau possibiliste *garanti*. Nous montrons qu'il est possible de représenter logiquement ce réseau garanti à travers une base possibiliste Δ . Une équivalence syntactique est aussi établie entre les réseaux possibilistes garantis et les réseaux possibilistes standards afin de ré-utiliser les algorithmes de propagation existants.

Nous terminons cette thèse en présentant les conclusions de nos travaux et nos perspectives.

Chapitre 1

Théories de l'incertain

1.1 Introduction

Depuis toujours, les êtres humains ont été fascinés par le hasard tel qu'il se manifeste dans un lancer de dés ou dans les prévisions météo : un désordre indéfinissable qui échappe à leurs connaissances. Ainsi, le hasard représente tout simplement notre incompetence à cerner tous les facteurs naturels qui contribuent à l'état de notre environnement, dans un lieu et à un instant donnés. Deux incertitudes sont rattachées à la notion du hasard : l'incertitude sur les causes possibles d'un événement hasardeux (s'il y en a), et sur les conséquences observables d'un tel événement (s'il y en a).

La théorie des probabilités [She97, Dur96] est apparue plus formellement au dix-septième siècle quand deux grands mathématiciens français, Blaise Pascal et Pierre de Fermat, ont tenté de résoudre deux problèmes liés aux jeux de hasard. La résolution de tels problèmes a incité d'autres chercheurs comme Huygens, Bernoulli, De Moivre et Bayes à mettre en place un fondement mathématique pour la théorie des probabilités et le calcul différentiel. Au vingtième siècle, on a assisté aux développements des théories des probabilités subjectives (dans les années 50) puis de l'évidence (dans les années 60). Cette théorie trouve aujourd'hui, une grande place dans divers domaines : les mathématiques, la physique, l'économie, l'informatique, la médecine, les prévisions météorologiques.

La logique classique permet de coder les issues d'un raisonnement à l'aide de système n'ayant que deux états *vrai* ou *faux*. À l'inverse de la logique classique, la logique floue [Zad65, BG02, H95, Nov06, Nov09] permet à une condition d'être dans un état autre que vrai ou faux : il existe des degrés dans la vérification de cette condition. La fonction d'appartenance, caractéristique d'un sous-ensemble flou, permet de quantifier les degrés d'appartenance d'un élément à un domaine défini, sur l'intervalle $[0, 1]$. Les fonctions ordinales conditionnelles de Spohn [Spo88b], appe-

lées fonctions "kappa", traitent l'incertitude en associant des entiers ordinaux aux interprétations. Ainsi, toutes les interprétations sont considérées, dans une certaine mesure, possibles. Issue de la théorie des ensembles flous, la théorie des possibilités [Zad78, DP04b, Kli99, DP98] est vue comme un modèle simple et naturel pour le traitement des informations incertaines. À travers ses mesures d'incertitude (degré de possibilité et degré de nécessité) [Fab96], la théorie de possibilité permet de mesurer l'incertitude dans deux cadres de raisonnement selon la nature des connaissances initiales détenues : un cadre quantitatif et un cadre qualitatif [DP98]. La logique possibiliste [BP06, DP04b, BDGP02] est une extension de la logique classique qui permet de traiter des informations totalement ordonnées. Les priorités entre ces informations sont représentées à l'aide de formules pondérées, où chaque poids représente la borne inférieure d'une mesure de nécessité.

Dans ce chapitre, nous ferons un tour d'horizon sur l'essentiel des théories de l'incertain : nous commencerons dans un premier temps par un rappel sur la théorie des probabilités (section 1.2), suivi d'une présentation de la théorie des ensembles flous (section 1.3). Une brève description des fonctions ordinales de Spohn est présentée dans la section 1.4. La section 1.5 expose le Modèle des Croyances Transférables *MCT*. Enfin, les sections 1.6 et 1.7 présentent, respectivement, les notions de base de la théorie des possibilités ainsi que sa contre-partie la logique possibiliste.

1.2 La théorie des probabilités

La théorie des probabilités fournit des modèles mathématiques permettant l'étude d'expériences dont le résultat ne peut être prévu avec une totale certitude. Bien que le résultat précis de chacune de ces expériences soit imprévisible, l'observation et l'intuition laissent penser que ces phénomènes obéissent à certaines lois. Intuitivement, la probabilité d'un événement correspond à sa fréquence de réalisation : c'est ce que l'on appelle l'approche fréquentiste des probabilités. Cette approche présume que si on considère une série de tentatives d'un même événement, la probabilité que l'évènement se réalise correspond au rapport du nombre de réalisations par le nombre de tentatives. La théorie des probabilités permet donc de donner un sens précis à ces considérations un peu vagues. La statistique permet de confronter les modèles probabilistes avec la réalité observée afin de les valider ou de les invalider.

1.2.1 Espace d'état et événements

On appelle espace d'état l'ensemble non vide Ω de tous les résultats possibles d'une expérience aléatoire ; ses éléments ω sont appelés épreuves ou événements élé-

mentaires. L'espace d'état Ω peut être fini, dénombrable ou infini non dénombrable. Dans le cadre de cette thèse, l'ensemble Ω est considéré fini. On appelle événement A un sous ensemble de Ω . Si un élément $\omega \in A$, alors ω est un résultat possible de A ; cela signifie que ω satisfait A ou bien que A se réalise dans ω . A chaque événement on associe une variable aléatoire qui permet de le caractériser par sa valeur.

Lorsque l'on se limite aux variables booléennes, les opérateurs logiques «et», «ou» et «négation», définis sur les événements, se traduisent par des opérations ensemblistes intersection, réunion et passage au complémentaire. Le tableau suivant trace la correspondance entre les deux vocabulaires le premier est ensembliste et le second est probabiliste :

| Notation | Vocabulaire ensembliste | Vocabulaire probabiliste |
|------------------------|----------------------------|---|
| \emptyset | ensemble vide | événement impossible |
| Ω | ensemble plein | événement certain |
| ω | élément de Ω | événement élémentaire |
| A | sous-ensemble de Ω | événement |
| $\omega \in A$ | ω appartient à A | ω est une des réalisations possible de A |
| $A \subset B$ | A est inclus dans B | A satisfait B |
| $A \cup B$ | réunion de A et B | A ou B , réalisation de A ou de B |
| $A \cap B$ | intersection de A et B | A et B , réalisation de A et de B |
| A^c | complémentaire de A | événement contraire de A |
| $A \cap B = \emptyset$ | A et B sont disjoints | A et B sont incompatibles |

TAB. 1.1 – Correspondance entre théorie des ensembles et Probabilités

1.2.2 Mesure de probabilités

Une distribution de probabilités (appelée aussi une loi de probabilités) est une application P qui associe à tout événement A un nombre $P(A)$, appelé probabilité de A . P doit satisfaire les axiomes suivants :

- i) $P(A) \in [0, 1]$
- ii) $P(\Omega) = 1$
- iii) Si A et B sont deux événements incompatibles, alors $P(A \cup B) = P(A) + P(B)$.

Une distribution de probabilités satisfait la condition de normalisation suivante :

$$\sum_{\omega \in \Omega} p(\omega) = 1 \quad (1.1)$$

En effet, si on pose p_i la probabilité que ω_i soit réalisé, on peut calculer la probabilité de tout événement A par la formule suivante :

$$P(A) = \sum_{i, \omega_i \in A} p_i \quad (1.2)$$

Nous appelons espace probabilisé tout couple (Ω, P) où Ω est un espace d'état et P est une distribution de probabilités sur Ω .

En plus des trois axiomes précédents, toute probabilité P sur (Ω, P) vérifie les propriétés suivantes :

- iv) $P(\emptyset) = 0$
- v) $P(A^c) = 1 - P(A)$
- vi) Si A satisfait B , alors $P(A) \leq P(B)$ et $P(B \setminus A) = P(B) - P(A)$
- vii) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, et par conséquent :
 $P(A \cup B) \leq P(A) + P(B)$

1.2.3 Probabilité conditionnelle et indépendance

Un concept fondamental en calcul des probabilités, qui permet de tenir compte des informations complémentaires, est celui de probabilité conditionnelle [She97]. Dans un espace probabilisé (Ω, P) , la probabilité conditionnelle d'un événement A , sachant qu'un autre événement B de probabilité non nulle s'est réalisé, est le degré $P(A|B)$ défini par :

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (1.3)$$

La probabilité de A sachant B ou probabilité de A conditionnellement à B est très utile pour calculer les probabilités d'intersection des événements (ou probabilité jointe des événements) :

$$P(A \cap B) = P(A, B) = P(A|B).P(B) = P(B|A).P(A) \quad (1.4)$$

Une conséquence directe de cette formule du conditionnement est la règle de Bayes énoncée comme suit :

$$P(A|B) = \frac{P(B|A).P(A)}{P(B)} \quad (1.5)$$

Cette règle de Bayes peut être interprétée de la manière suivante : le terme $P(A)$ est la probabilité *a priori* de A . Elle est « antérieure » au sens qu'elle précède toute information sur B . $P(A)$ est aussi appelée la probabilité marginale de A . Le terme $P(A|B)$ est appelée la probabilité *a posteriori* de A sachant B (ou de A conditionnellement à B). Elle est « postérieure », au sens qu'elle dépend directement

de B . Le terme $P(B|A)$, pour un B connu, est appelée *la fonction de vraisemblance* de A . De même, le terme $P(B)$ est appelé la probabilité marginale ou *a priori* de B .

La formule de conditionnement peut être généralisée sur N événements A_1, A_2, \dots, A_N :

$$\begin{aligned} P(A_1, A_2, \dots, A_N) &= P(A_1).P(A_2, \dots, A_N|A_1) \\ &= P(A_1).P(A_2|A_1).P(A_3, \dots, A_N|A_1, A_2) \\ &= P(A_1).P(A_2|A_1).P(A_N|A_1, A_2, \dots, A_{N-1}) \\ &= \prod_{i=1}^n P(A_i|A_1, A_2, \dots, A_{i-1}) \end{aligned} \quad (1.6)$$

Dans un espace probabilisé, deux événements A et B sont indépendants si et seulement si :

$$P(A, B) = P(A).P(B) \quad (1.7)$$

Autrement dit, A et B sont indépendants si et seulement si :

$$P(A|B) = P(A) \Leftrightarrow P(B|A) = P(B) \quad (1.8)$$

Cette définition peut être étendue afin de tenir compte du contexte. A est indépendant de B conditionnellement à C si et seulement si :

$$P(A|B, C) = P(A|C) \Leftrightarrow P(B|A, C) = P(B|C) \quad (1.9)$$

Exemple 1.1 soient A et B deux variables aléatoires dont la probabilité jointe est donnée dans le tableau suivant :

| $P(A, B)$ | b_1 | b_2 |
|-----------|-------|-------|
| a_1 | 0.25 | 0.12 |
| a_2 | 0.33 | 0.3 |

TAB. 1.2 – Probabilités jointes sur A et B

La probabilité marginale de chaque variable est calculée comme suit :

$$P(A = a_1) = P(A = a_1, B = b_1) + P(A = a_1, B = b_2) = 0.25 + 0.12 = 0.37, \text{ d'où :}$$

$P(A = a_1, B = b_1) = 0.25 \neq P(A = a_1).P(B = b_1) = 0.21$, On peut déduire que A et B sont dépendants.

| | | | | | |
|--------|-------|-------|--------|-------|-------|
| | a_1 | a_2 | | b_1 | b_2 |
| $P(A)$ | 0.37 | 0.63 | $P(B)$ | 0.58 | 0.42 |

TAB. 1.3 – Probabilités marginales sur A et B

1.3 La théorie des ensembles flous

1.3.1 Définition

La théorie des ensembles flous [Zad65, Kle84, Bal79, KF88, DP80] constitue une extension des ensembles classiques définie, en 1965, par Lotfi Zadeh dans un objectif de modélisation du langage naturel. Le langage naturel est caractérisé par la manipulation de termes vagues, faisant référence à des classes d'objets ne possédant pas de critères d'appartenance précis, comme par exemple : chaud, grand... La théorie des ensembles flous permet donc la prise en compte d'ensembles définis de façon imprécise.

Contrairement à la logique classique, où les variables gérées sont booléennes, la logique floue s'appuie sur la théorie des ensembles flous et a pour but de raisonner à partir de connaissances imprécises et se propose de remplacer les variables booléennes par des variables flous.

Un sous-ensemble flou F est défini sur un domaine U par une fonction d'appartenance μ_F telle que :

$$\forall u \in U : \mu_F(u) \in [0, 1]. \quad (1.10)$$

μ_F quantifie le degré d'appartenance de chaque élément de U à F .

Les degrés d'appartenance $\mu_F(u) = 0$ et $\mu_F(u) = 1$ signifient respectivement que u appartient et n'appartient pas à F . Dans le cas où les degrés d'appartenance sont toujours 0 ou 1 (valeur binaire), on se ramène à un ensemble classique. Si au contraire ils peuvent prendre des valeurs intermédiaires, l'ensemble est dit flou et dans ce cas une valeur comprise entre 0 et 1 estime le degré d'appartenance de u à F .

1.3.2 Propriétés

Un sous-ensemble flou F est formellement défini par sa fonction d'appartenance μ_F caractérisée par les propriétés suivantes :

- i) Le noyau d'un sous-ensemble flou F est l'ensemble des éléments qui appartiennent totalement à F c'est-à-dire dont le degré d'appartenance à F est égal à 1.

$$n(F) = \{u \in U : \mu_F(u) = 1\}. \quad (1.11)$$

- ii) Le support d'un sous-ensemble flou F est l'ensemble des éléments appartenant à F avec des degrés divers non nuls, c'est-à-dire dont le degré d'appartenance, d'un élément, à F est strictement supérieur à 0.

$$s(F) = \{u \in U : \mu_F(u) > 0\}. \quad (1.12)$$

- iii) La hauteur d'un sous-ensemble flou F est la borne supérieure de la fonction d'appartenance même si elle n'est pas nécessairement atteinte.

$$h(F) = \sup\{\mu_F(u) : u \in U\}. \quad (1.13)$$

Un sous-ensemble flou est dit normalisé s'il est de hauteur 1.

- iv) Une α -coupe d'un sous-ensemble flou F est l'ensemble des éléments ayant un degré d'appartenance supérieur ou égal à α .

$$\alpha - \text{coupe}(F) = \{u \in U : \mu_F(u) \geq \alpha\}. \quad (1.14)$$

Exemple 1.2 Soit un sous-ensemble flou B décrivant les personnes jeunes. La fonction d'appartenance μ_B associée B représente le fait qu'une personne appartient à la jeunesse selon son âge. Elle est définie de la manière suivante :

$$\mu_B(\text{age}) = \begin{cases} 1 & \text{si } 0 \leq \text{age} \leq 20, \\ \frac{30-\text{age}}{10} & \text{si } 20 < \text{age} \leq 30, \\ 0 & \text{si } \text{age} > 30. \end{cases}$$

La figure 1.1 représente graphiquement la fonction d'appartenance μ_B en fonction de l'âge d'une personne.

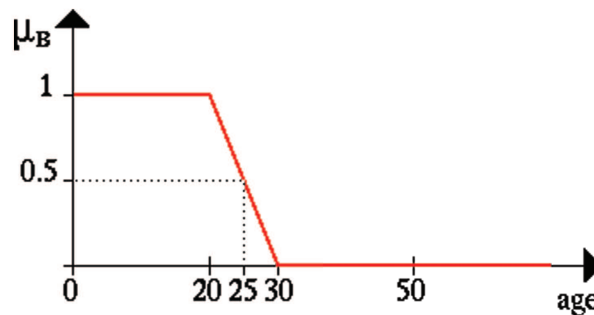


FIG. 1.1 – Représentation de la fonction d'appartenance μ_B en logique floue.

- Le noyau $n(B)$ représente les personnes ayant un âge inférieur ou égal à 20.

- Le support $s(B)$ représente les personnes ayant un âge strictement inférieur à 30.
- La hauteur $h(B)$ est égal à 1. Le sous-ensemble B est ainsi normalisé.
- En prenant, par exemple, $\alpha = 0.5$ alors α -coupe(B) représente les personnes ayant un âge inférieur ou égal à 25.

1.3.3 Opérations sur les ensembles flous

La théorie des ensembles flous se pose comme une extension de la théorie des ensembles standards. Elle redéfinit donc toutes ces notions, ainsi que ces opérations sur les ensembles que sont l'intersection, la réunion et le complémentaire.

• **Intersection :**

La fonction d'appartenance de l'intersection de deux sous-ensembles flous A et B , représentée respectivement par μ_A et μ_B , est définie par le minimum de ces deux fonctions d'appartenance comme suit :

$$\mu_{A \cap B} = \min(\mu_A, \mu_B). \quad (1.15)$$

En théorie des ensembles flous, cet opérateur d'intersection est équivalent à l'opérateur ET dans la logique classique dans le sens où, pour chaque élément de $A \cap B$, la plus petite valeur assurée simultanément par les fonctions μ_A et μ_B sera attribuée à la fonction d'appartenance résultante.

• **Réunion :**

La fonction d'appartenance de la réunion de deux ensembles flous A et B , représentée respectivement par μ_A et μ_B , est définie par le maximum de ces deux fonctions d'appartenance comme suit :

$$\mu_{A \cup B} = \max(\mu_A, \mu_B). \quad (1.16)$$

En théorie des ensembles flous, cet opérateur de réunion est équivalent à l'opérateur OU dans la logique classique dans le sens où, pour chaque élément de $A \cup B$, la plus grande valeur assurée parmi les fonctions μ_A et μ_B sera attribuée à la fonction d'appartenance résultante.

• **Complémentaire :**

La fonction d'appartenance du complémentaire d'un ensemble flou A , représentée par μ_A , est définie par la fonction d'appartenance $1 - \mu_A$.

En théorie des ensembles flous, cet opérateur du complémentaire est équivalent à l'opérateur de négation dans la logique classique.

Ces définitions sont les plus courantes mais, pour certains cas, d'autres sont plus appropriées. Par exemple, l'intersection peut être définie par le produit des fonctions d'appartenance et l'union par la moyenne arithmétique des fonctions d'appartenance. Ces différentes techniques de calcul engendrent une énorme capacité d'adaptation des raisonnements flous.

Exemple 1.3 *On désire représenter la température de l'eau d'un récipient par les qualificatifs : Froide : F, Tiède : T, Chaude : C.*

Dans la logique classique, uniquement les valeurs 0 et 1 sont utilisées, par les fonctions d'appartenance associées à chaque qualificatif, pour évaluer la température de l'eau. Ainsi, l'eau est totalement froide pour une température entre 0 et 20 °C puis à 20 °C, l'eau n'est plus froide et devient brusquement tiède et ceci jusqu'à la température 45 °C. À partir de 45 °C l'eau devient chaude après avoir cessée d'être tiède. Elle en restera chaude jusqu'à la température 100 °C. La figure 1.2 illustre ces qualificatifs de l'eau.

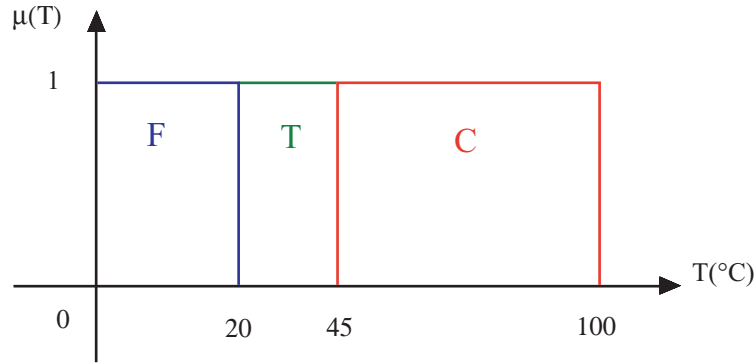


FIG. 1.2 – Représentation de la température de l'eau par la logique classique.

Dans la théorie des ensembles flous, la représentation graphique des trois fonctions d'appartenance Froid, Tiède et Chaud permet de superposer sur des plages de température donnée les qualificatifs froid et tiède ainsi que tiède et chaud. Par exemple, pour une température entre 20 et 45 °C, l'eau est considérée, à la fois, froide et tiède. On se rapproche, donc, du raisonnement humain. La figure 1.3 montre l'allure des fonctions d'appartenance des 3 qualificatifs.

1.4 Les fonctions conditionnelles ordinales de Spohn

Les fonctions conditionnelles ordinales [Spo88b, Spo88a], appelées aussi fonctions Kappa, modélisent l'incertitude associée à l'état du monde réel par une fonction κ

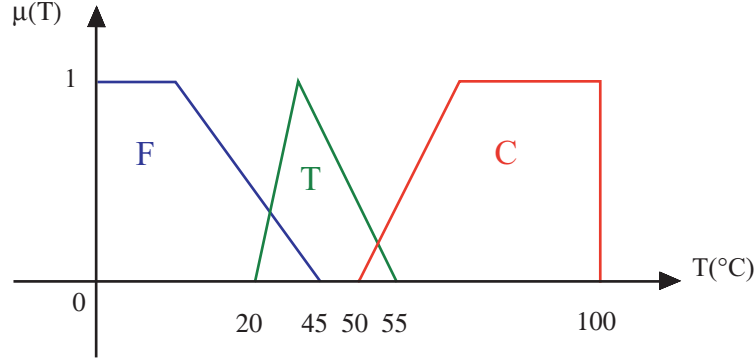


FIG. 1.3 – Représentation de la température de l'eau par la logique floue.

définie de l'ensemble des interprétations Ω sur l'ensemble des entiers naturels. Pour une interprétation $\omega \in \Omega$, $\kappa(\omega)$ représente un degré mesurant l'incrédulité (disbelief) de ω dans l'état du monde réel décrit par une fonction Kappa.

Dans une distribution kappa κ , il existe au moins une interprétation ω telle que $\kappa(\omega) = 0$, et dans ce cas l'interprétation (ou l'élément élémentaire) ω est dite alors acceptée. D'une manière générale, plus $\kappa(\omega)$ est petit, plus ω est préférée par rapport aux autres interprétations.

Ainsi, il est possible de définir un rang pour un événement $\phi \subset \Omega$ de la manière suivante :

$$\kappa(\phi) = \min\{\kappa(\omega) : \omega \in \phi\}. \quad (1.17)$$

$\omega \in \phi$ signifie que ω est un modèle de ϕ (ω satisfait ϕ).

Spohn définit le conditionnement d'une interprétation ω par un événement ϕ comme suit :

$$\kappa(\omega|\phi) = \begin{cases} \kappa(\omega) - \kappa(\phi) & \text{si } \omega \in \phi, \\ +\infty & \text{sinon.} \end{cases} \quad (1.18)$$

De plus, pour deux événements quelconques ϕ et ψ ,

$$\kappa(\phi|\psi) = \min\{\kappa(\omega|\phi) : \omega \in \phi \wedge \psi\} = \kappa(\phi \wedge \psi) - \kappa(\phi),$$

d'où l'axiome de conjonction :

$$\kappa(\phi \wedge \psi) = \kappa(\phi|\psi) + \kappa(\phi). \quad (1.19)$$

Les fonctions ordinales conditionnelles de Spohn représentent un modèle "qualitatif" pour la représentation de l'incertitude. En effet, elles permettent l'ordonnancement des événements selon leur rang exprimé par le degré d'incrédulité. Les fonctions Kappa sont utilisées pour la révision des croyances [FH99] et le raisonnement par

défaut [PGA87]. La section suivante montre qu'il est possible de représenter l'incertitude non pas sur l'ensemble des entiers naturels mais plutôt sur l'échelle unitaire.

Exemple 1.4 *On désire exprimer notre croyance (ou incrédulité) sur les conditions de voyage d'un passager à bord d'un bateau. Les trois conditions, à tenir compte, sont matérialisées par trois variables binaires A , B et C comme suit :*

- *La variable A exprime la nature de l'hébergement sur le bateau selon deux états :*
 - * *l'état a_1 qui exprime dormir en cabine,*
 - * *l'état a_2 qui exprime dormir en couchette.*
- *La variable B exprime le niveau de confort à choisir à bord du bateau :*
 - * *l'état b_1 qui exprime voyager en première classe,*
 - * *l'état b_2 qui exprime voyager en classe économique.*
- *La variable C exprime l'option de nourriture à choisir pendant le voyage :*
 - * *l'état c_1 qui exprime voyager en pension complète,*
 - * *l'état c_2 qui exprime voyager en demi-pension (petit-déjeuner + dîner).*

Le tableau 1.4 donne les valeurs de la distribution κ associée à l'univers Ω .

| A | B | C | κ |
|-------|-------|-------|----------|
| a_1 | b_1 | c_1 | 0 |
| a_1 | b_1 | c_2 | 2 |
| a_1 | b_2 | c_1 | 2 |
| a_1 | b_2 | c_2 | 7 |
| a_2 | b_1 | c_1 | 5 |
| a_2 | b_1 | c_2 | 8 |
| a_2 | b_2 | c_1 | 11 |
| a_2 | b_2 | c_2 | 19 |

TAB. 1.4 – Distribution κ associée à Ω

L'interprétation (a_1, b_1, c_1) est la formule de voyage qui sera préférée par un voyageur (mais à quel prix!), son degré κ sera donc égal à 0.

Selon le choix effectué par le passager, les valeurs de la distribution κ associée à chaque interprétation prennent des valeurs positives (exprimant ainsi l'incrédulité vis à vis de ces interprétations). Par exemple, l'interprétation $\omega = a_2 b_1 c_1$, qui exprime voyager en dormant dans une couchette en première classe et en pension complète ($\kappa(\omega) = 5$), est préférée à l'interprétation $\omega' = a_1 b_2 c_2$ qui exprime voyager avec une cabine en classe économique en demi-pension ($\kappa(\omega') = 7$).

Soit la formule $\phi = a_1 b_2$ qui exprime voyager en choisissant une cabine en classe économique. selon l'équation (1.17), $\kappa(\phi) = \min(2, 7) = 2$.

Soit les deux interprétations suivantes : $\omega_1 = a_1b_1c_1$ et $\omega_2 = a_1b_2c_2$. Sachant la formule $\phi = a_1b_2$, nous pouvons calculer la valeur de kappa en chaque ω conditionnellement à ϕ , selon l'équation (1.18) :

- $\kappa(\omega_1|\phi) = +\infty$, car ω_1 ne satisfait pas ϕ .
- Par contre, ω_2 satisfait ϕ donc, $\kappa(\omega_2|\phi) = \kappa(\omega_2) - \kappa(\phi) = 7 - 2 = 5$.

1.5 Le Modèle des Croyances Transférables

Le Modèle des Croyances Transférables MCT (*Transferable Belief Model*) est un cadre formel générique développé par [SK94] pour la représentation et la combinaison des connaissances. Ce modèle permet de prendre en compte les imperfections des connaissances, aussi bien les incertitudes que les imprécisions. Le MCT est basé sur la définition de fonctions de croyance [Dem68, Sha76] fournies par des sources d'information pouvant être complémentaires et redondantes. Il propose un ensemble d'opérateurs permettant de combiner ces fonctions. Il est donc naturellement employé dans le cadre de la fusion d'informations pour améliorer l'analyse et l'interprétation de données issues de sources d'informations multiples. Dans le cadre du MCT, deux niveaux sont distingués :

- Le niveau **crédal**, où sont représentées (la partie statique) et manipulées (la partie dynamique) les informations disponibles,
- Le niveau **pignistique** où la décision est prise.

1.5.1 Représentation des connaissances

L'état du système est supposé prendre des valeurs discrètes appelées hypothèses ω_k et l'ensemble des K hypothèses possibles pouvant expliquer l'état du système est appelé cadre de discernement (ou univers de discours) généralement noté par : $\Omega = \{\omega_1, \dots, \omega_k, \dots, \omega_K\}$. Le nombre d'hypothèses $|\Omega| = K$ composant Ω est appelé cardinal et les hypothèses sont considérées exclusives.

Une fonction de masse de croyance sur Ω est une application m de 2^Ω dans $[0, 1]$ qui vérifie :

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

Chaque sous-ensemble (appelé proposition) $A \subseteq \Omega$ tel que $m(A) > 0$ est appelé élément focal de m . Ainsi, la masse $m(A)$ représente la part de croyance allouée à la proposition A mais à aucun de ses sous-ensembles stricts. $m(\Omega)$ représente le degré d'ignorance tandis que $m(\emptyset)$ représente le degré du conflit.

A partir de la fonction de masse m , une autre fonction équivalente peut être déduite.

C'est la fonction de croyance bel qui est une mesure de 2^Ω dans $[0, 1]$ définie par :

$$bel(A) = \sum_{\emptyset \neq B \subseteq A} m(B) \quad \forall A \subseteq \Omega.$$

$bel(A)$ s'interprète comme le degré de croyance en A .

Les fonctions de croyance sont reconnues pour la modélisation des informations imprécises et incertaines (allant de l'ignorance totale à la connaissance complète). Ainsi, une situation d'ignorance totale correspond à la fonction de croyance définie par $m(\Omega) = 1$. La connaissance complète sera représentée par une fonction de croyance certaine et précise, c'est-à-dire une fonction où la totalité de la masse est allouée à un singleton unique de Ω .

La seconde étape au niveau crédal correspond à la révision des croyances. Considérons une masse $m(A)$ strictement positive allouée à un sous-ensemble A de Ω . Si nous apprenons avec certitude que la vérité se situe dans un sous-ensemble B de Ω , la masse initialement allouée à A devra alors être transférée à $A \cap B$. Cette règle correspond à la règle de conditionnement non normalisée de Dempster. Le terme *modèle des croyances transférables* est issu de ce transfert de masse créditée à des sous-ensembles en fonction des informations disponibles. Issu de ce principe de conditionnement, la règle de combinaison conjonctive permet de fusionner deux fonctions de masse, m_1 et m_2 , issues de sources d'informations distinctes et fiables. Cette règle est définie par :

$$m_1 \cap m_2(A) = \sum_{B \cap C = A} m_1(B)m_2(C) \quad \forall A \subseteq \Omega.$$

1.5.2 Niveau pignistique

Les modèles des croyances transférables, en tant que systèmes de fusion d'informations, ont pour finalité la prise de décision. Prendre une décision consiste à choisir une hypothèse sur le cadre de discernement. La phase de décision s'appuie sur la distribution de probabilités pignistiques, notée $BetP$, obtenue à partir de la fonction de masse m de la manière suivante : $\forall \omega_k \in \Omega$,

$$BetP(\omega_k) = \frac{1}{1 - m(\emptyset)} \sum_{A \subseteq \Omega, \omega_k \in A} \frac{m(A)}{|A|},$$

où $|A|$ représente la cardinalité de $A \subseteq \Omega$. Dans cette transformation pignistique, la masse de croyance $m(A)$ est uniformément distribuée sur les hypothèses contenues dans A . La décision est généralement prise en choisissant l'élément ω_k possédant la plus grande probabilité pignistique.

Exemple 1.5 Soient $\omega_1, \omega_2, \omega_3$ trois hypothèses, sur l'état d'un système, formant

le cadre de discernement Ω . La fonction de masse m associée à chaque proposition est donnée dans la colonne 2 du tableau 1.5. La 3^e colonne présente les degrés de croyance bel de chaque proposition.

Supposons maintenant qu'une autre source fiable (colonne 4) fournisse les informations représentées par la fonction de masse m_s . La fusion, par combinaison conjonctive, des fonctions de masses m et m_s est donnée dans la colonne 5. Le calcul de la distribution pignistique $BetP$ pour chaque hypothèse donne : $BetP(\omega_1) = 0.25$, $BetP(\omega_2) = 0.25$ et $BetP(\omega_3) = 0.5$. La décision sur l'état du système ω_0 pourrait être $\omega_0 = \omega_3$ signifiant que l'état réel du système est l'état ω_3 .

| A | $m(A)$ | $bel(A)$ | $m_s(A)$ | $m \cap m_s(A)$ |
|--------------------------|--------|----------|----------|-----------------|
| \emptyset | 0 | 0 | 0 | 0.45 |
| $\{\omega_1\}$ | 0 | 0 | 0 | 0 |
| $\{\omega_2\}$ | 0 | 0 | 0.8 | 0.40 |
| $\{\omega_1, \omega_2\}$ | 0.5 | 0.5 | 0 | 0.05 |
| $\{\omega_3\}$ | 0.5 | 0.5 | 0.1 | 0.1 |
| $\{\omega_1, \omega_3\}$ | 0 | 0.5 | 0 | 0 |
| $\{\omega_2, \omega_3\}$ | 0 | 0.5 | 0 | 0 |
| Ω | 0 | 1 | 0.1 | 0 |

TAB. 1.5 – Exemple de fonction de masse $m(A)$, de croyance $bel(A)$ et combinaison conjonctive entre m et m_s

1.6 La théorie des possibilités

La théorie des possibilités introduite par Zadeh [Zad78] et développée par Dubois et Prade [DP04b, DP98, DP88] traite l'incertitude sur l'intervalle $[0, 1]$, appelé échelle possibiliste. Les fondements de la théorie des possibilités se basent sur les notions de distribution de possibilités, mesures de possibilité et de nécessité, conditionnement et indépendance possibiliste. La définition de certaines notions peut différer selon le cadre du raisonnement utilisé : qualitatif ou quantitatif.

1.6.1 Distribution de possibilités

La théorie des possibilités est basée sur les distributions de possibilités. Une distribution de possibilités, notée π , est une application de Ω (l'univers de discours) vers l'échelle $[0, 1]$ et représente l'état du monde réel en donnant un ordre de préférences entre les différentes interprétations de ce monde. L'échelle possibiliste peut être interprétée de deux manières différentes [BABMS07] :

- d'une manière ordinale, lorsque les valeurs affectées ne reflètent que l'ordre de classement entre les différents états possibles.
- d'une manière cardinale, lorsque les valeurs affectées prennent un sens sur l'intervalle.

Soit $V = \{A_1, A_2, \dots, A_n\}$ un ensemble de variables dont chacune est munie d'un domaine fini, noté D_{A_i} , formé d'instances a_i . L'univers du discours Ω s'écrit alors comme le produit cartésien des domaines de toutes les variables ; $\Omega = \times_{A_i \in V} D_{A_i}$. Une interprétation $\omega \in \Omega$ s'écrit comme un tuple (a_1, a_2, \dots, a_n) tel que $\omega[A_i] = a_i$. Une formule (ou événement) φ est un sous ensemble de Ω combinant les instances de certains A_i par les connecteurs logiques : conjonction ' \wedge ', disjonction ' \vee ' et négation ' \neg '. \models exprime la relation de satisfaction de la manière suivante : $\omega \models \varphi$ signifie que ω est un modèle de φ (ou ω satisfait φ).

Les points suivants résument quelques interprétations associées à une distribution de possibilités π :

- $\pi(\omega) > \pi(\omega')$ signifie que ω est préférée à ω' pour décrire l'état du monde réel.
- $\pi(\omega) = 0$ signifie que ω est impossible.
- $\pi(\omega) = 1$ signifie que ω est totalement possible.
- $0 < \pi(\omega) < 1$ signifie qu'il existe une information qui contredise notre croyance vis à vis de l'interprétation ω .

• *Normalisation* : Une distribution de possibilités π est dite **α -normalisée** si son degré de normalisation, noté par $h(\pi)$, est égal à α . Autrement dit :

$$\alpha = h(\pi) = \max_{\omega \in \Omega} \pi(\omega). \quad (1.20)$$

Lorsque $\alpha = 1$, alors la distribution π est dite normalisée. Cela veut dire qu'il existe au moins une interprétation cohérente avec les informations disponibles.

• *Marginalisation* : Soit une distribution de possibilités π sur Ω , la **distribution marginale** sur un sous-ensemble de variables $X \subseteq V$ est donnée en utilisant l'opérateur maximum : $\forall X \subseteq V, \forall x \in D_X$,

$$\Pi(x) = \max_{\omega \in \Omega} \{\pi(\omega) : \omega[X] = x\}, \quad (1.21)$$

avec D_X le produit cartésien des domaines des variables constituant le sous-ensemble X et $\omega[X] = x$ est la configuration de X dans ω .

1.6.2 Mesure de possibilité et de nécessité

L'incertitude sur un événement φ peut être décrite par deux mesures duales : la mesure de possibilité Π et la mesure de nécessité N [DP88].

1.6.2.1 Mesure de possibilité Π

La mesure (ou degré) de possibilité Π d'un événement φ , notée par $\Pi(\varphi)$, est obtenue par :

$$\Pi(\varphi) = \max \{ \pi(\omega) : \omega \models \varphi \}. \quad (1.22)$$

Le degré de possibilité $\Pi(\varphi)$ décrit la situation la plus normale dans laquelle φ est vraie. Il évalue le niveau de compatibilité de φ avec les informations disponibles. $\Pi(\varphi) = 1$ signifie que φ est totalement possible et $\Pi(\varphi) = 0$ signifie que φ est impossible.

Dans le cas d'une distribution de possibilités normalisée, la mesure de possibilité Π vérifie les propriétés suivantes :

- $\Pi(\Omega) = 1$.
- $\Pi(\emptyset) = 0$.
- $\max(\Pi(\varphi), \Pi(\neg\varphi)) = 1$.
- $\Pi(\varphi \vee \psi) = \max(\Pi(\varphi), \Pi(\psi))$.

Ce dernier axiome, appelé de disjonction, est l'axiome principal de la théorie des possibilités sur lequel les deux cadres quantitatif et qualitatif concordent.

1.6.2.2 Mesure de nécessité N

La mesure de nécessité, notée N , représente une mesure duale (complémentaire) à la mesure de possibilité Π permettant de compléter nos connaissances sur un événement φ . La mesure de nécessité N d'un événement φ est définie comme suit :

$$N(\varphi) = 1 - \Pi(\neg\varphi). \quad (1.23)$$

$N(\varphi)$ correspond au degré de certitude associé à φ dans le contexte des connaissances disponibles. la mesure de nécessité N vérifie les propriétés suivantes :

- $N(\Omega) = 1$.
- $N(\emptyset) = 0$.
- $\min(N(\varphi), N(\neg\varphi)) = 0$.
- Si $N(\varphi) > 0$ alors $\Pi(\varphi) = 1$.
- $\forall \varphi, N(\varphi) \leq \Pi(\varphi)$.
- $N(\varphi \wedge \psi) = \min(\Pi(\varphi), \Pi(\psi))$.

La mesure de possibilité Π ne permet pas de décrire toute l'incertitude sur un événement φ . Par exemple, si φ est totalement possible ($\Pi(\varphi) = 1$) cela n'exclut pas que $\neg\varphi$ peut se réaliser ($\Pi(\neg\varphi) \neq 0$). Par contre, si $\Pi(\neg\varphi) = 0$ alors seul l'événement φ peut se réaliser avec certitude ($N(\varphi) = 1$).

Exemple 1.6 Soient A et B deux variables binaires. La distribution de possibilités

jointe de A et B est donnée dans le tableau suivant :

| A | B | $\Pi(A, B)$ |
|----------|----------|-------------|
| $\neg a$ | $\neg b$ | 0.8 |
| $\neg a$ | b | 0.3 |
| a | $\neg b$ | 0.5 |
| a | b | 1 |

TAB. 1.6 – Distribution de possibilités jointe sur A et B

L'interprétation $\neg a \wedge \neg b$ est préférée à celle $a \wedge \neg b$ car $\Pi(\neg a \wedge \neg b) > \Pi(a \wedge \neg b)$.

La distribution marginale sur A est calculée comme suit :

$$\Pi(a) = \max(\Pi(a, b), \Pi(a, \neg b)) = \max(1, 0.5) = 1$$

$$\Pi(\neg a) = \max(\Pi(\neg a, b), \Pi(\neg a, \neg b)) = \max(0.3, 0.8) = 0.8$$

La mesure de nécessité $N(a) = 1 - \Pi(\neg a) = 1 - 0.8 = 0.2$.

De même, $N(\neg a) = 1 - \Pi(a) = 1 - 1 = 0$.

1.6.3 Conditionnement possibiliste

Le conditionnement possibiliste consiste à modifier la distribution de possibilités π en observant l'arrivée d'une nouvelle information certaine. La nouvelle distribution sera obtenue en changeant nos croyances sur l'état du monde réel en considérant l'information observée ψ . Par conséquent, la mesure de l'incertitude changera au niveau de chaque interprétation. La distribution initiale π sera remplacée par une autre $\pi' = \pi(\bullet|\psi)$. Plusieurs définitions du conditionnement possibiliste ont été proposées dans la littérature [BMCM02, DCK94, Fon97, His78, DP88, Van03]. Dans le cadre possibiliste, le conditionnement diffère selon la nature de l'information disponible : quantitatif ou qualitatif.

1.6.3.1 Conditionnement possibiliste quantitatif

Le conditionnement dans un cadre de raisonnement possibiliste quantitatif est défini comme suit :

$$\Pi(\phi \wedge \psi) = \Pi(\phi|\psi) \cdot \Pi(\psi). \quad (1.24)$$

Ainsi, l'effet d'un événement ψ sur nos connaissances associées à ϕ est calculé comme suit :

$$\Pi(\phi|\psi) = \frac{\Pi(\phi \wedge \psi)}{\Pi(\psi)}. \quad (1.25)$$

L'effet d'un événement ψ sur nos connaissances sur un état ω du monde est donné par [Sha76, DC97, His78] :

$$\pi(\omega|\psi) = \begin{cases} \frac{\pi(\omega)}{\Pi(\psi)} & \text{si } \omega \models \psi \\ 0 & \text{sinon.} \end{cases} \quad (1.26)$$

Cette formule correspond au conditionnement possibiliste basé sur le produit.

Exemple 1.7 *Considérons la même distribution jointe donnée par la table 1.6. L'arrivée d'une information certaine $B = \neg b$ va modifier la distribution. Le tableau 1.7 montre la nouvelle distribution sur A et B après conditionnement :*

| A | B | $\Pi(A B = \neg b)$ |
|----------|----------|-----------------------|
| $\neg a$ | $\neg b$ | $\frac{0.8}{0.8} = 1$ |
| $\neg a$ | b | 0 |
| a | $\neg b$ | $\frac{0.5}{0.8}$ |
| a | b | 0 |

TAB. 1.7 – Distribution de possibilités sur A et B après conditionnement quantitatif par $B = \neg b$

1.6.3.2 Conditionnement possibiliste qualitatif

Le conditionnement dans un cadre de raisonnement possibiliste qualitatif est défini comme suit :

$$\Pi(\phi \wedge \psi) = \min(\Pi(\phi|\psi), \Pi(\psi)). \quad (1.27)$$

$\Pi(\phi|\psi)$ représente la solution la plus grande de cette équation. Ce qui implique :

$$\Pi(\phi|\psi) = \begin{cases} 1 & \text{si } \Pi(\phi \wedge \psi) = \Pi(\psi) \\ \Pi(\phi \wedge \psi) & \text{si } \Pi(\phi \wedge \psi) < \Pi(\psi). \end{cases} \quad (1.28)$$

Lorsqu'il s'agit de calculer l'effet d'un événement ψ sur nos connaissances sur un état ω , l'équation précédente peut aussi s'écrire sous la forme [His78, DP88] :

$$\pi(\omega|\psi) = \begin{cases} 1 & \text{si } \pi(\omega) = \Pi(\psi) \text{ et } \omega \models \psi \\ \pi(\omega) & \text{si } \pi(\omega) < \Pi(\psi) \text{ et } \omega \models \psi \\ 0 & \text{sinon.} \end{cases} \quad (1.29)$$

Cette formule correspond au conditionnement possibiliste *basé sur le min*.

Dans un cadre qualitatif, le degré de possibilité maximal est affecté aux meilleurs éléments de ψ (ce qui correspond aux états les plus plausibles dans ψ); puis on calcule le degré de possibilité des autres connaissances disponibles.

Exemple 1.8 *On considère de nouveau la distribution jointe de A et B donnée par le tableau 1.6. Le conditionnement possibiliste qualitatif appliqué à cette distribution suite à l'observation de $B = \neg b$ est donné par le tableau 1.8 :*

| A | B | $\Pi(A B = \neg b)$ |
|----------|----------|---------------------|
| $\neg a$ | $\neg b$ | 1 |
| $\neg a$ | b | 0 |
| a | $\neg b$ | 0.5 |
| a | b | 0 |

TAB. 1.8 – Distribution de possibilités sur A et B après conditionnement qualitatif par $B = \neg b$

1.6.4 Indépendance possibiliste

Étant donné que le conditionnement possibiliste est défini selon diverses façons et dépend du cadre du raisonnement considéré, il en est de même pour l'indépendance possibiliste [dCH99a,dCH99b,BABM02,BABD⁺00,BABM00,BABM01a]. Deux événements ϕ et ψ sont indépendants conditionnellement à φ , noté $\phi \perp \psi \mid \varphi$, si et seulement si ces événements vérifient :

$$\phi \perp \psi \mid \varphi \Leftrightarrow \Pi(\phi \mid \psi \wedge \varphi) = \Pi(\phi \mid \varphi) \text{ et } \Pi(\psi \mid \phi \wedge \varphi) = \Pi(\psi \mid \varphi). \quad (1.30)$$

L'indépendance entre les variables est vérifiée lorsque les égalités sont vérifiées pour toutes les réalisations de ces variables.

Un cas particulier d'indépendance possibiliste conditionnelle est retrouvé lorsque φ est un événement vide : c'est l'indépendance marginale :

$$\phi \perp \psi \Leftrightarrow \Pi(\phi \mid \psi) = \Pi(\phi) \text{ et } \Pi(\psi \mid \phi) = \Pi(\psi). \quad (1.31)$$

La notion d'indépendance peut aussi être exprimée en termes de mesures de nécessité. Un événement ϕ est indépendant de ψ si et seulement si [DFnDCHP94] :

$$N(\phi \mid \psi) = N(\phi) > 0. \quad (1.32)$$

1.6.4.1 Indépendance possibiliste quantitative

L'indépendance dans le cadre quantitatif est obtenue en utilisant le conditionnement basé sur le *produit* (voir équation 1.26).

Ainsi ϕ et ψ sont indépendants conditionnellement à φ si et seulement si [DMP98, DP91, Spo88b] :

$$\Pi(\phi \mid \psi \wedge \varphi) = \Pi(\phi \mid \varphi). \quad (1.33)$$

Cette équation est aussi équivalente à :

$$\Pi(\phi \wedge \psi \mid \varphi) = \Pi(\phi \mid \varphi) \cdot \Pi(\psi \mid \varphi). \quad (1.34)$$

Si φ est un événement vide alors l'équation précédente peut s'écrire :

$$\Pi(\phi \wedge \psi) = \Pi(\phi) \cdot \Pi(\psi). \quad (1.35)$$

Exemple 1.9 Reprenons les variables A et B de l'exemple 1.6. A et B sont dites indépendantes si toutes leurs réalisations vérifient l'équation (1.34), particulièrement l'équation (1.35). Or il existe au moins une réalisation dans laquelle cette dernière n'est pas vérifiée : par exemple, $\Pi(\neg a \wedge b) \neq \Pi(\neg a) \cdot \Pi(b)$

$$0.3 \neq 0.8 \cdot 1$$

Nous concluons que A et B sont dépendants.

1.6.4.2 Indépendance possibiliste qualitative

L'indépendance dans un raisonnement qualitatif est obtenue en utilisant le conditionnement basé sur l'opérateur *minimum* (voir équation 1.29). Ici on utilise la propriété de non-interactivité de Zadeh [Zad75]. ϕ et ψ sont indépendants dans le sens de la non-interactivité si et seulement si :

$$\Pi(\phi \wedge \psi) = \min(\Pi(\phi), \Pi(\psi)). \quad (1.36)$$

ϕ et ψ sont indépendants conditionnellement à φ et dans le sens de la non-interactivité si et seulement si :

$$\Pi(\phi \wedge \psi \mid \varphi) = \min(\Pi(\phi \mid \varphi), \Pi(\psi \mid \varphi)). \quad (1.37)$$

Exemple 1.10 D'après la distribution de possibilités des variables A et B donnée dans le tableau 1.6, on remarque que A et B sont dépendants au sens de la non-interactivité. En effet, il existe des réalisations de A et de B telles que $\Pi(\phi \wedge \psi) \neq \min(\Pi(\phi), \Pi(\psi))$. Par exemple : $\Pi(a \wedge \neg b) = 0.5 \neq \min(\Pi(a), \Pi(\neg b)) = \min(1, 0.8) = 0.8$

1.7 Logique possibiliste

La logique possibiliste [BP06, DP04b, BDGP02] peut être vue comme une extension de la logique classique qui permet un raisonnement en présence d'incertitude et de connaissances partiellement incohérentes. À chaque formule est associé un poids déterminant son niveau de priorité (de certitude) par rapport aux autres formules. Ainsi, les formules sont ordonnées selon leur poids, matérialisé par le degré de nécessité de chaque formule.

Une formule possibiliste est une paire (ϕ, α) où ϕ est une formule propositionnelle et α est le poids de la formule ϕ . (ϕ, α) interprète que ϕ est certaine avec au moins un degré α (c'est-à-dire que α est la borne inférieure du degré de nécessité de ϕ : $N(\phi) \geq \alpha$).

Une base de connaissances possibilistes [DLP94], notée \mathcal{B} , est un ensemble de formules logiques pondérées, défini comme suit :

$$\mathcal{B} = \{(\phi_i, \alpha_i) : i = 1, \dots, n\}.$$

Il est à noter qu'une formule possibiliste de poids 0 ($\alpha = 0$) n'est pas exprimée directement dans la base \mathcal{B} puisque $\forall \phi, N(\phi) \geq 0$.

Les bases de connaissances possibilistes sont des représentations réduites des distributions de possibilités. Plus précisément, on associe à chaque base de connaissances une distribution unique de possibilités [DLP94], notée $\pi_{\mathcal{B}}$, définie par : $\forall \omega \in \Omega$,

$$\pi_{\mathcal{B}}(\omega) = \begin{cases} 1 & \text{si } \forall (\phi_i, \alpha_i) \in \mathcal{B}, \omega \models \phi_i, \\ 1 - \max\{\alpha_i : (\phi_i, \alpha_i) \in \mathcal{B}, \omega \not\models \phi_i\} & \text{sinon.} \end{cases} \quad (1.38)$$

L'estimation du degré de possibilité associée à chaque interprétation revient à le maximiser à 1 si l'interprétation satisfait toutes les formules, ou bien à calculer le poids maximum des formules qui ne sont pas satisfaites par cette interprétation.

Exemple 1.11 Soit $\mathcal{B} = \{(a \vee \neg b, 0.7), (\neg a \vee b, 0.5), (a \vee b, 0.4)\}$ une base de connaissances possibiliste. La distribution de possibilités $\pi_{\mathcal{B}}$ associée à la base \mathcal{B} est :

$$\pi_{\mathcal{B}}(ab) = 1$$

$$\pi_{\mathcal{B}}(a\neg b) = 1 - \max\{0.5\} = 0.5$$

$$\pi_{\mathcal{B}}(\neg ab) = 1 - \max\{0.7\} = 0.3$$

$$\pi_{\mathcal{B}}(\neg a\neg b) = 1 - \max\{0.4\} = 0.6$$

Seule l'interprétation $\omega = ab$ satisfait toutes les formules de \mathcal{B} : son degré de possibilité vaut 1. Elle est compatible avec toute information disponible dans \mathcal{B} .

L'interprétation $\neg a\neg b$ est préférée à $a\neg b$ puisque la première falsifie la formule

$(a \vee b, 0.4)$ qui est moins certaine que la formule $(\neg a \vee b, 0.5)$ falsifiée quant à elle par l'interprétation $a \neg b$.

L'annexe A détaille davantage la logique possibiliste en présentant certaines de ces propriétés ainsi que l'opération de fusion entre les bases logiques.

1.8 Conclusion

Dans ce chapitre, nous avons présenté une synthèse des principales théories de l'incertain utilisées en Intelligence Artificielle. Nous nous sommes focalisés essentiellement sur deux théories de l'incertitude : la théorie des probabilités et la théorie des possibilités. Les représentations probabilistes de l'incertitude tirent profil de toute la puissance des fondements mathématiques sur lesquels elles se sont basées. Toutefois, une distribution de probabilités permet uniquement de représenter la notion d'indécision mais reste inadéquate lorsqu'il s'agit de représenter la notion d'ignorance totale. La théorie des possibilités représente un modèle naturel pour le traitement des informations incertaines et imprécises. Nous avons présenté les deux cadres du raisonnement possibiliste : le cadre qualitatif et le cadre quantitatif. La logique possibiliste, qui est une extension de la logique classique, constitue une alternative pour une représentation plus réduite des distributions de possibilités. Plusieurs notions cruciales, telles que le conditionnement, l'indépendance ont été présentées dans le cadre de la théorie des possibilités. Le chapitre suivant abordera les modèles graphiques probabilistes et possibilistes comme étant des structures graphiques et numériques pour la représentation de l'incertitude. Nous détaillerons, en particulier, les réseaux Bayésiens probabilistes et les réseaux possibilistes.

Chapitre 2

Les modèles graphiques probabilistes et possibilistes

2.1 Introduction

Les modèles graphiques [Pea88, Jen96, BK02, Lau96, Bis06, CGH96b], issus d'une combinaison entre la théorie de graphes [CGH96b, Bol02] et la théorie des probabilités [She97, Dur96], permettent l'acquisition, la représentation et l'utilisation des connaissances incertaines dans le contexte d'intelligence artificielle et/ou système expert. Plusieurs facteurs expliquent le succès de ces modèles. Premièrement, ces modèles graphiques représentent **visuellement** le contenu du système à modéliser ; ce qui permet aux intervenants de différents domaines (statisticiens, informaticiens, médecins, économistes,...) de collaborer ensemble et par suite de favoriser une meilleure communication et entente entre eux. Deuxièmement, les modèles graphiques expriment **naturellement** les relations de dépendance et d'indépendance, à travers des «liens», entre les variables constituant le système à modéliser. Troisièmement, ces modèles sont **modulaires** ; c'est-à-dire que certains problèmes et systèmes complexes peuvent être décrits et manipulés en les divisant en composants élémentaires.

Les réseaux bayésiens [Pea88, LS90, Jen01, NWL⁺07] constituent l'un des modèles graphiques probabilistes les plus utilisés et développés aussi bien du point de vue théorique que du point de vue pratique. Les réseaux bayésiens sont considérés comme l'une des techniques les plus intéressantes en intelligence artificielle, car ils permettent la représentation de la connaissance par un diagramme compréhensible dit d'influence [HJ84, Pea88, Sha86b, Sha88]. De plus, comme ils se basent sur des probabilités, le raisonnement dans les réseaux bayésiens intègre et prend en charge l'incertitude. Cependant, ces modèles ne permettent pas un raisonnement en présence de données numériques manquantes.

Afin de remédier à ce problème, les réseaux possibilistes [GK97, Fon94, BDGP99, BGK98] ont été développés comme une contrepartie des réseaux bayésiens dans le cadre de la théorie des possibilités. Les réseaux possibilistes se basent sur la théorie des possibilités pour représenter et raisonner, avec des informations incertaines ou imprécises ou même manquantes et s'adaptent aux deux cadres de raisonnement qualitatif et quantitatif.

Nous commençons ce chapitre par une présentation des modèles graphiques, leurs principales utilisations et les exemples les plus connus (la section 2.2). Deux exemples, en particulier, seront détaillés : les réseaux Bayésiens dans la section 2.3 et Les réseaux possibilistes dans la section 2.4. Un rappel des principales notions et définitions concernant la théorie des graphes est mis en annexe B.

2.2 Les modèles graphiques

D'une manière générale, un modèle graphique englobe une composante graphique qui est, selon le type du modèle, un graphe orienté acyclique DAG ou non orienté, et une composante numérique (par exemple, une distribution de probabilités dans le cas des modèles graphiques probabilistes). La composante graphique met en évidence les variables du domaine et les relations d'influence (causalité, dépendance conditionnelle, etc.) qu'elles entretiennent. La composante numérique sert, quant à elle, à quantifier, dans la théorie d'incertitude utilisée, les relations d'influences dans le modèle. Selon que le graphe du modèle est orienté ou non, on distingue les deux catégories suivantes des modèles graphiques :

- **Modèles graphiques orientés** : Dans ces modèles, les arcs sont orientés mais ne représentent pas obligatoirement des relations de causalité ; ils peuvent aussi représenter des dépendances conditionnelles entre les variables. Si les arcs représentent des relations de causalité, le modèle est alors appelé modèle graphique causal. Les réseaux Bayésiens [Pea88, LS90, Jen01, NWL⁺07], les réseaux possibilistes [GK97, BDGP99, BGK98, BS07b, BABM01b] et les modèles de Markov cachés [CMT05, Rab90, EM02] sont des exemples connus de modèles graphiques orientés.
- **Modèles graphiques non orientés** : également appelés modèles de Markov [Fin07, Pie03, BK02], les modèles graphiques non orientés sont très utilisés dans certains domaines comme la physique, la reconnaissance des formes, etc. L'impossibilité de former des cycles dans les modèles orientés est l'une des raisons justifiant l'usage des modèles non orientés. En effet, les arêtes dans les modèles de Markov expriment des dépendances conditionnelles et peuvent former n'importe quelle clique, ce qui permet une plus grande facilité de mo-

délisation dans certains problèmes.

2.2.1 Exemples d'utilisation des modèles graphiques

Il existe plusieurs domaines où les modèles graphiques servent comme moyen de modélisation. Parmi eux, nous citons la représentation de connaissances imparfaites, l'extraction de connaissances, la recherche d'information, la fusion des données, la détection d'intrusion et le raisonnement et l'inférence sous incertitude.

1. **La représentation des connaissances :** Les modèles graphiques conviennent particulièrement pour la représentation des connaissances incertaines [GH96, Coo99] où les variables du domaine établissent entre-elles des relations d'influence (de dépendance essentiellement). Les principaux avantages des représentations graphiques sont la modularité, le caractère intuitif et interprétable et la nature compacte de ces représentations. Dans l'exemple de la Figure 2.3, les connaissances concernant les influences entre les différents facteurs déterminants dans la maladie de l'obésité ainsi que ses conséquences, sont représentées de manière simple, intuitive et compréhensible.

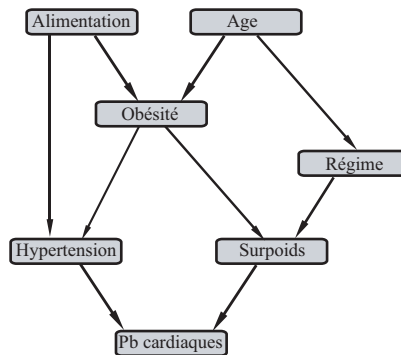


FIG. 2.1 – Exemple de modèle graphique représentant des relations d'influences

Cet exemple met en évidence les relations d'influence entre les différentes variables (quelques-unes) jouant un rôle essentiel dans la maladie de l'obésité. Selon le formalisme utilisé, ces influences sont quantifiées (numériquement) selon la théorie de l'incertitude utilisée par ce formalisme. Ainsi, dans le cadre de la théorie des probabilités, ce sont des distributions de probabilités locales qui sont utilisées pour quantifier l'effet exercé sur chaque variable par ses parents directs (par exemple, quantifier l'influence de la nature de l'alimentation et l'âge d'une personne sur les risques d'être atteint par l'obésité).

2. **L'extraction de connaissances et apprentissage automatique :** L'extraction de connaissances [Fay00, Kod95] est l'une des applications la plus

répandue qui se base sur les modèles graphiques. Elle permet l'exploitation de grandes quantités d'informations contenues dans de larges bases de données afin d'en extraire des connaissances implicites, précédemment inconnues et potentiellement intéressantes. Les classifieurs basés sur les modèles graphiques sont l'illustration la plus évidente de l'utilisation des modèles graphiques pour l'extraction de connaissances. L'existence en particulier, de plusieurs plates-formes de fouille de données et d'apprentissage automatique a permis de construire automatiquement des classifieurs bayésiens [BTN07, Zha04, FGG97, AOBY09] sur des données empiriques et de les utiliser directement pour classifier de nouvelles données.

L'apprentissage automatique [GWP88, DM83] est à l'origine d'un nombre important d'algorithmes permettant de découvrir des connaissances. Même si ces algorithmes sont variés, ils cherchent tous, en apprenant parmi un espace possible de concepts, à trouver une généralisation acceptable au sujet lié à l'extraction. Le problème est alors doublement complexe : éviter l'explosion combinatoire due à l'immensité du champ des généralisations possibles et sélectionner les bonnes généralisations [FPSM91]. Sans ces algorithmes d'apprentissage automatique, il est difficile pour un expert humain, dans plusieurs domaines, d'analyser et d'extraire des connaissances implicites et enfouies dans d'énormes quantités de données.

3. **La recherche d'information** : Dans un contexte où la quantité d'information disponible sur le réseau Internet ne cesse d'augmenter et vu la popularité du service web, les Systèmes de Recherche d'Information SRI constituent les principaux moyens d'accès aux informations. Un SRI [MRS08] regroupe l'ensemble des procédures et des opérations permettant la gestion, la représentation, l'interrogation, la recherche, le stockage et la sélection des informations répondant aux besoins d'un utilisateur. Dans un contexte documentaire, l'interrogation à l'aide d'une requête nécessite la représentation de cette dernière sous une forme compatible avec celle des documents. Les différents modèles standards de la RI (booléen, vectoriel, probabiliste) représentent les documents et les requêtes sous forme de listes de termes pondérés puis mesurent une valeur de pertinence en se basant sur ces termes et leurs poids. Pour être en mesure d'offrir aux utilisateurs les informations correspondant à leurs attentes, un SRI doit s'appuyer sur une technique de calcul de pertinence. La pertinence d'un document par rapport à une requête est un score calculé pour exprimer la corrélation entre ce document et la requête de l'utilisateur. Les informations qui auront le meilleur score de pertinence seront présentées aux utilisateurs par ordre décroissant. Il existe d'autres SRI qui modélisent les do-

cuments, les termes d'indexation, les requêtes et la pertinence en se basant sur les modèles graphiques. La première utilisation des réseaux bayésiens en Recherche d'Information est apparu dans les années 80 [Fri87] puis, elle s'est largement développée dans les travaux [SRNC⁺00, TC89, dCFLH03]. Des travaux plus récents [BBD05, BBD09, EEZA08, EEZA07, EEZA09] ont adopté une approche possibiliste pour la RI en proposant un SRI basé sur les réseaux possibilistes. Pour augmenter les performances de ces SRI, les auteurs de ces travaux utilisent les mesures de possibilité et de nécessité pour une meilleure détermination de la pertinence d'un document en réponse à une requête.

4. **La fusion des données** : La fusion de données [Tor03, Sap02, MBK⁺09] correspond à la volonté d'utiliser simultanément plusieurs sources de données différentes, ou de grouper des informations hétérogènes afin d'obtenir une nouvelle information de meilleure qualité. Cette nouvelle information est souvent destinée à la prise de décision.

Le problème de la combinaison et de l'utilisation simultanée de données et d'informations à partir de plusieurs sources se rencontre dans de nombreux champs d'application souvent liés au besoin de vouloir percevoir un environnement à partir de capteurs plus ou moins fiables et plus ou moins précis. La fusion de données est sollicitée pour des applications telles que le diagnostic médical, le commerce, la robotique, l'interprétation d'images (fusion d'images satellites ou médicales), le contrôle et le monitoring de processus et le contrôle et le commandement militaire. Tous ces domaines ont en commun le fait de devoir manipuler de grandes quantités de données de nature et de type très variés, afin d'obtenir une information de meilleure qualité.

Trois principales méthodes de fusion sont utilisées dans les systèmes complexes. L'une d'entre elles est basée sur l'utilisation de techniques de classification et plus généralement d'intelligence artificielle : modèles probabilistes, et théorie de l'information.

Les modèles graphiques forment un cadre particulièrement efficace pour la fusion de données. Ils permettent de modéliser un problème et d'utiliser des données en provenance de multiples sources, qu'elles soient qualitatives ou quantitatives. Ces données servent à mettre à jour la connaissance et les croyances que l'on a sur le problème en appliquant un formalisme d'inférence et d'avoir ainsi une nouvelle vue du problème, sachant les nouvelles données.

5. **La détection d'intrusion** : Une intrusion est une tentative non autorisée pour accéder ou manipuler des informations ou rendre un système non opérationnel ou inaccessible. La détection d'intrusions [Eva04, Shi00] concerne

l'analyse des événements ayant lieu dans un système dans le but de trouver en temps réel, en quasi temps réel ou en différé des tentatives d'accès non autorisé et aussitôt de notifier ces tentatives à l'administrateur du système. La détection d'intrusions est un problème qualifié par la complexité et l'incertitude : il se caractérise d'une part par sa grande complexité (complexité de modéliser le comportement d'une application, utilisateur, réseau, attaque, etc.) et d'autre part, par l'incertitude associée aux caractéristiques des données analysées qui peuvent être incomplètes ou incertaines. De plus, la nécessité de détecter en temps réel et d'anticiper la détection requiert des mécanismes à la fois rapides et capables de représenter et raisonner avec des observations incertaines ou manquantes. C'est ainsi que plusieurs modélisations du problème de détection d'intrusions et de sécurité informatique sont basées sur les modèles graphiques aussi bien pour la modélisation des connaissances que pour le raisonnement [BT08b,BT08a,JZA07,JZA09].

6. **Le traitement d'images** : Les modélisations probabilistes et les traitements statistiques peuvent s'avérer d'une remarquable efficacité dans le traitement de certains problèmes se posant en imagerie notamment dans la segmentation statistique des images. La segmentation d'image est une opération qui a pour but de rassembler des pixels entre eux suivant des critères pré-définis. Les pixels sont ainsi regroupés en régions ou zones constituant des parties de l'image. De manière générale, on peut envisager de faire appel à des méthodes probabilistes de traitement lorsque l'on se trouve dans une situation réunissant les deux conditions suivantes :

(i) on souhaite déterminer des caractéristiques de l'image, non observables directement et représentées par un élément x d'un espace X , à partir des caractéristiques de l'image observables et représentées par un élément y d'un espace Y ;

(ii) il n'existe pas de liens déterministes entre les y et les x (en d'autres termes, plusieurs x possibles peuvent correspondre à un y observé).

Les modèles de Markov cachés classiques et les méthodes bayésiennes de segmentation d'une part [BKZ96,BL95,BP00,FDP⁺03,Pie03] et la logique floue d'autre part [NAAB08,APBM10] sont appliquées avec toute leurs puissances dans le cadre de traitement statistique d'images.

7. **Le raisonnement et prise de décision** : La principale utilisation des modèles graphiques demeure le raisonnement et l'inférence. Cela concerne la déduction de certaines informations d'intérêt étant données d'autres informations disponibles (des observations par exemple). En effet, les modèles graphiques

sont dotés de processus d'inférence permettant de répondre à des questions d'intérêt afin de réaliser certaines tâches comme l'explication, le diagnostic et la classification.

D'une manière générale, l'inférence [Sha88, Mac03, DLN⁺90] consiste à propager une ou plusieurs informations certaines dans un modèle graphique afin de réviser les croyances concernant l'incertitude exprimée à travers le modèle. Dans le cadre probabiliste, il s'agit de calculer la probabilité a posteriori des variables X_i (dites variables d'intérêt) étant donnée une évidence E (ensemble de variables observées) avec $X_i \notin E$. Le résultat obtenu est la probabilité $P(X_i|E)$. Dans l'exemple de la Figure 2.3, on pourrait calculer la probabilité qu'une personne donnée soit obèse (la variable Obèse est à son état vrai) sachant la nature de son alimentation, son âge, etc. Ainsi, la valeur de probabilité obtenue servira pour une prise de décision dans un diagnostic médical. Le processus d'inférence permettra aussi d'expliquer ou de trouver la meilleure explication de l'évidence E , en fonction des états des autres variables. Dans le modèle de la Figure 2.3, on peut s'intéresser par exemple à déterminer l'âge, l'alimentation et le poids d'une personne sachant les valeurs des autres variables.

2.2.2 Exemples de modèles graphiques

Nous donnons, dans ce qui suit, une brève présentation des modèles graphiques les plus connus et les plus utilisés en pratique. Seuls les réseaux bayésiens et les réseaux possibilistes feront l'objet d'une analyse détaillée dans les prochaines sections.

2.2.2.1 Les réseaux Bayésiens

Appelés aussi réseaux de croyances, les réseaux bayésiens représentent un formalisme de modélisation et de raisonnement avec des connaissances incertaines et complexes. Formellement, un réseau Bayésien est un graphe orienté acyclique (*DAG* pour *Directed Acyclic Graph*) permettant de représenter les relations d'indépendances conditionnelles entre les variables du domaine. Chaque variable est munie d'une table de probabilités conditionnelles permettant de quantifier l'incertitude de ses relations d'influence. Un réseau Bayésien met en jeu donc deux aspects, le premier qualitatif (graphique) et l'autre quantitatif (numérique) :

- **La composante qualitative** : consiste en un graphe orienté acyclique noté $G(X, E)$ où X est l'ensemble des nœuds (chaque nœud représente une variable aléatoire X_i décrivant le domaine), et E est l'ensemble des arcs (un arc de X_i vers X_j représente une relation de dépendance directe entre eux).

- **La composante quantitative** : engendre une collection de tables de probabilités locales associées aux variables aléatoires X_i . Pour chaque variable X_i , un expert du domaine spécifie les probabilités conditionnelles de chaque instance de X_i sachant toutes les valeurs possibles de chaque parent de X_i .

L'inférence permet de propager des informations certaines (l'évidence) dans les réseaux bayésiens et de déduire n'importe quelle probabilité de variables d'intérêt sachant cette évidence. Il existe deux familles d'algorithmes d'inférence : les algorithmes d'inférence exacts qui fournissent des valeurs de probabilités exacts et les algorithmes d'inférence approximatifs qui fournissent des approximations des valeurs de probabilités réelles.

Les propriétés des réseaux bayésiens seront détaillées dans les sections suivantes ; quant à leurs algorithmes d'inférence, ils seront largement abordés dans le prochain chapitre.

2.2.2.2 Les réseaux possibilistes

Comme les réseaux Bayésiens, les réseaux possibilistes sont des graphes orientés mais utilisant la théorie des possibilités pour représenter et raisonner avec des informations incertaines, imprécises ou manquantes. A cause des différentes définitions du conditionnement possibiliste [BMCM02,DCK94,Fon97,His78,DP88, Van03], il existe deux types de réseaux possibilistes : les réseaux qualitatifs (basés sur le conditionnement utilisant l'opérateur minimum) et les réseaux quantitatifs (basés sur le conditionnement utilisant l'opérateur produit). L'inférence dans les réseaux possibilistes utilise le conditionnement possibiliste, la règle de chaînage et la marginalisation. La plupart des algorithmes d'inférence possibilistes sont des adaptations directes des algorithmes du cadre probabiliste. Le détail des réseaux possibilistes sera développé dans les prochaines sections. Leurs algorithmes d'inférence seront l'objet principal du chapitre suivant.

2.2.2.3 Les Chaînes de Markov

Supposons qu'on dispose d'un ensemble d'états $S = \{s_1, s_2, \dots, s_r\}$. Une chaîne de Markov (appelée aussi réseau de Markov) est représentée par un graphe orienté dont les noeuds correspondent aux r états du système observé et les arcs correspondent aux relations de dépendances conditionnelles entre les couples ordonnés d'états (i, j) . Le processus markovien démarre à partir de l'un de ces états, puis change successivement d'un état vers un autre. Chaque changement correspond à un pas. Si la chaîne est actuellement à un état s_i elle peut changer au prochain pas, vers l'état s_j avec une probabilité notée p_{ij} et appelée probabilité de transition. C'est la probabilité que le

système soit dans l'état s_j à l'instant $t + 1$ (prochain état) sachant qu'il se trouve dans l'état s_i à l'instant t (état actuel). Cette probabilité ne dépend en aucun cas de l'état qui précède s_i .

Ainsi nous pouvons décrire un tel système en donnant l'ensemble des r états possibles et une matrice P de dimensions $m \times m$ dont le terme p_{ij} est la probabilité que le processus soit dans l'état j à l'instant $t + 1$ étant donné qu'il se trouve dans l'état i à l'instant t . P est appelée matrice de transition du système. L'inférence dans les réseaux de Markov est semblable à l'inférence dans les réseaux bayésiens.

Exemple 2.2 La figure 2.4 représente une chaîne de Markov décrivant quelques états du ciel. Pour cet exemple on a considéré que 3 états du ciel $S = \{P, C, N\}$ correspondant respectivement aux états Pluvieux, Clair et Neigeux du ciel. L'observation de l'état du ciel se fait chaque jour. A chaque pas est associé une probabilité de transition marquée à côté de l'arc correspondant. Sachant que l'état actuel du ciel est Pluvieux (P), il y a une chance sur deux (probabilité de transition = 0.5) pour que l'état du ciel du lendemain soit aussi pluvieux (transition réflexive), un chance sur quatre (probabilité de transition = 0.25) qu'il soit neigeux ou clair. Le processus markovien associé à cette chaîne est pessimiste : il n'y a aucune chance qu'il y ait deux jours de suite où l'état du ciel est clair (pas de transition réflexive à l'état C). Donc, sachant que l'état actuel du ciel est Clair (C), l'état du ciel pour demain doit changer avec une probabilité équiprobable soit vers pluvieux ou neigeux. La matrice de transition P regroupe ces probabilités de transition :

$$P = \begin{pmatrix} 0.5 & 0.25 & 0.25 \\ 0.5 & 0 & 0.5 \\ 0.25 & 0.25 & 0.5 \end{pmatrix}$$

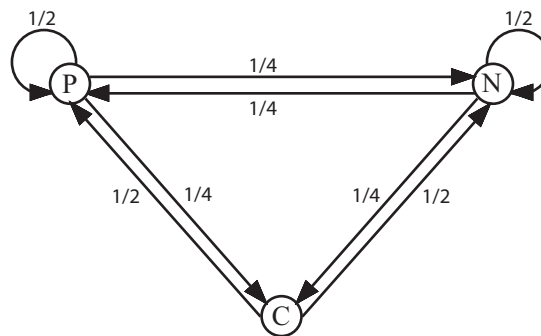


FIG. 2.2 – Exemple d'une chaîne de Markov à trois états

2.2.2.4 Les modèles de Markov cachés

Les modèles de Markov cachés [Rab90], abrégés dans la littérature par HMM (Hidden Markov Models), représentent les formes les plus simples des réseaux bayésiens dynamiques. Le qualificatif dynamique renvoie ici au fait que ces modèles servent à modéliser des processus dynamiques. Les HMM conviennent pour modéliser des processus séquentiels Markoviens, où l'état futur du système est indépendant de son passé étant donné l'état présent de ce système.

Les HMM sont des modèles orientés où il y a deux types de nœuds : les nœuds correspondant aux observations et les nœuds correspondant aux états du système qui sont, quant à eux, cachés. La variable d'état est inconnue (cachée) mais peut générer une sortie ou observation. Un état donné induit une observation selon une distribution de probabilités. De même, les transitions entre deux états successifs se font selon une distribution de probabilités.

2.3 Les réseaux Bayésiens

Dans cette section, nous nous intéressons de près au modèle des réseaux bayésiens [Pea88, LS90, Jen01, NWL⁺07]. Le rôle des graphes dans les modèles probabilistes est triple :

- Fournir un moyen efficace pour exprimer les relations de dépendance,
- Donner une représentation compacte des distributions de probabilités jointes,
- Faciliter l'inférence à partir d'observations.

Avant de définir à proprement dit les réseaux bayésiens, commençons par définir les critères qui régissent les relations de dépendance entre les variables (les nœuds) d'un réseau bayésien.

2.3.1 Les relations de dépendance

La circulation de l'information entre les nœuds d'un graphe dépend de la topologie de celui-ci. Trois cas de figure sont envisagés.

Connexion en série : Soit la situation de la figure 2.6. Le nœud A a une influence sur le nœud B qui a une influence sur le nœud C . L'information peut circuler de A vers C ou de C vers A à travers B , dans les deux sens, lorsqu'aucune variable n'est observée, en particulier B . Par contre, si B est connue, l'information ne circule plus entre A et C et ils deviennent indépendants. On dit dans ce cas, que A et C sont d-séparés étant donnée B , ou lorsque B est instanciée.



FIG. 2.3 – Connexion en série

Par exemple, si le temps était suffisamment ensoleillé (A), la récolte sera abondante (B). Si la récolte est abondante alors le prix du blé (C) sera bas. Si on connaît déjà que la récolte est abondante (B), le fait de connaître, en plus, que le temps a été ensoleillé (A) ne change pas notre idée sur le prix du blé (C). Le prix du blé (C) est alors indépendant de l'ensoleillement (A) sachant la bonne récolte (B).

Connexion divergente : Dans ce cas le nœud A est le parent des deux nœuds B et C . L'information peut passer entre les enfants de A lorsque celui-ci n'est pas instancié. Lorsque A est connu, ses deux enfants sont indépendants. Dans la figure 2.7, les enfants B et C sont dits d-séparés étant donnée A .

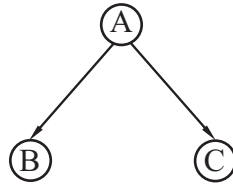


FIG. 2.4 – Connexion divergente

Par exemple, si je découvre, le matin, que la pelouse de mon jardin (B) est humide, j'ai tendance à croire qu'il a plu pendant la nuit (A) et par suite que la pelouse du jardin de mon voisin sera aussi humide (C). En ignorant la météo de la nuit précédente (A), l'état de la pelouse du voisin (C) dépend de l'état de celle de mon jardin (B).

Maintenant si je suis informé qu'il a plu cette nuit (A), l'idée (la pelouse est humide) que je me fait sur l'état de la pelouse de mon voisin (C) ne dépend plus de l'état de ma pelouse (B).

Connexion convergente : Dans ce cas de connexion telle que décrite dans la figure 2.8, lorsque le noeud fils (C) n'est pas instancié, les parents sont indépendants. Par contre, si l'état du fils est connu alors l'un des parents pourra passer l'information vers l'autre. L'information peut circuler dans une connexion convergente uniquement lorsque la variable de la connexion ou un de ses descendants est instancié.

Par exemple, le fait qu'un tremblement de terre se produit (A) est a priori indépendant que ma maison soit cambriolée (B). Le déclenchement de mon alarme (C)

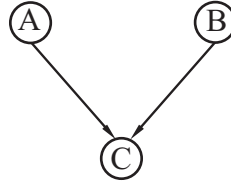


FIG. 2.5 – Connexion convergente

signifie que très probablement je viens d'être cambriolé (et je commence à m'inquiéter!). Le fait d'apprendre, à l'instant, qu'il y a eu un tremblement de terre (A) à proximité, me rassure (diminue mon inquiétude) sur l'éventualité de mon cambriolage (B). Étant donné que l'alarme s'est déclenchée (C), le tremblement de terre (A) et le cambriolage (B) sont dépendants l'un de l'autre.

Les situations décrites ci-dessus recouvrent les différentes manières possibles de transmettre l'information à travers un réseau bayésien. Elles ont été regroupées et généralisées dans un critère appelé **d-séparation** [Pea88, Jen96, GVP89, GVP90].

Définition 2.1 : (d-séparation) *Un chemin p est dit d-séparé (ou bloqué) par un ensemble Z de nœuds si et seulement si :*

1. *p contient une connexion de la forme $A_i \rightarrow A_k \rightarrow A_j$ ou $A_i \leftarrow A_k \rightarrow A_j$ telle que $A_k \in Z$,*
2. *ou p contient une connexion convergente de la forme $A_i \rightarrow A_k \leftarrow A_j$ telle que $A_k \notin Z$ ainsi que tous ces descendants.*

Un ensemble Z d-sépare X de Y si et seulement si Z bloque chaque chemin partant d'un nœud quelconque de X à un nœud quelconque de Y .

Dans une connexion en série $A_i \rightarrow A_k \rightarrow A_j$ ou divergente $A_i \leftarrow A_k \rightarrow A_j$, Si A_k est observé, alors les variables A_i et A_j qui étaient dépendantes conditionnellement à A_k deviennent indépendantes. Conditionner A_k bloque le flux d'information allant de A_i vers A_j c'est-à-dire qu'une nouvelle connaissance sur A_i ne pourra plus changer l'état de A_k puisque ce dernier est maintenant connu, et par suite A_i n'aura plus d'influence sur A_k . Constatation : A_i à travers A_k n'a plus d'influence sur A_j . Dans le cas d'une connexion convergente $A_i \rightarrow A_k \leftarrow A_j$ représentant deux parents ayant le même fils, le problème est inversé. Les deux parents sont indépendants jusqu'à ce qu'on observe un des états de leur fils commun. Ils deviennent alors dépendants.

La vérification du critère de la **d-séparation** entre les variables d'un réseau bayésien a pour but de déterminer les indépendances entre elles. En effet, deux variables sont d-séparées par un sous-ensemble de variables Z dans un réseau bayésien implique qu'elles sont indépendantes dans le contexte de Z .

Exemple 2.4 *Considérons le graphe de la figure 2.3. En supposant que la variable Obésité est **observée**, nous déduisons les constatations suivantes :*

- *Le fait de connaître l'âge ou la nature de l'alimentation de la personne obèse ne change rien à notre opinion sur son surpoids. Donc, la variable Alimentation (ou Age) est indépendante du Surpoids conditionnellement à la variable Obésité.*
- *Dans le même contexte, le fait de savoir que la personne souffre d'une hypertension artérielle ne modifie pas notre croyance quant à son surpoids. Donc, les variables hypertension et surpoids sont indépendantes conditionnellement à la variable Obésité.*
- *Pour une personne obèse, le fait de s'informer sur la nature de l'alimentation consommée par cette personne influence notre opinion sur son âge. Donc, Alimentation et Age sont dépendants conditionnellement à la variable Obésité.*

2.3.2 Les réseaux Bayésiens probabilistes

Définition 2.2 : *Soit p une distribution de probabilités jointe sur un ensemble de variables V , et $G = (V, E)$ un graphe acyclique orienté. Le couple (G, p) est appelé un réseau Bayésien si chaque variable $A_i \in V$ est conditionnellement indépendante de ses non-descendants dans le contexte de ses parents U_i . Ainsi, pour chaque variable V du graphe, les probabilités conditionnelles doivent satisfaire les contraintes de normalisation suivantes :*

- *Si $U_i = \emptyset$ (A_i est un nœud racine) alors la distribution marginale associée à A_i doit satisfaire :*

$$\forall a_i \in D_{A_i}, \sum_{a_i} P(a_i) = 1. \quad (2.1)$$

- *Si $U_i \neq \emptyset$ alors les distributions conditionnelles associées à A_i doivent satisfaire :*

$$\forall u_{a_i} \in \times D_{A_j \in U_{A_i}}, \sum_{a_i} P(a_i | u_{a_i}) = 1. \quad (2.2)$$

Un réseau bayésien est un formalisme qui permet de représenter de façon compacte une distribution jointe de probabilités sur un ensemble de variables en utilisant la notion d'indépendance. Les indépendances entre les variables permettent de décomposer cette large distribution en un ensemble de distributions locales associées à un petit nombre de variables. Cette approche permet une économie en espace mémoire et surtout le traitement de problèmes de tailles importantes.

Dans les réseaux bayésiens, l'équation du conditionnement (voir formule 1.6) se

transforme en une **règle de chaînage** permettant de calculer $P(V)$ d'une manière plus rapide et plus compacte vu les dépendances existantes entre les variables. Ainsi, la probabilité jointe est donnée par (la règle de chaînage) :

$$P(A_1, A_2, \dots, A_n) = \prod_{i=1}^n P(A_i | U_i). \quad (2.3)$$

Exemple 2.5 Soit le réseau bayésien composé de quatre nœuds : Age (A), Sexe (S), Fumeur (F) et Cancer (C), représenté dans la figure 2.9. La variable Age contient 3 états correspondants aux trois tranches d'âges : entre 15 et 40 ans (t_1), entre 40 et 60 ans (t_2) et plus de 60 ans (t_3). Les autres variables sont toutes binaires : Sexe = {masculin ou féminin}; Fumeur = {fumeur ou non-fumeur} et Cancer = {Cancéreux ou non-cancéreux}.

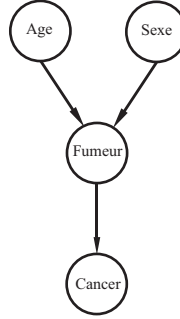


FIG. 2.6 – Exemple d'un réseau bayésien

Les relations d'influence dans ce réseau bayésien peuvent s'expliquer comme suit : pour une personne, la tranche d'âge à laquelle elle appartient ainsi que son sexe ont chacune une influence sur le fait qu'elle soit fumeuse ou non. De même, le fait d'être fumeur (ou non) a un effet direct sur la probabilité que la personne soit atteinte par la maladie du cancer. Les distributions des probabilités conditionnelles initiales de ce réseau sont données par la table 2.1.

| | | | | | | | | | | | | | | | | | |
|-------|------|---|-------|-----|----------|-----|-----|-----|----------|-----|-----|-----------|-----------|-----------|-----------|-----------|-----------|
| t_1 | 0.45 | <table><tr><td>s_1</td><td>0.5</td></tr><tr><td>s_2</td><td>0.5</td></tr></table> | s_1 | 0.5 | s_2 | 0.5 | | f | $\neg f$ | | | $t_1 s_1$ | $t_1 s_2$ | $t_2 s_1$ | $t_2 s_2$ | $t_3 s_1$ | $t_3 s_2$ |
| s_1 | 0.5 | | | | | | | | | | | | | | | | |
| s_2 | 0.5 | | | | | | | | | | | | | | | | |
| t_2 | 0.30 | c | 0.9 | 0.2 | f | 0.7 | 0.6 | 0.5 | 0.2 | 0.4 | 0.1 | | | | | | |
| t_3 | 0.25 | $\neg c$ | 0.1 | 0.8 | $\neg f$ | 0.3 | 0.4 | 0.5 | 0.8 | 0.6 | 0.9 | | | | | | |

TAB. 2.1 – Distributions de probabilités conditionnelles initiales

La distribution de probabilités jointe associée au réseau (selon la formule 2.3) est donnée par :

$$P(A, S, F, C) = P(A).P(S).P(F|A, S).P(C|F).$$

Ainsi le degré de probabilité associé à l'interprétation : "un individu de sexe masculin (s_1), de tranche d'âge (t_2), fumeur (f) et n'ayant pas de cancer ($\neg c$)" formulé en $t_2 \wedge s_1 \wedge f \wedge \neg c$ sera égal à :

$$P(t_2, s_1, f, \neg c) = P(t_2).P(s_1).P(f|t_2, s_1).P(\neg c|f) = 0.3 * 0.5 * 0.5 * 0.1 = 0.0075.$$

2.4 Les réseaux de croyances possibilistes

Les réseaux possibilistes [GK97, Fon94, BDGP99, BGK98, BS07b] sont des modèles graphiques orientés utilisant la théorie des possibilités pour représenter l'incertitude. Les travaux existant sur les réseaux possibilistes sont soit des adaptations directes de l'approche probabiliste [BDGP99], ou bien des méthodes d'apprentissage à partir de données imprécises [BGeRK00]. Commençons par donner la définition de base d'un réseau possibiliste.

2.4.1 Définition

Un réseau possibiliste, noté par Π_G , est caractérisé par une composante graphique et une composante numérique. La première est un graphe acyclique orienté (DAG) sur un ensemble de variables (nœuds) V . La structure du graphe représente l'ensemble des nœuds ainsi que les relations d'indépendance conditionnelle entre ces nœuds. La deuxième composante quantifie les liens du graphe en utilisant des distributions de possibilités conditionnelles de chaque nœud dans le contexte de ses parents. Cette distribution locale correspond à une distribution de possibilités marginale lorsque le nœud est une racine et de distribution de possibilités conditionnelles lorsque le nœud admet un ensemble non vide de parents.

Ces distributions de possibilités doivent vérifier la contrainte de normalisation. Pour chaque variable A_i :

- Si A_i est un nœud racine ($U_{A_i} = \emptyset$), la distribution marginale associée à A_i doit satisfaire :

$$\forall a_i \in D_{A_i}, \max_{a_i} \Pi(a_i) = 1. \quad (2.4)$$

- Si A_i n'est pas un nœud racine ($U_{A_i} \neq \emptyset$), la distribution conditionnelle de A_i dans le contexte de ses parents doit satisfaire :

$$\forall u_{a_i} \in \times D_{A_j \in U_{A_i}}, \max_{a_i} \Pi(a_i|u_{a_i}) = 1. \quad (2.5)$$

Selon [Fon94], le critère de d-séparation reste aussi valable pour les réseaux possibilistes. Cela permet à ces derniers de tenir compte des indépendances conditionnelles

entre les variables afin d'optimiser les calculs des distributions de possibilités jointes.

Exemple 2.6 La figure 2.10 représente un exemple d'un réseau possibiliste noté Π_G . Le tableau 2.2 fournit les distributions de possibilités conditionnelles initiales associées aux variables binaires A, B, C, D et E .

Il est facile de vérifier que les contraintes de normalisation (les équations 2.4 et 2.5) sont respectées dans les distributions de possibilités conditionnelles initiales. Par exemple, $\max(\pi(c|a), \pi(\neg c|a)) = \max(0.2, 1) = 1$.

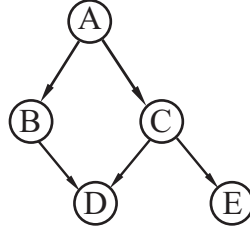


FIG. 2.7 – Exemple de réseau possibiliste Π_G

| | | |
|----------|-----|----------|
| | a | $\neg a$ |
| a | 1 | |
| $\neg a$ | 0.3 | |

| | | |
|----------|-----|----------|
| | a | $\neg a$ |
| b | 1 | 0.4 |
| $\neg b$ | 0 | 1 |

| | | |
|----------|-----|----------|
| | a | $\neg a$ |
| c | 0.2 | 1 |
| $\neg c$ | 1 | 0.3 |

| | | | | |
|----------|------|-----------|-----------|----------------|
| | bc | $b\neg c$ | $\neg bc$ | $\neg b\neg c$ |
| d | 1 | 0 | 0.8 | 0.1 |
| $\neg d$ | 1 | 1 | 1 | 1 |

| | | |
|----------|-----|----------|
| | c | $\neg c$ |
| e | 1 | 0.1 |
| $\neg e$ | 0.4 | 1 |

TAB. 2.2 – Distributions de possibilités conditionnelles associées à Π_G

Le chapitre précédent évoque plusieurs définitions proposées pour le conditionnement possibiliste. Parmi elles, on distingue deux définitions sur lesquelles deux types de réseaux possibilistes se sont basés : les réseaux possibilistes basés quantitatifs et les réseaux possibilistes qualitatifs.

2.4.2 Réseaux possibilistes quantitatifs

Les réseaux possibilistes basés sur le produit [Fon94,BGK98,GK97] (appelés également réseaux quantitatifs) sont très comparables en termes de conditionnement et d'indépendance aux réseaux bayésiens. Un réseau possibiliste basé sur le produit, noté par Π_{G_p} , est un DAG dont les distributions de possibilités conditionnelles associées sont obtenues par le conditionnement *produit* (formule 1.26). La distribution des possibilités jointe des réseaux possibilistes basés sur le produit, notée par π_{G_p} , sur

un ensemble de variables $V = \{A_1, A_2, \dots, A_n\}$ est obtenue par la règle de chaînage :

$$\pi_{G_p}(A_1, A_2, \dots, A_n) = \prod_{i=1}^n \pi(A_i | U_i). \quad (2.6)$$

Notons que cette règle de chaînage est très similaire à celle des réseaux bayésiens (équation 2.3) dans le sens où les deux équations utilisent l'opérateur produit. Les réseaux possibilistes quantitatifs et les réseaux bayésiens probabilistes se différencient au niveau de la marginalisation. Dans le cas des réseaux possibilistes, la marginalisation est réalisée en utilisant l'opérateur *maximum* (voir équations 2.4 et 2.5) tandis que l'opérateur *somme* (voir équations 2.1 et 2.2) est utilisé dans les réseaux probabilistes.

Exemple 2.7 *Considérons maintenant le réseau possibiliste de l'exemple 2.6 comme un réseau quantitatif, alors la distribution de possibilités jointe associée à ce réseau est définie par : $\pi_{G_p}(A, B, C, D, E) = \Pi(A) \cdot \Pi(B|A) \cdot \Pi(C|A) \cdot \Pi(D|B, C) \cdot \Pi(E|C)$. Les valeurs de la distribution de possibilités jointe π_{G_p} sont données dans le tableau 2.3.*

| A | B | C | D | E | π_{G_p} | A | B | C | D | E | π_{G_p} |
|-----|----------|----------|----------|----------|-------------|----------|----------|----------|----------|----------|-------------|
| a | b | c | d | e | 0.2 | $\neg a$ | b | c | d | e | 0.12 |
| a | b | c | d | $\neg e$ | 0.08 | $\neg a$ | b | c | d | $\neg e$ | 0.048 |
| a | b | c | $\neg d$ | e | 0.2 | $\neg a$ | b | c | $\neg d$ | e | 0.12 |
| a | b | c | $\neg d$ | $\neg e$ | 0.08 | $\neg a$ | b | c | $\neg d$ | $\neg e$ | 0.048 |
| a | b | $\neg c$ | d | e | 0 | $\neg a$ | b | $\neg c$ | d | e | 0 |
| a | b | $\neg c$ | d | $\neg e$ | 0 | $\neg a$ | b | $\neg c$ | d | $\neg e$ | 0 |
| a | b | $\neg c$ | $\neg d$ | e | 0.1 | $\neg a$ | b | $\neg c$ | $\neg d$ | e | 0.0036 |
| a | b | $\neg c$ | $\neg d$ | $\neg e$ | 1 | $\neg a$ | b | $\neg c$ | $\neg d$ | $\neg e$ | 0.036 |
| a | $\neg b$ | c | d | e | 0 | $\neg a$ | $\neg b$ | c | d | e | 0.24 |
| a | $\neg b$ | c | d | $\neg e$ | 0 | $\neg a$ | $\neg b$ | c | d | $\neg e$ | 0.096 |
| a | $\neg b$ | c | $\neg d$ | e | 0 | $\neg a$ | $\neg b$ | c | $\neg d$ | e | 0.3 |
| a | $\neg b$ | c | $\neg d$ | $\neg e$ | 0 | $\neg a$ | $\neg b$ | c | $\neg d$ | $\neg e$ | 0.12 |
| a | $\neg b$ | $\neg c$ | d | e | 0 | $\neg a$ | $\neg b$ | $\neg c$ | d | e | 0.0009 |
| a | $\neg b$ | $\neg c$ | d | $\neg e$ | 0 | $\neg a$ | $\neg b$ | $\neg c$ | d | $\neg e$ | 0.009 |
| a | $\neg b$ | $\neg c$ | $\neg d$ | e | 0 | $\neg a$ | $\neg b$ | $\neg c$ | $\neg d$ | e | 0.009 |
| a | $\neg b$ | $\neg c$ | $\neg d$ | $\neg e$ | 0 | $\neg a$ | $\neg b$ | $\neg c$ | $\neg d$ | $\neg e$ | 0.09 |

TAB. 2.3 – Distribution de possibilités jointe associée à Π_{G_p}

2.4.3 Réseaux possibilistes qualitatifs

Un réseau possibiliste basé sur le minimum [BABK03,BABM01b] (appelé réseau qualitatif), noté par Π_{G_m} , est un DAG dont les distributions de possibilités conditionnelles associées sont obtenues par le conditionnement *minimum* (formule 1.29). La distribution de possibilités jointe des réseaux possibilistes basées sur le minimum, notée par π_{G_m} , sur un ensemble de variables $V = \{A_1, A_2, \dots, A_n\}$ est obtenue par la règle de chaînage suivante :

$$\pi_{G_m}(A_1, A_2, \dots, A_n) = \min_{i=1, \dots, n} \pi(A_i | U_i). \quad (2.7)$$

Exemple 2.8 Si on considère le réseau possibiliste de l'exemple 2.6 comme un réseau qualitatif, alors la distribution de possibilités jointe associée à ce réseau est définie par : $\pi_{G_m}(A, B, C, D, E) = \min(\Pi(A), \Pi(B|A), \Pi(C|A), \Pi(D|B, C), \Pi(E|C))$.

Le tableau 2.4 donne les valeurs de la distribution de possibilités jointe π_{G_m} correspondante aux différentes interprétations de l'univers Ω .

| A | B | C | D | E | π_{G_m} | A | B | C | D | E | π_{G_m} |
|-----|----------|----------|----------|----------|-------------|----------|----------|----------|----------|----------|-------------|
| a | b | c | d | e | 0.2 | $\neg a$ | b | c | d | e | 0.3 |
| a | b | c | d | $\neg e$ | 0.2 | $\neg a$ | b | c | d | $\neg e$ | 0.3 |
| a | b | c | $\neg d$ | e | 0.2 | $\neg a$ | b | c | $\neg d$ | e | 0.3 |
| a | b | c | $\neg d$ | $\neg e$ | 0.2 | $\neg a$ | b | c | $\neg d$ | $\neg e$ | 0.3 |
| a | b | $\neg c$ | d | e | 0 | $\neg a$ | b | $\neg c$ | d | e | 0 |
| a | b | $\neg c$ | d | $\neg e$ | 0 | $\neg a$ | b | $\neg c$ | d | $\neg e$ | 0 |
| a | b | $\neg c$ | $\neg d$ | e | 0.1 | $\neg a$ | b | $\neg c$ | $\neg d$ | e | 0.1 |
| a | b | $\neg c$ | $\neg d$ | $\neg e$ | 1 | $\neg a$ | b | $\neg c$ | $\neg d$ | $\neg e$ | 0.3 |
| a | $\neg b$ | c | d | e | 0 | $\neg a$ | $\neg b$ | c | d | e | 0.3 |
| a | $\neg b$ | c | d | $\neg e$ | 0 | $\neg a$ | $\neg b$ | c | d | $\neg e$ | 0.3 |
| a | $\neg b$ | c | $\neg d$ | e | 0 | $\neg a$ | $\neg b$ | c | $\neg d$ | e | 0.3 |
| a | $\neg b$ | c | $\neg d$ | $\neg e$ | 0 | $\neg a$ | $\neg b$ | c | $\neg d$ | $\neg e$ | 0.3 |
| a | $\neg b$ | $\neg c$ | d | e | 0 | $\neg a$ | $\neg b$ | $\neg c$ | d | e | 0.1 |
| a | $\neg b$ | $\neg c$ | d | $\neg e$ | 0 | $\neg a$ | $\neg b$ | $\neg c$ | d | $\neg e$ | 0.1 |
| a | $\neg b$ | $\neg c$ | $\neg d$ | e | 0 | $\neg a$ | $\neg b$ | $\neg c$ | $\neg d$ | e | 0.1 |
| a | $\neg b$ | $\neg c$ | $\neg d$ | $\neg e$ | 0 | $\neg a$ | $\neg b$ | $\neg c$ | $\neg d$ | $\neg e$ | 0.3 |

TAB. 2.4 – Distribution de possibilités jointe associée à Π_{G_m}

2.5 Conclusion

Dans ce chapitre nous avons présenté les modèles graphiques comme outils puissants pour la représentation et le raisonnement avec des connaissances incertaines et complexes. L'une des caractéristiques essentielles des modèles graphiques est la modélisation intuitive et modulaire des relations d'indépendance (influence, causale, etc.) entre les variables constituant le graphe. De plus, les modèles graphiques offrent un cadre adéquat pour faire des raisonnements, appliquer l'inférence, réaliser une fusion des données, ou extraire des connaissances, etc.

La plupart des modèles graphiques se basent sur la théorie des probabilités pour exprimer et manipuler l'incertitude dans les données modélisées ; d'autres se fondent sur la théorie des possibilités. Parmi les modèles graphiques les plus utilisés, nous citons les réseaux bayésiens probabilistes et les réseaux possibilistes. Chacun de ces deux réseaux comporte une composante graphique orientée sans cycle et une composante numérique représentant l'incertitude selon la théorie (de probabilité ou de possibilité) mise en œuvre.

L'inférence consiste à propager l'effet d'une information observée dans le réseau (appelée évidence) sur le reste du graphe. Dans un cadre possibiliste, l'inférence est, en général, un problème de décision, de la forme "est-ce qu'il y a $\Pi(A = a) = x$?", appartenant à la classe des problèmes NP-difficile [Coo90]. Le problème d'inférence cherche à conclure, à partir des informations observées dans le réseau, dans divers sens : observer les effets pour remonter jusqu'aux "causes", ou bien observer les "causes" et en déduire les conséquences. Le prochain chapitre va clarifier comment se propage une évidence dans les réseaux probabilistes et possibilistes.

Chapitre 3

Les algorithmes de propagation probabilistes et possibilistes

3.1 Introduction

L'inférence est l'une des composantes principales de la logique. Littéralement, l'inférence sert à faire connaître l'existence d'une chose qu'on ne perçoit pas directement, en prenant pour moyen de raisonnement un indice perceptible ayant pour spécificité incontestable d'être en concordance réelle avec cette chose. Par exemple, la fumée a incontestablement pour spécificité d'être en concordance invariable avec le feu.

La plupart des modèles graphiques probabilistes ont été développés pour résoudre certains problèmes de prédiction et rétrodiction en Intelligence Artificielle. Dans le cas d'un réseau Bayésien, le problème général d'inférence (de propagation) est de calculer la probabilité conditionnelle $P(X_i \mid Y)$ avec $Y \subset V$ et $X_i \notin Y$. Généralement, Y est un ensemble d'observations (Evidence) et X_i est une variable d'intérêt décrivant le problème étudié. Deux classes principales d'algorithmes d'inférence ont été développées [GH02] : l'approche exacte et l'approche approximative.

L'approche exacte est basée sur une technique de propagation de messages et elle était, au départ, limitée aux réseaux sous la forme d'un polyarbre (graphe orienté sans boucles). Dans cette technique, on associe à chaque nœud un mécanisme qui permet d'envoyer des messages de façon asynchrone à ses voisins jusqu'à ce qu'un équilibre soit atteint, en un nombre fini d'étapes. Cette méthode a été depuis étendue à n'importe quelle topologie de réseaux pour donner l'algorithme de l'Arbre de jonction introduit par les auteurs de [LS90] puis amélioré par [Jen96]. Cet algorithme transforme, d'abord, le réseau initial en supprimant les boucles afin d'obtenir une structure en polyarbre de cliques. Ensuite, il procède à un échange de messages entre les cliques, identique à celui de propagation des messages. Le coût de l'algo-

l'arbre de jonction est déterminé par la taille de la plus grande clique et est exponentiel en espace par rapport à la taille de cette plus grande clique. Dans certains cas où les graphes sont à connexions multiples, l'inférence exacte peut s'avérer impraticable.

L'inférence approximative est une alternative de l'approche exacte lorsque celle-ci devient inapplicable. Les méthodes approximatives de propagation fournissent des valeurs rapprochées aux probabilités calculées.

Au cours de ce chapitre nous allons aborder, dans le cadre probabiliste et possibiliste, le principe de base des algorithmes exacts et approximatifs les plus abordés dans la littérature. Il est à noter que la plupart des algorithmes de propagation possibilistes sont des adaptations directes des algorithmes probabilistes tenant compte du cadre de raisonnement qualitatif ou quantitatif. En conséquence, le contenu de ce chapitre est organisé en deux sections : La première aborde les algorithmes probabilistes exacts et approximatifs. La deuxième traite de la propagation dans les réseaux possibilistes.

3.2 Propagation dans les réseaux probabilistes

Dans cette section nous faisons le tour d'horizon de l'ensemble des principaux algorithmes, exacts et approximatifs, de propagation probabilistes abordés dans la littérature.

3.2.1 Les algorithmes exacts de propagation

Dans cette section, on s'intéresse particulièrement aux algorithmes exacts de propagation probabilistes, les plus utilisés dans le cas des DAGs simplement connectés (polyarbres) et les DAGs à connexions multiples.

3.2.1.1 Propagation dans les polyarbres

La méthode de propagation de l'évidence (dite algorithme de Pearl) dans les réseaux bayésiens, dont la structure est un polyarbre, est la méthode de propagation exacte la plus ancienne [Pea88, CGH96b, YFW03, Wei00]. Elle exploite le fait que les polyarbres sont des graphes *simplement connectés* pour propager l'évidence dans le réseau selon une manière unique. En effet, dans les polyarbres il existe un chemin unique entre deux nœuds quelconques, ce qui implique que chaque nœud A_i du graphe divise le polyarbre en deux sous-graphes disjoints : le premier englobe les nœuds accessibles à travers les parents du nœud A_i et le deuxième englobe les nœuds accessibles à travers les enfants du nœud A_i .

Dans ce qui suit, nous allons présenter la version centralisée de l'algorithme de Pearl [Pea88, CGH96b] dans les polyarbres probabilistes. Dans cet algorithme l'impact de l'arrivée d'une partie de l'évidence sur un nœud est vu comme une perturbation qui a pour conséquence l'échange de messages entre ce nœud et ses voisins (parents et enfants). Il existe deux types de messages : les μ -messages envoyés des parents vers leurs enfants et les λ -messages envoyés des enfants vers leurs parents. Un μ -message représente l'effet de l'évidence observée sur un nœud, collecté à partir du sous-graphe accessible depuis le parent de ce nœud. De même, un λ -message représente l'effet de l'évidence observée sur un nœud, collecté à partir du sous-graphe accessible depuis l'enfant de ce nœud.

Ainsi, la propagation de l'évidence peut s'appliquer d'une manière efficace en combinant, au niveau d'un nœud A_i , les informations provenant de chaque sous-graphe puis de les envoyer vers l'autre sous-graphe par un échange local de messages.

• Décomposition de l'évidence :

Dans un polyarbre, il existe au plus un chemin entre deux nœuds. Ce qui signifie qu'un nœud quelconque A_i divise le polyarbre en deux sous-polyarbres disjoints : l'un contient les nœuds accessibles à travers les parents de A_i , l'autre contenant les nœuds accessibles à travers les enfants de A_i . Soit E un ensemble de nœuds observés (évidence) instancié par la valeur e . Notons par $U_{A_i} = \{U_1, U_2, \dots, U_n\}$ l'ensemble des parents de A_i et par $Y_{A_i} = \{Y_1, Y_2, \dots, Y_m\}$ l'ensemble des enfants de A_i . Pour un nœud quelconque A_i l'ensemble E peut être partitionné en deux sous-ensembles :

- $E_i^+ = \{E_{U_1 A_i}^+, \dots, E_{U_n A_i}^+\}$: correspond au sous-ensemble de E que l'on peut atteindre à partir des n parents $\{U_1, \dots, U_n\}$ de A_i et instancié par e_i^+ .
- $E_i^- = E/E_i^+ = \{E_{A_i}, E_{A_i Y_1}^-, \dots, E_{A_i Y_m}^-\}$: correspond au sous-ensemble de E que l'on peut atteindre à partir des m enfants $\{Y_1, \dots, Y_m\}$ de A_i , instancié par e_i^- et à lui même (A_i) instancié par e_{A_i} .

La figure 3.1 illustre les différentes partitions obtenues précédemment.

• Expressions des messages échangés :

Étant donné la valeur de l'évidence e , nous souhaitons calculer $\forall a_i \in D_{A_i}$ la probabilité conditionnelle $P(a_i|e)$ qu'on dénotera par $Bel(a_i)$. Cette probabilité dépend essentiellement des deux expressions notées $\lambda(a_i)$ et $\mu(a_i)$. Les valeurs de $\lambda(a_i)$ et $\mu(a_i)$ sont calculées à leur tour respectivement à partir des λ -messages reçus des enfants du nœud A_i et des μ -messages reçus des parents du nœud A_i .

Ainsi, $\forall a_i \in D_{A_i}$

$$Bel(a_i) = P(a_i | e) = P(a_i | e_i^+, e_i^-) = \alpha \cdot \lambda(a_i) \cdot \mu(a_i). \quad (3.1)$$

avec α est un coefficient de normalisation.

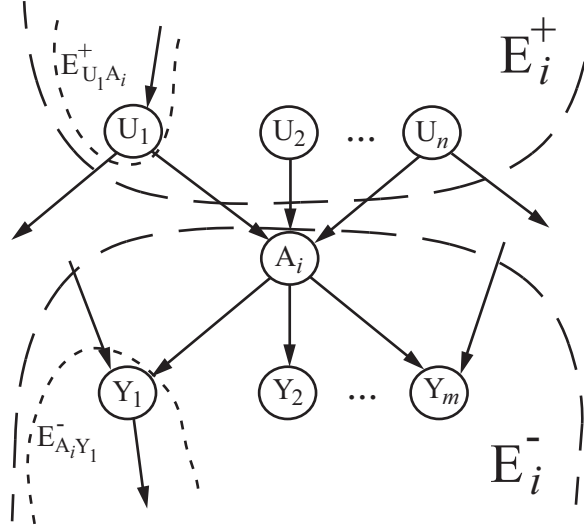


FIG. 3.1 – Partition de l'ensemble E en deux sous-ensembles E_i^+ et E_i^-

L'expression $\lambda(a_i)$ est calculée relativement à chaque instance $\forall a_i \in D_{A_i}$ de la manière suivante :

$$\lambda(a_i) = P(a_i \mid e_i^-) = \lambda_{A_i}(a_i) \cdot \prod_{j=1}^m \lambda_{Y_j}(a_i). \quad (3.2)$$

avec $\lambda_{A_i}(a_i)$ correspond à l'évidence locale associée au nœud A_i et $\lambda_{Y_j}(a_i)$ est le message reçu par A_i de la part de son enfant Y_j .

La valeur de μ -message relative à chaque instance de $a_i \in D_{A_i}$ est définie comme suit :

$$\mu(a_i) = P(a_i \mid e_i^+) = \sum_u P(a_i \mid u) \cdot \prod_{k=1}^n \mu_{A_i}(u_k). \quad (3.3)$$

où, $\mu_{A_i}(u_k)$ est le message reçu par le nœud A_i de son parent U_k

A chaque réception d'un μ -message de son parent U_k , $k \in \{1, \dots, n\}$, le nœud A_i lui réplique par un λ -message correspondant à :

$$\lambda_{A_i}(u_k) = P(e_{U_k A_i}^- \mid u_k) = \beta \cdot \sum_{a_i} \lambda(a_i) \left[\sum_{u_l: l \neq k} P(a_i \mid u) \cdot \prod_{l \neq k} \mu_{A_i}(u_l) \right], \quad (3.4)$$

De la même façon, le nœud A_i répond au message envoyé de la part de son enfant Y_j par un μ -message :

$$\mu_{Y_j}(a_i) = P(a_i \mid e_{A_i Y_j}^+) = \alpha \cdot \lambda_{A_i}(a_i) \cdot \prod_{r=1, r \neq j}^m \lambda_{Y_r}(a_i). \quad (3.5)$$

• **L'algorithme de propagation :**

Les auteurs de [PS91] ont proposé une version centralisée de l'algorithme de propagation dans les polyarbres. En effet, dans l'algorithme proposé par Pearl [Pea88], chaque nœud est visité au plus une fois pour chaque nœud observé. Cette version améliorée diminue le nombre de messages puisque chaque nœud est visité au plus deux fois quel que soit le nombre de nœuds observés. Le nombre de messages est moins que deux fois le nombre de nœuds. Soit s la taille de la plus grande table de probabilité conditionnelle, et n le nombre des nœuds alors la complexité théorique de l'algorithme est $O(s * n)$.

Les principales étapes de cet algorithme sont résumées succinctement à travers les points suivants :

- *Choix du nœud pivot* : Dans cette étape on doit choisir, arbitrairement, un nœud pivot qui déterminera le sens d'envoi des messages pendant les phases de collecte et de distribution d'information.
- *Initialisation* : Tous les messages sont initialisés à 1. Pour chaque nœud racine A , $\forall a \in D_A$ les messages $\mu(a)$ sont initialisés par $P(a)$. Pour chaque nœud observé A , si A est observé avec la valeur a alors $\lambda_A(a)$ est initialisé à 1, sinon à 0.
- *La collecte d'informations* : Durant cette phase, chaque nœud (autre que pivot) calcule d'abord ses λ et μ valeurs à l'aide des formules (3.2) et (3.3) puis passe un message à son nœud voisin en direction du pivot. Si ce voisin est un parent, le message à envoyer sera calculé selon la formule (3.4) sinon il sera calculé selon la formule (3.5). Cette phase commence avec le nœud le plus éloigné du pivot choisi.
- *La distribution d'informations* : Durant cette phase, chaque nœud calcule d'abord ses λ et μ expressions à travers les formules (3.2) et (3.3); ensuite, il envoie à chacun de ses voisins (selon la direction du Pivot vers les feuilles du polyarbre) un message dont le type dépend de ce voisin. Si le voisin est un parent le message est calculé selon la formule (3.4) sinon il sera calculé selon la formule (3.5).
- *Marginalisation* : Pour chaque nœud non observé A , calculer $Bel(a) = P(a | e)$ selon la formule (3.1).

Exemple 3.1 Soit le graphe donné par la figure 3.2. On suppose que le nœud A est observé à son état a_1 , c'est à dire que $A = a_1$. Le nœud A est choisi comme pivot. Dans la même figure on montre, aussi, les différents messages qui circulent au niveau de ce polyarbre afin de propager l'évidence $E = \{A\}$. Les messages sont échangés dans l'ordre suivant : $\lambda_D(C)$ puis $\mu_C(B)$ et $\lambda_C(A)$ (pendant la phase de collecte d'information) et $\mu_C(A)$ puis $\lambda_C(B)$ et $\mu_D(C)$ (pendant la phase de distribution

d'information).

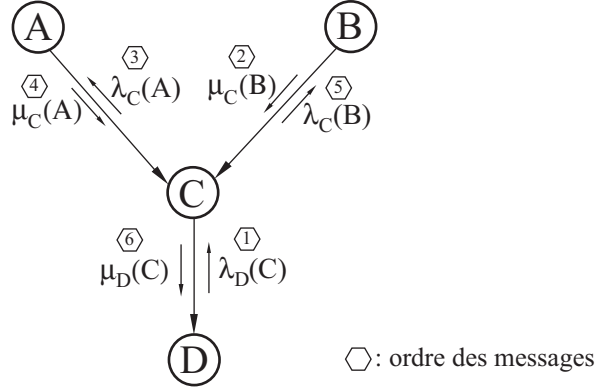


FIG. 3.2 – Messages échangés dans le polyarbre lors de la propagation d'évidence

3.2.1.2 Propagation dans les DAGs à connexions multiples

Dans les graphes à connexions multiples, il peut exister plusieurs chaînes entre deux nœuds (nous sommes dans le cas de présence de boucles). L'application de l'algorithme de Pearl sur ce type de réseaux n'est plus appropriée : les messages risquent de circuler indéfiniment dans le graphe à cause des boucles, ce qui ne garantit plus l'exactitude des résultats. L'idée dans ce cas est de transformer le graphe initial en une structure équivalente simplement connectée afin de pouvoir appliquer l'algorithme de propagation. Cette structure secondaire est un arbre de cliques, appelé arbre de jonction. Chaque nœud dans l'arbre de jonction est un ensemble de variables appelé clique, et chaque lien est étiqueté par l'intersection de deux cliques adjacentes, appelé séparateur. L'une des propriétés essentielles que vérifie un arbre de jonction est la *propriété d'intersection courante* [CDLS03] définie comme suit :

Définition 3.1 (Propriété d'intersection courante) : Soit Z_1, Z_2, \dots, Z_k des sous-ensembles de $N = \{1, 2, \dots, n\}$, tel que $N = \bigcup_{i=1}^k Z_i$ et AJ un arbre de jonction. Pour tout sous-ensembles de variables X_{Z_i}, X_{Z_j} tel que $X_{Z_i} \cap X_{Z_j} = T_{ij} \neq \emptyset$, il existe un chemin dans AJ qui connecte X_{Z_i} et X_{Z_j} tel que tous ses nœuds contiennent tous les indices de T_{ij} .

Dans cette définition, les $X_{Z_i}, i = 1, \dots, k$ représentent les cliques de l'arbre de jonction et les T_{ij} les groupes (séparateurs et cliques) intermédiaires séparant X_{Z_i} de X_{Z_j} . La propriété d'intersection courante indique que si une variable est commune entre deux cliques (pas nécessairement adjacentes), alors il existe un chemin entre ces deux cliques dans lequel la variable commune apparaît dans les cliques et séparateurs

(intermédiaires) de ce chemin. Cette propriété est importante pour la propagation dans les arbres de jonction.

Exemple 3.2 Soit l'arbre de jonction de la figure 3.3 composé des trois cliques C_1 , C_2 et C_3 . Les deux cliques C_1 et C_3 ne sont pas adjacentes avec $T_{13} = C_1 \cap C_3 = \{C\}$. Si cet arbre de jonction vérifie la propriété d'intersection courante, cela signifie que chaque séparateur ou clique situé sur le chemin qui relie les deux cliques C_1 et C_3 contiendra $T_{13} = \{C\}$. En effet, sur la même figure, $\{C\}$ vérifie bien qu'il appartient au séparateur $\{BC\}$, à la clique C_2 et au séparateur $\{CD\}$.

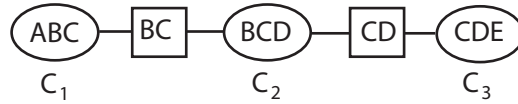


FIG. 3.3 – Exemple d'un arbre de jonction formé par trois cliques

Dans ce qui suit nous présentons les étapes de transformation d'un DAG à connexions multiples en un arbre de jonction. Quant à la propagation dans les arbres de jonction elle sera détaillée ultérieurement dans son adaptation aux réseaux possibilistes (voir la sous-section 3.3.1.1).

• **Les étapes de construction de l'arbre de jonction :**

Les principales étapes de transformation d'un DAG à connexions multiples vers un arbre de jonction sont : la moralisation, la triangulation et la construction de l'arbre de jonction.

- *La moralisation* : La moralisation d'un graphe consiste à passer d'un graphe orienté à un graphe non orienté. Pour moraliser le graphe, on doit marier les parents d'un même sommet, puis supprimer l'orientation du graphe. Cette étape peut s'effectuer en temps linéaire $O(n)$. Dans la figure 3.4 tous les arcs du graphe orienté sont devenus des arêtes. De plus les deux parents du nœud E ont été liés par une nouvelle arête.

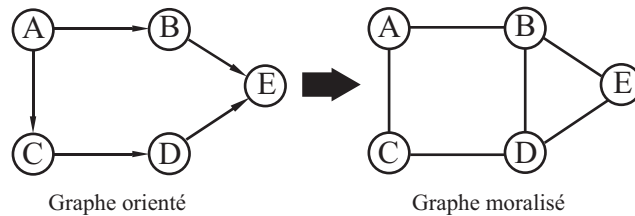


FIG. 3.4 – Etape de moralisation d'un graphe

- *La triangulation* : Un graphe est triangulé si tout cycle de longueur supérieur à 3 admet une "corde" (une arête qui relie deux nœuds non adjacents du cycle pour former un sous-graphe connexe). La triangulation du graphe consiste à rendre triangulé un graphe moralisé et non triangulé. La triangulation d'un graphe n'est pas unique et la recherche de la triangulation optimale (au sens du nombre minimum d'arêtes ajoutées) est un problème NP-Complet. Dans la figure 3.5 une corde a été ajoutée entre les nœuds A et D pour obtenir un graphe triangulé. Notons que cette corde aurait pu être ajoutée entre les nœuds B et C .

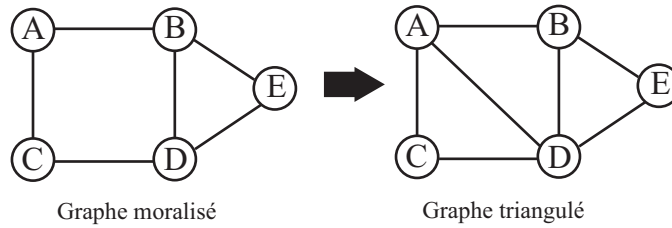


FIG. 3.5 – Etape de triangulation d'un graphe

- *Construction de l'arbre de jonction* : Cette étape consiste à transformer le graphe triangulé en un arbre de jonction. Une clique est un sous graphe connecté maximal. Suite à la triangulation, le graphe est constitué de cliques. La construction de l'arbre de jonction est réalisée en reliant ces cliques par des séparateurs. Dans la figure 3.6 le graphe triangulé peut se subdiviser en trois sous-graphes connectés maximaux. Chacun d'entre eux se transformera en clique dans l'arbre de jonction. Les séparateurs lient les cliques adjacentes. Les cliques sont dessinées comme des ellipses et les séparateurs comme des rectangles. Ainsi, nous pouvons vérifier la propriété d'intersection courante dans cet arbre de jonction. En effet, $C_1 \cap C_3 = \{D\}$ qui appartient respectivement à $S_{12} = \{A, D\}$, $C_2 = \{A, B, D\}$ et $S_{23} = \{BD\}$ situés sur le chemin entre C_1 et C_3 (voir figure 3.6).

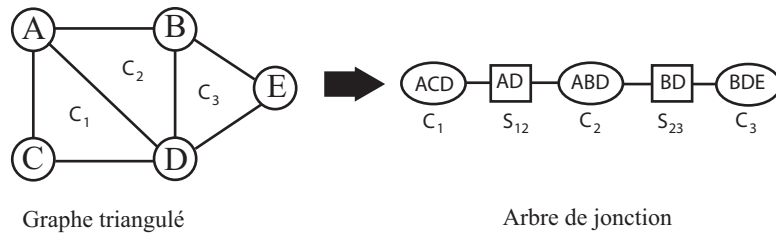


FIG. 3.6 – Etape de construction de l'arbre de jonction

La construction de l'arbre de jonction constitue une étape préliminaire afin de procéder efficacement à la propagation des observations. Cette propagation se fait en trois étapes :

- *Initialisation de l'arbre de jonction* : une fois l'arbre de jonction construit, cette étape permet de le quantifier en transformant les distributions de probabilités conditionnelles initiales en des distributions jointes locales rattachées aux cliques et aux séparateurs.
- *Propagation globale* : Une fois l'arbre de jonction initialisé, le processus de propagation globale permettra d'assurer sa "cohérence globale" à travers un passage de messages entre les cliques comparable à celui réalisé par l'algorithme de propagation dans les polyarbres.
- *Réponse aux requêtes* : À partir d'un arbre de jonction globalement cohérent, il est possible de calculer la distribution de probabilités associée à toute variable $A \in V$, en utilisant toute clique contenant A . La complexité de cette étape a un coût linéaire.

3.2.1.3 Autres méthodes d'inférence exactes

Ensemble de coupe

L'algorithme *Loop Cutset Conditioning* a été introduit par [Pea86]. Dans cette méthode, la connectivité du réseau est changée en instantiant un certain sous-ensemble de variables appelé l'ensemble de coupe (*loop cutset*). Dans le réseau résultant, l'inférence est effectuée en utilisant l'algorithme des messages locaux. Puis les résultats de toutes les instantiations sont combinés par leurs probabilités *a priori*. La complexité de cet algorithme augmente exponentiellement en fonction de la taille de l'ensemble de coupe.

Inversion d'arcs

Cette méthode a été introduite par [Sha86a]. Il s'agit de changer la direction des arcs en utilisant la règle de Bayes jusqu'à ce que le réseau soit transformé de telle manière que les nœuds de la requête soient des enfants des nœuds observés [Hen90]. Cette méthode est à privilégier soit lorsque l'ensemble des nœuds de la requête est restreint soit lorsque nous savons par avance que les mêmes nœuds vont être interrogés assez souvent en ayant à chaque fois les mêmes nœuds observés.

Élimination de variables

L'algorithme élimination de variables [ZP94] supprime les variables une par une après avoir sommé sur celles-ci. Cette méthode a été généralisée dans [Dec98] par l'algorithme *Bucket Elimination*. Un ordre des variables doit être donné en entrée et sera alors l'ordre d'élimination des variables. Le nombre de calculs dépend alors de

cet ordre puisqu'il influe sur la taille des facteurs futurs. Trouver le meilleur ordre équivaut au problème de trouver l'arbre de plus petite largeur dans le réseau [Dec98], ce qui est un problème NP-difficile. Cette méthode est avantageuse lorsqu'un ordre d'élimination des variables est déjà connu ou si le réseau est peu dense mais avec de nombreux circuits.

Méthodes symboliques

L'inférence probabiliste symbolique (*Symbolic Probabilistic Inference*) a été introduite dans [SDD90,LD94]. Cette méthode est orientée par un but : n'effectuer que les calculs nécessaires pour répondre à la requête. Des expressions symboliques peuvent être obtenues en remettant à plus tard l'évaluation des expressions, et en les gardant sous forme symbolique.

Par ailleurs, les auteurs de [CGH96a] ont proposé une autre technique d'inférence symbolique en modifiant les méthodes existantes d'inférences numériques et en remplaçant les paramètres initiaux par des paramètres symboliques.

Ces méthodes ont le désavantage qu'il est difficile de calculer et de simplifier automatiquement des expressions symboliques mais, l'avantage qu'elles orientent les calculs intelligemment.

Méthodes différentielles

Les méthodes différentielles transforment un réseau Bayésien en polynôme multivarié [Dar00]. Elles calculent ensuite les dérivées partielles de ce polynôme. Il est alors possible d'utiliser ces dérivées pour calculer les réponses à de nombreuses requêtes, et cela en temps constant. Cette méthode est très utile lorsque régulièrement les mêmes requêtes sont effectuées.

3.2.2 Les algorithmes approximatifs de propagation

Les algorithmes approximatifs sont des alternatives des algorithmes exacts lorsque ces derniers deviennent inapplicables. Cette incapacité est due essentiellement aux structures complexes des réseaux à inférer d'une part, et aux valeurs infinitésimales (proches de 0) qui peuvent être attribuées a priori aux nœuds d'autre part. Parmi les algorithmes approximatifs, on distingue :

- Les algorithmes qui utilisent les méthodes exactes appliquées sur des topologies de réseaux approchées (modifiées) ;
- Les algorithmes qui utilisent des méthodes stochastiques, méthodes qui effectuent des estimations en générant des échantillons pour calculer des probabilités.

Nous nous intéressons essentiellement dans ce qui suit aux algorithmes approximatifs d'inférence qui se basent sur des méthodes stochastiques.

3.2.2.1 Les algorithmes d'échantillonnage par importance

La plupart des algorithmes stochastiques peuvent être généralisés dans le cadre de l'échantillonnage par importance [Sri02, CD00] (*Importance Sampling*) dans lequel les échantillons sont indépendants les uns des autres.

L'échantillonnage par importance fournit une alternative aux les méthodes exactes dans le calcul des intégrales multidimensionnelles ou des sommes complexes. C'est une technique de réduction de variance dans l'estimation des intégrales, employée dans les méthodes de *Monte Carlo*. L'idée de base dans l'échantillonnage par importance est que certaines valeurs des variables aléatoires ont plus d'effet sur la valeur de l'intégrale estimée. Si ces valeurs sont tirées assez fréquemment alors la variance de l'intégrale estimée peut être réduite.

► Fondements mathématiques

Explicitement, soit $f(X)$ une fonction à n variables avec $X = (X_1, \dots, X_n)$ sur un domaine Ω . Si on considère le problème d'estimation de l'intégrale :

$$I = \int_{\Omega} f(X) dX,$$

en supposant que l'intégrale I existe, l'échantillonnage par importance aborde ce problème en réécrivant l'intégrale I comme :

$$I = \int_{\Omega} \frac{f(X)}{g(X)} g(X) dX,$$

où $g(X)$, appelée la fonction d'importance, est une fonction de densité de probabilité sur le domaine Ω satisfaisant la condition que pour tout $\forall X f(X) \neq 0 \Rightarrow g(X) \neq 0$. Ainsi, on génère n échantillons indépendants s_1, s_2, \dots, s_n appartenant tous à Ω à partir de la fonction de densité de probabilité $g(X)$ pour estimer l'intégrale I par :

$$\hat{I}_n = \frac{1}{n} \sum_{i=1}^n \frac{f(s_i)}{g(s_i)},$$

et la variance de \hat{I} est estimée par :

$$\hat{\sigma}^2(\hat{I}_n) = \frac{1}{n(n-1)} \sum_{i=1}^n \left(\frac{f(s_i)}{g(s_i)} - \hat{I}_n \right)^2.$$

Rubinstein [Rub81] prouve que si $f(X) > 0$ alors la fonction optimale d'importance est $g(X) = \frac{f(X)}{I}$.

Dans la pratique, on ne cherche pas à déterminer avec exactitude la fonction optimale d'importance car trouver cette fonction revient à calculer I . Cependant si on arrive à trouver une fonction qui est très proche de la fonction optimale, on pourra

s'attendre à de bons taux de convergence dans l'estimation de I .

Les résultats précédents restent aussi valides pour des variables discrètes et dans ce cas, l'intégrale sera remplacée par une sommation.

► **Un algorithme générique d'échantillonnage par importance pour réseau bayésien**

La méthode échantillonnage par importance est appliquée dans les réseaux probabilistes de structure complexe afin d'estimer les probabilités a posteriori. Etant donné un réseau bayésien sur un ensemble de nœuds V , la marginalisation de l'ensemble E des nœuds observés instancié par e nécessite le calcul de la somme de toutes les probabilités $P(V \setminus E, E = e)$:

$$P(E = e) = \sum_{V \setminus E} P(V \setminus E, E = e)$$

Cette équation est similaire à celle de I : l'intégrale est remplacée par l'opérateur somme et l'univers Ω est remplacé par $V \setminus E$.

Comparé au calcul des intégrales multidimensionnelles, l'application de l'échantillonnage par importance pour le calcul de l'inférence dans les réseaux bayésiens, possède plusieurs avantages : tout d'abord, la distribution jointe $P(X)$ est décomposable et peut être factorisée en des composants distincts. Ensuite, le réseau lui-même offre une structure claire montrant les dépendances conditionnelles entre les nœuds.

La probabilité a posteriori $P(X = x | E = e)$ peut être obtenue en calculant, respectivement, $P(x, e)$ et $P(e)$ puis le rapport peut être fait selon la formule de probabilité conditionnelle :

$$P(x | e) = \frac{P(x, e)}{P(e)}.$$

En calculant la probabilité a posteriori et pour augmenter la précision du résultat obtenu par la méthode échantillonnage par importance, il est recommandé de choisir deux fonctions d'importance pour estimer chacune des probabilités $P(x, e)$ et $P(e)$. Cependant, ce choix engendre des complexités supplémentaires pendant le calcul. En pratique, c'est la même fonction d'importance qui est utilisée pour l'estimation des deux probabilités [CD00].

L'algorithme 3.1 résume les principales étapes d'un algorithme générique d'échantillonnage par importance dont la plupart des algorithmes existants basés sur l'échantillonnage, s'inspirent. Cette version d'algorithme est appelée *forward sampling*, dans laquelle les échantillons sont générés selon l'ordre topologique des nœuds du réseau (les parents d'un nœud sont placés avant lui-même).

La complexité d'un tel algorithme est fortement liée à celle du choix de la fonction d'importance. Cette complexité se développe quadratiquement en fonction de

1. Classer les nœuds selon leur ordre topologique.
2. Initialisation de la fonction d'échantillonnage $P^0(V \setminus E)$, le nombre souhaité d'échantillons m , l'intervalle de mise à jour l et le vecteur score de chaque nœud.
3. $k = 0, T \leftarrow \emptyset$.
4. **Pour** i de 1 à m **faire**
5. **Si** $(i \bmod l == 0)$ **alors**
6. $k \leftarrow k + 1$
7. Mettre à jour la fonction d'importance $P^k(V \setminus E)$ selon T
8. **FinSi**
9. $s_i \leftarrow$ générer un échantillon selon $P^k(V \setminus E)$.
10. $T = T \cup \{s_i\}$.
11. Calculer le score $(s_i, P(V \setminus E, e), P^k(V \setminus E))$ et l'ajouter dans l'entrée correspondante de chaque vecteur de score selon les états instanciés.
12. **FinPour**
13. Normaliser le vecteur score de chaque nœud.

Algorithme 3.1 - Algorithme générique d'échantillonnage par importance

la somme des degrés estimés [YPSPL05].

► Classification des algorithmes d'échantillonnage par importance pour les Réseaux Bayésiens

La différence principale entre les variantes existantes de l'algorithme d'échantillonnage stochastique résident dans la manière dont ils implémentent les étapes 2, 7 et 8 de l'algorithme générique présenté plus haut. Une classification de ces algorithmes en trois familles a été établie par [Cha06] :

- Première famille : utilise la distribution a priori du réseau bayésien comme fonction d'importance. Cette famille inclut les algorithmes : *Probabilistic logic sampling* [Hen88] et *Likelihood weighting* [FC89, SP89].
- Deuxième famille : calcule la fonction d'importance selon la combinaison de la distribution a priori et l'évidence. Cette famille inclut les algorithmes : *Backward sampling* [FdF94], *Importance Sampling* [SP89] et *Evidence Pre-propagation Importance Sampling algorithm (EPIS-BN) algorithm* [CD03].
- Troisième famille : utilise des méthodes d'apprentissage et d'heuristique pour concevoir et mettre à jour la fonction d'importance. Cette famille inclut *Self-importance sampling* [SP89] et *Adaptive importance sampling (AIS-BN) algorithm* [CD00].

3.2.2.2 Les méthodes de Monte-carlo par chaînes de Markov

Les méthodes MCMC (pour *Markov Chains Monte Carlo*) [GRS96] permettent d'échantillonner des variables aléatoires en construisant une chaîne de Markov. Ces méthodes statistiques approchées ne calculent pas exactement les lois marginales, mais en donnent une estimation. Elles permettent donc de traiter des applications de grande taille en temps raisonnable, ce qui n'est pas le cas des méthodes exactes. Le principe est ici d'effectuer un certain nombre de tirages aléatoires compatibles avec une loi. Pour cela, la décomposition de la loi jointe en produit de lois conditionnelles permet de mener les calculs. En pratique, chaque variable dont tous les parents sont connus est tirée aléatoirement, jusqu'à ce que toutes les variables aient été simulées. Pour que l'estimation soit fine, il faut alors effectuer un très grand nombre de simulations.

Par ailleurs, ces méthodes sont inefficaces quand certaines probabilités sont très faibles mais elles possèdent l'avantage d'être facilement implémentables et de ne pas être fermées : plus l'algorithme est arrêté tard, plus de cas seront simulés et plus l'évaluation des résultats sera précise.

Les principaux algorithmes des méthodes MCMC sont :

- L'algorithme de Metropolis-Hastings [Has70] : il produit, selon une démarche aléatoire, une série de valeurs selon une « loi de proposition » puis utilise une méthode pour accepter ou rejeter les valeurs « des mouvements » proposées.
- L'échantillonneur de Gibbs [GG84] : il exige la connaissance de toutes les lois de la distribution conditionnelle de chaque variable afin de construire une base d'exemples à partir de laquelle il sera possible d'évaluer les probabilités recherchées.

3.2.2.3 L'algorithme *Loopy Belief Propagation*

Belief Propagation [Pea88] (ou algorithme de propagation de messages) est une méthode exacte d'inférence appliquée sur des modèles graphiques simplement connectés. Dans un réseau bayésien, son rôle consiste à déterminer la probabilité a posteriori de n'importe quel nœud A_i conditionnellement à un ensemble E des nœuds observés par une valeur $E = e$. Cette probabilité conditionnelle, notée $\forall a_i \in D_{A_i}, Bel(a_i) = P(A_i = a_i \mid E = e)$, se calcule en effectuant un échange de messages entre les nœuds voisins afin de propager l'évidence sur l'ensemble des nœuds du réseau.

Toutefois, cet algorithme de propagation ne s'applique que sur des polyarbres. Avec une légère modification dans les formules de propagation, il est possible d'appliquer l'algorithme *Belief propagation* sur des réseaux à connexions multiples (avec boucles). L'algorithme ainsi obtenu est appelé *Loopy Belief Propagation* [MWJ99,

Wei00, AOBY07] puisqu'il s'applique sur des réseaux avec boucles (*loops*).

L'idée de base de l'algorithme *Loopy Belief Propagation* (LBP) est de laisser les messages circuler dans le réseau à connexions multiples jusqu'à atteindre un état de convergence (s'il y en a). Autrement dit, l'algorithme LBP va itérer les étapes de propagation des messages à travers le réseau jusqu'à atteindre un état où cette propagation des messages ne modifie plus aucune probabilité dans chaque nœud : c'est l'état de convergence. Arrivé à cet état, l'algorithme LBP ne fournit pas (toujours) les valeurs exactes des probabilités a posteriori, mais donne une approximation de celle-ci.

En pratique, plusieurs applications réelles ont prouvé l'utilisation efficace de l'algorithme LBP dans le cas d'inférence et des problèmes d'optimisation. Parmi les applications les plus connues, nous citons :

- Analyse d'images [TIT03] : étant donné des mesures sur la coloration et le voisinage de certains pixels d'une image, on doit classifier les pixels de celle-ci en calculant la probabilité qu'un pixel appartient à une forme donnée sachant les mesures précédentes.
- Code de correction d'erreur ou Turbo-code [BGT93, MMC98] : Dans le contexte de transmission d'informations sur des supports sensibles aux bruits et perturbations, l'algorithme LBP est utilisé comme technique pour décoder les codes LDPC (*Low Density Parity Check codes* pour codes à faible densité).

En ce qui nous concerne, l'algorithme *Loopy Belief Propagation* sera largement détaillé dans le prochain chapitre et une adaptation de son principe au cadre possibiliste quantitatif et qualitatif sera proposée par nos soins.

3.2.2.4 Les méthodes de simplification de modèle

Ces méthodes simplifient d'abord le réseau probabiliste jusqu'à la faisabilité d'un algorithme exact puis applique ce dernier pour calculer les probabilités a posteriori. Ces méthodes sont classées en deux groupes :

- Le premier groupe exploite le fait que certaines dépendances du réseau sont faibles et qu'il est possible de les négliger. En d'autres termes, s'il existe un arc entre deux variables, exprimant une dépendance faible, nous pouvons le supprimer. L'idée de ces méthodes est de supprimer de tels arcs rendant les deux variables indépendantes : le calcul sera accéléré et la taille des tables des probabilités conditionnelles (CPT) sera réduite. Parmi ces méthodes, on distingue :
 - celles qui introduisent des transformations : *Variational approach* [JJ99] ;
 - celles qui réduisent la complexité du modèle : *Annihilating small probabilities* [JLO90] ;

- celles qui réduisent les dépendances faibles : *Removal of weak dependencies* [Kja94].
 - celles qui réduisent la cardinalité de la CPT : *State space abstraction* [WIL94].
- Le deuxième groupe opère un peu différemment : il s'agit de conserver le graphe d'origine, mais de n'effectuer la propagation de l'évidence que sur une partie du réseau ou bien sur tout le réseau et de ne considérer qu'une partie des CPT. Parmi ces méthodes nous citons :
- *Bounded conditioning* [HSC89] : qui place une limite inférieure pour les probabilités ;
 - *Incremental algorithm* [D'A93] : qui évalue partiellement le CPT du réseau ;
 - *Context specific approximation* [Poo98] : qui tient compte de la structure contextuelle du réseau.

3.2.2.5 Autres méthodes d'inférence approchées

Méthodes variationnelles

Les méthodes variationnelles [Bea03] ont été introduites pour approcher l'apprentissage par maximum de vraisemblance en présence d'une base d'exemples incomplète. Elles sont alors utilisées pour approcher les intégrales nécessaires pour l'inférence bayésienne ou encore l'apprentissage bayésien.

Le principe de l'inférence variationnelle consiste à approcher la loi de probabilité $p(x)$ par une famille de lois $q(x|\mu)$, puis trouver μ^* tel que $p(x)$ et $q(x|\mu^*)$ soient les plus proches possibles. On réalise alors l'inférence sur la loi $q(x|\mu^*)$ plutôt que sur $p(x)$. La difficulté dans cette méthode est comment choisir $q(x|\mu)$.

Méthodes de recherche de masse

Ces méthodes supposent qu'une petite partie de l'espace de définition des variables contient une grande partie de la masse de probabilité. Elles recherchent alors les instantiations de haute probabilité et les utilisent pour obtenir une bonne approximation [Hen90, Poo93]. Cette méthode est à utiliser lorsque nous voulons négliger les événements de probabilités faibles.

3.3 Propagation dans les réseaux possibilistes

Dans cette section nous présentons les algorithmes de propagation dans les réseaux possibilistes quantitatifs et qualitatifs. La plupart de ces algorithmes sont des adaptations directes des algorithmes de propagation probabilistes. D'autres algorithmes ont été développés entièrement pour certains réseaux possibilistes. Parmi ces algorithmes nous citons celui développé pour le cas des réseaux hybrides (quan-

titatifs) qui combinent les réseaux possibilistes avec la logique possibiliste pour des représentations compactes des distributions de possibilités au niveau des nœuds. Un autre algorithme de propagation, dit "anytime", est appliqué sur les réseaux possibilistes qualitatifs. Cet algorithme transforme le graphe initial en une structure secondaire (dont la complexité est beaucoup moins coûteuse que l'arbre de jonction), puis applique une procédure de stabilité pour atteindre un état de cohérence entre les cliques du graphe obtenu.

3.3.1 Propagation exacte dans les réseaux possibilistes quantitatifs

3.3.1.1 Adaptation de l'arbre de jonction aux réseaux quantitatifs

L'algorithme arbre de jonction adapté aux réseaux possibilistes quantitatifs [Sma07] se réalise en quatre phases mentionnées ci-dessous :

- Transformation du graphe initial en un arbre de jonction,
- Initialisation de l'arbre de jonction,
- Propagation globale,
- Réponse aux requêtes.

Vu que la première phase (de transformation) a été abordée dans le cadre probabiliste, nous ne détaillerons dans cette section que les trois autres. Pour une bonne compréhension de chacune de ces étapes, considérons l'exemple suivant :

Exemple 3.3 Soit le réseau possibiliste quantitatif à connexions multiples ΠG_p présenté dans la figure 3.7. Les distributions initiales de possibilités conditionnelles associées aux cinq nœuds binaires sont données dans les tableaux 3.1 et 3.2.

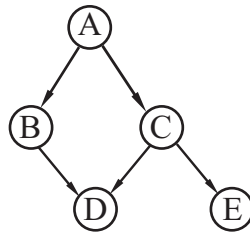


FIG. 3.7 – Exemple de réseau possibiliste quantitatif à connexions multiples

La distribution jointe de possibilités liée au réseau ΠG_p est calculée en utilisant la formule de la règle de chaînage basée sur le produit (formule 2.7). Le tableau 3.3 donne les valeurs de cette distribution jointe.

| A | $\Pi(A)$ | B | A | $\Pi(B A)$ | C | A | $\Pi(C A)$ | E | C | $\Pi(E C)$ |
|----------|----------|----------|----------|------------|----------|----------|------------|----------|----------|------------|
| a | 1 | b | a | 1 | c | a | 0.5 | e | c | 0.5 |
| $\neg a$ | 0.2 | b | $\neg a$ | 0.2 | c | $\neg a$ | 1 | e | $\neg c$ | 1 |
| | | $\neg b$ | a | 0 | $\neg c$ | a | 1 | $\neg e$ | c | 1 |
| | | $\neg b$ | $\neg a$ | 1 | $\neg c$ | $\neg a$ | 0.4 | $\neg e$ | $\neg c$ | 0.5 |

TAB. 3.1 – Distributions initiales de possibilités conditionnelles associées à ΠG_p

| D | B | C | $\Pi(D B, C)$ |
|----------|----------|----------|---------------|
| d | b | c | 1 |
| d | b | $\neg c$ | 0.5 |
| d | $\neg b$ | c | 0.1 |
| d | $\neg b$ | $\neg c$ | 0.2 |
| $\neg d$ | b | c | 0.3 |
| $\neg d$ | b | $\neg c$ | 1 |
| $\neg d$ | $\neg b$ | c | 1 |
| $\neg d$ | $\neg b$ | $\neg c$ | 1 |

TAB. 3.2 – Distributions initiales de possibilités conditionnelles associées à ΠG_p

• **Phase 1 : Transformation en arbre de jonction.** La première phase de l'algorithme qui assure la transformation du graphe initial ΠG_p en un arbre de jonction, notée AJ , est identique à celle du cadre probabiliste mentionnée dans la section (3.2.1.2). Elle se décompose en trois étapes [LS90, Jen96] :

- La moralisation : supprimer l'orientation du graphe puis marier le parents communs d'un nœud.
- La triangulation : rendre triangulée toute boucle de longueur supérieure à trois en lui ajoutant une arête.
- La construction de l'arbre de jonction : relier les cliques (sous-graphes connectés maximaux) adjacentes entre elles par des séparateurs.

Exemple 3.4 Les différentes phases de transformation du réseau possibiliste ΠG_p de l'exemple 3.3 vers un arbre de jonction AJ sont représentées dans la figure 3.8.

• **Phase 2 : Initialisation de l'arbre de jonction.** Une fois l'arbre de jonction construit, on passe à l'initialisation des cliques C_i et des séparateurs S_{ij} par des distributions de possibilités initiales. Cette initialisation s'obtient par la trans-

| A | B | C | D | E | $\pi_{\Pi G_p}(ABCDE)$ | A | B | C | D | E | $\pi_{\Pi G_p}(ABCDE)$ |
|-----|----------|----------|----------|----------|------------------------|----------|----------|----------|----------|----------|------------------------|
| a | b | c | d | e | 0.25 | $\neg a$ | b | c | d | e | 0.02 |
| a | b | c | d | $\neg e$ | 0.5 | $\neg a$ | b | c | d | $\neg e$ | 0.04 |
| a | b | c | $\neg d$ | e | 0.075 | $\neg a$ | b | c | $\neg d$ | e | 0.006 |
| a | b | c | $\neg d$ | $\neg e$ | 0.15 | $\neg a$ | b | c | $\neg d$ | $\neg e$ | 0.012 |
| a | b | $\neg c$ | d | e | 0.5 | $\neg a$ | b | $\neg c$ | d | e | 0.008 |
| a | b | $\neg c$ | d | $\neg e$ | 0.25 | $\neg a$ | b | $\neg c$ | d | $\neg e$ | 0.004 |
| a | b | $\neg c$ | $\neg d$ | e | 1 | $\neg a$ | b | $\neg c$ | $\neg d$ | e | 0.016 |
| a | b | $\neg c$ | $\neg d$ | $\neg e$ | 0.5 | $\neg a$ | b | $\neg c$ | $\neg d$ | $\neg e$ | 0.008 |
| a | $\neg b$ | c | d | e | 0 | $\neg a$ | $\neg b$ | c | d | e | 0.01 |
| a | $\neg b$ | c | d | $\neg e$ | 0 | $\neg a$ | $\neg b$ | c | d | $\neg e$ | 0.02 |
| a | $\neg b$ | c | $\neg d$ | e | 0 | $\neg a$ | $\neg b$ | c | $\neg d$ | e | 0.1 |
| a | $\neg b$ | c | $\neg d$ | $\neg e$ | 0 | $\neg a$ | $\neg b$ | c | $\neg d$ | $\neg e$ | 0.2 |
| a | $\neg b$ | $\neg c$ | d | e | 0 | $\neg a$ | $\neg b$ | $\neg c$ | d | e | 0.016 |
| a | $\neg b$ | $\neg c$ | d | $\neg e$ | 0 | $\neg a$ | $\neg b$ | $\neg c$ | d | $\neg e$ | 0.008 |
| a | $\neg b$ | $\neg c$ | $\neg d$ | e | 0 | $\neg a$ | $\neg b$ | $\neg c$ | $\neg d$ | e | 0.08 |
| a | $\neg b$ | $\neg c$ | $\neg d$ | $\neg e$ | 0 | $\neg a$ | $\neg b$ | $\neg c$ | $\neg d$ | $\neg e$ | 0.04 |

TAB. 3.3 – Distribution de possibilités jointe $\pi_{\Pi G_p}$

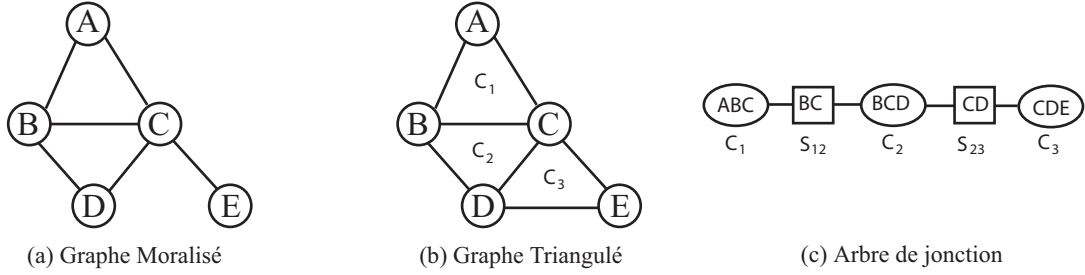


FIG. 3.8 – Étapes de transformation du réseau ΠG_p en un arbre de jonction

formation de la distribution de possibilités conditionnelles initiales des nœuds vers une autre associée aux cliques et séparateurs.

En effet, la distribution de possibilités jointe associée à l'arbre de jonction AJ , notée π_{AJ}^t , se définit à partir des distributions de possibilités (appelées également potentiels) associées aux cliques $\pi_{C_i}^t$ et aux séparateurs $\pi_{S_{ij}}^t$ de la manière suivante :

$$\pi_{AJ}^t(A_1, A_2, \dots, A_n) = \frac{\prod_{i=1}^m \pi_{C_i}^t(A_1, A_2, \dots, A_n)}{\prod_{i=1}^{m-1} \pi_{S_{ij}}^t(A_1, A_2, \dots, A_n)} \quad (3.6)$$

ou t est relatif à l'étape de propagation : $t = I$ pour l'étape d'initialisation, $t = C$ pour l'étape de cohérence globale.

La procédure d'initialisation peut être résumée comme suit :

- (i) Pour chaque clique C_i affecter une distribution de possibilités dont les éléments sont égaux à 1 : $\pi_{C_i}^I \leftarrow 1$.
- (ii) Pour chaque séparateur S_{ij} affecter une distribution de possibilités dont les éléments sont égaux à 1 : $\pi_{S_{ij}}^I \leftarrow 1$.
- (iii) Pour chaque variable $A_k \in V$, choisir une clique C_i contenant $\{A_k\} \cup U_{A_k}$ et mettre à jour sa distribution de possibilités locale : $\pi_{C_i}^I \leftarrow \pi_{C_i}^I \cdot \Pi(A_k | U_{A_k})$.

L'exemple suivant permet d'illustrer l'étape d'initialisation.

Exemple 3.5 *On reprend le réseau possibiliste de l'exemple 3.3 qui est transformé en un arbre de jonction donné dans la figure 3.8(c). L'étape d'initialisation permet d'associer aux cliques et séparateurs les potentiels suivants :*

- $\pi_{C_1}^I(A, B, C) = \Pi(A) \cdot \Pi(B|A) \cdot \Pi(C|A)$
- $\pi_{C_2}^I(B, C, D) = \Pi(D|BC)$
- $\pi_{C_3}^I(C, D, E) = \Pi(E|C)$
- $\pi_{S_{12}}^I(B, C) = 1$
- $\pi_{S_{23}}^I(C, D) = 1$

Plus précisément, le tableau 3.4 donne les valeurs initiales des distributions associées aux cliques C_1 , C_2 et C_3 .

| A | B | C | $\pi_{C_1}^I$ | D | B | C | $\pi_{C_2}^I$ | D | E | C | $\pi_{C_3}^I$ |
|----------|----------|----------|---------------|----------|----------|----------|---------------|----------|----------|----------|---------------|
| a | b | c | 0.5 | d | b | c | 1 | d | e | c | 0.5 |
| a | b | $\neg c$ | 1 | d | b | $\neg c$ | 0.5 | d | e | $\neg c$ | 1 |
| a | $\neg b$ | c | 0 | d | $\neg b$ | c | 0.1 | d | $\neg e$ | c | 1 |
| a | $\neg b$ | $\neg c$ | 0 | d | $\neg b$ | $\neg c$ | 0.2 | d | $\neg e$ | $\neg c$ | 0.5 |
| $\neg a$ | b | c | 0.04 | $\neg d$ | b | c | 0.3 | $\neg d$ | e | c | 0.5 |
| $\neg a$ | b | $\neg c$ | 0.016 | $\neg d$ | b | $\neg c$ | 1 | $\neg d$ | e | $\neg c$ | 1 |
| $\neg a$ | $\neg b$ | c | 0.2 | $\neg d$ | $\neg b$ | c | 1 | $\neg d$ | $\neg e$ | c | 1 |
| $\neg a$ | $\neg b$ | $\neg c$ | 0.08 | $\neg d$ | $\neg b$ | $\neg c$ | 1 | $\neg d$ | $\neg e$ | $\neg c$ | 0.5 |

TAB. 3.4 – Distributions initiales de possibilités associées aux cliques

Après l'étape d'initialisation, il est facile de vérifier que la distribution de possibilités jointe π_{AJ} associée à l'arbre de jonction AJ est équivalente à la distribution de possibilités jointe $\pi_{\Pi G_p}$ associée au graphe initiale ΠG_p .

En effet, après l'initialisation, toutes les distributions de possibilités associées aux séparateurs sont instanciées par des 1, ce qui simplifie la formule (3.6) en :

$$\begin{aligned}\pi_{AJ}^I(A_1, A_2, \dots, A_n) &= \prod_{i=1}^m \pi_{C_i}^I(A_1, A_2, \dots, A_n) \\ &= \prod_{k=1}^n \Pi(A_k | U_{A_k}) \\ &= \pi_{\Pi G_p}(A_1, A_2, \dots, A_n)\end{aligned}$$

Exemple 3.6 Soit l'arbre de jonction AJ de la figure 3.8(c). Après initialisation, la distribution de possibilités jointe π_{AJ}^I peut s'écrire :

$$\begin{aligned}\pi_{AJ}^I(A, B, C, D, E) &= \pi_{C_1}^I(A, B, C) \cdot \pi_{C_2}^I(B, C, D) \cdot \pi_{C_3}^I(C, D, E) \\ &= \Pi(A) \cdot \Pi(B|A) \cdot \Pi(C|A) \cdot \Pi(D|BC) \cdot \Pi(E|C) \\ &= \pi_{\Pi G_p}(A, B, C, D, E)\end{aligned}$$

• **Phase 3 : Propagation globale.**

Le but de la propagation dans un arbre de jonction est de mettre à jour chaque potentiel des cliques afin qu'il devienne la distribution jointe a posteriori des variables de la clique. Cette actualisation des potentiels sera calculée par un échange de messages locaux entre cliques adjacentes. La cohérence globale de potentiels (voir définition 3.2) permet de s'assurer que la marginalisation d'une variable à partir des différents potentiels de cliques à laquelle elle appartient doit être la même. C'est à dire qu'on peut obtenir la distribution marginale d'une variable en effectuant la marginalisation du potentiel de n'importe quelle clique à laquelle elle appartient.

Avant de détailler les étapes de propagation, nous définissons d'abord la cohérence globale dans les arbres de jonction.

Définition 3.2 : Soient deux cliques adjacentes C_i et C_j dans un arbre de jonction AJ et soit S_{ij} leur séparateur. Le lien entre C_i et C_j est dit stable ou cohérent si :

$$\max_{C_i \setminus S_{ij}} \pi_{C_i} = \pi_{S_{ij}} = \max_{C_j \setminus S_{ij}} \pi_{C_j} \quad (3.7)$$

Où $\max_{C_i \setminus S_{ij}} \pi_{C_i}$ est la distribution marginale de S_{ij} définie à partir de C_i .

Un arbre de jonction AJ est dit globalement cohérent si tous ces liens sont cohérents. Si un arbre de jonction est globalement cohérent, alors le potentiel de chaque clique C_i satisfait :

$$\forall c_i, \pi_{C_i}(c_i) = \pi_{\Pi G_p}(c_i) \quad (3.8)$$

$\pi_{\Pi G_p}(c_i)$ est la distribution de possibilités marginale des variables appartenant à la clique C_i lorsque $C_i = c_i$ calculée à partir du graphe initial ΠG_p .

L'arbre de jonction AJ initialisé peut ne pas être globalement cohérent vu que les distributions initiales risquent de ne pas satisfaire la condition de cohérence globale (équation 3.7). L'exécution de la phase de propagation globale assurera donc cette cohérence globale dans l'arbre de jonction.

La phase de propagation globale démarre par le choix arbitraire d'une clique pivot définissant ainsi le sens d'envoi de messages dans l'arbre. Cet échange de message se fait en deux étapes distinctes :

- étape de *collecte d'information* : commençant par la clique la plus éloignée du pivot, chaque clique envoie un message à sa clique adjacente dans la direction du pivot (il y a une seule clique). Une clique ne peut envoyer un message à sa clique adjacente dans la direction du pivot que si elle a reçu tout les messages de ses autres cliques adjacentes (qui ne sont pas dans la direction du pivot). Cette condition garantit la cohérence globale dans l'arbre de jonction une fois la propagation globale achevée.
- étape de *distribution d'information* : maintenant les messages vont être envoyés dans le sens contraire du pivot commençant par lui-même. Chaque clique envoie un message à ces cliques adjacentes jusqu'à atteindre les extrémités (les feuilles) de l'arbre de jonction.

Exemple 3.7 La figure 3.9 résume les deux étapes de la phase de propagation globale. La clique C_3 est choisie comme pivot. La clique C_1 , étant la plus éloignée du pivot, commence l'étape de collecte d'information par l'envoi d'un message (n°1) à C_2 qui le reçoit puis à son tour envoie un message (n°2) à C_3 (le pivot). Ce dernier entame l'étape de distribution d'information par l'envoi d'un message (n°3) à C_2 qui envoie le dernier message (n°4) à C_1 .

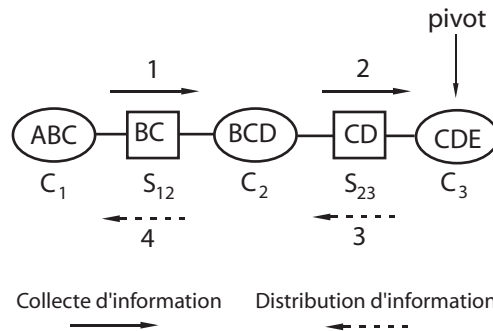


FIG. 3.9 – L'ordre des messages échangés pendant la phase de propagation globale

Explicitement, dans les deux étapes de propagation, si une clique C_i envoie un

message à sa clique adjacente C_j , alors les potentiels des cliques C_i et C_j et de leur séparateur S_{ij} sont calculés (actualisés) selon les formules suivantes :

1. Le potentiel de la clique C_i demeure inchangé :

$$\pi_{C_i}^{t+1} \leftarrow \pi_{C_i}^t. \quad (3.9)$$

2. Le potentiel du séparateur S_{ij} est mis à jour :

$$\pi_{S_{ij}}^{t+1} \leftarrow \max_{C_i \setminus S_{ij}} \pi_{C_i}. \quad (3.10)$$

3. Le potentiel de la clique C_j est mis à jour :

$$\pi_{C_j}^{t+1} \leftarrow \pi_{C_j}^t \cdot \frac{\pi_{S_{ij}}^{t+1}}{\pi_{S_{ij}}^t}. \quad (3.11)$$

La proposition suivante montre qu'après la phase de propagation globale la distribution de possibilités jointe π_{AJ} (calculée selon la formule 3.6) associée à l'arbre de jonction AJ est équivalente à la distribution de possibilités jointe $\pi_{\Pi G_p}$ associée au graphe initiale ΠG_p .

Proposition 3.1 *Soit π_{AJ}^t la distribution de possibilités jointe associée à l'arbre de jonction AJ et obtenue après la phase de propagation globale. Soit $\pi_{\Pi G_p}$ la distribution de possibilités jointe associée au graphe initiale ΠG_p . Nous pouvons écrire que :*

$$\pi_{AJ}^t(A_1, \dots, A_n) = \pi_{\Pi G_p}(A_1, \dots, A_n) \quad (3.12)$$

• **Phase 4 : réponse aux requêtes.**

A partir d'un arbre de jonction globalement cohérent, il est possible de calculer la distribution de possibilités associée à toute variable $A \in V$ à partir de n'importe quelle clique C_i contenant cette variable. La marginalisation du potentiel de la clique C_i contenant A permet d'obtenir la distribution de possibilités de celle-ci par :

$$\Pi(A) = \max_{C_i \setminus A} \pi_{C_i}^C. \quad (3.13)$$

Exemple 3.8 *considérons de nouveau l'arbre de jonction AJ (voir figure 3.8(c)) obtenu à partir du graphe initial ΠG_p de l'exemple 3.3. Après la phase de propagation globale, l'arbre de jonction est globalement consistant. Les distributions de possibilités associées aux cliques et séparateurs issues de cette phase sont données par les tableaux 3.5 et 3.6. En utilisant la formule 3.6, nous pouvons vérifier que, par exemple pour l'interprétation $\omega = abc \neg de$,*

$$\pi_{AJ}^t(abc \neg de) = \frac{\pi_{C_1}^t(abc) \cdot \pi_{C_2}^t(bc \neg d) \cdot \pi_{C_3}^t(c \neg de)}{\pi_{S_{12}}^t(\neg bc) \cdot \pi_{S_{12}}^t(c \neg d)} = \frac{0.5 * 0.15 * 0.5}{0.5 * 1} = 0.075 = \pi_{\Pi G_p}(abc \neg de).$$

| A | B | C | $\pi_{C_1}^C$ | D | B | C | $\pi_{C_2}^C$ | D | E | C | $\pi_{C_3}^C$ |
|----------|----------|----------|---------------|----------|----------|----------|---------------|----------|----------|----------|---------------|
| a | b | c | 0.5 | d | b | c | 0.5 | d | e | c | 0.25 |
| a | b | $\neg c$ | 1 | d | b | $\neg c$ | 0.5 | d | e | $\neg c$ | 0.2 |
| a | $\neg b$ | c | 0 | d | $\neg b$ | c | 0.02 | d | $\neg e$ | c | 0.5 |
| a | $\neg b$ | $\neg c$ | 0 | d | $\neg b$ | $\neg c$ | 0.0024 | d | $\neg e$ | $\neg c$ | 0.1 |
| $\neg a$ | b | c | 0.04 | $\neg d$ | b | c | 0.15 | $\neg d$ | e | c | 0.5 |
| $\neg a$ | b | $\neg c$ | 0.016 | $\neg d$ | b | $\neg c$ | 1 | $\neg d$ | e | $\neg c$ | 0.008 |
| $\neg a$ | $\neg b$ | c | 0.2 | $\neg d$ | $\neg b$ | c | 0.2 | $\neg d$ | $\neg e$ | c | 1 |
| $\neg a$ | $\neg b$ | $\neg c$ | 0.08 | $\neg d$ | $\neg b$ | $\neg c$ | 0.008 | $\neg d$ | $\neg e$ | $\neg c$ | 0.004 |

TAB. 3.5 – Distributions de possibilités associées aux cliques après propagation

| B | C | $\pi_{S_{12}}^C$ | C | D | $\pi_{S_{23}}^C$ |
|----------|----------|------------------|----------|----------|------------------|
| b | c | 0.5 | c | d | 0.5 |
| b | $\neg c$ | 1 | $\neg c$ | d | 1 |
| $\neg b$ | c | 0.2 | c | $\neg d$ | 0.2 |
| $\neg b$ | $\neg c$ | 0.008 | $\neg c$ | $\neg d$ | 0.008 |

TAB. 3.6 – Distributions de possibilités associées aux séparateurs après propagation

Ce qui confirme bien la proposition 3.1.

En appliquant la formule 3.13, nous pouvons, par exemple, déterminer la distribution marginale de la variable B en choisissant une clique à laquelle elle appartient (ici on a le choix entre C_1 ou C_2). En choisissant C_2 , $\Pi(B) = \max_{C_2 \setminus B} \pi_{DBC}^C$ ce qui nous donne $\Pi(b) = 1$ et $\Pi(\neg b) = 0.2$.

Traitement de l'évidence

La phase de propagation globale peut être généralisée en considérant l'entrée (l'observation) d'une nouvelle information (appelée évidence) à travers une variable dans l'arbre de jonction. Ainsi, le calcul de $\Pi(A)$ pour chaque variable $A \in V$ sera remplacé par $\Pi(A \wedge e)$ où e représente l'évidence globale observée. Cette évidence est définie par $\Lambda_A(a)$:

$$\Lambda_A(a) = \begin{cases} 1 & \text{si } A \text{ n'est pas observée} \\ 1 & \text{si } A \text{ est observée à } a \\ 0 & \text{si } A \text{ est observée mais pas à } a. \end{cases} \quad (3.14)$$

Pour tenir compte de cette évidence, il faut que la phase d'initialisation ajoute

à sa procédure les deux instructions suivantes :

- Pour toute variable observée A , l'observation $A = a$ est matérialisée par $\Lambda_A(a)$.
- Identifier une clique contenant A puis mettre à jour son potentiel par :

$$\pi_{C_i}^I \leftarrow \pi_{C_i}^I \cdot \Lambda_A$$

En intégrant l'évidence, le potentiel d'une clique C_i est représenté par $\Pi(C_i \wedge e)$. Au cours de la marginalisation de chaque potentiel de clique sur une variable A appartenant à une clique C_i , nous obtenons la mesure de possibilité de $A \wedge e$:

$$\Pi(A \wedge e) = \max_{C_i \setminus A} \pi_{C_i}^C. \quad (3.15)$$

A partir de cette valeur calculée, on peut déterminer $\Pi(A|e)$ en appliquant la formule de conditionnement possibiliste :

$$\Pi(A|e) = \frac{\Pi(A \wedge e)}{\Pi(e)} = \frac{\Pi(A \wedge e)}{\max_{A'} \Pi(A' \wedge e)}. \quad (3.16)$$

Exemple 3.9 *Considérons le réseau possibiliste présenté dans la figure 3.8(c). Supposons qu'on observe une évidence dans le nœud B (c-à-d $B = b$). La distribution locale au niveau de la clique C_1 est initialisée par :*

$$\pi_{C_2}(ABC) = \begin{cases} \Pi(A) \cdot \Pi(B|A) \cdot \Pi(C|A) & \text{si } B = b \\ 0 & \text{sinon.} \end{cases}$$

Pour terminer, l'algorithme arbre de jonction est considéré comme efficace car il ne nécessite que l'envoi de $2 * (m - 1)$ messages entre les m cliques de l'arbre. Cependant, il est assez exigeant en espace mémoire (exponentiel en fonction de la taille de la clique) ; en particulier pour des graphes à connexions multiples où il est parfois impossible d'initialiser l'arbre de jonction. L'algorithme arbre de jonction, comme dans les réseaux bayésiens, appartient à la famille des problèmes NP-difficile [Coo90].

3.3.1.2 Algorithme de propagation dans les réseaux possibilistes hybrides

Une autre façon de représenter l'incertitude en théorie des possibilités est la logique possibiliste qui, à travers un poids associé aux formules, définit un ordre d'importance entre celles-ci. L'ensemble de ces formules pondérées constitue une base de connaissances possibilistes. Les réseaux possibilistes hybrides permettent une représentation compacte de l'incertitude dans les réseaux possibilistes en substituant les distributions de possibilités par des bases de connaissances possibilistes. Dans cette section nous allons introduire les réseaux possibilistes hybrides, suivis d'une adaptation de l'algorithme de propagation dans un cadre quantitatif.

► Les réseaux possibilistes hybrides

Les réseaux possibilistes hybrides [BS07b], notés HG , ne diffèrent des réseaux possibilistes standards que par leurs composantes numériques. En effet, dans les HG les distributions de possibilités locales de chaque nœud A_i sont remplacées par des bases de connaissances codifiant l'information concernant A_i dans le contexte de ses parents U_{A_i} . A partir de ces bases de connaissances locales B_{A_i} il est possible de calculer la distribution de possibilités associée $\pi_{B_{A_i}}$ en utilisant la formule 1.38. Ainsi, la distribution de possibilités jointe, exprimée par la règle de chaînage, peut être obtenue en combinant les distributions de possibilités locales $\pi_{B_{A_i}}$ selon le cadre de raisonnement choisi (ici on s'intéresse au cadre quantitatif) :

$$\forall \omega \in \Omega, \pi_{HG}(\omega) = \prod_{i=1..n} \pi_{B_{A_i}}(\omega). \quad (3.17)$$

La base de connaissances possibiliste, notée B_{HG} et associée à HG , résultante de la fusion des n bases de connaissances locales B_{A_i} , s'obtient par l'équivalent de l'opérateur syntaxique de fusion *produit* B^* (voir annexe A).

Le principal avantage de la représentation hybride des réseaux possibiliste concerne la complexité spatiale. En effet, la spécification des distributions conditionnelles initiales peut devenir impossible, surtout pour les variables qui ont plusieurs parents et dans ce cas, la nouvelle représentation compacte présente une alternative intéressante.

► Adaptation de l'algorithme arbre de jonction pour les réseaux possibilistes hybrides

Maintenant, nous présentons un résumé de la propagation dans les réseaux possibilistes hybrides. Pour plus de détails, voir [BS07b].

Étape 1 : Construction de l'arbre de jonction

Vu que les réseaux possibilistes hybrides sont des DAGs, la construction de l'arbre de jonction hybride, noté $H AJ$, suit la même démarche que celle suivie pour la construction des arbres de jonction classique.

Étape 2 : Initialisation de l'arbre de jonction

Une fois l'arbre de jonction $H AJ$ obtenu, une base de connaissances $B_{C_i}^t$ est associée à chaque clique C_i et une autre, $B_{S_{ij}}^t$, est associée à chaque séparateur S_{ij} . t correspond à l'étape de propagation initialisation ($t = I$) ou cohérence ($t = C$). La base de connaissances associée à l'arbre de jonction $H AJ$ résulte de la réunion des m bases de connaissances $B_{C_i}^t$ des cliques.

Étape 3 : Propagation globale

Dans cette étape de propagation, les messages sont échangés entre les cliques adjacentes jusqu'à atteindre un état de cohérence globale. En recevant un message, une clique C_i ne modifie pas directement sa base de connaissances mais utilise le message reçu pour envoyer à son tour ses messages aux autres cliques adjacentes. La base de connaissances de la clique C_i ne sera calculée qu'après envoi de tous les messages.

Étape 4 : Réponses aux requêtes

Après l'étape de propagation globale, la marginalisation de chaque variable $A_k \in V$ est calculée à partir de chaque clique C_i contenant A_k .

Traitement de l'évidence

Le traitement de l'évidence consiste à calculer pour chaque variable $A_k \in V$ le degré de possibilité $\Pi(A_k \wedge e)$ où e est l'évidence totale. Afin de prendre en compte les informations observées, l'étape d'initialisation doit être modifiée par l'ajout des instructions suivantes :

- Modifier la base de connaissances de chaque variable A_k observée comme suit :

$$B_{A_k} = \begin{cases} (a_k, 1) & \text{si } A_k = a_k \\ \emptyset & \text{sinon.} \end{cases}$$

- Choisir une clique C_i contenant la variable observée A_k et mettre à jour la base de connaissances de la clique comme suit : $B_{C_i} = B^*(B_{C_i}, B_{A_k})$.

3.3.2 Algorithme de propagation *anytime* pour les réseaux qualitatifs

Dans ce qui suit, nous allons présenter un algorithme de propagation pour les réseaux possibilistes qualitatifs qui n'est pas une adaptation directe de la propagation probabiliste. En effet, cet algorithme évite la transformation du graphe initiale en un arbre de jonction dont la construction est un problème difficile [Coo90].

Étant donné un réseau possibiliste qualitatif, noté ΠG_m , les auteurs de l'article [BABK03] proposent un algorithme de propagation permettant de calculer localement, pour chaque instance a_i d'une variable d'intérêt A_i , le degré de possibilité $\Pi_m(a_i)$. Cet algorithme est composé de plusieurs étapes qui, progressivement, donnent une approximation des distributions de possibilités à calculer.

L'algorithme proposé est dit algorithme "anytime" puisque plus le temps alloué à son exécution est long, plus le rapprochement est meilleur entre les distributions

calculées et celles exactes.

L'idée de base de cet algorithme est de calculer, pour une instance a_i d'une variable A_i , le degré de possibilité $\Pi_m(a_i)$. Pour cela, à partir de la distribution de possibilités π_m , on définit une nouvelle distribution π_{a_i} comme suit :

$$\pi_{a_i}(\omega) = \begin{cases} \pi_m(\omega) & \text{si } \omega[A_i] = a_i \\ 0 & \text{sinon.} \end{cases} \quad (3.18)$$

où a_i est l'instance de A_i dans l'interprétation ω .

Une fois la distribution π_{a_i} calculée, il est simple de déduire le degré de possibilité $\Pi_m(a_i)$ de la manière suivante : $\Pi_m(a_i) = h(\pi_{a_i}) = \max_{\omega} \pi_{a_i}(\omega)$.

Les différentes étapes de l'algorithme sont résumées comme suit :

• **Étape 1 : Initialisation**

Dans cette étape, on commence par transformer le graphe initial en une structure équivalente, appelée graphe moral et notée \mathcal{MG} , dans laquelle chaque nœud est rassemblé avec ses nœuds parents et l'orientation des arcs sera supprimée. Un nœud avec ses parents forment ensemble une *clique* notée C_i (c'est à dire $C_i = \{A_i\} \cup U_{A_i}$). Deux cliques C_i et C_j , dont l'intersection n'est pas vide, sont reliées par une arête étiquetée par cette intersection, appelée *séparateur* et notée S_{ij} . On associe à chaque clique C_i la distribution de possibilités jointe locale de ses nœuds, appelée *potentiel* et notée π_{C_i} , comme suit : $\pi_{C_i}(A_i \wedge U_{A_i}) = \Pi(A_i|U_{A_i})$. De même on associe à \mathcal{MG} la distribution de possibilités jointe définie par : $\pi_{MG}(A_1, \dots, A_n) = \min_{i=1, \dots, n} \pi_{C_i}$. Une fois le graphe moral \mathcal{MG} initialisé, la seconde phase de l'étape de l'initialisation consiste à introduire dans \mathcal{MG} une instance a_i d'une variable d'intérêt A_i (c'est la variable à qui on veut déterminer son $\Pi_m(a_i)$). Ceci peut être obtenu en modifiant le potentiel des cliques C_k comme suit :

$$\pi_{C_k}^I(c_k) = \begin{cases} \pi_{C_k}^I(c_k) & \text{si } c_k[A_i] = a_i \\ 0 & \text{sinon.} \end{cases} \quad (3.19)$$

• **Étape 2 : Stabilité vis-à-vis d'un parent**

Au cours de cette étape, l'algorithme vérifie que chaque clique est stable avec ses cliques voisines (adjacentes). Plus précisément, chaque séparateur S_{ij} entre deux cliques C_i et C_j doit être stable ; c'est à dire que la distribution marginale de S_{ij} issue de chacune des deux cliques doit être identique.

Cette stabilité est atteinte par échange de messages entre les différentes cliques et leurs séparateurs. Chaque séparateur S_{ij} collecte les informations à partir de ses cliques voisines puis leur distribue sa mise à jour afin qu'à leur tour elles puissent mettre à jour leur potentiels. Cet envoi de messages est répété jusqu'à ce qu'aucune

clique ne modifie son potentiel (c'est l'état de Stabilité vis-à-vis d'un parent). Ces deux étapes sont répétées jusqu'à ce que toutes les cliques soient stabilisées vis-à-vis d'un parent. Cet état de stabilité est atteint après un nombre fini de messages échangés ; d'où sa complexité théorique polynomiale. Lorsqu'un graphe moral \mathcal{MG} est stabilisé vis-à-vis d'un parent alors, tous les potentiels de ces cliques possèdent la même valeur maximale α . Autrement dit : $\alpha = \max \pi_{C_i}$.

• **Étape 3 : Stabilité vis-à-vis de n parent**

L'étape de stabilité vis-à-vis d'un parent ne permet pas, à elle seule, de garantir une meilleure approximation du degré de possibilité $\Pi_m(a_i)$ pour une variable d'intérêt A_i . Afin d'améliorer davantage cette approximation en la faisant converger vers la valeur exacte, l'algorithme propose une troisième étape permettant de stabiliser le graphe moral en considérant progressivement n parents au lieu d'un. On commence par stabiliser le graphe moral en considérant deux parents puis trois jusqu'à n parents, où n est le plus grand nombre des parents des cliques.

3.4 Conclusion

Dans ce chapitre, nous avons présenté plusieurs algorithmes de propagation dans les réseaux probabilistes et possibilistes. Tous ces algorithmes ont pour fonction de déterminer d'une manière soit exacte soit approximative, l'impact de l'arrivée d'une nouvelle information certaine (appelée évidence) sur les autres variables du réseau appelées quant à elles variables d'intérêts.

Dans un premier temps, nous avons abordé la propagation probabiliste exacte dans des réseaux ayant comme structure un polyarbre. Cet algorithme se base sur une propagation locale de messages entre un nœud et ses voisins afin de "faire connaître" l'évidence à tous les nœuds du polyarbre. Bien que sa complexité théorique soit linéaire, cet algorithme n'est pas applicable dans des réseaux à connexions multiples à cause des boucles. L'idée dans ce cas est de transformer le réseau initial en une structure équivalente simplement connectée (appelée arbre de jonction) dans laquelle sera appliqué l'algorithme de propagation de messages.

Une autre méthode a été proposée par les auteurs de [BNG04], qui ont présenté une unification des algorithmes de propagation de messages et de l'arbre de jonction. En effet, ces auteurs proposent une méthode permettant d'obtenir une structure secondaire orientée, très liée à celle de l'arbre de jonction, et ont montré que des calculs similaires à ceux de l'algorithme de propagation de messages menés dans cette nouvelle structure atteignent des performances identiques à celles de l'algorithme arbre de jonction.

Dans un second temps, nous avons présenté quelques algorithmes approximatifs de propagation probabilistes. Ce sont des alternatives aux algorithmes exacts lorsque l'exécution de ces derniers demeurent impossible. Parmi les algorithmes approximatifs, nous prêtons un intérêt particulier à l'algorithme 'Loopy Belief Propagation' qui est une variante de l'algorithme de propagation de messages itérée sur un réseau à connexions multiples. Cet algorithme ainsi que son adaptation aux réseaux possibilistes quantitatifs et qualitatifs seront largement détaillés dans le prochain chapitre. Dans la troisième section de ce chapitre nous avons présenté certains algorithmes de propagation dans les réseaux possibilistes :

- Le premier est une adaptation directe de l'algorithme arbre de jonction probabiliste pour les réseaux possibilistes quantitatifs.
- Le deuxième algorithme s'applique sur des réseaux possibilistes (quantitatifs) dit hybrides où l'incertitude au niveau des nœuds est représentée par des bases de connaissances possibilistes au lieu des distributions de possibilités.
- Le troisième algorithme de propagation s'applique sur des réseaux possibilistes qualitatifs. Cet algorithme approche par plusieurs étapes de stabilité les valeurs exactes à calculer.

Bien que efficaces et précis, les algorithmes exacts de propagation (probabilistes et possibilistes) n'arrivent pas toujours à s'achever normalement surtout lorsqu'il sont appliqués dans des réseaux à connexions multiples. Dans ce contexte, les algorithmes approximatifs de propagation peuvent être une alternative à mettre en œuvre pour approcher les valeurs exactes à déterminer.

Le prochain chapitre présentera l'algorithme de propagation 'Loopy Belief Propagation' possibiliste comme une alternative d'inférence exacte dans les deux cadres quantitatif et qualitatif. À notre connaissance, aucun algorithme approximatif d'inférence probabiliste n'a été adapté aux réseaux possibilistes quantitatifs et qualitatifs.

Chapitre 4

Un nouvel algorithme approximatif pour les réseaux possibilistes

4.1 Introduction

L'algorithme exact, *arbre de jonction*, est célèbre pour être efficace. Cependant, cette efficacité a un prix coûteux : la construction et la propagation dans cet arbre sont considérées comme des problèmes appartenant à la classe NP-difficile [Coo90]. En effet, le problème d'optimalité de formation des cliques et la taille de celles-ci rendent l'algorithme arbre de jonction dans plusieurs cas inapplicable.

Une alternative à ces algorithmes exacts, lorsque ces derniers deviennent inapplicables, est l'exécution des algorithmes de propagation approximatifs afin d'approcher les valeurs exactes. Parmi les algorithmes approximatifs, nous portons un intérêt particulier, dans cette thèse, à l'algorithme 'Loopy Belief Propagation' (LBP) pour les raisons suivantes : d'abord, il trouve ses origines dans l'algorithme de propagation de messages (efficace et simple à implémenter), ensuite, il a prouvé que, dans divers domaines [MMC98, Wei00, TIT03], il génère une bonne approximation des probabilités à calculer.

Dans ce chapitre, nous allons proposer une adaptation de l'algorithme 'Loopy Belief Propagation' aux réseaux possibilistes quantitatifs et qualitatifs. À notre connaissance, aucun algorithme approximatif d'inférence probabiliste n'a été adapté aux réseaux possibilistes quantitatifs et qualitatifs.

Le reste de ce chapitre est organisé comme suit : La section 4.2 expose la version probabiliste de l'algorithme LBP. La section 4.3 détaille la propagation dans un polyarbre possibiliste quantitatif. La section 4.4 explique l'adaptation de l'algorithme de propagation de messages aux réseaux possibilistes quantitatifs à connexion multiples. La section 4.5 présente l'application de LBP possibiliste au réseau quantitatifs certains. La section 4.6 propose, quant à elle, l'adaptation aux réseaux qualitatifs de

l'algorithme LBP. La section 4.7 présente l'application de LBP possibiliste aux réseaux qualitatifs certains. Les résultats expérimentaux qui évaluent les performances de ce nouvel algorithme approximatif seront présentés dans le chapitre 5.

Les principaux résultats de ce chapitre sont publiés dans [ABOY08, AOBY08].

4.2 L'algorithme approximatif '*Loopy Belief Propagation*'

Loopy Belief Propagation [MWJ99, Wei00] est un algorithme approximatif de propagation d'inférence qui applique le principe de l'algorithme de propagation de messages (voir sous-section 3.2.1.1) sur des graphes à connexions multiples. En effet, l'algorithme de propagation de messages (dit aussi de Pearl [Pea88]), appliqué sur un réseau bayésien ayant la structure d'un polyarbre, fournit un résultat exact. Malheureusement, dans la réalité, la plupart des réseaux probabilistes ont des structures plus complexes et présentent des boucles. L'algorithme de Pearl appliqué tel qu'il est sur un graphe à connexions multiples n'est plus adéquat : l'information est susceptible de passer via deux ou plusieurs chaînes vers une même destination qui va devoir l'incorporer plusieurs fois. De plus, les messages échangés entre les nœuds vont circuler indéfiniment à travers les boucles et les résultats ne seront, en aucun cas, exacts ni même valides.

Pour continuer à appliquer l'inférence exacte sur des modèles graphiques à connexions multiples, la majorité des algorithmes transforment le graphe initial vers une structure équivalente ayant la forme d'un polyarbre appelée *arbre de jonction*, puis exécute l'algorithme de propagation de messages. Toutefois, cette transformation en arbre de jonction est très coûteuse et dans la plupart des applications du monde réel, les cliques formant cet arbre de jonction ont une taille énorme ce qui empêche, parfois, l'algorithme de propagation de s'initialiser.

L'idée de base de l'algorithme '*Loopy Belief Propagation*' est la suivante : sans aucune transformation de la structure initiale du graphe, l'algorithme va appliquer la propagation des messages et laisser les messages circuler dans le réseau à connexions multiples jusqu'à atteindre un état de convergence (s'il y en a). Dans l'algorithme LBP, il n'y a pas d'ordre d'envoi de messages : tous les nœuds calculent et envoient leurs messages d'une manière parallèle. L'algorithme LBP va itérer la propagation des messages à travers le réseau à connexions multiples jusqu'à atteindre un état où cette propagation des messages ne modifie plus aucune probabilité pour tous les nœuds : c'est l'état de convergence. Arrivé à cet état (s'il existe), l'algorithme LBP ne fournit pas les valeurs exactes des probabilités a posteriori, mais présente une approximation de celles-ci.

Un état de convergence est donc atteint lorsque, pour tous les nœuds non observés,

la différence entre les valeurs de probabilité issues de deux itérations successives est inférieure à un seuil fixé à l'avance arbitrairement. Cette condition peut être renforcée par une limitation du nombre maximal d'itérations afin de ne pas itérer indéfiniment lorsque l'algorithme LBP ne converge pas.

Dans ce qui suit, nous allons présenter les différentes formules et expressions des messages calculés et échangés au cours de l'exécution de l'algorithme LBP. Ces formules et expressions sont les mêmes que l'algorithme de propagation de messages auxquelles l'aspect temps, symbolisé par t , a été ajouté afin de représenter la notion d'itération entre les étapes de l'algorithme.

Étant donné un ensemble d'évidence E instancié par la valeur e , la probabilité conditionnelle d'un nœud non observé A_i sachant e est approximé comme suit :

$$\forall a_i \in D_{A_i} \text{ Bel}(a_i) = p(a_i | e) \approx \alpha \lambda(a_i) \mu(a_i), \quad (4.1)$$

avec α une constante de normalisation et $\lambda(a_i)$ et $\mu(a_i)$ sont calculés itérativement. À l'itération t , $\lambda^{(t)}(a_i)$ et $\mu^{(t)}(a_i)$ traduisent les messages entrants au nœud A_i :
- $\lambda^{(t)}(a_i)$ représente la quantité résultante des messages reçus par A_i envoyés par ses fils Y_k : $k = 1..m$.

$$\lambda^{(t)}(a_i) = \lambda_{A_i}(a_i) \prod_{k=1}^m \lambda_{Y_k}^{(t)}(a_i), \quad (4.2)$$

avec $\lambda_{Y_k}^{(t)}(a_i)$ est le message envoyé par le fils Y_k à son parent A_i à l'itération t .

- De même, $\mu^{(t)}(a_i)$ représente la quantité résultante des messages reçus par A_i envoyés par ses parents U_j : $j = 1..n$.

$$\mu^{(t)}(a_i) = \sum_u p(a_i | u) \cdot \prod_{j=1}^n \mu_{A_i}^{(t)}(u_j), \quad (4.3)$$

avec $\mu_{A_i}^{(t)}(u_j)$ est le message envoyé par le parent U_j vers A_i à l'itération t .

À chaque itération t , tous les nœuds calculent, parallèlement, leur $\lambda^{(t)}(a_i)$ et $\mu^{(t)}(a_i)$ ce qui permet d'approximer, graduellement, $p(a_i | u)$. Si la condition de convergence n'est pas satisfaite, alors chaque nœud calcule les messages sortants pour la prochaine itération $t + 1$ en fonction des messages entrants à l'itération t . Ainsi,

$$\lambda_{A_i}^{(t+1)}(u_j) = \beta \cdot \sum_{a_i} \lambda^{(t)}(a_i) \left[\sum_{u_l, l \neq j} p(a_i | u) \cdot \prod_{l \neq j} \mu_{A_i}^{(t)}(u_l) \right], \quad (4.4)$$

avec β est une constante de normalisation.

$\lambda_{A_i}^{(t+1)}(u_j)$ est le message à envoyer au parent U_j de la part du nœud A_i à la prochaine

itération $t + 1$. Ce message dépend essentiellement de $\lambda^{(t)}(a_i)$ et de tous les messages $\mu_{A_i}^{(t)}(u_l)$, envoyés par tout les parents de A_i sauf U_j , calculés à l'itération t . De même,

$$\mu_{Y_k}^{(t+1)}(a_i) = \alpha \mu^{(t)}(a_i) \lambda_{A_i}(a_i) \prod_{j \neq k} \lambda_{Y_j}^{(t)}(a_i), \quad (4.5)$$

avec α est une constante de normalisation.

$\mu_{Y_k}^{(t+1)}(a_i)$ est le message à envoyer de la part de A_i à son fils Y_k à la prochaine itération $t + 1$. Ce message dépend essentiellement de $\mu^{(t)}(a_i)$ et de tous les messages $\lambda_{Y_j}^{(t)}(a_i)$, envoyés par tous les fils de A_i sauf Y_k , calculés à l'itération t .

Ce processus est itéré jusqu'à satisfaire la condition de convergence. Lorsque l'algorithme LBP converge, les probabilités a posteriori ainsi obtenues sont des bonnes approximations pour les valeurs exactes de ces mêmes probabilités.

Exemple 4.1 La figure 4.1 représente un réseau probabiliste à connexions multiples. Le tableau 4.1 fournit les distributions de probabilités conditionnelles et a priori associées aux variables binaires A , B , C et D .

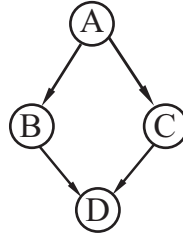


FIG. 4.1 – Réseau probabiliste à connexions multiples

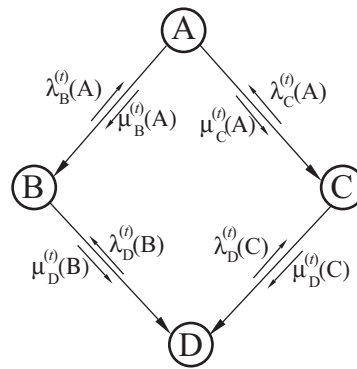


FIG. 4.2 – Messages sortants de chaque nœud pendant l'itération $t=0$ et $t=1$

On suppose qu'on a observé comme évidence le nœud D à son état $D = d$. L'algorithme LBP propose de calculer, pour les nœuds non observés A , B et C , une

| | | B | A | $P(B A)$ | C | A | $P(C A)$ |
|----------|--------|----------|----------|----------|----------|----------|----------|
| A | $P(A)$ | $\neg b$ | $\neg a$ | 0.3 | $\neg c$ | $\neg a$ | 0.6 |
| $\neg a$ | 0.5 | $\neg b$ | a | 0.9 | $\neg c$ | a | 0.2 |
| a | 0.5 | b | $\neg a$ | 0.7 | c | $\neg a$ | 0.4 |
| | | b | a | 0.1 | c | a | 0.8 |

| D | C | B | $P(D B, C)$ | D | C | B | $P(D B, C)$ |
|----------|----------|----------|-------------|-----|----------|----------|-------------|
| $\neg d$ | $\neg c$ | $\neg b$ | 1 | d | $\neg c$ | $\neg b$ | 0 |
| $\neg d$ | $\neg c$ | b | 0 | d | $\neg c$ | b | 1 |
| $\neg d$ | c | $\neg b$ | 0 | d | c | $\neg b$ | 1 |
| $\neg d$ | c | b | 0 | d | c | b | 1 |

TAB. 4.1 – Distributions de probabilités conditionnelles initiales

approximation de leurs probabilités conditionnelles a posteriori. Les étapes de l'algorithme s'exécutent comme suit : (Les messages sortants de chaque nœud pendant chaque itération sont représentés dans la figure 4.2.)

À l'itération $t = 0$

Calcul de $\mu^{(0)}$ et $\lambda^{(0)}$ pour chaque nœud

| A | $\mu^{(0)}$ | $\lambda^{(0)}$ | Bel | B | $\mu^{(0)}$ | $\lambda^{(0)}$ | Bel | C | $\mu^{(0)}$ | $\lambda^{(0)}$ | Bel | D | $\mu^{(0)}$ | $\lambda^{(0)}$ |
|----------|-------------|-----------------|-------|----------|-------------|-----------------|-------|----------|-------------|-----------------|-------|----------|-------------|-----------------|
| a | 0.5 | 1 | 0.5 | b | 0.8 | 1 | 0.4 | c | 1.2 | 1 | 0.6 | d | 3 | 1 |
| $\neg a$ | 0.5 | 1 | 0.5 | $\neg b$ | 1.2 | 1 | 0.6 | $\neg c$ | 0.8 | 1 | 0.4 | $\neg d$ | 1 | 0 |

Calcul des messages normalisés sortants de chaque nœud

| $\mu_B^{(1)}(A)$ | $\mu_C^{(1)}(A)$ | $\mu_D^{(1)}(B)$ | $\lambda_B^{(1)}(A)$ | $\mu_D^{(1)}(C)$ | $\lambda_C^{(1)}(A)$ | $\lambda_D^{(1)}(B)$ | $\lambda_D^{(1)}(C)$ |
|------------------|------------------|------------------|----------------------|------------------|----------------------|----------------------|----------------------|
| 0.5 | 0.5 | 0.4 | 0.5 | 0.6 | 0.5 | 0.6667 | 0.6667 |
| 0.5 | 0.5 | 0.6 | 0.5 | 0.4 | 0.5 | 0.3333 | 0.3333 |

À l'itération $t = 1$

Calcul de $\mu^{(1)}$ et $\lambda^{(1)}$ pour chaque nœud

| A | $\mu^{(1)}$ | $\lambda^{(1)}$ | Bel | B | $\mu^{(1)}$ | $\lambda^{(1)}$ | Bel | C | $\mu^{(1)}$ | $\lambda^{(1)}$ | Bel | D | $\mu^{(1)}$ | $\lambda^{(1)}$ |
|----------|-------------|-----------------|-------|----------|-------------|-----------------|-------|----------|-------------|-----------------|-------|----------|-------------|-----------------|
| a | 0.5 | 0.25 | 0.5 | b | 0.4 | 1 | 0.526 | c | 0.6 | 1 | 0.79 | d | 0.76 | 1 |
| $\neg a$ | 0.5 | 0.25 | 0.5 | $\neg b$ | 0.6 | 0.6 | 0.474 | $\neg c$ | 0.4 | 0.4 | 0.21 | $\neg d$ | 0.24 | 0 |

Calcul des messages normalisés sortants de chaque nœud

| $\mu_B^{(2)}(A)$ | $\mu_C^{(2)}(A)$ | $\mu_D^{(2)}(B)$ | $\lambda_B^{(2)}(A)$ | $\mu_D^{(2)}(C)$ | $\lambda_C^{(2)}(A)$ | $\lambda_D^{(2)}(B)$ | $\lambda_D^{(2)}(C)$ |
|------------------|------------------|------------------|----------------------|------------------|----------------------|----------------------|----------------------|
| 0.5 | 0.5 | 0.4 | 0.4 | 0.6 | 0.6 | 0.625 | 0.625 |
| 0.5 | 0.5 | 0.6 | 0.6 | 0.4 | 0.4 | 0.375 | 0.375 |

À l'itération $t = 2$

Calcul de $\mu^{(2)}$ et $\lambda^{(2)}$ pour chaque nœud

| A | $\mu^{(2)}$ | $\lambda^{(2)}$ | Bel | B | $\mu^{(2)}$ | $\lambda^{(2)}$ | Bel | C | $\mu^{(2)}$ | $\lambda^{(2)}$ | Bel | D | $\mu^{(2)}$ | $\lambda^{(2)}$ |
|----------|-------------|-----------------|-------|----------|-------------|-----------------|-------|----------|-------------|-----------------|-------|----------|-------------|-----------------|
| a | 0.5 | 0.264 | 0.5 | b | 0.4 | 1 | 0.526 | c | 0.6 | 1 | 0.79 | d | 0.76 | 1 |
| $\neg a$ | 0.5 | 0.264 | 0.5 | $\neg b$ | 0.6 | 0.6 | 0.474 | $\neg c$ | 0.4 | 0.4 | 0.21 | $\neg d$ | 0.24 | 0 |

D'après ce dernier calcul, on ne constate aucune variation entre ces Bel et ceux de l'itération précédente. Ce qui conduit l'algorithme à un état de convergence. D'où les valeurs approchées des probabilités a posteriori conditionnellement à l'évidence $D = d$:

| A | $P(A D = d)$ | B | $P(B D = d)$ | C | $P(C D = d)$ |
|----------|--------------|----------|--------------|----------|--------------|
| a | 0.5 | b | 0.526 | c | 0.79 |
| $\neg a$ | 0.5 | $\neg b$ | 0.474 | $\neg c$ | 0.21 |

4.3 Propagation dans les réseaux quantitatifs simplement connectés

Dans cette section, nous fournissons une présentation détaillée de l'adaptation de l'algorithme de propagation de messages aux réseaux possibilistes quantitatifs simplement connectés (ayant la structure d'un polyarbre). Le critère d-séparation défini sur les réseaux bayésiens reste aussi valide sur les réseaux possibilistes quantitatifs [Fon94]. Le but de la propagation dans les réseaux quantitatifs est de calculer la mesure de possibilité conditionnelle $\Pi(A_i = a_i|E = e)$ où $a_i \in D_{A_i}$ et E représente l'ensemble des variables observées instanciées par la valeur e . Cet ensemble E peut être subdivisé, par rapport à un nœud A_i , en deux sous-ensembles disjoints E_i^+ et E_i^- puisque A_i d-sépare E_i^+ de E_i^- . (La figure 4.3 illustre ces différentes partitions) :

- $E_i^+ = \{E_{U_1 A_i}^+, \dots, E_{U_n A_i}^+\}$: correspond au sous-ensemble de E que l'on peut atteindre à partir des n parents $\{U_1, \dots, U_n\}$ de A_i et instancié par e_i^+ .
- $E_i^- = E/E_i^+ = \{E_{A_i}, E_{A_i Y_1}^-, \dots, E_{A_i Y_m}^-\}$: correspond au sous-ensemble de E que l'on peut atteindre à partir des m fils $\{Y_1, \dots, Y_m\}$ de A_i , instancié par e_i^- et à lui même (A_i) instancié par e_{A_i} . e_{A_i} ce définie comme suit :

$$e_{A_i} = \begin{cases} a_i & \text{si } A_i \text{ est observé à } a_i \\ 0 & \text{sinon.} \end{cases}$$

e_i^+ et e_i^- représentent respectivement les valeurs observées associées aux ensembles des nœuds E_i^+ et E_i^- .

$e_{U_k A_i}^+$ avec $k \in \{1, \dots, n\}$ représente la valeur observée associée à $E_{U_k A_i}^+$ provenant des ancêtres communs de A_i et U_k et circulant vers A_i à travers U_k . De même, $e_{A_i Y_j}^-$ avec $j \in \{1, \dots, m\}$ représente la valeur observée associée à $E_{A_i Y_j}^-$ provenant des descendants communs de A_i et Y_j et circulant vers A_i à travers Y_j .

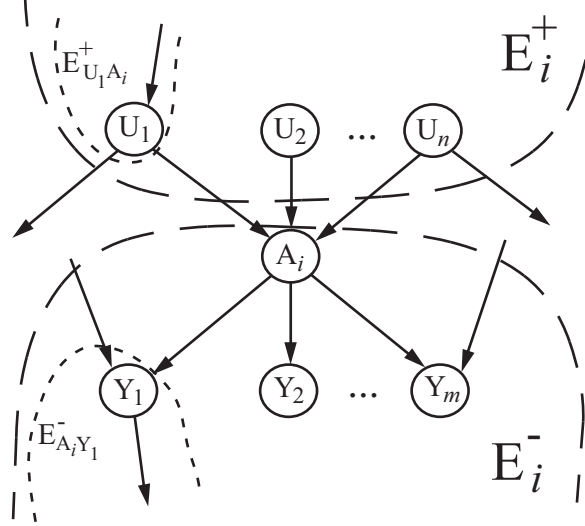


FIG. 4.3 – Partition de l'ensemble E en deux sous-ensembles E_i^+ et E_i^-

Par conséquent, étant donné la valeur de l'évidence e , la mesure de possibilité conditionnelle $\Pi(a_i|e)$ noté souvent par $Bel(a_i)$ peut s'écrire :

$$Bel(a_i) = \Pi(a_i | e) = \Pi(a_i | e_i^-, e_i^+). \quad (4.6)$$

En appliquant la règle du conditionnement possibiliste, la formule précédente s'écrit :

$$\Pi(a_i | e_i^-, e_i^+) = \frac{1}{\Pi(e_i^-, e_i^+)} \cdot \Pi(e_i^-, e_i^+ | a_i) \cdot \Pi(a_i). \quad (4.7)$$

Dans un polyarbre, chaque nœud de l'ensemble E_i^+ est indépendant de chaque nœud de E_i^- conditionnellement à A_i puisque celui-ci d-sépare les nœuds des deux ensembles E_i^+ et E_i^- . Donc l'équation 4.12 devient :

$$\Pi(a_i | e_i^-, e_i^+) = \frac{1}{\Pi(e_i^-, e_i^+)} \cdot \Pi(e_i^- | a_i) \cdot \Pi(e_i^+ | a_i) \cdot \Pi(a_i). \quad (4.8)$$

En remplaçant $\Pi(e_i^+ | a_i) \cdot \Pi(a_i)$ par $\Pi(e_i^+, a_i)$, on obtient :

$$\Pi(a_i | e_i^-, e_i^+) = \frac{1}{\Pi(e_i^-, e_i^+)} \cdot \Pi(e_i^- | a_i) \cdot \Pi(e_i^+, a_i). \quad (4.9)$$

ou encore,

$$Bel(a_i) = \Pi(a_i | e) = \Pi(a_i | e_i^-, e_i^+) = \alpha \cdot \lambda(a_i) \cdot \mu(a_i). \quad (4.10)$$

avec $\alpha = \frac{1}{\Pi(e_i^-, e_i^+)}$ une constante de normalisation ;

$$\lambda(a_i) = \Pi(e_i^- | a_i), \quad (4.11)$$

exprime la portion de l'évidence provenant des fils de A_i ; et

$$\mu(a_i) = \Pi(a_i, e_i^+), \quad (4.12)$$

exprime la portion de l'évidence provenant des parents de A_i .

• Calcul de $\mu(a_i)$:

Commençons par calculer l'expression $\mu(a_i)$. Pour cela, Considérons $u = \{u_1, \dots, u_n\}$ une instantiation des n parents du nœud A_i . Ainsi, $\mu(a_i)$ peut s'écrire :

$$\mu(a_i) = \Pi(a_i, e_i^+) = \max_u \pi(a_i, e_i^+, u). \quad (4.13)$$

En appliquant la formule du conditionnement possibiliste, nous obtenons :

$$\mu(a_i) = \max_u \pi(a_i | e_i^+, u) \cdot \pi(e_i^+, u). \quad (4.14)$$

Du fait que dans un réseau possibiliste quantitatif un nœud A_i ne dépend conditionnellement que de l'ensemble de ces parents U , on peut simplifier la formule précédente en :

$$\mu(a_i) = \max_u \pi(a_i | u) \cdot \pi(e_i^+, u), \quad (4.15)$$

plus précisément, on remplace e_i^+ et u par leurs expressions :

$$\mu(a_i) = \max_u \pi(a_i | u) \cdot \pi(u_1, e_{U_1 A_i}^+, \dots, u_n, e_{U_n A_i}^+). \quad (4.16)$$

En exploitant le fait que dans un polyarbre $\{u_j, e_{U_j A_i}^+\}$ est indépendant de $\{u_k, e_{U_k A_i}^+\}$ pour tout $j \neq k$, nous pouvons écrire :

$$\mu(a_i) = \max_u \pi(a_i | u) \cdot \prod_{j=1}^n \Pi(u_j, e_{U_j A_i}^+), \quad (4.17)$$

ou encore,

$$\mu(a_i) = \max_u \pi(a_i | u) \cdot \prod_{j=1}^n \mu_{A_i}(u_j), \quad (4.18)$$

avec $\mu_{A_i}(u_j) = \Pi(u_j, e_{U_j A_i}^+)$ est le μ -message envoyé par le parent U_j à son fils A_i . Ce message ne dépend que de l'information circulant à travers le lien $U_j \rightarrow A_i$. Si un parent U_j est observé, alors le message $\mu_{A_i}(u_j)$ se réduit à la fonction suivante :

$$\mu_{A_i}(u_j) = \begin{cases} 1 & \text{si } u_j = e_j \\ 0 & \text{si } u_j \neq e_j. \end{cases}$$

• Calcul de $\lambda(a_i)$:

L'expression $\lambda(a_i)$ se calcule quant à elle de la manière suivante :

$$\lambda(a_i) = \Pi(e_i^- | a_i) = \Pi(e_{A_i}, e_{A_i Y_1}^-, \dots, e_{A_i Y_m}^- | a_i). \quad (4.19)$$

Puisque le nœud A_i d-sépare $E_{A_i Y_j}^-$ de $E_{A_i Y_k}^-$ pour tout $j \neq k$, nous avons :

$$\lambda(a_i) = \Pi(e_{A_i} | a_i) \prod_{k=1}^m \Pi(e_{A_i Y_k}^- | a_i), \quad (4.20)$$

finalemt, on a :

$$\lambda(a_i) = \lambda_{A_i}(a_i) \prod_{k=1}^m \lambda_{Y_k}(a_i), \quad (4.21)$$

avec $\lambda_{Y_k}(a_i) = \Pi(e_{A_i Y_k}^- | a_i)$ le λ -message que le nœud Y_k envoie à son parent A_i , et $\lambda_{A_i}(a_i)$ signifie l'évidence locale associée au nœud A_i définie comme suit :

$$\lambda_{A_i}(a_i) = \begin{cases} 0 & \text{si } e_{A_i} \neq a_i \text{ (} A_i \text{ est instancié par } e_{A_i} \neq a_i \text{)} \\ 1 & \text{sinon (} A_i \text{ est instancié par } e_{A_i} = a_i \text{ ou } A_i \text{ n'est pas instancié).} \end{cases}$$

À partir de l'équation 4.23 nous constatons que le nœud A_i ne peut calculer l'expression $\mu(a_i)$ qu'après la réception de tous les μ -messages provenant de ses parents. De même, l'expression de $\lambda(a_i)$ (dans la formule 4.26) ne pourra être calculée qu'après que A_i ait reçu tout les λ -messages de ses fils. La figure 4.6 montre les différents messages reçus et envoyés par un nœud A_i quelconque.

Pour un lien entre un parent U_j et son fils A_i , le message $\lambda_{A_i}(u_j)$, envoyé de A_i vers U_j , est considéré comme la réponse au message $\mu_{A_i}(u_j)$ envoyé précédemment de U_j vers A_i . Si on considère un lien $A_i \rightarrow Y_k$, le message $\mu_{Y_k}(a_i)$, envoyé de A_i vers Y_k , représente aussi la réponse au message déjà envoyé $\lambda_{Y_k}(a_i)$ de Y_k vers A_i .

Après avoir reçu tout les μ -messages de ses parents (dont l'expression de chacun est $\mu_{A_i}(u_j)$), le nœud A_i calcule les λ -messages (dont l'expression de chacun est $\lambda_{A_i}(u_j)$) pour les envoyer à chacun de ses parents $\{U_1, \dots, U_n\}$ lorsque $U_j = u_j$. De même, le nœud A_i calcule les μ -messages (dont l'expression de chacun est $\mu_{Y_k}(a_i)$)

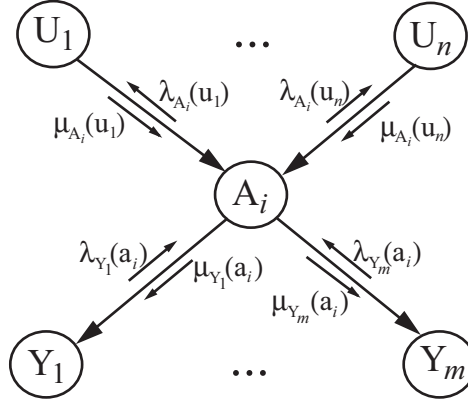


FIG. 4.4 – Les différents messages reçus et envoyés par le nœud A_i

pour les envoyer à chacun de ses fils $\{Y_1, \dots, Y_m\}$.

Commençons par calculer l'expression de $\mu_{Y_k}(a_i)$.

• Calcul de $\mu_{Y_k}(a_i)$:

Considérons de nouveau le nœud A_i et un de ses fils Y_k auquel le message $\mu_{Y_k}(a_i)$ est envoyé de la part de son parent A_i . Le message $\mu_{Y_k}(a_i)$ traduit l'évidence atteignant le nœud Y_k à travers son parent A_i :

$$\mu_{Y_k}(a_i) = \Pi(a_i, e_{A_i Y_k}^+), \quad (4.22)$$

avec $e_{A_i Y_k}^+$ correspondant à la valeur observée de l'évidence provenant du nœud A_i (s'il est observé), des nœuds qu'on peut atteindre à partir de ses parents ainsi que les nœuds qu'on peut atteindre à partir de tous les fils de A_i sauf le nœud Y_k . Ainsi :

$$E_{A_i Y_k}^+ = E_i^+ \cup E_{A_i} \bigcup_{j \neq k} E_{A_i Y_j}^-. \quad (4.23)$$

Nous pouvons écrire :

$$\begin{aligned} \mu_{Y_k}(a_i) &= \Pi(a_i, e_{A_i Y_k}^+) \\ &= \Pi(a_i, e_i^+, e_{A_i}, \bigcup_{j \neq k} e_{A_i Y_j}^-) \\ &= \Pi(e_i^+ | a_i, e_{A_i}, \bigcup_{j \neq k} e_{A_i Y_j}^-) \cdot \Pi(a_i, e_{A_i}, \bigcup_{j \neq k} e_{A_i Y_j}^-). \end{aligned}$$

Du fait que chaque nœud de E_i^+ est indépendant de chaque nœud de E_i^- conditionnellement à A_i , le premier terme de $\mu_{Y_k}(a_i)$ peut se réduire en $\Pi(e_i^+ | a_i)$ d'où :

$$\mu_{Y_k}(a_i) = \Pi(e_i^+ | a_i) \cdot \Pi(e_{A_i}, \bigcup_{j \neq k} e_{A_i Y_j}^- | a_i) \cdot \Pi(a_i). \quad (4.24)$$

En tenant compte que chaque $E_{A_i Y_j}^-$ est indépendant de $E_{A_i Y_l}^-$ pour tout $j \neq l$ conditionnellement à A_i et en remplaçant $\Pi(e_i^+ | a_i)$ par $\frac{\Pi(a_i | e_i^+) \cdot \Pi(e_i^+)}{\Pi(a_i)}$, nous obtenons :

$$\begin{aligned} \mu_{Y_k}(a_i) &= \alpha \cdot \Pi(a_i | e_i^+) \cdot \Pi(e_{A_i} | a_i) \cdot \prod_{j \neq k} \Pi(e_{A_i Y_j}^- | a_i) \\ &= \alpha \cdot \mu(a_i) \cdot \lambda_{A_i}(a_i) \cdot \prod_{j \neq k} \lambda_{Y_j}(a_i), \end{aligned} \quad (4.25)$$

avec α est une constante de normalisation.

Il est à noter que dans le cas où le nœud A_i est observé, $\Pi(a_i | e_i^+)$ se réduit à la fonction suivante :

$$\Pi(a_i | e_i^+) = \begin{cases} 1 & \text{si } a_i = e_{A_i} \\ 0 & \text{si } a_i \neq e_{A_i}. \end{cases}$$

• Calcul de $\lambda_{A_i}(u_j)$:

Le message $\lambda_{A_i}(u_j)$ est envoyé par le nœud A_i à son parent U_j . Il exprime l'évidence parvenant à U_j à travers son fils A_i à partir de l'ensemble des nœuds observées $E_{U_j A_i}^- = E_i^- \bigcup_{l \neq j} e_{U_l A_i}^+$. Cet ensemble engendre, respectivement, les nœuds observés que l'on peut atteindre à partir des fils de A_i , y compris le nœud A_i s'il est observé, ainsi que les nœuds que l'on peut atteindre à partir de tous les parents de A_i sauf le parent U_j . Ainsi, nous avons :

$$\lambda_{A_i}(u_j) = \Pi(e_{U_j A_i}^- | u_j) = \max_{a_i, u_l, l \neq j} \pi(e_{U_j A_i}^-, a_i, u_l | u_j). \quad (4.26)$$

En remplaçant l'évidence $e_{U_j A_i}^-$ par sa valeur :

$$\lambda_{A_i}(u_j) = \max_{a_i, u_l, l \neq j} \pi(e_i^-, \bigcup_{l \neq j} e_{U_l A_i}^+, a_i, u_l | u_j). \quad (4.27)$$

Cette expression peut se décomposer en trois termes comme suit :

$$\lambda_{A_i}(u_j) = \max_{a_i, u_l, l \neq j} \pi(e_i^- | a_i, u_l, u_j, \bigcup_{l \neq j} e_{U_l A_i}^+) \cdot \pi(a_i | u_l, u_j, \bigcup_{l \neq j} e_{U_l A_i}^+) \cdot \pi(u_l, \bigcup_{l \neq j} e_{U_l A_i}^+ | u_j) \quad (4.28)$$

En utilisant les différentes propriétés d'indépendance conditionnelle dans le poly-arbre, cette somme peut se simplifier en :

$$\lambda_{A_i}(u_j) = \beta \cdot \max_{a_i} \pi(e_i^- | a_i) [\max_{u_l, l \neq j} \pi(a_i | u) \cdot \prod_{l \neq j} \pi(u_l, e_{U_l A_i}^+)]. \quad (4.29)$$

Finalement, on obtient :

$$\lambda_{A_i}(u_j) = \beta \cdot \max_{a_i} \lambda(a_i) [\max_{u_l, l \neq j} \pi(a_i | u) \cdot \prod_{l \neq j} \mu_{A_i}(u_l)], \quad (4.30)$$

avec β une constante de normalisation.

À partir des formules 4.23, 4.26, 4.30 et 4.35 nous pouvons constater que :

- La formule 4.23 montre que la quantité $\mu(a_i)$ peut être calculée dès la réception du nœud A_i de tous les μ -messages envoyés par ses parents.
- La formule 4.26 montre que la quantité $\lambda(a_i)$ peut être calculée dès la réception du nœud A_i de tous les λ -messages envoyés par ses fils.
- La formule 4.30 montre que le nœud A_i peut envoyer le message $\mu_{Y_k}(a_i)$ à son fils Y_k dès le calcul de son $\mu(a_i)$ et la réception de tous les λ -messages de ses autres fils.
- La formule 4.35 montre que le nœud A_i peut envoyer le message $\lambda_{A_i}(u_j)$ à son parent U_j dès le calcul de son $\lambda(a_i)$ et la réception de tous les μ -messages de ses autres parents.

La version centralisée de l'algorithme de propagation probabiliste dans les polyarbres introduite par [PS91] peut être aussi prise en compte dans son adaptation au possibiliste, où chaque nœud est visité au plus deux fois, quel que soit le nombre de nœuds observés.

Les principales étapes de l'algorithme de propagation, dans sa version centralisée, sont traduites dans l'algorithme 4.1. Il est à noter que dans l'étape de *collecte de l'évidence*, le sens selon lequel les messages sont envoyés s'appelle **postordre**. Par ailleurs, le sens d'envoi des messages dans l'étape de *distribution de l'évidence* s'appelle, quant à lui, **préordre**. Ces deux sens déterminent l'ordre de traitement des nœuds dans chaque étape.

Exemple 4.2 Soit Π_{G_p} un réseau possibiliste quantitatif formé par 5 nœuds binaires. La structure de ce graphe est un polyarbre illustré dans la figure 4.5. Le tableau 4.2 donne les distributions de possibilités conditionnelles initiales associées à Π_{G_p} .

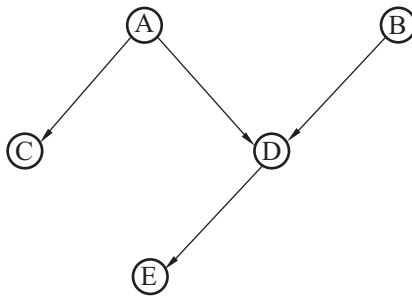


FIG. 4.5 – Réseau possibiliste quantitatif en polyarbre

On suppose, l'observation comme évidence du nœud A à son état a , c'est à dire que $E = \{A\}$ est observé avec la valeur $e = \{A = a\}$. Nous allons appliquer l'algorithme de propagation de messages possibiliste pour déterminer les mesures de

Algorithme : Propagation dans les polyarbres.

• **Entrée :** Un polyarbre formé par un ensemble de variables V muni d'une distribution de possibilités π et un ensemble de variables observées E avec la valeur $E = e$.

• **Sortie :** La distribution de possibilités conditionnelles $\pi(a_i|e)$ de chaque variable A_i non observée, avec $a_i \in D_{A_i}$.

Initialisation :

Pour chaque variable non observée $A_i : \forall a_i \in D_{A_i}$,

- $\mu(a_i) = 1$.

- $\lambda(a_i) = 1$.

Pour chaque variable A_i ayant $U_{A_i} = \emptyset : \forall a_i \in D_{A_i}, \mu(a_i) = \pi(a_i)$.

Pour chaque variable A_i observée à l'état a_i :

- $\mu(a_i) = 1$ et $\mu(a_j) = 0 \forall a_j \in D_{A_i} \setminus \{a_i\}$

- $\lambda(a_i) = 1$ et $\lambda(a_j) = 0 \forall a_j \in D_{A_i} \setminus \{a_i\}$

Collecte de l'évidence :

Pour i **de** 1 **à** $\text{long}(\text{Postordre})$ **faire**

$A \leftarrow \text{Postordre}[i]$

$B \leftarrow$ nœud adjacent à A dans Postordre

calculer $\lambda(A)$ en utilisant (4.26), calculer $\mu(A)$ en utilisant (4.23).

Si B est un parent de A **alors** calculer et envoyer à B un λ -message en utilisant (4.35).

Sinon calculer et envoyer un μ -message en utilisant (4.30).

Fin Pour

Distribution de l'évidence :

Pour i **de** 1 **à** $\text{long}(\text{Préordre})$ **faire**

$A \leftarrow \text{Préordre}[i]$

$\text{Adj} \leftarrow$ ensemble des nœuds adjacents à A dans Préordre

calculer $\lambda(A)$ en utilisant (4.26), calculer $\mu(A)$ en utilisant (4.23).

Pour j **de** 1 **à** $\text{long}(\text{Adj})$ **faire**

$B \leftarrow \text{Adj}(j)$

Si B est un parent de A **alors** calculer et envoyer à B un λ -message en utilisant (4.35).

Sinon calculer et envoyer un μ -message en utilisant (4.30).

Fin Pour

Fin Pour

Marginalisation :

Pour chaque nœud non observée A_i , calculer $\pi(a_i|e)$ en utilisant (4.15).

Algorithme 4.1 - Propagation exacte dans les polyarbres dans le cadre quantitatif

possibilité conditionnelle $\Pi(A_i|E = e)$ pour chaque nœud A_i non observé. L'algorithme suit les étapes suivantes :

• **Initialisation**

L'étape de l'initialisation permet d'initialiser les quantités μ et λ de chaque nœud du graphe. Pour les nœuds A , B , C , D et E , nous obtenons :

- Le nœud A est observé à l'état $A = a$, donc :

| A | $\Pi(A)$ | B | $\Pi(B)$ | A | C | $\Pi(C A)$ | D | E | $\Pi(E D)$ |
|----------|----------|----------|----------|----------|----------|------------|----------|----------|------------|
| a | 1 | b | 1 | a | c | 1 | d | e | 0.2 |
| $\neg a$ | 0.3 | $\neg b$ | 0.6 | a | $\neg c$ | 0.5 | d | $\neg e$ | 1 |
| | | | | $\neg a$ | c | 1 | $\neg d$ | e | 0.75 |
| | | | | $\neg a$ | $\neg c$ | 0.25 | $\neg d$ | $\neg e$ | 1 |

| A | B | D | $\Pi(D A, B)$ | A | B | D | $\Pi(D A, B)$ |
|-----|----------|----------|---------------|----------|----------|----------|---------------|
| a | b | d | 1 | $\neg a$ | b | d | 1 |
| a | b | $\neg d$ | 0.2 | $\neg a$ | b | $\neg d$ | 0.3 |
| a | $\neg b$ | d | 1 | $\neg a$ | $\neg b$ | d | 0.6 |
| a | $\neg b$ | $\neg d$ | 0.1 | $\neg a$ | $\neg b$ | $\neg d$ | 1 |

TAB. 4.2 – Distributions de possibilités conditionnelles initiales associées à Π_{G_p}

- $\mu(a) = 1$ $\lambda(a) = 1$
 $\mu(\neg a) = 0$ $\lambda(\neg a) = 0$
- Le nœud B est un nœud racine (sans parents), d'où : $\mu(B) = \Pi(B)$
 - $\mu(b) = 1$ $\lambda(b) = 1$
 $\mu(\neg b) = 0.6$ $\lambda(\neg b) = 1$
- Le nœud C est initialisé comme suit :
 - $\mu(c) = 1$ $\lambda(c) = 1$
 $\mu(\neg c) = 1$ $\lambda(\neg c) = 1$
- Le nœud D est initialisé comme suit :
 - $\mu(d) = 1$ $\lambda(d) = 1$
 $\mu(\neg d) = 1$ $\lambda(\neg d) = 1$
- Le nœud E est initialisé comme suit :
 - $\mu(e) = 1$ $\lambda(e) = 1$
 $\mu(\neg e) = 1$ $\lambda(\neg e) = 1$

• Collecte de l'évidence

Le **Postordre** selon lequel les nœuds seront traités pendant l'étape de collecte de l'évidence est : le nœud E puis B puis D et enfin C .

* Le nœud E :

- Etant donné que le nœud E n'a pas de fils, son $\lambda(E)$ reste à sa valeur initiale.
- La quantité $\mu(E)$ ne pourra pas être calculée, car le nœud E n'a pas encore reçu le μ -message de son parent D .
- Vu que le nœud E possède un unique parent D et que son $\lambda(E)$ est déterminé,

le message $\lambda_E(D)$ peut être calculé en utilisant (4.35) et sera envoyé à D , ce qui donne : $\forall d \in D_D$,

$$\lambda_E(d) = \max_e \lambda(e)\pi(e|d).$$

En développant cette formule, nous obtenons :

$$\begin{aligned} \lambda_E(d) &= \max\{\lambda(e)\pi(e|d), \lambda(\neg e)\pi(\neg e|d)\} &= \max\{1 \times 0.2, 1 \times 1\} &= 1 \\ \lambda_E(\neg d) &= \max\{\lambda(e)\pi(e|\neg d), \lambda(\neg e)\pi(\neg e|\neg d)\} &= \max\{1 \times 0.75, 1 \times 1\} &= 1 \end{aligned}$$

* Le nœud B :

- Vu que B est un nœud racine, la quantité $\mu(B)$ prend la valeur déjà calculée lors de l'étape de l'initialisation.

- Le nœud B , n'ayant pas reçu le λ -message de D , ne peut pas calculer sa quantité $\lambda(B)$.

- $\mu(B)$ étant calculée, la racine B peut envoyer à son fils D un message $\mu_D(B)$ en utilisant (4.30). Vu que D est l'unique fils de B , nous pouvons écrire :

$\forall b \in D_B$,

$$\mu_D(b) = \mu(b)\lambda_B(b).$$

Où encore,

$$\begin{aligned} \mu_D(b) &= \mu(b)\lambda_B(b) &= 1 \times 1 &= 1 \\ \mu_D(\neg b) &= \mu(\neg b)\lambda_B(\neg b) &= 0.6 \times 1 &= 0.6 \end{aligned}$$

* Le nœud D :

- Le nœud D a reçu le λ -message de son unique fils E , donc il peut calculer sa quantité $\lambda(D)$ selon (4.26). D'où, $\forall d \in D_D$

$$\lambda(d) = \lambda_D(d)\lambda_E(d).$$

Ce qui nous donne :

$$\begin{aligned} \lambda(d) &= \lambda_D(d)\lambda_E(d) &= 1 \times 1 &= 1 \\ \lambda(\neg d) &= \lambda_D(\neg d)\lambda_E(\neg d) &= 1 \times 1 &= 1 \end{aligned}$$

- La quantité $\mu(D)$ ne peut pas être calculée puisque le nœud D n'a pas encore reçu tout les μ -messages de ses parents (il n'a reçu uniquement que le μ -message de son parent B).

- Etant donné que le nœud D a calculé son $\lambda(D)$ et qu'il a reçu un μ -message de son parent B , il peut calculer et envoyé un λ -message à son autre parent A . Ce message s'exprime selon la formule (4.35) : $\forall a \in D_A$

$$\lambda_D(a) = \max_d \lambda(d) [\max_b \pi(d|a, b) \cdot \mu_D(b)].$$

Ce qui nous donne :

$$\begin{aligned}\lambda_D(a) &= \max_d \lambda(d) [\max\{\pi(d|a, b) \times \mu_D(b), \pi(d|a, \neg b) \times \mu_D(\neg b)\}] \\ &= \max\{1 \times [\max\{1 \times 1, 1 \times 0.6\}], 1 \times [\max\{0.2 \times 1, 0.1 \times 0.6\}]\} \\ &= \max\{1 \times 1, 1 \times 0.2\} = 1\end{aligned}$$

$$\begin{aligned}\lambda_D(\neg a) &= \max_d \lambda(d) [\max\{\pi(d|\neg a, b) \times \mu_D(b), \pi(d|\neg a, \neg b) \times \mu_D(\neg b)\}] \\ &= \max\{1 \times [\max\{1 \times 1, 0.6 \times 0.6\}], 1 \times [\max\{0.3 \times 1, 0.1 \times 0.6\}]\} \\ &= \max\{1 \times 1, 1 \times 0.6\} = 1\end{aligned}$$

* Le nœud C :

- Etant donné que le nœud C n'a pas de fils, son $\lambda(C)$ reste à sa valeur initiale.
- La quantité $\mu(C)$ ne pourra pas être calculée, car le nœud C n'a pas encore reçu le μ -message de son parent A .
- Vu que le nœud C possède un unique parent A et que son $\lambda(C)$ est déterminé, le message $\lambda_C(A)$ peut être calculé en utilisant (4.35) et sera envoyé à A , ce qui donne : $\forall a \in D_A$,

$$\lambda_C(a) = \max_c \lambda(c) \pi(c|a).$$

En développant cette formule, nous obtenons :

$$\begin{aligned}\lambda_C(a) &= \max\{\lambda(c) \pi(c|a), \lambda(\neg c) \pi(\neg c|a)\} = \max\{1 \times 1, 1 \times 0.5\} = 1 \\ \lambda_C(\neg a) &= \max\{\lambda(c) \pi(c|\neg a), \lambda(\neg c) \pi(\neg c|\neg a)\} = \max\{1 \times 1, 1 \times 0.25\} = 1\end{aligned}$$

La figure 4.6 illustre les messages envoyés ainsi que leur ordre d'émission pendant l'étape de collecte de l'évidence.

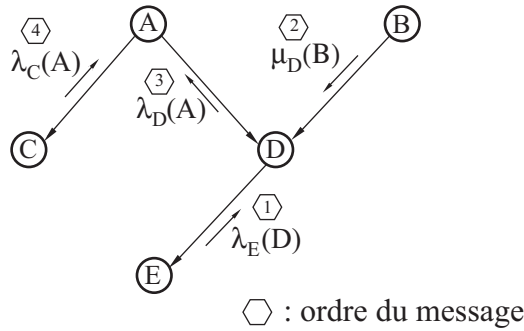


FIG. 4.6 – Messages envoyés lors de l'étape de collecte de l'évidence

• Distribution de l'évidence

Le **Préordre** selon lequel les nœuds seront traités pendant cette étape de distribution de l'évidence est : le nœud A avec son $Adj = \{C, D\}$, le nœud C avec son $Adj = \emptyset$, le nœud D avec son $Adj = \{B, E\}$, le nœud B avec son $Adj = \emptyset$ et le nœud E avec son $Adj = \emptyset$.

* *Le nœud A :*

- *Le nœud A est un nœud racine, sa quantité $\mu(A)$ restera sur sa valeur initiale.*
- *Le nœud A a reçu les λ -messages de ses fils C et D et par suite il peut calculer sa quantité $\lambda(A)$ en utilisant (4.26) comme suit : $\forall a \in D_A$*

$$\lambda(a) = \lambda_A(a) \cdot \lambda_C(a) \cdot \lambda_D(a).$$

Ce qui nous donne :

$$\lambda(a) = \lambda_A(a) \cdot \lambda_C(a) \cdot \lambda_D(a) = 1 \times 1 \times 1 = 1$$

$$\lambda(\neg a) = \lambda_A(\neg a) \cdot \lambda_C(\neg a) \cdot \lambda_D(\neg a) = 0 \times 1 \times 1 = 0$$

- *Le nœud C se trouve dans l'ensemble des adjacents de A. Puisqu'il est le fils de A, le nœud C recevra de A le message $\mu_C(a)$ calculé comme suit : $\forall a \in D_A$*

$$\mu_C(a) = \mu(a) \cdot \lambda_A(a) \cdot \lambda_D(a).$$

Ce qui nous donne :

$$\mu_C(a) = \mu(a) \cdot \lambda_A(a) \cdot \lambda_D(a) = 1 \times 1 \times 1 = 1$$

$$\mu_C(\neg a) = \mu(\neg a) \cdot \lambda_A(\neg a) \cdot \lambda_D(\neg a) = 0 \times 0 \times 1 = 0$$

De même, pour le nœud D. Il recevra un message $\mu_D(a)$ de son parent A dont l'expression est $\forall a \in D_A \mu_D(a) = \mu(a) \cdot \lambda_A(a) \cdot \lambda_C(a)$:

$$\mu_D(a) = \mu(a) \cdot \lambda_A(a) \cdot \lambda_C(a) = 1 \times 1 \times 1 = 1$$

$$\mu_D(\neg a) = \mu(\neg a) \cdot \lambda_A(\neg a) \cdot \lambda_C(\neg a) = 0 \times 0 \times 1 = 0$$

* *Le nœud C :*

- *Ayant reçu le message $\mu_C(a)$ de son parent A, le nœud C peut calculer sa quantité $\mu(C)$ en utilisant (4.23) comme suit : $\forall c \in D_C$,*

$$\mu(c) = \max_a \pi(c|a) \cdot \mu_C(a).$$

Ce qui nous donne après calcul :

$$\mu(c) = \max\{\pi(c|a) \times \mu_C(a), \pi(c|\neg a) \times \mu_C(\neg a)\}$$

$$= \max\{1 \times 1, 1 \times 0\} = 1$$

$$\mu(\neg c) = \max\{\pi(\neg c|a) \times \mu_C(a), \pi(\neg c|\neg a) \times \mu_C(\neg a)\}$$

$$= \max\{0.5 \times 1, 0.25 \times 0\} = 0.5$$

* *Le nœud D :*

- *Le nœud D a reçu les deux messages provenant de ces deux parents A et B. Il peut maintenant calculer sa quantité $\mu(D)$ en utilisant (4.23). D'où $\forall d \in D_D$,*

$$\mu(d) = \max_{ab} \pi(c|a, b) \cdot \mu_D(a) \cdot \mu_D(b).$$

En développant cette expression, nous obtenons :

$$\begin{aligned}
\mu(d) &= \max_{ab} \pi(d|a, b) \times \mu_D(a) \times \mu_D(b) \\
&= \max\{(1 \times 1 \times 1), (1 \times 1 \times 0.6), (1 \times 0 \times 1), (0.60 \times 0.6)\} = 1 \\
\mu(\neg d) &= \max_{ab} \pi(\neg d|a, b) \times \mu_D(a) \times \mu_D(b) \\
&= \max\{(0.2 \times 1 \times 1), (0.1 \times 1 \times 0.6), (0.3 \times 0 \times 1), (1 \times 0 \times 0.6)\} = 0.2
\end{aligned}$$

- Le nœud B se trouve dans l'ensemble des adjacents de D . Puisqu'il est le parent de D , le nœud B recevra le message $\lambda_D(B)$ calculé comme suit : $\forall b \in D_B$,

$$\lambda_D(b) = \max_d \lambda(d) [\max_a \pi(d|a, b) \cdot \mu_D(a)].$$

Ce qui nous donne :

$$\begin{aligned}
\lambda_D(b) &= \max_d \lambda(d) [\max\{\pi(d|a, b) \times \mu_D(a), \pi(d|\neg a, b) \times \mu_D(\neg a)\}] \\
&= \max\{1 \times [\max\{1 \times 1, 1 \times 0\}], 1 \times [\max\{0.2 \times 1, 0.3 \times 0\}]\} \\
&= \max\{1 \times 1, 1 \times 0.2\} = 1 \\
\lambda_D(\neg b) &= \max_d \lambda(d) [\max\{\pi(d|a, \neg b) \times \mu_D(a), \pi(d|\neg a, \neg b) \times \mu_D(\neg a)\}] \\
&= \max\{1 \times [\max\{1 \times 1, 0.6 \times 0\}], 1 \times [\max\{0.1 \times 1, 1 \times 0\}]\} \\
&= \max\{1 \times 1, 1 \times 0.1\} = 1
\end{aligned}$$

- Le nœud E se trouve dans l'ensemble des adjacents de D . Comme il est le fils de ce dernier, le nœud E recevra de D le message $\mu_E(D)$ calculé en utilisant (4.30) comme suit : $\forall d \in D_D$,

$$\mu_E(d) = \mu(d) \cdot \lambda_D(d).$$

Ce qui donne en développant la formule précédente :

$$\begin{aligned}
\mu_E(d) &= \mu(d) \cdot \lambda_D(d) = 1 \times 1 = 1 \\
\mu_E(\neg d) &= \mu(\neg d) \cdot \lambda_D(\neg d) = 0.2 \times 1 = 0.2
\end{aligned}$$

* Le nœud B :

- Le nœud B peut calculer sa quantité $\lambda(B)$ après la réception du message $\lambda_D(B)$ de son fils D . Ainsi, en utilisant (4.26), nous écrivons $\forall b \in D_B$,

$$\lambda(b) = \lambda_B(b) \cdot \lambda_D(b).$$

Où encore :

$$\begin{aligned}
\lambda(b) &= \lambda_B(b) \cdot \lambda_D(b) = 1 \times 1 = 1 \\
\lambda(\neg b) &= \lambda_B(\neg b) \cdot \lambda_D(\neg b) = 1 \times 1 = 1
\end{aligned}$$

* Le nœud E :

- Le nœud E reçoit le message $\mu_E(D)$ de son parent D . De ce fait, il peut calculer sa quantité $\mu(E)$ en utilisant (4.23) comme suit : $\forall e \in D_E$,

$$\mu(e) = \max_d \pi(e|d) \cdot \mu_E(d).$$

ce qui donne après calcul :

$$\begin{aligned}\mu(e) &= \max\{\pi(e|d) \times \mu_E(d), \pi(e|\neg d) \times \mu_E(\neg d)\} \\ &= \max\{0.2 \times 1, 0.75 \times 0.2\} = 0.2\end{aligned}$$

$$\begin{aligned}\mu(\neg e) &= \max\{\pi(\neg e|d) \times \mu_E(d), \pi(\neg e|\neg d) \times \mu_E(\neg d)\} \\ &= \max\{1 \times 1, 1 \times 0.2\} = 1\end{aligned}$$

La figure 4.7 illustre les messages envoyés ainsi que leur ordre d'émission pendant l'étape de distribution de l'évidence.

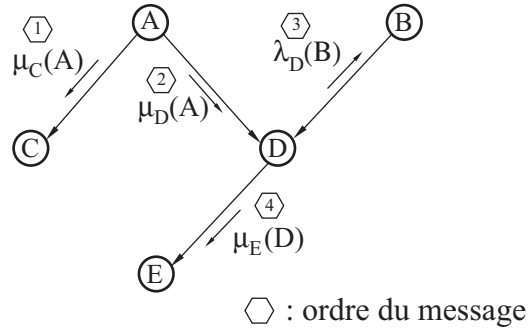


FIG. 4.7 – Messages envoyés lors de l'étape de distribution de l'évidence

• Marginalisation

Pendant cette étape, chacun des nœuds non observés B , C , D et E calcule sa mesure de possibilité conditionnellement à l'évidence $A = a$ selon la formule (4.15).

* Le nœud B :

$$\begin{aligned}\pi(b|a) &= \mu(b)\lambda(b) = 1 \times 1 = 1 \\ \pi(\neg b|a) &= \mu(\neg b)\lambda(\neg b) = 0.6 \times 1 = 0.6\end{aligned}$$

* Le nœud C :

$$\begin{aligned}\pi(c|a) &= \mu(c)\lambda(c) = 1 \times 1 = 1 \\ \pi(\neg c|a) &= \mu(\neg c)\lambda(\neg c) = 0.5 \times 1 = 0.5\end{aligned}$$

* Le nœud D :

$$\begin{aligned}\pi(d|a) &= \mu(d)\lambda(d) = 1 \times 1 = 1 \\ \pi(\neg d|a) &= \mu(\neg d)\lambda(\neg d) = 0.2 \times 1 = 0.2\end{aligned}$$

* Le nœud E :

$$\begin{aligned}\pi(e|a) &= \mu(e)\lambda(e) = 0.2 \times 1 = 0.2 \\ \pi(\neg e|a) &= \mu(\neg e)\lambda(\neg e) = 1 \times 1 = 1\end{aligned}$$

4.4 Adaptation de l'algorithme *LBP* aux réseaux quantitatifs

Cette section présente l'adaptation directe de l'algorithme *Loopy Belief Propagation* à partir de l'algorithme de propagation exact quantitatif appliqué sur les polyarbres. Le principe de l'algorithme adapté LBP reste le même (que sa version probabiliste), mais il sera appliqué sur des réseaux possibilistes quantitatifs (basés sur le produit). L'aspect temps, symbolisé par t , est ajouté afin de représenter la notion d'itération entre les étapes de l'algorithme LBP possibiliste.

Etant donné un ensemble de nœuds observés E instancié par la valeur e , la mesure de possibilité d'un nœud non observé A_i conditionnellement à e , notée par $\Pi(a_i | e)$, est approximée par :

$$\forall a_i \in D_{A_i} \text{ Bel}(a_i) = \Pi(a_i | e) \approx \alpha \lambda(a_i) \mu(a_i), \quad (4.31)$$

avec α une constante de normalisation et $\lambda(a_i)$ et $\mu(a_i)$ sont calculés itérativement à partir des messages reçus et envoyés par le nœud A_i à chaque itération.

Pendant une itération t , l'effet des messages reçus, respectivement par les m fils et n parents de A_i , sont représentés par :

$$\lambda^{(t)}(a_i) = \lambda_{A_i}(a_i) \prod_{k=1}^m \lambda_{Y_k}^{(t)}(a_i), \quad (4.32)$$

avec $\lambda_{Y_k}^{(t)}(a_i)$ le message envoyé par le fils Y_k à son parent A_i à l'itération t .

$$\mu^{(t)}(a_i) = \max_u \pi(a_i | u) \prod_{j=1}^n \mu_{A_i}^{(t)}(u_j), \quad (4.33)$$

avec $\mu_{A_i}^{(t)}(u_j)$ le message envoyé par le parent U_j vers A_i à l'itération t .

La figure 4.8 résume les messages entrants au nœud A_i à l'itération t .

À l'itération $t + 1$, le nœud A_i envoie à chacun de ses parents U_j , $j = 1..n$ le λ -message :

$$\lambda_{A_i}^{(t+1)}(u_j) = \beta \cdot \max_{a_i} \lambda^{(t)}(a_i) [\max_{u_l, l \neq j} \pi(a_i | u) \cdot \prod_{l \neq j} \mu_{A_i}^{(t)}(u_l)], \quad (4.34)$$

avec β une constante de normalisation.

De même, à l'itération $t + 1$, le nœud A_i envoie à chacun de ses fils Y_k , $k = 1..m$ le μ -message :

$$\mu_{Y_k}^{(t+1)}(a_i) = \alpha \mu^{(t)}(a_i) \lambda_{A_i}(a_i) \prod_{j \neq k} \lambda_{Y_j}^{(t)}(a_i), \quad (4.35)$$

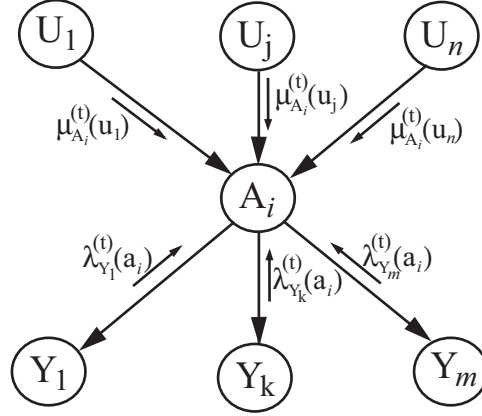


FIG. 4.8 – Messages entrants au nœud A_i à l'itération t

avec α une constante de normalisation.

La figure 4.9 résume les messages sortants du nœud A_i pour l'itération $t + 1$.

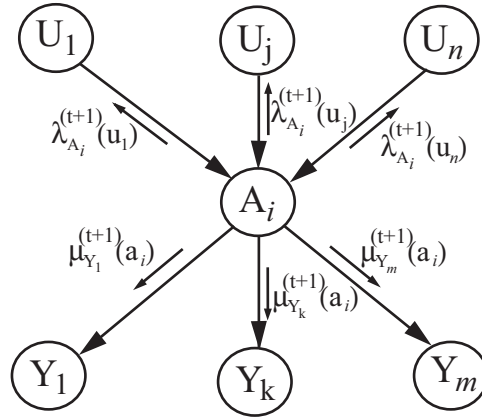


FIG. 4.9 – Messages sortants du nœud A_i pour l'itération $t + 1$

Pendant chaque itération t , tous les nœuds calculent, parallèlement, leur $\lambda^{(t)}(a_i)$ et $\mu^{(t)}(a_i)$ ce qui permet d'approximer, progressivement, $\pi(a_i | u)$ jusqu'à atteindre la condition de convergence.

La version possibiliste de l'algorithme LBP est décrite par l'algorithme 4.2.

Exemple 4.3 Soit le réseau possibiliste quantitatif à connexions multiples ayant une structure graphique identique à celle présentée dans la figure 4.1. Les distributions de possibilités initiales et conditionnelles des nœuds A , B , C et D sont données dans le tableau 4.3.

On suppose l'observation d'une évidence $D = d$. Les degrés de possibilité associés

Algorithme : Loopy Belief Propagation possibiliste.

Initialisation :

- N = l'ensemble de tous les nœuds.
- M = l'ensemble des nœuds non observés.
- Tous les messages sont initialisés par le vecteur 1.
- Pour chaque variable non observée $A_i \in M$: $\forall a_i \in D_{A_i}, Bel(a_i) = 0$.
- Pour chaque variable A_i ayant $U_{A_i} = \emptyset$: $\forall a_i \in D_{A_i}, \mu^{(0)}(a_i) = \pi(a_i)$.
- Pour chaque variable A_i observée à l'état a_i :
 - $\mu^{(0)}(a_i) = 1$ et $\mu^{(0)}(a_j) = 0 \forall a_j \in D_{A_i} \setminus \{a_i\}$.
 - $\lambda^{(0)}(a_i) = 1$ et $\lambda^{(0)}(a_j) = 0 \forall a_j \in D_{A_i} \setminus \{a_i\}$.
- $t = 0$.
- convergence = *false*.
- maxitr = 100.
- tol = 0.0001.

Propagation des messages :

```

Tant que NOT convergence and t < max_itr faire
  Pour i de 1 à long(N) faire
    calculer  $\lambda^{(t)}(a_i)$  en utilisant (4.37);
    calculer  $\mu^{(t)}(a_i)$  en utilisant (4.38);
  Fin Pour
  Pour i de 1 à long(M) faire
    OldBel( $A_i$ )  $\leftarrow$  Bel( $A_i$ );
    calculer Bel( $A_i$ ) en utilisant (4.36);
  Fin Pour
  Si  $\forall m \in \{1, \dots, long(M)\}, Bel(A_m) - OldBel(A_m) < tol$  alors
    convergence = true;
  Fin Si
  Si NOT convergence alors
    Pour i de 1 à long(N) faire
      calculer pour chaque parent  $U_j$  de  $A_i$  ;  $\lambda_{A_i}^{(t+1)}(u_j)$  en utilisant (4.39);
      calculer pour chaque fils  $Y_k$  de  $A_i$  ;  $\mu_{Y_k}^{(t+1)}(a_i)$  en utilisant (4.40);
    Fin Pour
    t = t + 1;
  Fin Si
Fin Tant que
    
```

Algorithme 4.2 - Algorithme LBP adapté au cadre possibiliste quantitatif

au nœuds non observés A , B et C après propagation de l'évidence sont approchés selon les étapes suivantes :

► **Initialisation**

- Tous les messages sont initialisés par le vecteur 1.

| | | B | A | $\Pi(B A)$ | C | A | $\Pi(C A)$ |
|----------|----------|----------|----------|------------|----------|----------|------------|
| A | $\Pi(A)$ | $\neg b$ | $\neg a$ | 0.2 | $\neg c$ | $\neg a$ | 1 |
| $\neg a$ | 0.3 | $\neg b$ | a | 1 | $\neg c$ | a | 0.4 |
| a | 1 | b | $\neg a$ | 1 | c | $\neg a$ | 0 |
| | | b | a | 0.3 | c | a | 1 |

| D | C | B | $\Pi(D B, C)$ | D | C | B | $\Pi(D B, C)$ |
|----------|----------|----------|---------------|-----|----------|----------|---------------|
| $\neg d$ | $\neg c$ | $\neg b$ | 1 | d | $\neg c$ | $\neg b$ | 1 |
| $\neg d$ | $\neg c$ | b | 0 | d | $\neg c$ | b | 1 |
| $\neg d$ | c | $\neg b$ | 0.8 | d | c | $\neg b$ | 1 |
| $\neg d$ | c | b | 0.1 | d | c | b | 1 |

TAB. 4.3 – Distributions de possibilités conditionnelles initiales

- A est un nœud racine donc :

$$\mu^{(0)}(a) = \pi(a) = 1$$

$$\mu^{(0)}(\neg a) = \pi(\neg a) = 0.3$$

- Le nœud D est observé à l'état $D = d$, donc :

$$\lambda^{(0)}(d) = 1$$

$$\lambda^{(0)}(\neg d) = 0$$

► Propagation de messages

À l'itération $t = 0$

Calcul de $\mu^{(0)}$ et $\lambda^{(0)}$ pour chaque nœud

* Le nœud A :

$$\mu^{(0)}(a) = 1 \quad \lambda^{(0)}(a) = \lambda_A(a)\lambda_B^{(0)}(a)\lambda_C^{(0)}(a) = 1$$

$$\mu^{(0)}(\neg a) = 0.3 \quad \lambda^{(0)}(\neg a) = \lambda_A(\neg a)\lambda_B^{(0)}(\neg a)\lambda_C^{(0)}(\neg a) = 1$$

* Le nœud B :

$$\mu^{(0)}(b) = \max_a \pi(b|a)\mu_B^{(0)}(a) = 1 \quad \lambda^{(0)}(b) = \lambda_B(b)\lambda_D^{(0)}(b) = 1$$

$$\mu^{(0)}(\neg b) = \max_a \pi(\neg b|a)\mu_B^{(0)}(a) = 1 \quad \lambda^{(0)}(\neg b) = \lambda_B(\neg b)\lambda_D^{(0)}(\neg b) = 1$$

* Le nœud C :

$$\mu^{(0)}(c) = \max_a \pi(c|a)\mu_C^{(0)}(a) = 1 \quad \lambda^{(0)}(c) = \lambda_C(c)\lambda_D^{(0)}(c) = 1$$

$$\mu^{(0)}(\neg c) = \max_a \pi(\neg c|a)\mu_C^{(0)}(a) = 1 \quad \lambda^{(0)}(\neg c) = \lambda_C(\neg c)\lambda_D^{(0)}(\neg c) = 1$$

* Le nœud D :

$$\begin{aligned}\mu^{(0)}(d) &= \max_{b,c} \pi(d|b,c) \mu_D^{(0)}(b) \mu_D^{(0)}(c) &= 1 & \quad \lambda^{(0)}(d) = \lambda_D(d) &= 1 \\ \mu^{(0)}(\neg d) &= \max_{b,c} \pi(\neg d|b,c) \mu_D^{(0)}(b) \mu_D^{(0)}(c) &= 1 & \quad \lambda^{(0)}(\neg d) = \lambda_D(\neg d) &= 0\end{aligned}$$

Calcul de Bel pour chaque nœud

| A | $Bel(A)$ | B | $Bel(B)$ | C | $Bel(C)$ |
|----------|----------|----------|----------|----------|----------|
| a | 1 | b | 1 | c | 1 |
| $\neg a$ | 0.3 | $\neg b$ | 1 | $\neg c$ | 1 |

Calcul des messages normalisés sortants de chaque nœud

* Le nœud A :

$$\begin{aligned}\mu_B^{(1)}(a) &= \mu^{(0)}(a) \lambda_A(a) \lambda_C^{(0)}(a) &= 1 \\ \mu_B^{(1)}(\neg a) &= \mu^{(0)}(\neg a) \lambda_A(\neg a) \lambda_C^{(0)}(\neg a) &= 0.3 \\ \mu_C^{(1)}(a) &= \mu^{(0)}(a) \lambda_A(a) \lambda_B^{(0)}(a) &= 1 \\ \mu_C^{(1)}(\neg a) &= \mu^{(0)}(\neg a) \lambda_A(\neg a) \lambda_B^{(0)}(\neg a) &= 0.3\end{aligned}$$

* Le nœud B :

$$\begin{aligned}\mu_D^{(1)}(b) &= \mu^{(0)}(b) \lambda_B(b) &= 1 \\ \mu_D^{(1)}(\neg b) &= \mu^{(0)}(\neg b) \lambda_B(\neg b) &= 1 \\ \lambda_B^{(1)}(a) &= \max_b \lambda^{(0)}(b) \pi(b|a) &= 1 \\ \lambda_B^{(1)}(\neg a) &= \max_b \lambda^{(0)}(b) \pi(b|\neg a) &= 1\end{aligned}$$

* Le nœud C :

$$\begin{aligned}\mu_D^{(1)}(c) &= \mu^{(0)}(c) \lambda_C(c) &= 1 \\ \mu_D^{(1)}(\neg c) &= \mu^{(0)}(\neg c) \lambda_C(\neg c) &= 1 \\ \lambda_C^{(1)}(a) &= \max_c \lambda^{(0)}(c) \pi(c|a) &= 1 \\ \lambda_C^{(1)}(\neg a) &= \max_c \lambda^{(0)}(c) \pi(c|\neg a) &= 1\end{aligned}$$

* Le nœud D :

$$\begin{aligned}\lambda_D^{(1)}(b) &= \max_d \lambda^{(0)}(d) [\max_c \pi(d|b,c) \mu_D^{(0)}(c)] &= 1 \\ \lambda_D^{(1)}(\neg b) &= \max_d \lambda^{(0)}(d) [\max_c \pi(d|\neg b,c) \mu_D^{(0)}(c)] &= 1 \\ \lambda_D^{(1)}(c) &= \max_d \lambda^{(0)}(d) [\max_b \pi(d|b,c) \mu_D^{(0)}(b)] &= 1 \\ \lambda_D^{(1)}(\neg c) &= \max_d \lambda^{(0)}(d) [\max_b \pi(d|b,\neg c) \mu_D^{(0)}(b)] &= 1\end{aligned}$$

La figure 4.10 illustre les messages sortants de chaque nœud, suite à l'itération $t = 0$.

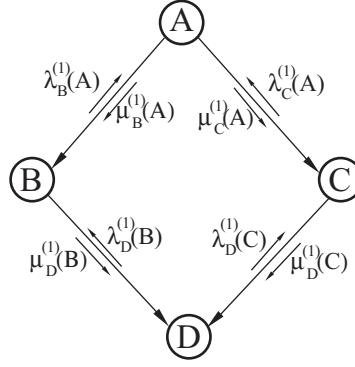


FIG. 4.10 – Messages sortants de chaque nœud à la fin de l'itération $t=0$

À l'itération $t = 1$

Calcul de $\mu^{(1)}$ et $\lambda^{(1)}$ pour chaque nœud

* Le nœud A :

$$\begin{aligned}\mu^{(1)}(a) &= 1 & \lambda^{(1)}(a) &= \lambda_A(a)\lambda_B^{(1)}(a)\lambda_C^{(1)}(a) = 1 \\ \mu^{(1)}(\neg a) &= 0.3 & \lambda^{(1)}(\neg a) &= \lambda_A(\neg a)\lambda_B^{(1)}(\neg a)\lambda_C^{(1)}(\neg a) = 1\end{aligned}$$

* Le nœud B :

$$\begin{aligned}\mu^{(1)}(b) &= \max_a \pi(b|a)\mu_B^{(1)}(a) = 0.3 & \lambda^{(1)}(b) &= \lambda_B(b)\lambda_D^{(1)}(b) = 1 \\ \mu^{(1)}(\neg b) &= \max_a \pi(\neg b|a)\mu_B^{(1)}(a) = 1 & \lambda^{(1)}(\neg b) &= \lambda_B(\neg b)\lambda_D^{(1)}(\neg b) = 1\end{aligned}$$

* Le nœud C :

$$\begin{aligned}\mu^{(1)}(c) &= \max_a \pi(c|a)\mu_C^{(1)}(a) = 1 & \lambda^{(1)}(c) &= \lambda_C(c)\lambda_D^{(1)}(c) = 1 \\ \mu^{(1)}(\neg c) &= \max_a \pi(\neg c|a)\mu_C^{(1)}(a) = 0.4 & \lambda^{(1)}(\neg c) &= \lambda_C(\neg c)\lambda_D^{(1)}(\neg c) = 1\end{aligned}$$

* Le nœud D :

$$\begin{aligned}\mu^{(1)}(d) &= \max_{b,c} \pi(d|b,c)\mu_D^{(1)}(b)\mu_D^{(1)}(c) = 1 & \lambda^{(1)}(d) &= \lambda_D(d) = 1 \\ \mu^{(1)}(\neg d) &= \max_{b,c} \pi(\neg d|b,c)\mu_D^{(1)}(b)\mu_D^{(1)}(c) = 1 & \lambda^{(1)}(\neg d) &= \lambda_D(\neg d) = 0\end{aligned}$$

Calcul de Bel pour chaque nœud

| A | $Bel(A)$ | B | $Bel(B)$ | C | $Bel(C)$ |
|----------|----------|----------|----------|----------|----------|
| a | 1 | b | 0.3 | c | 1 |
| $\neg a$ | 0.3 | $\neg b$ | 1 | $\neg c$ | 0.4 |

Calcul des messages normalisés sortants de chaque nœud

* Le nœud A :

$$\begin{aligned}\mu_B^{(2)}(a) &= \mu^{(1)}(a)\lambda_A(a)\lambda_C^{(1)}(a) &= 1 \\ \mu_B^{(2)}(\neg a) &= \mu^{(1)}(\neg a)\lambda_A(\neg a)\lambda_C^{(1)}(\neg a) &= 0.3\end{aligned}$$

$$\begin{aligned}\mu_C^{(2)}(a) &= \mu^{(1)}(a)\lambda_A(a)\lambda_B^{(1)}(a) &= 1 \\ \mu_C^{(2)}(\neg a) &= \mu^{(1)}(\neg a)\lambda_A(\neg a)\lambda_B^{(1)}(\neg a) &= 0.3\end{aligned}$$

* Le nœud B :

$$\begin{aligned}\mu_D^{(2)}(b) &= \mu^{(1)}(b)\lambda_B(b) &= 0.3 \\ \mu_D^{(2)}(\neg b) &= \mu^{(1)}(\neg b)\lambda_B(\neg b) &= 1\end{aligned}$$

$$\begin{aligned}\lambda_B^{(2)}(a) &= \max_b \lambda^{(1)}(b)\pi(b|a) &= 1 \\ \lambda_B^{(2)}(\neg a) &= \max_b \lambda^{(1)}(b)\pi(b|\neg a) &= 1\end{aligned}$$

* Le nœud C :

$$\begin{aligned}\mu_D^{(2)}(c) &= \mu^{(1)}(c)\lambda_C(c) &= 1 \\ \mu_D^{(2)}(\neg c) &= \mu^{(1)}(\neg c)\lambda_C(\neg c) &= 0.4\end{aligned}$$

$$\begin{aligned}\lambda_C^{(2)}(a) &= \max_c \lambda^{(1)}(c)\pi(c|a) &= 1 \\ \lambda_C^{(2)}(\neg a) &= \max_c \lambda^{(1)}(c)\pi(c|\neg a) &= 1\end{aligned}$$

* Le nœud D :

$$\begin{aligned}\lambda_D^{(2)}(b) &= \max_d \lambda^{(1)}(d)[\max_c \pi(d|b, c)\mu_D^{(1)}(c)] &= 1 \\ \lambda_D^{(2)}(\neg b) &= \max_d \lambda^{(1)}(d)[\max_c \pi(d|\neg b, c)\mu_D^{(1)}(c)] &= 1 \\ \lambda_D^{(2)}(c) &= \max_d \lambda^{(1)}(d)[\max_b \pi(d|b, c)\mu_D^{(1)}(b)] &= 1 \\ \lambda_D^{(2)}(\neg c) &= \max_d \lambda^{(1)}(d)[\max_b \pi(d|b, \neg c)\mu_D^{(1)}(b)] &= 1\end{aligned}$$

La figure 4.11 illustre les messages sortants de chaque nœud, suite à l'itération $t = 1$.

À l'itération $t = 2$

Calcul de $\mu^{(2)}$ et $\lambda^{(2)}$ pour chaque nœud

* Le nœud A :

$$\begin{aligned}\mu^{(2)}(a) &= 1 & \lambda^{(2)}(a) &= \lambda_A(a)\lambda_B^{(2)}(a)\lambda_C^{(2)}(a) &= 1 \\ \mu^{(2)}(\neg a) &= 0.3 & \lambda^{(2)}(\neg a) &= \lambda_A(\neg a)\lambda_B^{(2)}(\neg a)\lambda_C^{(2)}(\neg a) &= 1\end{aligned}$$

* Le nœud B :

$$\begin{aligned}\mu^{(2)}(b) &= \max_a \pi(b|a)\mu_B^{(2)}(a) &= 0.3 & \lambda^{(2)}(b) &= \lambda_B(b)\lambda_D^{(2)}(b) &= 1 \\ \mu^{(2)}(\neg b) &= \max_a \pi(\neg b|a)\mu_B^{(2)}(a) &= 1 & \lambda^{(2)}(\neg b) &= \lambda_B(\neg b)\lambda_D^{(2)}(\neg b) &= 1\end{aligned}$$

* Le nœud C :

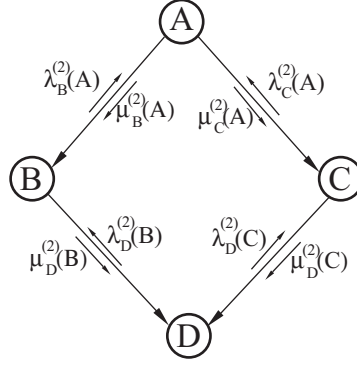


FIG. 4.11 – Messages sortants de chaque nœud à la fin de l'itération $t=1$

$$\begin{aligned}\mu^{(2)}(c) &= \max_a \pi(c|a) \mu_C^{(2)}(a) = 1 & \lambda^{(2)}(c) &= \lambda_C(c) \lambda_D^{(2)}(c) = 1 \\ \mu^{(2)}(\neg c) &= \max_a \pi(\neg c|a) \mu_C^{(2)}(a) = 0.4 & \lambda^{(2)}(\neg c) &= \lambda_C(\neg c) \lambda_D^{(2)}(\neg c) = 1\end{aligned}$$

* Le nœud D :

$$\begin{aligned}\mu^{(2)}(d) &= \max_{b,c} \pi(d|b,c) \mu_D^{(2)}(b) \mu_D^{(2)}(c) = 1 & \lambda^{(2)}(d) &= \lambda_D(d) = 1 \\ \mu^{(2)}(\neg d) &= \max_{b,c} \pi(\neg d|b,c) \mu_D^{(2)}(b) \mu_D^{(2)}(c) = 0.4 & \lambda^{(2)}(\neg d) &= \lambda_D(\neg d) = 0\end{aligned}$$

Calcul de Bel pour chaque nœud

| A | $Bel(A)$ | B | $Bel(B)$ | C | $Bel(C)$ |
|----------|----------|----------|----------|----------|----------|
| a | 1 | b | 0.3 | c | 1 |
| $\neg a$ | 0.3 | $\neg b$ | 1 | $\neg c$ | 0.4 |

Ce dernier calcul montre qu'il n'y a aucune modification des Bel entre les deux dernières itérations. Ce qui implique un état de convergence de l'algorithme. Ainsi, les valeurs approchées des degrés de possibilité a posteriori conditionnellement à l'évidence $D = d$ sont :

| A | $\Pi(A D = d)$ | B | $\Pi(B D = d)$ | C | $\Pi(C D = d)$ |
|----------|----------------|----------|----------------|----------|----------------|
| a | 1 | b | 0.3 | c | 1 |
| $\neg a$ | 0.3 | $\neg b$ | 1 | $\neg c$ | 0.4 |

4.5 L'algorithme LBP possibiliste appliqué sur des réseaux quantitatifs certains

Maintenant, nous allons aborder la propagation approximative sur des réseaux quantitatifs certains avec l'algorithme LBP possibiliste. Nous commençons, d'abord,

par définir les réseaux certains puis nous expliquerons l'effet de propagation de l'évidence sur l'évolution des valeurs des degrés de possibilités des instances des variables non observées.

4.5.1 Les réseaux certains

La théorie de possibilité est un cadre adéquat pour représenter et manipuler des informations incertaines. Elle peut aussi, naturellement, coder les informations certaines à travers des réseaux certains. Nous définissons un réseau possibiliste certain par :

- Une composante graphique identique à celle des réseaux possibilistes standards.
- Une composante numérique qui traduit les distributions de possibilités associées à chaque variable. Ces distributions de possibilités ne contiennent que deux valeurs possibles :

- 0 pour indiquer qu'une instance d'une variable est impossible.
- 1 pour indiquer qu'une instance d'une variable est totalement possible.

Selon le cadre du raisonnement, les réseaux certains peuvent être quantitatifs ou qualitatifs. La règle de chaînage associée à un réseau certain indique son type. L'exemple suivant illustre un réseau quantitatif certain.

Exemple 4.4 *Considérons un réseau quantitatif certain G_{pc} dont la composante graphique est identique à celle présentée dans la figure 4.1. Le tableau 4.4 montre les distributions de possibilités conditionnelles initiales associées au réseau certain.*

| | | B | A | $\Pi(B A)$ | | | | C | A | $\Pi(C A)$ |
|----------|----------|----------|----------|------------|--|--|--|----------|----------|------------|
| A | $\Pi(A)$ | $\neg b$ | $\neg a$ | 0 | | | | $\neg c$ | $\neg a$ | 1 |
| $\neg a$ | 0 | $\neg b$ | a | 1 | | | | $\neg c$ | a | 0 |
| a | 1 | b | $\neg a$ | 1 | | | | c | $\neg a$ | 0 |
| | | b | a | 0 | | | | c | a | 1 |

| D | C | B | $\Pi(D B, C)$ | D | C | B | $\Pi(D B, C)$ |
|----------|----------|----------|---------------|-----|----------|----------|---------------|
| $\neg d$ | $\neg c$ | $\neg b$ | 1 | d | $\neg c$ | $\neg b$ | 1 |
| $\neg d$ | $\neg c$ | b | 0 | d | $\neg c$ | b | 1 |
| $\neg d$ | c | $\neg b$ | 0 | d | c | $\neg b$ | 1 |
| $\neg d$ | c | b | 0 | d | c | b | 1 |

TAB. 4.4 – Distributions de possibilités conditionnelles initiales certaines

4.5.2 Application de LBP possibiliste sur des réseaux quantitatifs certains

Dans ce qui suit, nous allons étudier l'effet de l'application de l'algorithme LBP possibiliste sur les degrés de possibilités des instances des variables d'un réseau quantitatif certain. Plus précisément, nous allons expliquer cet effet par l'étude des deux situations suivantes : lorsqu'une information est totalement possible (son degré de possibilité = 1) peut-elle devenir impossible (son degré de possibilité = 0) suite à la propagation de l'évidence ? Inversement, est-ce-que une information impossible peut-elle devenir totalement possible ? Commençons par la première situation.

► **Une instance totalement possible peut-elle devenir impossible ?**

Soit A_i une variable du réseau quantitatif certain et soit a_i une instance de A_i . Pendant chaque itération, l'algorithme LBP possibiliste calcule pour chaque $a_i \in D_{A_i}$ la quantité $Bel(a_i)$ qui estimera $\Pi(A_i = a_i|e)$. On veut savoir est-il envisageable que, entre deux itérations consécutives, un $Bel(a_i) = 1$ peut devenir égal à 0 ? La réponse est *OUI*. En effet, selon la formule (4.31) $Bel(a_i)$ se calcule par le produit des quantités $\lambda^{(t)}(a_i)$ et $\mu^{(t)}(a_i)$. Plus précisément, la quantité $\mu^{(t)}(a_i)$ (dont l'expression est donnée par la formule 4.33) peut changer de valeur (de 1 à 0) en fonction des messages envoyés par les parents de la variable A_i . De même, la quantité $\lambda^{(t)}(a_i)$ (dont l'expression est donnée par la formule 4.32) peut changer de valeur (de 1 à 0) en fonction des messages envoyés par les fils de la variable A_i .

Considérons l'exemple suivant : Reprenons le réseau quantitatif certain de l'exemple 4.4 et supposons l'observation de l'évidence $D = d$ puis regardons l'évolution de la valeur de Bel pour la variable B par exemple. Après une première itération ($t = 0$), les valeurs de Bel pour les variables A , B et C sont les suivantes :

| A | $Bel(A)$ | B | $Bel(B)$ | C | $Bel(C)$ |
|----------|----------|----------|----------|----------|----------|
| a | 1 | b | 1 | c | 1 |
| $\neg a$ | 0 | $\neg b$ | 1 | $\neg c$ | 1 |

Les messages sortants de la variable A respectivement vers les variables B et C ont pour valeur :

$$\begin{aligned} \mu_B^{(1)}(a) &= \mu^{(0)}(a)\lambda_A(a)\lambda_C^{(0)}(a) &= 1 & \quad \mu_C^{(1)}(a) &= \mu^{(0)}(a)\lambda_A(a)\lambda_B^{(0)}(a) &= 1 \\ \mu_B^{(1)}(\neg a) &= \mu^{(0)}(\neg a)\lambda_A(\neg a)\lambda_C^{(0)}(\neg a) &= 0 & \quad \mu_C^{(1)}(\neg a) &= \mu^{(0)}(\neg a)\lambda_A(\neg a)\lambda_B^{(0)}(\neg a) &= 0 \end{aligned}$$

Pour l'itération $t = 1$, le calcul de la quantité $\mu^{(1)}$ pour la variable B donne :

$$\begin{aligned} \mu^{(1)}(b) &= \max_a \pi(b|a)\mu_B^{(1)}(a) &= \max\{\pi(b|a)\mu_B^{(1)}(a), \pi(b|\neg a)\mu_B^{(1)}(\neg a)\} \\ &= \max\{0 \times 1, 1 \times 0\} &= 0 \end{aligned}$$

De même pour l'instance $B = \neg b$:

$$\begin{aligned}\mu^{(1)}(\neg b) &= \max_a \pi(\neg b|a) \mu_B^{(1)}(a) = \max\{\pi(\neg b|a) \mu_B^{(1)}(a), \pi(\neg b|\neg a) \mu_B^{(1)}(\neg a)\} \\ &= \max\{1 \times 1, 0 \times 0\} = 1\end{aligned}$$

La quantité $\lambda^{(1)}$ pour la variable B demeure à 1 ce qui nous donne comme valeurs de Bel pour l'itération $t = 1$:

| A | $Bel(A)$ | B | $Bel(B)$ | C | $Bel(C)$ |
|----------|----------|----------|----------|----------|----------|
| a | 1 | b | 0 | c | 1 |
| $\neg a$ | 0 | $\neg b$ | 1 | $\neg c$ | 0 |

Nous remarquons que $Bel(B = b)$ est passé de la valeur 1 à 0 ; il en est de même pour $Bel(C = \neg c)$ qui est devenu égal à 0.

► **Une instance impossible peut-elle devenir totalement possible ?**

Ce cas concerne la situation suivante : Si un $Bel(A_i = a_i) = 0$, peut-il devenir égal à 1 au passage d'une itération à une autre de l'algorithme LBP possibiliste ? La réponse est a priori *NON*. En effet, différents cas se présentent :

- Si A_i est un nœud (variable) racine alors pour chacune de ses instances a_i , sa quantité $\mu^{(t)}(a_i)$, $\forall t$ prendra toujours la valeur de $\pi(a_i)$. Cela veut dire que si initialement $\pi(A_i = a_i) = 0$ alors $\mu^{(t)}(A_i = a_i) = 0 : t = 0, 1, 2, ..$ et par conséquent $Bel(A_i = a_i) = 0$ pendant chaque itération de l'algorithme LBP et jusqu'à la convergence.
- Si A_i est observé à l'état a_i alors $\lambda^{(t)}(A_i = a_i) = 1$ et $\lambda^{(t)}(A_i = a_j) = 0 : \forall a_j \in D_{A_i} \setminus \{a_i\}$. Par conséquent $Bel(A_i = a_j) : \forall a_j \in D_{A_i} \setminus \{a_i\}$ sera toujours égal à 0 jusqu'à la convergence de l'algorithme.
- Si A_i n'est pas un nœud racine et ni observé alors les quantités $\lambda^{(0)}(A_i = a_i)$ et $\mu^{(0)}(A_i = a_i)$, $\forall a_i \in D_{A_i}$ sont initialisées par 1 selon la phase d'initialisation de l'algorithme 4.2. Lorsque un $\lambda^{(t)}(A_i = a_i)$ ou un $\mu^{(t)}(A_i = a_i)$ devient égal à 0 (voir la première situation), il en restera car, selon les formules 4.32 et 4.33, le ou les messages reçus par A_i qui a causé le passage de 1 à 0 appliquera toujours son effet dans les formules pendant chaque itération et ceci jusqu'à la convergence de l'algorithme.

Prenons l'exemple suivant : considérons de nouveau le réseau quantitatif certain précédent. Après l'itération $t = 1$, le Bel de chaque variable non observée a pour valeur :

| A | $Bel(A)$ | B | $Bel(B)$ | C | $Bel(C)$ |
|----------|----------|----------|----------|----------|----------|
| a | 1 | b | 0 | c | 1 |
| $\neg a$ | 0 | $\neg b$ | 1 | $\neg c$ | 0 |

Vu qu'il y a eu changement de valeur par rapport à l'itération $t = 0$, l'algorithme va entamer une nouvelle itération $t = 2$. En particulier pour la variable B (qui est

ni racine ni observée), le message envoyé par son parent A vaut :

$$\begin{aligned}\mu_B^{(2)}(a) &= \mu^{(1)}(a)\lambda_A(a)\lambda_C^{(1)}(a) &= 1 \\ \mu_B^{(2)}(\neg a) &= \mu^{(1)}(\neg a)\lambda_A(\neg a)\lambda_C^{(1)}(\neg a) &= 0\end{aligned}$$

Ce qui donne pour la quantité $\mu^{(2)}(B)$:

$$\begin{aligned}\mu^{(2)}(b) &= \max_a \pi(b|a)\mu_B^{(2)}(a) &= \max\{\pi(b|a)\mu_B^{(2)}(a), \pi(b|\neg a)\mu_B^{(2)}(\neg a)\} \\ & &= \max\{0 \times 1, 1 \times 0\} = 0 \\ \mu^{(2)}(\neg b) &= \max_a \pi(\neg b|a)\mu_B^{(2)}(a) &= \max\{\pi(\neg b|a)\mu_B^{(2)}(a), \pi(\neg b|\neg a)\mu_B^{(2)}(\neg a)\} \\ & &= \max\{1 \times 1, 0 \times 0\} = 1\end{aligned}$$

Indépendamment de la valeur de $\lambda^{(2)}(B = b)$, $Bel(B = b)$ gardera sa valeur précédente qui est 0.

Cette étude, sur l'application de l'algorithme LBP possibiliste sur des réseaux quantitatifs certains, nous permet de constater que pour toute variable d'intérêt A_i , si $Bel(A_i = a_i) = 0$ pendant une itération t alors nous pouvons conclure que $\Pi(A_i = a_i) = 0$.

4.6 Adaptation de l'algorithme *LBP* aux réseaux qualitatifs

Dans cette section, nous présentons, respectivement, l'adaptation de l'algorithme exact de propagation de messages aux réseaux qualitatifs simplement connectés puis nous proposons notre adaptation de *LBP* aux réseaux qualitatifs à connexion multiples.

4.6.1 Propagation dans les réseaux qualitatifs simplement connectés

L'adaptation de l'algorithme de propagation de messages, lorsqu'il s'agit de raisonnement de type qualitatif, est légèrement différente de celle proposée pour les réseaux possibilistes quantitatifs [BA02]. Dans le cadre qualitatif, la distribution de possibilités conditionnelles locale $\Pi(A_i|U_{A_i})$ au niveau d'un nœud A_i est remplacée par une distribution de possibilités jointe $\Pi(A_i \wedge U_{A_i})$.

Nous rappelons que chaque nœud A_i divise un polyarbre en deux sous-polyarbres disjoints : l'un contient les nœuds accessibles à travers les n parents de A_i , l'autre contenant les nœuds accessibles à travers les m fils de A_i . En conséquence, l'ensemble E des nœuds observés peut être partitionné en deux sous-ensembles E_i^+ et E_i^- .

Tout nœud A_i d-sépare E_i^+ de E_i^- . Les différentes étapes de la propagation dans les réseaux possibilistes qualitatifs simplement connectés se présentent comme suit :
Etant donné la valeur de l'évidence e , la mesure de possibilité conditionnelle $\Pi(a_i|e)$ peut être obtenu à partir de la mesure de possibilité jointe $\Pi(a_i \wedge e)$ en utilisant :
 $\forall a_i \in D_{A_i}$,

$$Bel(a_i) = \Pi(a_i|e) = \begin{cases} \Pi(a_i \wedge e) & \text{si } \Pi(a_i \wedge e) < \Pi(e) \\ 1 & \text{sinon.} \end{cases} \quad (4.36)$$

$$\forall a_i \in D_{A_i}, \quad \Pi(a_i \wedge e) = \Pi(a_i \wedge (e_i^+, e_i^-)) = \Pi(a_i \wedge e_i^+, a_i \wedge e_i^-).$$

En profitant des relations d'indépendance dans un polyarbre, nous obtenons :

$$\Pi(a_i \wedge e) = \min(\Pi(a_i \wedge e_i^+), \Pi(a_i \wedge e_i^-)) = \min(\mu(a_i), \lambda(a_i)),$$

avec $\lambda(a_i)$ est définie par :

$$\lambda(a_i) = \Pi(a_i \wedge e_i^-) = \min_{j=1..m} \lambda_{Y_j}(a_i), \quad (4.37)$$

où $\lambda_{Y_j}(a_i)$ correspond au message envoyé par le fils Y_j à son parent A_i .

De même, $\mu(a_i)$ est définie par :

$$\mu(a_i) = \Pi(a_i \wedge e_i^+) = \max_{u_1, \dots, u_n} \min(\Pi(a_i \wedge u_1 \wedge \dots \wedge u_n), \min_{k=1..n} \mu_{A_i}(u_k)), \quad (4.38)$$

où $\mu_{A_i}(u_k)$ représente le message envoyé par le parent U_k à son fils A_i .

Après avoir reçu tout les messages de ses enfants, chaque nœud A_i peut calculer ses λ -messages pour les envoyer à chacun de ses parents $\{U_1, \dots, U_n\}$ lorsque $U_k = u_k$:

$$\begin{aligned} \lambda_{A_i}(u_k) &= \Pi(u_k \wedge e_{U_k A_i}^-) \\ &= \max_a \min[\lambda(a), \max_{u_k: k \neq i} (\min(\Pi(a_i \wedge u_1 \wedge \dots \wedge u_n), \min_{k \neq i} \mu_{A_i}(u_k)))] \end{aligned} \quad (4.39)$$

De la même manière, chaque nœud A_i calcule ses μ -messages à envoyer à chacun de ses fils $\{Y_1, \dots, Y_m\}$:

$$\mu_{Y_j}(a_i) = \Pi(a_i \wedge e_{A_i Y_j}^+) = \min(\lambda_{A_i}(a_i), \min_{l \neq j} \lambda_{Y_l}(a_i), \mu(a_i)). \quad (4.40)$$

Les principales étapes de cet algorithme de propagation sont résumées par l'algorithme 4.3. Deux étapes supplémentaires sont ajoutées : celle consistant à transformer les distributions de possibilités conditionnelles en distributions de possibilités jointes et celle permettant de calculer $Bel(A_i)$ à partir de $\Pi(A_i|e)$.

Algorithme : Propagation dans les polyarbres.

• **Entrée :** Un polyarbre formé par un ensemble de variables V muni d'une distribution de possibilités π et un ensemble de variables observées E avec la valeur $E = e$.

• **Sortie :** La distribution de possibilités conditionnelles $\pi(a_i|e)$ de chaque variable A_i non observée, avec $a_i \in D_{A_i}$.

Initialisation :

Pour chaque variable non observée $A_i : \forall a_i \in D_{A_i}$,

- $\mu(a_i) = 1$.

- $\lambda(a_i) = 1$.

Pour chaque variable A_i ayant $U_{A_i} = \emptyset : \forall a_i \in D_{A_i}, \mu(a_i) = \pi(a_i)$.

Pour chaque variable A_i observée à l'état a_i :

- $\mu(a_i) = 1$ et $\mu(a_j) = 0 \forall a_j \in D_{A_i} \setminus \{a_i\}$

- $\lambda(a_i) = 1$ et $\lambda(a_j) = 0 \forall a_j \in D_{A_i} \setminus \{a_i\}$

Remplacer les distributions de possibilités conditionnelles par des distributions de possibilités jointes locales. **Collecte de l'évidence :**

Pour i **de** 1 **à** $\text{long}(\text{Postordre})$ **faire**

$A \leftarrow \text{Postordre}[i]$

$B \leftarrow$ nœud adjacent à A dans Postordre

 calculer $\lambda(A)$ en utilisant (4.37), calculer $\mu(A)$ en utilisant (4.38).

Si B est un parent de A **alors** calculer et envoyer à B un λ -message en utilisant (4.39).

Sinon calculer et envoyer un μ -message en utilisant (4.40).

Fin Pour

Distribution de l'évidence :

Pour i **de** 1 **à** $\text{long}(\text{Préordre})$ **faire**

$A \leftarrow \text{Préordre}[i]$

$\text{Adj} \leftarrow$ ensemble des nœuds adjacents à A dans Préordre

 calculer $\lambda(A)$ en utilisant (4.37), calculer $\mu(A)$ en utilisant (4.38).

Pour j **de** 1 **à** $\text{long}(\text{Adj})$ **faire**

$B \leftarrow \text{Adj}(j)$

Si B est un parent de A **alors** calculer et envoyer à B un λ -message en utilisant (4.39).

Sinon calculer et envoyer un μ -message en utilisant (4.40).

Fin Pour

Fin Pour

Marginalisation :

Pour chaque nœud non observée A_i , calculer $\text{Bel}(a_i) = \pi(a_i|e)$ en utilisant (4.36).

Algorithme 4.3 - Propagation exacte dans les polyarbres dans le cadre qualitatif

4.6.2 LBP possibiliste pour les réseaux qualitatifs

Les formules utilisées dans l'algorithme *Loopy Belief Propagation* adapté aux réseaux possibilistes qualitatifs à connexion multiples se déduisent à partir des formules de l'algorithme de propagation exact qualitatif appliqué sur les polyarbres. Le symbole t dans ces formules désigne l'aspect temps introduit pour exprimer les itéra-

tions dans l'algorithme. Etant donné un ensemble de nœuds observés E instancié par la valeur e , la mesure de possibilité d'un nœud non observé A_i conditionnellement à e , notée par $\Pi(a_i | e)$, est approximée par :

$$\forall a_i \in D_{A_i}, \quad Bel(a_i) = \Pi(a_i | e) \approx \Pi(a_i \wedge e) = \min(\mu(a_i), \lambda(a_i)). \quad (4.41)$$

$\lambda(a_i)$ et $\mu(a_i)$ sont calculés itérativement à partir des messages reçus et envoyés par le nœud A_i à chaque itération.

Pendant une itération t , l'effet des messages entrants au nœud A_i respectivement par les m fils et n parents de A_i , sont représentés par (voir figure 4.10) :

$$\lambda^{(t)}(a_i) = \min_{j=1..m} \lambda_{Y_j}^{(t)}(a_i), \quad (4.42)$$

avec $\lambda_{Y_j}^{(t)}(a_i)$ le message envoyé par le fils Y_j à son parent A_i à l'itération t ,

et par :

$$\mu^{(t)}(a_i) = \max_{u_1, \dots, u_n} \min(\Pi(a_i \wedge u_1 \wedge \dots \wedge u_n), \min_{k=1..n} \mu_{A_i}^{(t)}(u_k)), \quad (4.43)$$

avec $\mu_{A_i}^{(t)}(u_k)$ le message envoyé par le parent U_k vers A_i à l'itération t .

À l'itération $t + 1$, les messages sortants du nœud A_i vers respectivement ces n parents et m fils sont représentés par (voir figure 4.11) : $\forall k = 1..n$,

$$\lambda_{A_i}^{(t+1)}(u_k) = \max_{a_i} \min[\lambda^{(t)}(a_i), \max_{u_k: k \neq i} (\min(\Pi(a_i \wedge u_1 \wedge \dots \wedge u_n), \min_{k \neq i} \mu_{A_i}^{(t)}(u_k)))] \quad (4.44)$$

et par : $\forall j = 1..m$,

$$\mu_{Y_j}^{(t+1)}(a_i) = \min(\lambda_{A_i}(a_i), \min_{l \neq j} \lambda_{Y_l}^{(t)}(a_i), \mu^{(t)}(a_i)). \quad (4.45)$$

A chaque itération t , tous les nœuds calculent, parallèlement, leur $\lambda^{(t)}(a_i)$ et $\mu^{(t)}(a_i)$ ce qui permet d'approximer, progressivement, $Bel(a_i)$ jusqu'à vérification de la condition de convergence.

Les étapes de l'algorithme LBP possibiliste sont décrites par l'algorithme 4.4.

Exemple 4.4 *Considérons le réseau possibiliste à connexion multiples dont l'allure est présentée dans la figure 4.1 et les distributions de possibilités conditionnelles initiales données par le tableau 4.3. On Suppose l'observation, comme évidence, du nœud D par la valeur d . Pour approximer les valeurs des degrés de possibilités des nœuds A , B et C , appliquons l'algorithme LBP possibiliste adapté au cadre qualitatif, décrit dans l'algorithme 4.4.*

► Initialisation

- Tous les messages sont initialisés par le vecteur 1.

Algorithme : Loopy Belief Propagation possibiliste.

Initialisation :

- N = l'ensemble de tous les nœuds.

- M = l'ensemble des nœuds non observés.

Remplacer les distributions de possibilités conditionnelles par des distributions de possibilités jointes locales.

Tous les messages sont initialisés par le vecteur 1.

Pour chaque variable non observée $A_i \in M$: $\forall a_i \in D_{A_i}, Bel(a_i) = 0$.

Pour chaque variable A_i ayant $U_{A_i} = \emptyset$: $\forall a_i \in D_{A_i}, \mu^{(0)}(a_i) = \pi(a_i)$.

Pour chaque variable A_i observée à l'état a_i :

- $\mu^{(0)}(a_i) = 1$ et $\mu^{(0)}(a_j) = 0 \forall a_j \in D_{A_i} \setminus \{a_i\}$.

- $\lambda^{(0)}(a_i) = 1$ et $\lambda^{(0)}(a_j) = 0 \forall a_j \in D_{A_i} \setminus \{a_i\}$.

- $t = 0$.

- convergence = *false*.

- maxitr = 100.

- tol = 0.0001.

Propagation des messages :

Tant que NOT convergence **and** t < max_itr **faire**

Pour i de 1 à long(N) **faire**

 calculer $\lambda^{(t)}(a_i)$ en utilisant (4.42) ;

 calculer $\mu^{(t)}(a_i)$ en utilisant (4.43) ;

Fin Pour

Pour i de 1 à long(M) **faire**

 OldBel(A_i) \leftarrow Bel(A_i) ;

 calculer Bel(A_i) en utilisant (4.41) ;

Fin Pour

Si $\forall m \in \{1, \dots, \text{long}(M)\}, \text{Bel}(A_m) - \text{OldBel}(A_m) < \text{tol}$ **alors**

 convergence = *true* ;

Fin Si

Si NOT convergence **alors**

Pour i de 1 à long(N) **faire**

 calculer pour chaque parent U_j de A_i ; $\lambda_{A_i}^{(t+1)}(u_j)$ en utilisant (4.44) ;

 calculer pour chaque fils Y_k de A_i ; $\mu_{Y_k}^{(t+1)}(a_i)$ en utilisant (4.45) ;

Fin Pour

 t = t + 1 ;

Fin Si

Fin Tant que

Algorithme 4.4 - Algorithme LBP adapté au cadre possibiliste qualitatif

- Remplacer la distribution $\Pi(B|A)$ par $\Pi(B \wedge A)$, $\Pi(C|A)$ par $\Pi(C \wedge A)$ et $\Pi(D|B, C)$ par $\Pi(D \wedge B \wedge C)$.

- A est un nœud racine donc :

$$\begin{aligned}\mu^{(0)}(a) &= \pi(a) = 1 \\ \mu^{(0)}(\neg a) &= \pi(\neg a) = 0.3\end{aligned}$$

- Le nœud D est observé à l'état $D = d$, donc :

$$\begin{aligned}\lambda^{(0)}(d) &= 1 \\ \lambda^{(0)}(\neg d) &= 0\end{aligned}$$

► Propagation de messages

À l'itération $t = 0$

Calcul de $\mu^{(0)}$ et $\lambda^{(0)}$ pour chaque nœud

* Le nœud A :

$$\begin{aligned}\mu^{(0)}(a) &= 1 & \lambda^{(0)}(a) &= \min(\lambda_B^{(0)}(a), \lambda_C^{(0)}(a)) = 1 \\ \mu^{(0)}(\neg a) &= 0.3 & \lambda^{(0)}(\neg a) &= \min(\lambda_B^{(0)}(\neg a), \lambda_C^{(0)}(\neg a)) = 1\end{aligned}$$

* Le nœud B :

$$\begin{aligned}\mu^{(0)}(b) &= \max_a \min(\pi(b \wedge a), \mu_B^{(0)}(a)) = 1 & \lambda^{(0)}(b) &= \lambda_D^{(0)}(b) = 1 \\ \mu^{(0)}(\neg b) &= \max_a \min(\pi(\neg b \wedge a), \mu_B^{(0)}(a)) = 1 & \lambda^{(0)}(\neg b) &= \lambda_D^{(0)}(\neg b) = 1\end{aligned}$$

* Le nœud C :

$$\begin{aligned}\mu^{(0)}(c) &= \max_a \min(\pi(c \wedge a), \mu_C^{(0)}(a)) = 1 & \lambda^{(0)}(c) &= \lambda_D^{(0)}(c) = 1 \\ \mu^{(0)}(\neg c) &= \max_a \min(\pi(\neg c \wedge a), \mu_C^{(0)}(a)) = 1 & \lambda^{(0)}(\neg c) &= \lambda_D^{(0)}(\neg c) = 1\end{aligned}$$

* Le nœud D :

$$\begin{aligned}\mu^{(0)}(d) &= \max_{b,c} \min(\pi(d \wedge b \wedge c), \min(\mu_D^{(0)}(b), \mu_D^{(0)}(c))) = 1 & \lambda^{(0)}(d) &= \lambda_D(d) = 1 \\ \mu^{(0)}(\neg d) &= \max_{b,c} \min(\pi(\neg d \wedge b \wedge c), \min(\mu_D^{(0)}(b), \mu_D^{(0)}(c))) = 1 & \lambda^{(0)}(\neg d) &= \lambda_D(\neg d) = 0\end{aligned}$$

Calcul de Bel pour chaque nœud

| A | $Bel(A)$ | B | $Bel(B)$ | C | $Bel(C)$ |
|----------|----------|----------|----------|----------|----------|
| a | 1 | b | 1 | c | 1 |
| $\neg a$ | 0.3 | $\neg b$ | 1 | $\neg c$ | 1 |

Calcul des messages normalisés sortants de chaque nœud

* Le nœud A :

$$\begin{aligned}\mu_B^{(1)}(a) &= \min(\mu^{(0)}(a), \lambda_A(a), \lambda_C^{(0)}(a)) = 1 \\ \mu_B^{(1)}(\neg a) &= \min(\mu^{(0)}(\neg a), \lambda_A(\neg a), \lambda_C^{(0)}(\neg a)) = 0.3\end{aligned}$$

$$\begin{aligned}\mu_C^{(1)}(a) &= \min(\mu^{(0)}(a), \lambda_A(a), \lambda_B^{(0)}(a)) &= 1 \\ \mu_C^{(1)}(\neg a) &= \min(\mu^{(0)}(\neg a), \lambda_A(\neg a), \lambda_B^{(0)}(\neg a)) &= 0.3\end{aligned}$$

* *Le nœud B :*

$$\begin{aligned}\mu_D^{(1)}(b) &= \min(\mu^{(0)}(b), \lambda_B(b)) &= 1 \\ \mu_D^{(1)}(\neg b) &= \min(\mu^{(0)}(\neg b), \lambda_B(\neg b)) &= 1 \\ \lambda_B^{(1)}(a) &= \max_b \min(\lambda^{(0)}(b), \pi(b \wedge a)) &= 1 \\ \lambda_B^{(1)}(\neg a) &= \max_b \min(\lambda^{(0)}(b), \pi(b \wedge \neg a)) &= 1\end{aligned}$$

* *Le nœud C :*

$$\begin{aligned}\mu_D^{(1)}(c) &= \min(\mu^{(0)}(c), \lambda_C(c)) &= 1 \\ \mu_D^{(1)}(\neg c) &= \min(\mu^{(0)}(\neg c), \lambda_C(\neg c)) &= 1 \\ \lambda_C^{(1)}(a) &= \max_c \min(\lambda^{(0)}(c), \pi(c \wedge a)) &= 1 \\ \lambda_C^{(1)}(\neg a) &= \max_c \min(\lambda^{(0)}(c), \pi(c \wedge \neg a)) &= 1\end{aligned}$$

* *Le nœud D :*

$$\begin{aligned}\lambda_D^{(1)}(b) &= \max_d \min(\lambda^{(0)}(d), [\max_c \min(\pi(d \wedge b \wedge c), \mu_D^{(0)}(c))]) &= 1 \\ \lambda_D^{(1)}(\neg b) &= \max_d \min(\lambda^{(0)}(d), [\max_c \min(\pi(d \wedge \neg b \wedge c), \mu_D^{(0)}(c))]) &= 1 \\ \lambda_D^{(1)}(c) &= \max_d \min(\lambda^{(0)}(d), [\max_b \min(\pi(d \wedge b \wedge c), \mu_D^{(0)}(b))]) &= 1 \\ \lambda_D^{(1)}(\neg c) &= \max_d \min(\lambda^{(0)}(d), [\max_b \min(\pi(d \wedge b \wedge \neg c), \mu_D^{(0)}(b))]) &= 1\end{aligned}$$

À l'itération $t = 1$

Calcul de $\mu^{(1)}$ et $\lambda^{(1)}$ pour chaque nœud

* *Le nœud A :*

$$\begin{aligned}\mu^{(1)}(a) &= 1 & \lambda^{(1)}(a) &= \min(\lambda_B^{(1)}(a), \lambda_C^{(1)}(a)) &= 1 \\ \mu^{(1)}(\neg a) &= 0.3 & \lambda^{(1)}(\neg a) &= \min(\lambda_B^{(1)}(\neg a), \lambda_C^{(1)}(\neg a)) &= 1\end{aligned}$$

* *Le nœud B :*

$$\begin{aligned}\mu^{(1)}(b) &= \max_a \min(\pi(b \wedge a), \mu_B^{(1)}(a)) &= 0.3 & \lambda^{(1)}(b) &= \lambda_D^{(1)}(b) &= 1 \\ \mu^{(1)}(\neg b) &= \max_a \min(\pi(\neg b \wedge a), \mu_B^{(1)}(a)) &= 1 & \lambda^{(1)}(\neg b) &= \lambda_D^{(1)}(\neg b) &= 1\end{aligned}$$

* *Le nœud C :*

$$\begin{aligned}\mu^{(1)}(c) &= \max_a \min(\pi(c \wedge a), \mu_C^{(1)}(a)) &= 1 & \lambda^{(1)}(c) &= \lambda_D^{(1)}(c) &= 1 \\ \mu^{(1)}(\neg c) &= \max_a \min(\pi(\neg c \wedge a), \mu_C^{(1)}(a)) &= 0.4 & \lambda^{(1)}(\neg c) &= \lambda_D^{(1)}(\neg c) &= 1\end{aligned}$$

* *Le nœud D :*

$$\begin{aligned}\mu^{(1)}(d) &= \max_{b,c} \min(\pi(d \wedge b \wedge c), \min(\mu_D^{(1)}(b), \mu_D^{(1)}(c))) = 1 & \lambda^{(1)}(d) &= \lambda_D(d) = 1 \\ \mu^{(1)}(\neg d) &= \max_{b,c} \min(\pi(\neg d \wedge b \wedge c), \min(\mu_D^{(1)}(b), \mu_D^{(1)}(c))) = 1 & \lambda^{(1)}(\neg d) &= \lambda_D(\neg d) = 0\end{aligned}$$

Calcul de Bel pour chaque nœud

| A | $Bel(A)$ | B | $Bel(B)$ | C | $Bel(C)$ |
|----------|----------|----------|----------|----------|----------|
| a | 1 | b | 0.3 | c | 1 |
| $\neg a$ | 0.3 | $\neg b$ | 1 | $\neg c$ | 0.4 |

Calcul des messages normalisés sortants de chaque nœud

* Le nœud A :

$$\begin{aligned}\mu_B^{(2)}(a) &= \min(\mu^{(1)}(a), \lambda_A(a), \lambda_C^{(1)}(a)) = 1 \\ \mu_B^{(2)}(\neg a) &= \min(\mu^{(1)}(\neg a), \lambda_A(\neg a), \lambda_C^{(1)}(\neg a)) = 0.3 \\ \mu_C^{(2)}(a) &= \min(\mu^{(1)}(a), \lambda_A(a), \lambda_B^{(1)}(a)) = 1 \\ \mu_C^{(2)}(\neg a) &= \min(\mu^{(1)}(\neg a), \lambda_A(\neg a), \lambda_B^{(1)}(\neg a)) = 0.3\end{aligned}$$

* Le nœud B :

$$\begin{aligned}\mu_D^{(2)}(b) &= \min(\mu^{(1)}(b), \lambda_B(b)) = 1 \\ \mu_D^{(2)}(\neg b) &= \min(\mu^{(1)}(\neg b), \lambda_B(\neg b)) = 1 \\ \lambda_B^{(2)}(a) &= \max_b \min(\lambda^{(1)}(b), \pi(b \wedge a)) = 1 \\ \lambda_B^{(2)}(\neg a) &= \max_b \min(\lambda^{(1)}(b), \pi(b \wedge \neg a)) = 1\end{aligned}$$

* Le nœud C :

$$\begin{aligned}\mu_D^{(2)}(c) &= \min(\mu^{(1)}(c), \lambda_C(c)) = 1 \\ \mu_D^{(2)}(\neg c) &= \min(\mu^{(1)}(\neg c), \lambda_C(\neg c)) = 1 \\ \lambda_C^{(2)}(a) &= \max_c \min(\lambda^{(1)}(c), \pi(c \wedge a)) = 1 \\ \lambda_C^{(2)}(\neg a) &= \max_c \min(\lambda^{(1)}(c), \pi(c \wedge \neg a)) = 1\end{aligned}$$

* Le nœud D :

$$\begin{aligned}\lambda_D^{(2)}(b) &= \max_d \min(\lambda^{(1)}(d), [\max_c \min(\pi(d \wedge b \wedge c), \mu_D^{(1)}(c))]) = 1 \\ \lambda_D^{(2)}(\neg b) &= \max_d \min(\lambda^{(1)}(d), [\max_c \min(\pi(d \wedge \neg b \wedge c), \mu_D^{(1)}(c))]) = 1 \\ \lambda_D^{(2)}(c) &= \max_d \min(\lambda^{(1)}(d), [\max_b \min(\pi(d \wedge b \wedge c), \mu_D^{(1)}(b))]) = 1 \\ \lambda_D^{(2)}(\neg c) &= \max_d \min(\lambda^{(1)}(d), [\max_b \min(\pi(d \wedge b \wedge \neg c), \mu_D^{(1)}(b))]) = 1\end{aligned}$$

À l'itération $t = 2$

Calcul de $\mu^{(2)}$ et $\lambda^{(2)}$ pour chaque nœud

* Le nœud A :

$$\begin{aligned}\mu^{(2)}(a) &= 1 & \lambda^{(2)}(a) &= \min(\lambda_B^{(2)}(a), \lambda_C^{(2)}(a)) = 1 \\ \mu^{(2)}(\neg a) &= 0.3 & \lambda^{(2)}(\neg a) &= \min(\lambda_B^{(2)}(\neg a), \lambda_C^{(2)}(\neg a)) = 1\end{aligned}$$

* Le nœud B :

$$\begin{aligned}\mu^{(2)}(b) &= \max_a \min(\pi(b \wedge a), \mu_B^{(2)}(a)) = 0.3 & \lambda^{(2)}(b) &= \lambda_D^{(2)}(b) = 1 \\ \mu^{(2)}(\neg b) &= \max_a \min(\pi(\neg b \wedge a), \mu_B^{(2)}(a)) = 1 & \lambda^{(2)}(\neg b) &= \lambda_D^{(2)}(\neg b) = 1\end{aligned}$$

* Le nœud C :

$$\begin{aligned}\mu^{(2)}(c) &= \max_a \min(\pi(c \wedge a), \mu_C^{(2)}(a)) = 1 & \lambda^{(2)}(c) &= \lambda_D^{(2)}(c) = 1 \\ \mu^{(2)}(\neg c) &= \max_a \min(\pi(\neg c \wedge a), \mu_C^{(2)}(a)) = 0.4 & \lambda^{(2)}(\neg c) &= \lambda_D^{(2)}(\neg c) = 1\end{aligned}$$

* Le nœud D :

$$\begin{aligned}\mu^{(2)}(d) &= \max_{b,c} \min(\pi(d \wedge b \wedge c), \min(\mu_D^{(2)}(b), \mu_D^{(2)}(c))) = 1 & \lambda^{(2)}(d) &= \lambda_D(d) = 1 \\ \mu^{(2)}(\neg d) &= \max_{b,c} \min(\pi(\neg d \wedge b \wedge c), \min(\mu_D^{(2)}(b), \mu_D^{(2)}(c))) = 1 & \lambda^{(2)}(\neg d) &= \lambda_D(\neg d) = 0\end{aligned}$$

Calcul de Bel pour chaque nœud

| A | $Bel(A)$ | B | $Bel(B)$ | C | $Bel(C)$ |
|----------|----------|----------|----------|----------|----------|
| a | 1 | b | 0.3 | c | 1 |
| $\neg a$ | 0.3 | $\neg b$ | 1 | $\neg c$ | 0.4 |

Ce dernier calcul montre qu'il n'y a aucune modification des Bel entre les deux dernières itérations. Ce qui implique un état de convergence de l'algorithme. Ainsi, les valeurs approchées des degrés de possibilité a posteriori conditionnellement à l'évidence $D = d$ sont :

| A | $\Pi(A D = d)$ | B | $\Pi(B D = d)$ | C | $\Pi(C D = d)$ |
|----------|----------------|----------|----------------|----------|----------------|
| a | 1 | b | 0.3 | c | 1 |
| $\neg a$ | 0.3 | $\neg b$ | 1 | $\neg c$ | 0.4 |

4.7 Application de **LBP** possibiliste sur des réseaux qualitatifs certains

Dans cette section, nous allons étudier, comme dans le cadre quantitatif, l'effet de l'application de l'algorithme LBP possibiliste sur les degrés de possibilités des instances des variables d'un réseau qualitatif certain. Nous allons considérer les deux situations suivantes : lorsqu'une information est totalement possible peut-elle

devenir impossible suite à la propagation de l'évidence ? Inversement, est-ce-que une information impossible peut-elle devenir totalement possible ? Commençons par la première situation.

► **Une instance totalement possible peut-elle devenir impossible ?**

Soit A_i une variable du réseau qualitatif certain. Pendant chaque itération, l'algorithme LBP possibiliste calcule pour chaque $a_i \in D_{A_i}$ la quantité $Bel(A_i = a_i)$, selon la formule 4.41, qui estimera $\Pi(a_i \wedge e)$. On veut savoir est-il envisageable que, entre deux itérations consécutives, un $Bel(a_i) = 1$ peut devenir égal à 0 ? Nous l'envisageons que *OUI*. En effet, $Bel(a_i)$ se calcule par le minimum des quantités $\lambda^{(t)}(a_i)$ et $\mu^{(t)}(a_i)$. Plus précisément, la quantité $\mu^{(t)}(a_i)$ (dont l'expression est donnée par la formule 4.43) peut changer de valeur (de 1 à 0) en fonction des valeurs des messages envoyés par les parents de la variable A_i . De même, la quantité $\lambda^{(t)}(a_i)$ (dont l'expression est donnée par la formule 4.42) peut changer de valeur (de 1 à 0) en fonction des messages envoyés par les fils de la variable A_i (on prendra le *min* de ces messages).

Prenons l'exemple suivant : Considérons le réseau certain de l'exemple 4.4 comme étant qualitatif et supposons l'observation de l'évidence $D = d$ puis regardons l'évolution de la valeur de Bel pour la variable B par exemple. Après une première itération ($t = 0$), les valeurs de Bel pour les variables A , B et C sont les suivantes :

| A | $Bel(A)$ | B | $Bel(B)$ | C | $Bel(C)$ |
|----------|----------|----------|----------|----------|----------|
| a | 1 | b | 1 | c | 1 |
| $\neg a$ | 0 | $\neg b$ | 1 | $\neg c$ | 1 |

Le message sortant de la variable A vers la variable B a pour valeur :

$$\begin{aligned}\mu_B^{(1)}(a) &= \min(\mu^{(0)}(a), \lambda_A(a), \lambda_C^{(0)}(a)) = 1 \\ \mu_B^{(1)}(\neg a) &= \min(\mu^{(0)}(\neg a), \lambda_A(\neg a), \lambda_C^{(0)}(\neg a)) = 0\end{aligned}$$

Pour l'itération $t = 1$, le calcul de la quantité $\mu^{(1)}$ pour la variable B donne :

$$\begin{aligned}\mu^{(1)}(b) &= \max_a \min(\pi(b \wedge a), \mu_B^{(1)}(a)) \\ &= \max\{\min(\pi(b \wedge a), \mu_B^{(1)}(a)), \min(\pi(b \wedge \neg a), \mu_B^{(1)}(\neg a))\} \\ &= \max\{\min(0, 1), \min(1, 0)\} = 0 \\ \mu^{(1)}(\neg b) &= \max_a \min(\pi(\neg b \wedge a), \mu_B^{(1)}(a)) \\ &= \max\{\min(\pi(\neg b \wedge a), \mu_B^{(1)}(a)), \min(\pi(\neg b \wedge \neg a), \mu_B^{(1)}(\neg a))\} \\ &= \max\{\min(1, 1), \min(0, 0)\} = 1\end{aligned}$$

La quantité $\lambda^{(1)}$ pour la variable B demeure à 1 ce qui nous donne comme valeurs

de Bel pour l'itération $t = 1$:

| A | $Bel(A)$ | B | $Bel(B)$ | C | $Bel(C)$ |
|----------|----------|----------|----------|----------|----------|
| a | 1 | b | 0 | c | 1 |
| $\neg a$ | 0 | $\neg b$ | 1 | $\neg c$ | 0 |

Nous remarquons que $Bel(B = b)$ est passé de la valeur 1 à 0 ; il en est de même pour $Bel(C = \neg c)$ qui est devenu égal à 0.

► **Une instance impossible peut-elle devenir totalement possible ?**

Ce cas concerne la situation suivante : Si un $Bel(A_i = a_i) = 0$, peut-il devenir égal à 1 au passage d'une itération à une autre de l'algorithme LBP possibiliste ? La réponse est a priori *NON*. En effet, différents cas se présentent :

- Si A_i est un nœud (variable) racine alors pour chacune de ses instances a_i , sa quantité $\mu^{(t)}(a_i)$, $\forall t$ prendra toujours la valeur de $\pi(a_i)$. Cela veut dire que si initialement $\pi(A_i = a_i) = 0$ alors $\mu^{(t)}(A_i = a_i) = 0 : t = 0, 1, 2, \dots$ et par conséquent $Bel(A_i = a_i) = 0$ pendant chaque itération de l'algorithme LBP et jusqu'à la convergence.
- Si A_i est observé à l'état a_i alors $\lambda^{(t)}(A_i = a_i) = 1$ et $\lambda^{(t)}(A_i = a_j) = 0 : \forall a_j \in D_{A_i} \setminus \{a_i\}$. Par conséquent $Bel(A_i = a_j) : \forall a_j \in D_{A_i} \setminus \{a_i\}$ sera toujours égal à 0 jusqu'à la convergence de l'algorithme.
- Si A_i n'est pas un nœud racine et ni observé alors les quantités $\lambda^{(0)}(A_i = a_i)$ et $\mu^{(0)}(A_i = a_i)$, $\forall a_i \in D_{A_i}$ sont initialisées par 1 selon la phase d'initialisation de l'algorithme 4.2. Lorsque un $\lambda^{(t)}(A_i = a_i)$ ou un $\mu^{(t)}(A_i = a_i)$ devient égal à 0 (voir la première situation), il en restera car, selon les formules 4.42 et 4.43, le ou les messages reçus par A_i qui a causé le passage de 1 à 0 appliquera toujours son effet dans les formules pendant chaque itération et ceci jusqu'à la convergence de l'algorithme.

Prenons l'exemple suivant : considérons de nouveau le réseau qualitatif certain précédent. Après l'itération $t = 1$, le Bel de chaque variable non observée a pour valeur :

| A | $Bel(A)$ | B | $Bel(B)$ | C | $Bel(C)$ |
|----------|----------|----------|----------|----------|----------|
| a | 1 | b | 0 | c | 1 |
| $\neg a$ | 0 | $\neg b$ | 1 | $\neg c$ | 0 |

Vu qu'il y a eu changement de valeur par rapport à l'itération $t = 0$, l'algorithme va entamer une nouvelle itération $t = 2$. En particulier pour la variable B (qui est

ni racine ni observée), le message envoyé par son parent A vaut :

$$\begin{aligned}\mu_B^{(2)}(a) &= \min(\mu^{(1)}(a), \lambda_A(a), \lambda_C^{(1)}(a)) = 1 \\ \mu_B^{(2)}(\neg a) &= \min(\mu^{(1)}(\neg a), \lambda_A(\neg a), \lambda_C^{(1)}(\neg a)) = 0\end{aligned}$$

Ce qui donne pour la quantité $\mu^{(2)}(B)$:

$$\begin{aligned}\mu^{(2)}(b) &= \max_a \min(\pi(b \wedge a), \mu_B^{(2)}(a)) \\ &= \max\{\min(\pi(b \wedge a), \mu_B^{(2)}(a)), \min(\pi(b \wedge \neg a), \mu_B^{(2)}(\neg a))\} \\ &= \max\{\min(0, 1), \min(1, 0)\} = 0\end{aligned}$$

$$\begin{aligned}\mu^{(2)}(\neg b) &= \max_a \min(\pi(\neg b \wedge a), \mu_B^{(2)}(a)) \\ &= \max\{\min(\pi(\neg b \wedge a), \mu_B^{(2)}(a)), \min(\pi(\neg b \wedge \neg a), \mu_B^{(2)}(\neg a))\} \\ &= \max\{\min(1, 1), \min(0, 0)\} = 1\end{aligned}$$

Indépendamment de la valeur de $\lambda^{(2)}(B = b)$, $Bel(B = b)$ gardera sa valeur précédente qui est 0 (puisque le \min entre $\mu^{(2)}(b)$ et $\lambda^{(2)}(b)$ est 0).

Cette étude, sur l'application de l'algorithme LBP possibiliste sur des réseaux quantitatifs certains, nous permet de constater que pour toute variable d'intérêt A_i , si $Bel(A_i = a_i) = 0$ pendant une itération t alors nous pouvons conclure que $\Pi(A_i = a_i) = 0$.

4.8 Conclusion

Dans ce chapitre, nous avons proposé un algorithme de propagation approximatif pour les réseaux possibilistes quantitatifs et qualitatifs. Toutes les expressions et les messages échangés, définis pour la propagation probabiliste, ont été redéfinis et prouvés dans le cadre possibiliste. Sans aucune transformation de la structure du réseau initial, le *LBP possibiliste* propage l'évidence à travers le réseau à connexions multiples par un échange local et itératif de messages entre les nœuds. Notre adaptation apparaît comme une alternative intéressante lorsque les algorithmes exacts possibilistes deviennent inapplicables sur des réseaux complexes. Tout comme la complexité des algorithmes exacts, les algorithmes approximatifs (probabilistes et possibilistes) sont considérés comme des problèmes NP-complet [DL93]. L'implémentation puis l'expérimentation de notre adaptation ont montré des résultats satisfaisants tant au niveau de convergence qu'au niveau d'exactitude. Une étude expérimentale des performances du LBP possibiliste sera présentée dans le chapitre suivant.

Chapitre 5

Implémentation et expérimentations

5.1 Introduction

Dans ce chapitre, nous présentons les résultats expérimentaux relatifs à l'adaptation de l'algorithme de propagation, "Loopy Belief Propagation" (LBP) aux réseaux possibilistes quantitatifs et qualitatifs, détaillé dans le chapitre 4. Dans la section 5.2, nous commençons par annoncer l'environnement logiciel dans lequel l'implémentation de l'algorithme LBP possibiliste a été expérimentée. La section 5.3 expose les différentes structures des modèles graphiques possibilistes sur lesquelles les expérimentations ont été réalisées. Les sections 5.4 et 5.5 présentent, respectivement, les résultats expérimentaux obtenus suite à l'exécution de l'algorithme LBP possibiliste sur des réseaux, respectivement, quantitatifs et qualitatifs exposés dans la section 5.3. Ces résultats expérimentaux engendrent l'étude de la convergence du LBP possibiliste sur des réseaux possibilistes standards et certains, ainsi que la conformité des distributions calculées par rapport à celles issues d'un algorithme exact possibiliste.

5.2 L'environnement expérimental

L'adaptation de l'algorithme "Loopy Belief Propagation" aux réseaux possibilistes quantitatifs et qualitatifs a été implémentée dans **Bayes Net Toolbox** (BNT) qui est une bibliothèque open source de fonctions Matlab pour les modèles graphiques directs [Mur01]. BNT permet la création, l'inférence et l'apprentissage des modèles graphiques probabilistes. Les principales caractéristiques de BNT sont :

- BNT supporte les réseaux bayésiens statiques et dynamiques.
- BNT supporte plusieurs types de distributions de probabilités conditionnelles : polynômiale (sous la forme de tableau), gaussienne, Noisy-or,...
- BNT intègre plusieurs algorithmes exacts et approximatifs d'inférence aussi

bien pour les réseaux bayésiens statiques que dynamiques.

- BNT supporte différentes méthodes d'apprentissage de structure et de paramètres pour les modèles probabilistes.
- BNT reste ouvert à toute extension pour intégrer des nouvelles fonctionnalités.

Profitant de cet environnement, nous avons développé et mise en place, sous BNT, un module (package) permettant :

- La création de la structure graphique d'un réseau possibiliste,
- L'attribution d'une distribution de possibilités conditionnelles initiales pour les nœuds du réseau possibiliste basé sur le produit ainsi crée,
- La propagation de l'évidence selon l'algorithme approximatif LBP possibiliste détaillé dans les sections 4.4 et 4.5,
- Le calcul pour chaque variable d'intérêt son degré de possibilité a posteriori.

5.3 Les réseaux expérimentaux

Dans cette section, nous allons décrire les composantes des réseaux utilisés durant la phase d'expérimentation de l'algorithme LBP possibiliste. Chacun de ces réseaux est décrit par deux composantes : une composante graphique qui est sa structure et une composante numérique qui est la distribution de possibilités initiales associées aux nœuds du réseau. Nos expérimentations ont été effectuées sur trois topologies différents de réseaux possibilistes décrits ci-dessous. Le point commun de ces réseaux est qu'ils sont tous à connexions multiples. Parmi ces trois réseaux possibilistes, nous avons choisi deux dont la structure est connue (réseau ALARM [BSCC89] et réseau QMR-DT [SMH⁺91]). La structure et distributions du troisième exemple de réseau sont générés aléatoirement pour montrer que notre implémentation peut s'appliquer sur n'importe quel réseau possibiliste quantitatif ou qualitatif.

5.3.1 Exemple 1 : réseau ALARM

D'origine, ALARM [BSCC89] (A Logical Alarm Reduction Mechanism) est un réseau bayésien pour la surveillance des patients en soins intensifs. Il représente une application de diagnostic médical permettant de calculer la probabilité de différents diagnostics en se basant sur des évidences observées. Il est composé de 37 nœuds répartis en 8 diagnostics, 16 symptômes et 13 variables intermédiaires. La cardinalité du domaine des nœuds varie entre deux et quatre. Nous avons choisi cet exemple de réseau car il est à connexions multiples. Pour le besoin de notre expérimentation, nous avons gardé la même structure du réseau ALARM, par contre, nous avons substitué

les distributions conditionnelles de probabilités par des distributions conditionnelles initiales de possibilités. Étant donné que les deux composantes graphique et numérique du réseau ALARM sont fixes nous allons faire varier, pour l'expérimentation, le nombre de nœuds observés afin d'étudier la convergence et l'exactitude de l'algorithme LBP adapté au possibiliste. La figure 5.1 montre la structure du réseau ALARM.

5.3.2 Exemple 2 : réseau aléatoire

Un réseau aléatoire est un réseau dont les composantes graphiques et numériques sont générées d'une manière aléatoire, en faisant varier trois paramètres : le nombre total des nœuds (30, 40 ou 50 nœuds), la cardinalité du domaine des nœuds (2 ou 3) et le nombre maximal de parents par nœud. Pendant les expérimentations, on génère une structure à plusieurs niveaux puis les distributions conditionnelles initiales de possibilités sont créées en fonction de cette structure. Le nombre de nœuds observés (évidence) varie selon le nombre total des nœuds (10, 13 et 17). La figure 5.3 montre l'allure d'un réseau aléatoire de 30 nœuds avec le nombre maximal de parents par nœud égal à 5.

5.3.3 Exemple 3 : réseau QMR-DT

Le réseau QMR-DT [SMH⁺91] (Quick Medical Reference) est un modèle graphique à deux niveaux (ou bipartite). Le niveau supérieur contient des nœuds représentant des maladies et le niveau inférieur contient des nœuds représentant des symptômes. La figure 5.2 montre l'allure d'un réseau QMR-DT dont la taille réelle est plus grande. En effet, le réseau original de QMR-DT contient 571 maladies et 4075 symptômes. Étant donné l'observation de certains symptômes, l'inférence dans ce réseau permet de déterminer les probabilités des maladies susceptibles d'être à l'origine des symptômes observés. Maladies et symptômes ont des états binaires. Il existe des suppositions sur l'indépendance conditionnelle dans la structure du réseau : par exemple, deux maladies ne sont pas mutuellement exclusives (un patient peut souffrir de plusieurs maladies simultanément).

Nous avons, particulièrement, choisi ce type de réseau car d'une part, il présente plusieurs boucles et d'autre part, l'exécution de l'algorithme LBP probabiliste a présenté des cas de non convergence sur ce type de réseau. Nos expérimentations vont porter sur des réseaux bipartites formés par, respectivement, 30, 40 et 50 nœuds. La structure et les distributions conditionnelles initiales de possibilités, de chaque réseau, seront générés d'une manière aléatoire. Le nombre de symptômes observés comme évidence varie selon le nombre total des nœuds (10, 13 et 17).

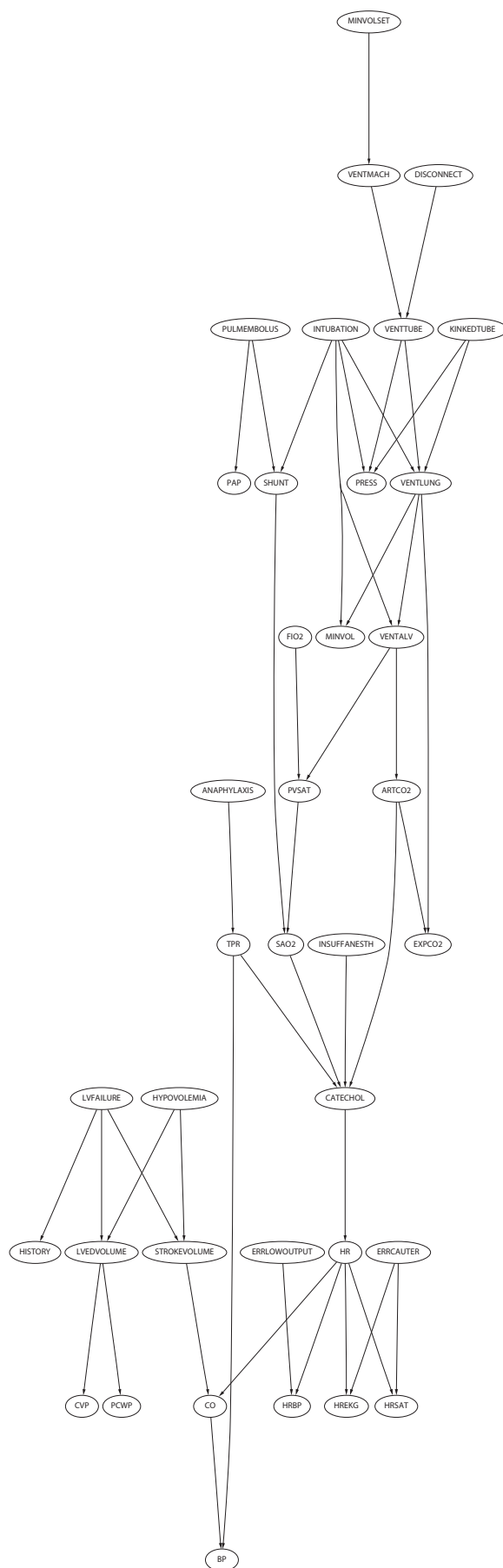


FIG. 5.1 – Structure du réseau ALARM

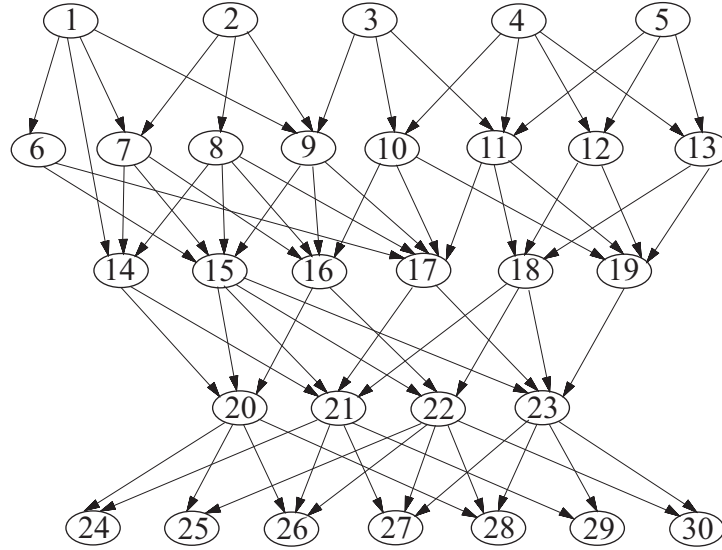


FIG. 5.2 – Structure d'un réseau aléatoire

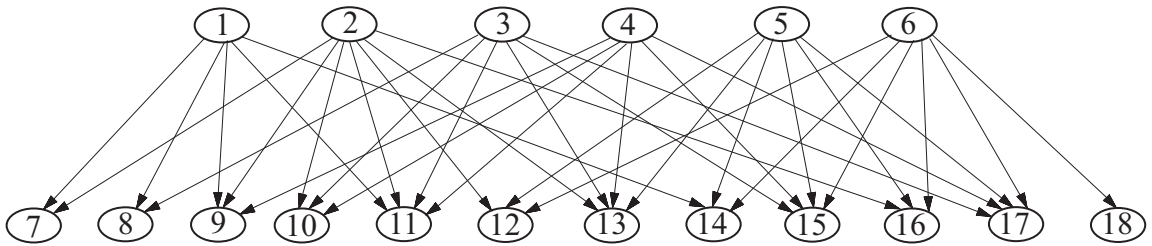


FIG. 5.3 – Structure d'un réseau QMR-DT

5.4 Expérimentations sur les réseaux quantitatifs

Dans cette section, nous allons présenter les résultats expérimentaux issus de l'application de LBP possibiliste sur les différents exemples de réseaux quantitatifs exposés dans la section précédente.

5.4.1 Evaluation des performances de l'algorithme LBP possibiliste

Pour évaluer les performances de l'algorithme LBP possibiliste ainsi que la qualité des valeurs générées par celui-ci, nous allons, respectivement, étudier la convergence de l'algorithme, puis comparer sa concordance par rapport à un algorithme exact

possibiliste et enfin, calculer un score d'évaluation, nommé MAR^1 . Ce score a été retenu et utilisé par la Conférence Internationale sur l'incertitude en Intelligence Artificielle (*Uncertainty in Artificial Intelligence UAI 2010*) pour évaluer les performances des modèles qui calculent et estiment les valeurs marginales des degrés de probabilité et les degrés de possibilité.

Il est à préciser que ces trois mesures d'évaluation seront considérées dans le cas des réseaux possibilistes quantitatifs et qualitatifs.

Étant donné que l'étude de la convergence de l'algorithme n'est, en aucun cas un problème de classification (on ne cherche pas à trouver parmi les cas réels de convergence quels sont les faux cas ou bien quels sont les cas qui sont réellement de convergence mais l'algorithme les a considéré comme des cas de non-convergence), cette étude concerne la détermination du nombre d'itérations nécessaire pour que l'algorithme atteigne un état de convergence.

L'évaluation est basée, aussi, sur le calcul d'un taux de précision (d'exactitude) qui mesure la concordance (à un ε près) entre les valeurs approchées (calculées par notre algorithme) et les valeurs fournies par un algorithme exact. Le taux de précision détermine, pour chaque exécution de l'algorithme, le pourcentage de concordance exacte (PCE_ε) qui est le rapport entre le nombre de cas où la valeur absolue de la différence entre la valeur approchée et la valeur exacte des instances de toutes les variables est inférieur à ε et le nombre total des valeurs calculées. Pour l'évaluation de notre algorithme, nous avons choisi $\varepsilon = 0.01$.

Le score d'évaluation MAR est calculé comme suit :

$$MAR = \frac{1}{N} \sum_{i=1}^N \sum_{x_i \in X_i} |\Pi_{app}(X_i = x_i) - \Pi_{exa}(X_i = x_i)|, \quad (5.1)$$

Où, $\Pi_{app}(X_i = x_i)$ est la valeur approchée du degré possibilité lorsque la variable X_i prend l'instance x_i , $\Pi_{exa}(X_i = x_i)$ est la valeur exacte du degré possibilité lorsque la variable X_i prend l'instance x_i et N le nombre des variables d'intérêt (non observés) du réseau. Numériquement, lorsque le score MAR est égale à 0, les valeurs approchées coïncident parfaitement avec les valeurs exactes. D'une manière générale, plus le score est grand plus l'écart entre certaines valeurs approchées et exactes est grand.

Commençons par présenter l'étude de la convergence.

5.4.2 Étude de la convergence

Cette étude s'intéresse au nombre nécessaire d'itérations pour que l'algorithme LBP possibiliste converge.

¹<http://www.cs.huji.ac.il/project/UAI10/eval.php>

5.4.2.1 Exemple 1 : réseau ALARM

Pendant cette première expérimentation, l'algorithme LBP possibiliste est exécuté, une seule fois, sur le réseau ALARM en variant, à chaque fois, le nombre de nœuds observés. Le tableau 5.1 présente le nombre d'itérations nécessaire pour atteindre la convergence selon les nœuds observés.

| nœuds observés | nombre d'itérations |
|----------------------|---------------------|
| aucun | 13 |
| 6 nœuds feuilles | 12 |
| 11 nœuds feuilles | 21 |
| 8 nœuds non-feuilles | 6 |

TAB. 5.1 – Nombre d'itérations nécessaire pour atteindre la convergence sur un réseau ALARM quantitatif

Nous constatons que la propagation de l'évidence dans le réseau ALARM nécessite plus d'itérations lorsque tout les nœuds feuilles sont observés (car l'évidence se trouve aux extrémités du réseau).

5.4.2.2 Exemple 2 : réseau aléatoire

On répète 3 fois l'expérimentation sur 1000 réseaux aléatoires. A chaque fois, le nombre total des nœuds était égal à 30, puis à 40, puis à 50. La structure et les distributions de possibilités conditionnelles sont choisies arbitrairement. La figure 5.4 montre le résultat de cette expérimentation.

On remarque que 46% des réseaux à 30 nœuds, 45% des réseaux à 40 nœuds et 55% des réseaux à 50 nœuds convergent en moins de 5 itérations. Ainsi, la plupart (aux environs de 90%) des réseaux générés ont besoin de moins de 20 itérations pour converger. Nous remarquons aussi qu'entre 50 et 100 itérations, le pourcentage de réseaux sur lesquels LBP possibiliste a convergé ne change pas.

5.4.2.3 Exemple 3 : réseau QMR-DT

Maintenant on répète 3 fois la même expérimentation sur 1000 réseaux bipartites générés aléatoirement. A Chaque fois, le réseau contient respectivement 30, 40 et 50 nœuds répartis sur deux niveaux. On associe, ensuite, les distributions conditionnelles initiales de possibilités aux différents nœuds de chaque réseau. La figure 5.5 présente le nombre d'itérations nécessaire pour que l'algorithme LBP possibiliste atteigne la convergence. Par exemple, 85% des réseaux à 30 nœuds, 82% des réseaux

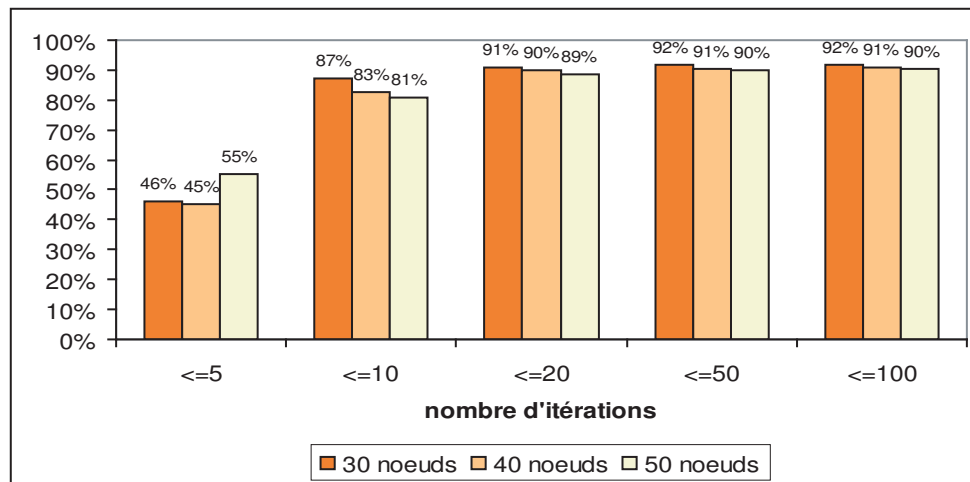


FIG. 5.4 – Nombre d'itérations effectuées pour la convergence dans le cas d'un réseau aléatoire quantitatif

à 40 nœuds et 81% seulement des réseaux à 50 nœuds nécessitent moins de 20 itérations pour converger. Notons que ce résultat est moins performant par rapport à la structure précédente. Ceci était prévisible puisque l'algorithme LBP (probabiliste) trouvait des difficultés à converger sur ce type de réseau.

Autre constatation, la convergence de LBP possibiliste est moins rapide sur les réseaux QMR-DT que sur les réseaux aléatoires.

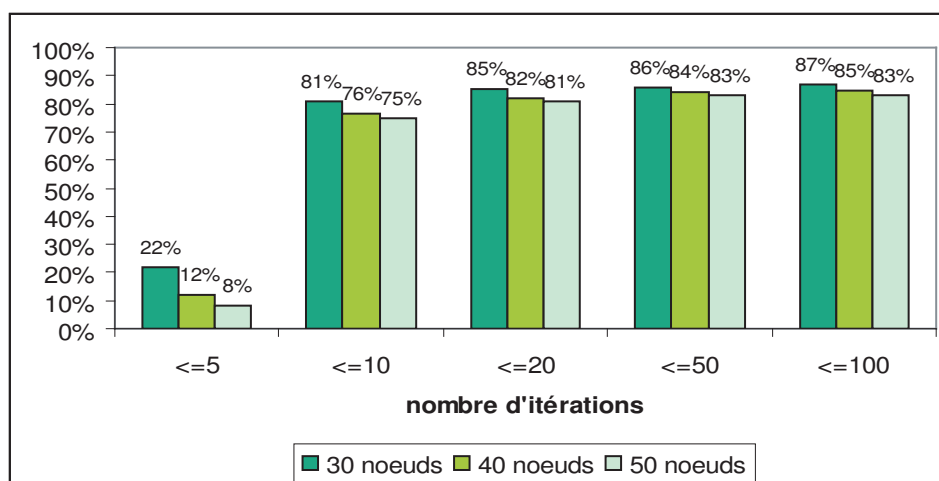


FIG. 5.5 – Nombre d'itérations effectuées pour la convergence dans le cas du réseau QMR-DT

5.4.3 Étude de l'exactitude

Maintenant, on s'intéresse à l'exactitude des valeurs générées par l'algorithme approximatif LBP possibiliste. Pour cela, on mesure la différence entre les valeurs exactes et approchées générées respectivement à partir d'une inférence exacte et de l'inférence approximative. Cette différence nous indique avec quelle précision notre algorithme peut approximer le résultat d'une inférence exacte. Cette étude expérimentale se déroule comme suit :

- Pour un réseau possibiliste, nous calculons les degrés de possibilité de chaque nœud non observé, en utilisant LBP possibiliste.
- Pour le même réseau possibiliste, nous calculons les degrés de possibilités de chaque nœud en utilisant l'algorithme exact de propagation possibiliste : arbre de jonction quantitatif, implémenté dans le PNT [BA02] (Possibilistic Net Toolbox).
- A ce moment, on peut calculer, pour chaque nœud, la différence entre le degré de possibilité exact et celui approximatif. La condition d'égalité sera vérifiée si cette différence est inférieure à un seuil fixé à l'avance (par exemple 10^{-2}).
- Pour terminer, on déduit le pourcentage PCE_{ϵ} .

5.4.3.1 Exemple 1 : réseau ALARM

En appliquant l'étude expérimentale citée plus haut, sur le réseau ALARM, nous obtenons les résultats montrés dans la figure 5.6. Pour cette exécution, nous avons considéré le cas où les 11 nœuds feuilles sont observés. L'expérimentation montre que 86% des valeurs approximatives sont en accord avec les valeurs exactes.

5.4.3.2 Exemple 2 : réseau aléatoire

L'expérimentation est effectuée 3 fois sur 1000 réseaux aléatoires. Pour chaque réseau généré de 30, 40 ou 50 nœuds, on applique notre algorithme approximatif puis l'algorithme exact arbre de jonction possibiliste basé sur le produit. La figure 5.7 montre que, par exemple, sur les 1000 réseaux de 30 nœuds, les distributions exactes et approximatives calculées coïncident dans 192 cas, contre 32 cas seulement pour les 1000 réseaux de 50 nœuds.

On remarque aussi que dans la majorité des cas (607 pour les réseaux à 30 nœuds, 648 pour les réseaux à 40 nœuds et 668 pour les réseaux à 50 nœuds), plus de 80% des valeurs approximatives concordent avec les valeurs exactes. Ceci prouve la bonne performance de LBP possibiliste même en cas de grands réseaux.

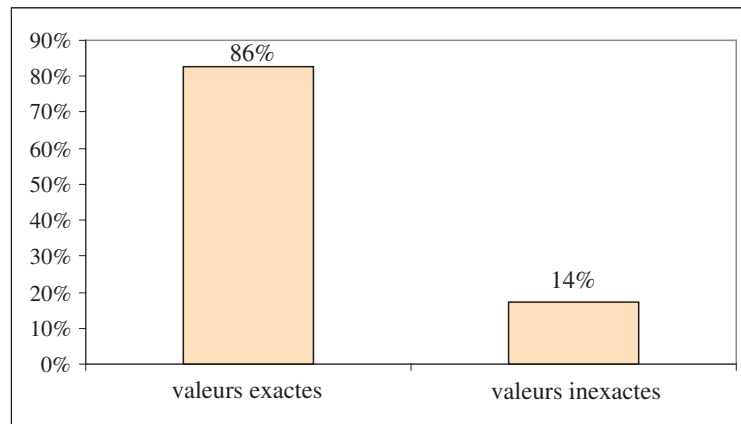


FIG. 5.6 – Mesure de l’exactitude des valeurs issues de l’arbre de jonction exacte et de LBP possibiliste approximatif pour le cas du réseau ALARM quantitatif

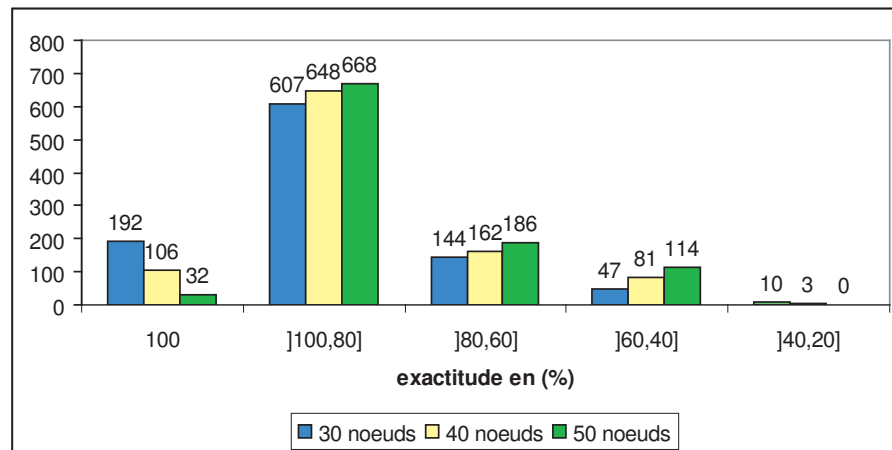


FIG. 5.7 – Mesure de l’exactitude des valeurs issues de l’arbre de jonction exacte et de LBP possibiliste approximatif pour le cas des réseaux aléatoires à 30, 40 et 50 noeuds

5.4.3.3 Exemple 3 : réseau QMR-DT

De nouveau, on répète 3 fois la même expérimentation précédente sur 1000 réseaux générés de topologie QMR-DT et composés respectivement par 30, 40 et 50 nœuds. La figure 5.8 illustre le nombre de cas d'exécution en fonction du pourcentage d'exactitude calculé. Par exemple, pour 200 cas d'exécution sur 1000 réalisés sur des réseaux à 30 nœuds, nous obtenons des valeurs approximatives qui coïncident avec les valeurs exactes, contre 180 cas sur des réseaux à 30 nœuds et 162 cas pour des réseaux à 30 nœuds.

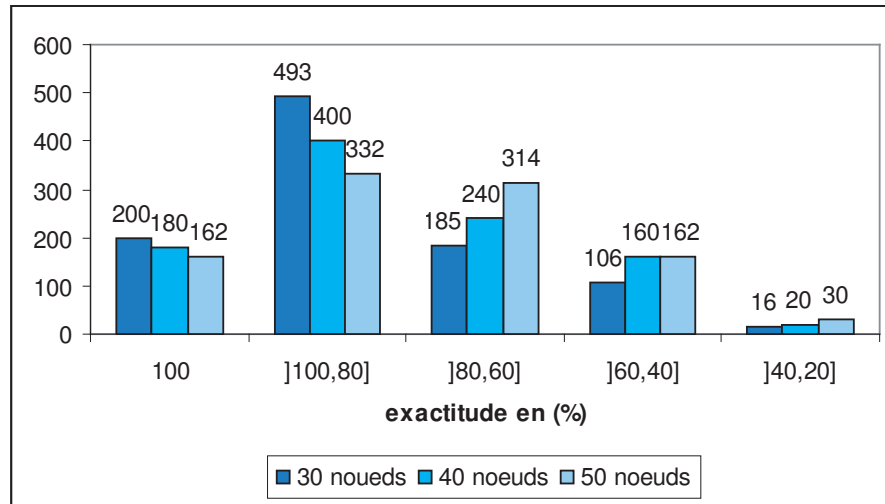


FIG. 5.8 – Mesure de l'exactitude des valeurs issues de l'arbre de jonction exact et de LBP possibiliste approximatif pour le cas de réseaux QMR-DT quantitatif formés par 30, 40 et 50 nœuds

En comparant, d'une part les résultats présentés dans les figures 5.7 et 5.8 et d'autre part les résultats présentés dans les figures 5.4 et 5.5, nous constatons que bien que l'algorithme LBP possibiliste, appliqué sur des réseaux quantitatifs, converge moins sur des réseaux QMR-DT que sur des réseaux aléatoires, il fournit des bonnes approximations des distributions a posteriori sur les réseaux biparties plutôt que sur les réseaux aléatoires.

5.4.4 Détermination du score *MAR*

Nous allons maintenant évaluer la précision des valeurs approchées générées par l'algorithme LBP possibiliste par rapport aux valeurs générées par un algorithme exact de propagation. Pour cela, nous avons effectué la même expérimentation (c-à-d 100 exécutions de LBP possibiliste et Arbre de jonction) sur des réseaux quantitatifs

aléatoires et de type QMR respectivement de taille 30, 40 et 50 nœuds. Pour chaque expérimentation, nous avons calculé le score MAR selon la formule 5.1. Le tableau 5.2 donne la répartition, en intervalles, des différentes valeurs du score MAR pour chaque chaque expérimentation. Par exemple, sur 100 exécutions de l'algorithme LBP possibiliste sur un réseau aléatoire de 30 nœuds, 22 exécutions ont donné des valeurs approchées qui coïncident parfaitement avec les valeurs exactes (le score $MAR=0$).

Ces résultats montrent que, sur des réseaux aléatoires, plus la taille est grande plus le score est grand (26 fois le score est supérieur à 0.05 pour des réseaux à 50 nœuds, contre 16 fois à 30 nœuds). Cette constatation reste valable pour les réseaux de type QMR ; de plus, les valeurs approchées sont très proches de celles exactes surtout pour des réseaux respectivement à 30 et 40 nœuds (35 fois le score est inférieur à 0.002 pour des réseaux à 40 nœuds).

| | Réseaux aléatoires | | | Réseaux de type QMR | | |
|------------------------|--------------------|-----|-----|---------------------|-----|-----|
| Taille du réseau : | 30n | 40n | 50n | 30n | 40n | 50n |
| Intervalles de valeurs | Rép. du score MAR | | | Rép. du score MAR | | |
| $=0$ | 22 | 8 | 0 | 6 | 5 | 8 |
| $]0, 0.002[$ | 6 | 10 | 8 | 25 | 30 | 10 |
| $]0.002, 0.005[$ | 8 | 15 | 7 | 11 | 5 | 4 |
| $]0.005, 0.01[$ | 11 | 8 | 13 | 8 | 9 | 6 |
| $]0.01, 0.02[$ | 18 | 20 | 18 | 6 | 10 | 16 |
| $]0.02, 0.03[$ | 9 | 10 | 14 | 10 | 6 | 5 |
| $]0.03, 0.05[$ | 10 | 16 | 14 | 11 | 15 | 14 |
| $]0.05, 0.1[$ | 11 | 8 | 12 | 14 | 13 | 13 |
| >0.1 | 5 | 5 | 14 | 9 | 7 | 24 |

TAB. 5.2 – Répartition du score MAR sur différents intervalles de valeurs pour 100 exécutions de LBP possibiliste sur des réseaux aléatoires et de type QMR

5.4.5 Expérimentations sur les réseaux quantitatifs certains

On va, maintenant, procéder à l'étude expérimentale de la convergence et de l'exactitude de l'algorithme LBP possibiliste appliqué sur des réseaux quantitatifs certains. Les tableaux 5.3 et 5.4 présentent, respectivement, les résultats obtenus suite aux expérimentations effectuées sur les réseaux aléatoires et de type QMR-DT. La première constatation, à partir de ces deux tableaux, est que sur les 500

exécutions effectuées sur chaque type de réseau certain (aléatoire ou QMR-DT) LBP possibiliste a toujours atteint un état de convergence. Par exemple, sur des réseaux aléatoires certains de 30 nœuds toutes les exécutions de LBP possibiliste ont nécessité seulement 4 itérations pour converger. Aussi, nous pouvons déduire de ces résultats que 73.33% des 500 exécutions effectuées sur des réseaux aléatoires certains de 50 nœuds ont nécessité 4 itérations pour converger et que toutes les exécutions ont convergé après 5 itérations. Nous pouvons aussi constater que LBP possibiliste converge mieux (avec des taux supérieurs) et en moins d'itérations sur des réseaux de type QMR-DT que sur des réseaux aléatoires.

| Nombre de nœuds du réseau | Pourcentage de convergence après 4 itérations | Pourcentage de convergence après 5 itérations |
|------------------------------|--|--|
| 30 | 100% | 100% |
| 40 | 82.66% | 100% |
| 50 | 73.33% | 100% |

TAB. 5.3 – Expérimentations de LBP possibiliste sur les réseaux quantitatifs certains aléatoires

| Nombre de nœuds du réseau | Pourcentage de convergence après 3 itérations | Pourcentage de convergence après 4 itérations |
|------------------------------|--|--|
| 30 | 100% | 100% |
| 40 | 94.66% | 100% |
| 50 | 88.33% | 100% |

TAB. 5.4 – Expérimentations de LBP possibiliste sur les réseaux quantitatifs certains de type QMR-DT

Concernant l'exactitude, les résultats expérimentaux ont montrés que les valeurs approchées coïcident parfaitement avec les valeurs exactes et que le score MAR est égal à 0 pour chaque exécutions effectuée, quelque soit le type et la taille du réseau certain utilisé.

5.5 Expérimentations sur les réseaux qualitatifs

Dans cette section, nous allons refaire les mêmes expérimentations précédentes de l'algorithme LBP possibiliste mais en considérant, cette fois-ci, des réseaux pos-

sibilistes qualitatifs. Commençons par l'étude de la convergence.

5.5.1 Étude de la convergence

Cette étude s'intéresse au nombre nécessaire d'itérations pour que l'algorithme LBP possibiliste converge.

5.5.1.1 Exemple 1 : réseau ALARM

Le tableau 5.5 présente le nombre d'itérations nécessaire pour atteindre la convergence selon les nœuds observés du réseau ALARM.

| nœuds observés | nombre d'itérations |
|----------------------|---------------------|
| aucun | 13 |
| 6 nœuds feuilles | 13 |
| 11 nœuds feuilles | 17 |
| 8 nœuds non-feuilles | 7 |

TAB. 5.5 – Nombre d'itérations nécessaire pour atteindre la convergence sur un réseau ALARM qualitatif

Nous constatons que la propagation de l'évidence dans le réseau ALARM nécessite plus d'itérations lorsque tout les nœuds feuilles sont observés (car l'évidence se trouve aux extrémités du réseau). En comparant les valeurs du tableau 5.1 et 5.5, nous constatons que LBP possibiliste converge mieux sur le réseau ALARM qualitatif que quantitatif.

5.5.1.2 Exemple 2 : réseau aléatoire

L'étude de la convergence de LBP possibiliste sur des réseaux aléatoires qualitatifs se fait de la même manière que dans le cadre quantitatif : On répète 3 fois l'expérimentation sur 1000 réseaux aléatoires à 30, 40 et à 50 nœuds. La figure 5.9 montre le résultat de cette expérimentation.

On remarque que 42% des réseaux à 30 nœuds, 40% des réseaux à 40 nœuds et 46% des réseaux à 50 nœuds convergent en moins de 5 itérations. D'une manière générale, la convergence est légèrement moins rapide sur des réseaux aléatoires qualitatifs que sur des réseaux quantitatifs.

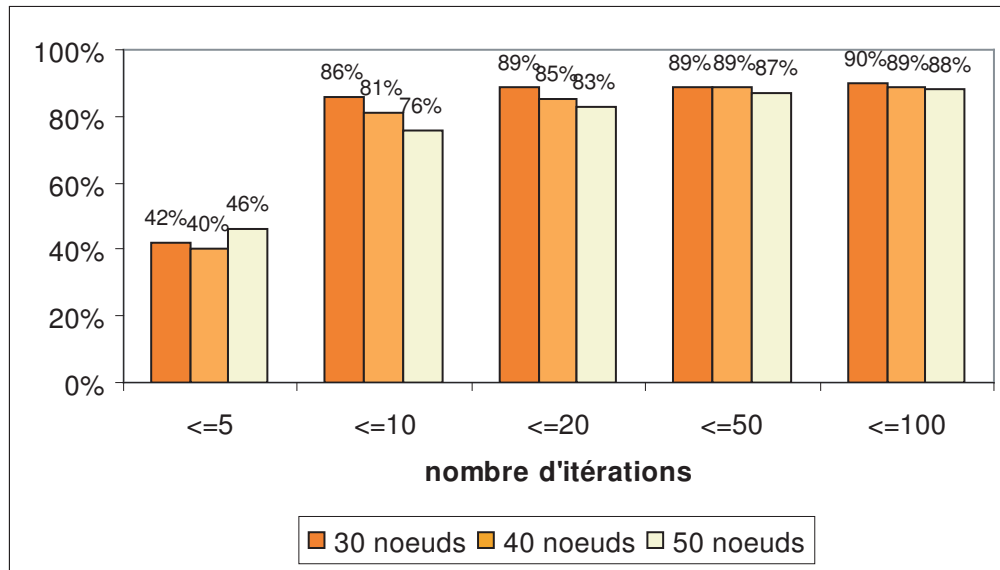


FIG. 5.9 – Nombre d'itérations effectuées pour la convergence dans le cas d'un réseau aléatoire qualitatif

5.5.1.3 Exemple 3 : réseau QMR-DT

Maintenant on répète 3 fois la même expérimentation sur 1000 réseaux bipartites générés aléatoirement contenant respectivement 30, 40 et 50 nœuds. La figure 5.10 présente le nombre d'itérations nécessaire pour que l'algorithme LBP possibiliste atteigne la convergence.

Nous remarquons, à partir de cette figure, que LBP possibiliste converge lentement sur les réseaux QMR-DT qualitatifs. Par exemple, sur des réseaux QMR-DT à 50 nœuds, seulement 6% des réseaux convergent en moins de 5 itérations. Progressivement cette valeur va augmenter pour atteindre 92% des réseaux nécessitant moins de 100 itérations pour converger : résultat relativement lent mais meilleur que sur les réseaux QMR-DT quantitatifs.

5.5.2 Étude de l'exactitude

Passons maintenant, à l'étude de l'exactitude des valeurs générées par l'algorithme LBP possibiliste. La différence entre les valeurs exactes et approchées générées respectivement à partir d'une inférence exacte et de l'inférence approximative est calculée de la même manière que pendant l'étude quantitative.

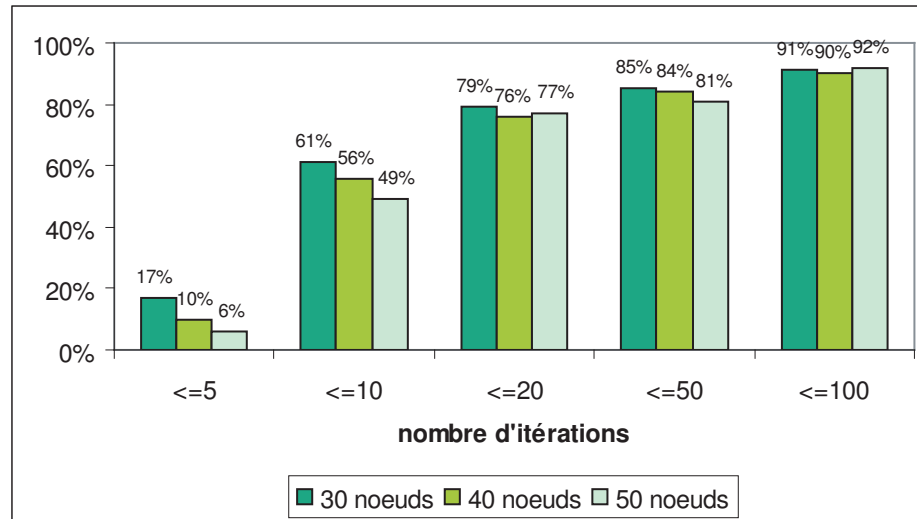


FIG. 5.10 – Nombre d'itérations effectuées pour la convergence dans le cas de réseau QMR-DT qualitatif

5.5.2.1 Exemple 1 : réseau ALARM

Le résultat de l'étude de l'exactitude des valeurs issu de l'application de LBP possibiliste sur le réseau ALARM qualitatif est présenté dans la figure 5.11. Pour cette exécution, nous avons considéré le cas où les 11 nœuds feuilles sont observés. L'expérimentation montre que 92% des valeurs approximatives sont en accord avec les valeurs exactes.

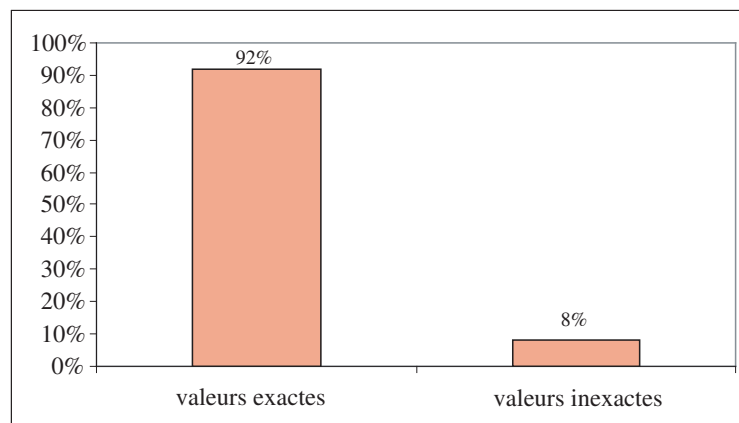


FIG. 5.11 – Mesure de l'exactitude des valeurs issues de l'arbre de jonction exacte et de LBP possibiliste approximatif pour le cas du réseau ALARM qualitatif

5.5.2.2 Exemple 2 : réseau aléatoire

L'expérimentation est effectuée 3 fois sur 1000 réseaux aléatoires. Pour chaque réseau généré de 30, 40 ou 50 nœuds, on applique notre algorithme approximatif puis l'algorithme exact arbre de jonction possibiliste basé sur le produit. La figure 5.12 montre que, par exemple, sur les 1000 réseaux générés de 30 nœuds, les distributions exactes et approximatives calculées coïncident dans 225 cas, 179 cas pour les 1000 réseaux de 40 nœuds et 111 cas pour les 1000 réseaux de 50 nœuds.

Ce résultat s'interprète par rapport à la figure 5.7 que les valeurs générées par LBP qualitatif sont plus proches des valeurs exactes issues de l'algorithme arbre de jonction qualitatif.

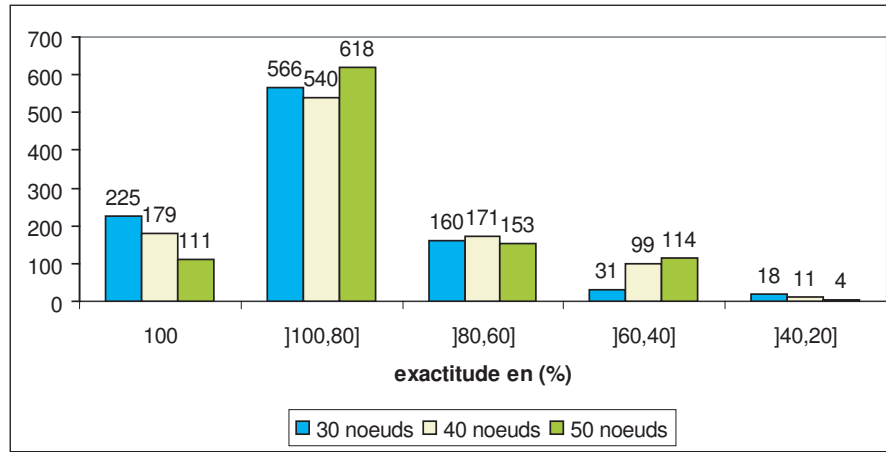


FIG. 5.12 – Mesure de l'exactitude des valeurs issues de l'arbre de jonction exacte et de LBP possibiliste approximatif pour le cas des réseaux aléatoires à 30, 40 et 50 nœuds

5.5.2.3 Exemple 3 : réseau QMR-DT

Finalement, on répète 3 fois la même expérimentation précédente sur 1000 réseaux générés de topologie QMR-DT et composés respectivement par 30, 40 et 50 nœuds. La figure 5.13 illustre les résultats obtenus. Le nombre total de nœuds a une conséquence sur l'exactitude des valeurs : 207 cas de concordance pour les réseaux à 30 nœuds, 167 cas pour les réseaux à 40 nœuds et 137 cas pour les réseaux à 50 nœuds. On remarque aussi une forte concentration des cas d'exactitude dans l'intervalle $]100\%,80\%]$ ce qui nous laisse déduire que LBP possibiliste qualitatif fournit des valeurs exactes mieux que la version quantitative (voir figure 5.8).

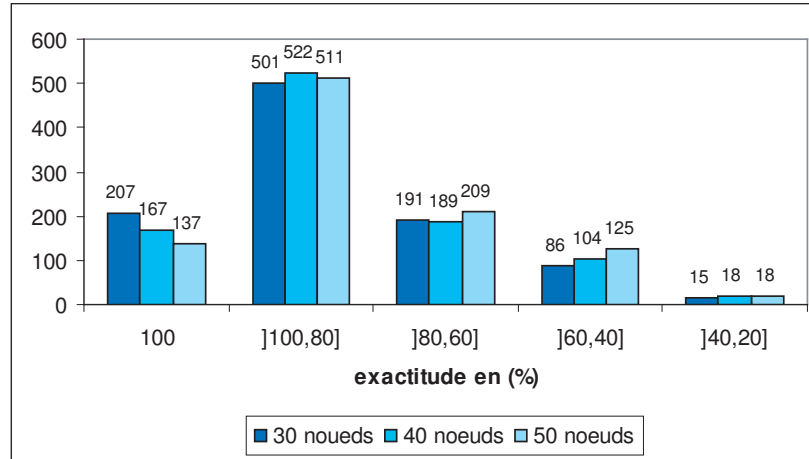


FIG. 5.13 – Mesure de l’exactitude des valeurs issues de l’arbre de jonction exact et de LBP possibiliste approximatif pour le cas de réseaux QMR-DT qualitatif formés par 30, 40 et 50 nœuds

5.5.3 Détermination du score *MAR*

Nous allons maintenant déterminer le score MAR (selon la formule 5.1) suite à l’application de LBP possibiliste sur des réseaux qualitatifs aléatoires et de type QMR respectivement de taille 30, 40 et 50 nœuds. Le tableau 5.6 donne la répartition, en intervalles, des différentes valeurs du score MAR pour chaque chaque expérimentation. Par exemple, sur 100 exécutions de l’algorithme LBP possibiliste sur un réseau aléatoire de 30 nœuds, 26 exécutions ont donnée des valeurs approchées qui coïncident parfaitement avec les valeurs exactes (le score MAR=0).

Ces résultats montrent que, sur des réseaux aléatoires, plus la taille est grande plus le score est grand (23 fois le score est supérieur à 0.05 pour des réseaux à 50 nœuds, contre 16 fois à 30 nœuds). Cette constatation reste valable pour les réseaux de type QMR ; de plus, les valeurs approchées sont très proches de celles exactes surtout pour des réseaux respectivement à 30 et 40 nœuds (36 fois le score est inférieur à 0.002 pour des réseaux à 30 nœuds).

5.5.4 Expérimentations sur les réseaux qualitatifs certains

On va, maintenant, procéder à l’évaluation des performances de l’algorithme LBP possibiliste appliqué sur des réseaux qualitatifs certains. Les tableaux 5.7 et 5.8 présentent, respectivement, les résultats obtenus suite aux expérimentations effectuées sur les réseaux aléatoires et de type QMR-DT. Tout comme les réseaux quantitatifs, sur les 500 exécutions effectuées sur chaque type de réseau qualitatif certain LBP possibiliste a toujours atteint un état de convergence. Par exemple, sur

| | Réseaux aléatoires | | | Réseaux de type QMR | | |
|------------------------|--------------------|-----|-----|---------------------|-----|-----|
| Taille du réseau : | 30n | 40n | 50n | 30n | 40n | 50n |
| Intervalles de valeurs | Rép. du score MAR | | | Rép. du score MAR | | |
| $=0$ | 26 | 14 | 2 | 9 | 10 | 9 |
| $]0, 0.002[$ | 5 | 13 | 11 | 27 | 24 | 18 |
| $]0.002, 0.005[$ | 10 | 18 | 8 | 14 | 8 | 2 |
| $]0.005, 0.01[$ | 7 | 11 | 16 | 5 | 7 | 5 |
| $]0.01, 0.02[$ | 9 | 15 | 14 | 5 | 6 | 15 |
| $]0.02, 0.03[$ | 13 | 9 | 13 | 9 | 8 | 6 |
| $]0.03, 0.05[$ | 9 | 10 | 13 | 12 | 16 | 14 |
| $]0.05, 0.1[$ | 8 | 7 | 12 | 11 | 14 | 12 |
| >0.1 | 5 | 3 | 11 | 8 | 7 | 19 |

TAB. 5.6 – Répartition du score MAR sur différents intervalles de valeurs pour 100 exécutions de LBP possibiliste sur des réseaux aléatoires et de type QMR

des réseaux aléatoires certains de 30 nœuds toutes les exécutions de LBP possibiliste ont nécessité seulement 4 itérations pour converger. d’une manière globale, nous pouvons constater que LBP possibiliste converge mieux et en moins d’itérations sur des réseaux qualitatifs de type QMR-DT que sur des réseaux aléatoires.

| Nombre de nœuds du réseau | Pourcentage de convergence après 4 itérations | Pourcentage de convergence après 5 itérations |
|------------------------------|--|--|
| 30 | 100% | 100% |
| 40 | 93.75% | 100% |
| 50 | 85.66% | 100% |

TAB. 5.7 – Expérimentations de LBP possibiliste sur les réseaux qualitatifs certains aléatoires

Concernant l’exactitude, les résultats expérimentaux ont montrés que les valeurs approchées coïcident parfaitement avec les valeurs exactes et que le score MAR est égal à 0 pour chaque exécutions effectuée, quelque soit le type et la taille du réseau certain utilisé.

| Nombre de nœuds du réseau | Pourcentage de convergence après 3 itérations | Pourcentage de convergence après 4 itérations |
|------------------------------|--|--|
| 30 | 100% | 100% |
| 40 | 99% | 100% |
| 50 | 97.16% | 100% |

TAB. 5.8 – Expérimentations de LBP possibiliste sur les réseaux qualitatifs certains de type QMR-DT

5.6 Conclusion

Au cours de ce chapitre nous avons procédé à l'expérimentation de l'implémentation de l'algorithme *Loopy Belief Propagation possibiliste* pour les réseaux quantitatifs et qualitatifs sous l'environnement Matlab. Ces expérimentations ont été effectuées sur trois topologies différentes de réseaux possibilistes pour varier l'évaluation des performances de l'algorithme étudié.

La convergence de l'algorithme n'est pas gravement affecté par la variation du nombre total des nœuds. Dans l'expérimentation quantitative sur des réseaux respectivement, aléatoires et biparties, le nombre de cas de convergence diminue respectivement de 2% et 4% en passant de 30 à 50 nœuds. Il en est de même pour l'expérimentation qualitative : le nombre de cas de convergence diminue respectivement de 2% et seulement 1% en passant de 30 à 50 nœuds.

Autre constatation, la convergence est meilleure dans le cas qualitatif que quantitatif sur les réseaux QMR-DT. Par exemple, 83% de cas de convergence sur des réseaux biparties quantitatifs à 50 nœuds contre 92% de cas de convergence sur des réseaux qualitatifs de même taille. Presque les mêmes taux de convergence sont observés sur les réseaux QMR-DT quantitatifs et qualitatifs.

L'exactitude des valeurs générés par LBP possibiliste et l'Arbre de jonction (quantitatif et qualitatif) confirme la performance de notre adaptation surtout dans le cadre qualitatif et sur les réseaux QMR-DT. Par exemple, les figures 5.7 et 5.12 montrent que parmi 1000 exécutions sur des réseaux aléatoires quantitatifs, respectivement, à 30, 40 et 50 nœuds, nous obtenons, respectivement, 192, 106 et 32 cas d'exactitude. Dans le cadre qualitatif, ces cas deviennent respectivement 222, 179 et 111.

Ces comparaisons nous laisse conclure que LBP possibiliste qualitatif s'adapte mieux aux réseaux QMR-DT, bien que sa convergence est moins rapide sur ceux-ci.

Maintenant nous allons proposer, dans le chapitre suivant, une représentation particulière de l'incertitude à travers un réseau possibiliste basé sur une mesure "récente" de possibilité garantie, notée par Δ . En effet et d'une manière générale,

les réseaux possibilistes sont des représentations compactes soit des connaissances incertaines soit des préférences utilisateurs à travers leurs distributions de possibilités associées. Plusieurs techniques ont été mises en œuvre pour modéliser la notion de préférence. Le chapitre suivant présentera ces principales techniques puis proposera un nouveau format pour la représentation des préférences : les réseaux possibilistes garantis.

Chapitre 6

Un nouveau format de représentation des préférences : les réseaux possibilistes garantis

6.1 Introduction

En théorie des possibilités, il existe deux grandeurs standards permettant de mesurer l'incertitude associée à une information exprimée sous la forme d'une formule logique. Ces deux mesures duales sont respectivement le degré de possibilité Π et le degré de nécessité N . Le premier degré mesure le niveau de cohérence d'une formule φ avec les informations disponibles, le deuxième correspond au degré de certitude associé à une formule φ dans le contexte des informations disponibles.

Il existe aussi une autre mesure proposée récemment, permettant l'évaluation de l'incertitude au niveau d'une information, basée sur la notion de mesure de possibilité garantie notée par Δ [DHP00, BK03]. Cette mesure représente le degré de possibilité minimal associé à chaque instance de cette information. Dans la logique possibiliste cela signifie que, pour une formule logique pondérée (φ, a) , $\Delta(\varphi) \geq a$.

La distribution de possibilités est une notion importante qui peut être dérivée de la mesure de possibilité. Une distribution de possibilités, notée π , est une représentation de l'état du monde Ω en donnant un ordre de préférences entre les différentes interprétations ω de ce monde.

Intuitivement, une distribution de possibilités peut représenter soit les connaissances que l'agent possède sur les états du monde réel, soit ses préférences. La distribution de possibilités distingue les mondes les plus plausibles de ceux qui le sont moins ; en donnant à chaque interprétation ω un degré de comptabilité avec les connaissances sur l'état du monde (respectivement elle distingue également les conséquences les

plus préférées de celles qui le sont moins lorsque la possibilité d'une conséquence peut être vue comme un degré auquel elle satisfait les préférences de l'agent).

Au cours de ce chapitre, nous allons aborder deux aspects : le premier concerne la définition des préférences et leurs techniques de modélisation basées, pour la plupart sur la théorie des possibilités. Le deuxième concerne l'utilisation de la mesure de possibilité garantie Δ pour définir un réseau possibiliste *garanti*. Le reste de ce chapitre sera donc organisé comme suit : nous introduisons dans la section 6.2 la notion des préférences. Les sections 6.3 et 6.4 présentent deux techniques de modélisation des préférences : Les CP-Nets et la logique *QCL*. La section 6.5 présente les variantes de la logique *QCL*. Les sections 6.6 et 6.7 définissent respectivement les bases possibilistes garanties et leurs utilisations pour la représentation logique des préférences. Dans la section 6.8 nous définirons les réseaux possibilistes, le conditionnement et la règle de chaînage basés tous sur la mesure Δ . La section 6.9 explique comment un réseau possibiliste garanti peut modéliser les préférences. La section 6.10 montre qu'un réseau possibiliste garanti peut se transformer (en un temps polynomial) en une base de connaissances possibilistes garantie. Enfin, la section 6.11 démontre qu'il est possible d'établir une équivalence syntactique linéaire entre un réseau possibiliste garanti et un réseau possibiliste qualitatif. Cette équivalence nous évitera de re-définir un algorithme de propagation pour les réseaux garantis et de ré-utiliser les algorithmes de propagation pour les réseaux qualitatifs comme ceux présentés dans le chapitre 3 (Arbre de jonction qualitatif) ou celui proposé dans le chapitre 4 (Loopy Belief Propagation possibiliste qualitatif).

6.2 Notion de préférences

La notion de préférences est importante du fait que souvent, on se retrouve confronté à des situations où la prise de décisions est nécessaire (Par exemple, quelle voiture acheter, où investir notre argent, quelle étude faire, où partir en vacances,...). Prendre une décision consiste à agir en fonction de plusieurs éléments à savoir nos objectifs, nos choix, nos préférences, ainsi que nos croyances et connaissances sur le monde. La représentation et la modélisation des préférences sont considérées comme un problème important sur lesquels de nombreux chercheurs de différents domaines se sont intéressés.

C'est la raison, pour laquelle, différentes disciplines telles que l'économie, la psychologie, l'intelligence artificielle et même la philosophie ont donné naissance à de nombreux formalismes (logiques, graphiques ou autres modes de représentation) permettant d'exprimer les préférences.

La préférence est mise en œuvre lorsqu'un agent se trouve confronté à un ensemble

d'alternatives à partir desquels il doit faire un choix. Pour que le choix effectué réponde au mieux à la préférence souhaitée, on doit procéder à une évaluation de ces alternatives en tenant compte des informations dont l'agent dispose (par exemple, l'issue de chaque alternative). Cette évaluation cherche à comparer, en termes de préférence, les différentes alternatives pour la prise de décision.

Par exemple, supposons qu'un agent doit acheter un billet d'avion le plus rapidement possible pour une mission de travail à l'étranger. Une agence de voyage ou un site sur le web peut lui offrir plusieurs possibilités (alternatives). Chaque alternative est caractérisée par différents "critères" à savoir le prix, l'horaire, le confort (première ou seconde classe), avec ou sans escale, etc. Si l'agent préfère un voyage moins cher, sans escale, plus confortable alors différentes méthodes peuvent être appliquées pour satisfaire ses préférences. Une première procédure toute simple consiste à éliminer certaines alternatives (alternatives non-buts ou non satisfaisantes) et considérer d'autres alternatives meilleures (alternatives satisfaisantes), il s'agit donc d'identifier un sous ensemble aussi restreint que possible, composé des alternatives jugées comme les meilleures. Il s'agit de *problématique de choix* dans un problème de décision. Une deuxième possibilité concerne le cas où chaque alternative peut être considérée indépendamment des autres en la positionnant sur une échelle ou en la comparant à des normes ou des niveaux de référence (chaque alternative sera classée dans une classe prédéfinie). Il s'agit de la *problématique de tri*. Une autre possibilité concerne la situation de comparer les alternatives entre elles, afin d'établir un classement constituant un ordre partiel. Le classement de l'ensemble des alternatives de la plus mauvaise à la meilleure, représente le niveau de satisfaction de chaque alternative, il s'agit ici de la *problématique de rangement*.

C'est ainsi donc que les alternatives sont évaluées de différentes manières selon la nature des jugements apportés (absolus ou relatifs). Concernant la problématique de choix et la problématique de rangement, les alternatives sont évaluées à base des jugements comparatifs ou relatifs. Il s'agit d'une *évaluation relationnelle, logique* ou *qualitative*. Lorsqu'il s'agit de la problématique de tri, l'évaluation des alternatives se base sur des jugements absolus. Il s'agit donc d'une *évaluation quantitative*.

6.3 Les CP-Nets

Les CP-Nets (*Conditional Preference Networks*) ont été introduits en 1999 par [BBHP99], comme outil de représentation compacte des relations de préférences qualitatives. Ce modèle graphique exploite l'indépendance préférentielle conditionnelle dans la structuration des préférences utilisateur sous l'hypothèse *ceteris-paribus*. Nous définissons, dans cette section, les notions de base liées au modèle CP-Net.

Soit $V = X_1, X_2, \dots, X_n$, un ensemble de variables (caractéristiques ou attributs) sur lesquelles les préférences utilisateur sont définies, étant donné un problème décisionnel fixé. On note par $O = Asst(V) = Dom(X_1) \times Dom(X_2) \times \dots \times Dom(X_n)$ l'ensemble de toutes les alternatives possibles, avec $Dom(X_i)$ est le domaine de valeurs de la variable X_i .

Relation de préférence : une relation de préférence, notée \succ , définie sur l'ensemble des alternatives O , est un préordre complet sur O , telle que :

$$\forall o, o' \in O, o \succ o' \text{ ou } o' \succ o,$$

où $o \succ o'$ signifie que l'alternative o est au moins aussi préférée que l'alternative o' .

Dépendance préférentielle : un ensemble de caractéristiques X est préférentiellement indépendant de son complément $Y = V - X$ si et seulement si :

$$\forall x_1, x_2 \in Asst(X), \forall y_1, y_2 \in Asst(Y), x_1 y_1 \succ x_2 y_1 \Leftrightarrow x_1 y_2 \succ x_2 y_2.$$

Si X est préférentiellement indépendant de son complément $Y = V - X$, noté par $PI(X, Y)$, ceci équivaut à dire que l'ordre de préférence sur les éléments x_1 et x_2 de X reste inchangé quelques soient les valeurs des éléments y_i de Y . On dit que x_1 est préférable *ceteris paribus* à x_2 .

Dépendance préférentielle conditionnelle : on dit que X et Y sont conditionnellement préférentiellement indépendants étant donné Z si et seulement si :

$$\forall x_1, x_2 \in Asst(X), \forall y_1, y_2 \in Asst(Y), \forall z \in Asst(Z) \quad x_1 y_1 z \succ x_2 y_1 z \Leftrightarrow x_1 y_2 z \succ x_2 y_2 z.$$

Un CP-Net est un graphe orienté acyclique $G = (V, E)$, où V est un ensemble de nœuds qui définissent les variables de préférence et E un ensemble d'arcs orientés entre les nœuds, traduisant des relations de dépendances préférentielles entre ces nœuds. Toute variable X_i du graphe est instanciable dans un domaine de valeurs $Dom(X_i) = \{x_{i1}, x_{i2}, \dots, x_{i3}\}$. Le prédécesseur d'un nœud X dans le graphe est dit son parent. On note $Pa(X)$ l'ensemble des parents de X . $X \cup Pa(X)$ constitue une famille du CP-Net. A chaque variable X du CP-Net, on associe une table de préférences conditionnelles $CPT(X)$ spécifiant un ordre de préférence total sur les valeurs x_i de X étant donné chaque instance de ses parents. Pour un nœud racine, la table CPT spécifie un ordre de préférence inconditionnel sur les valeurs du nœud. Un CP-Net induit un graphe partiel de préférences ordonné, construit sur l'ensemble de ses alternatives. Une alternative du CP-Net est un élément du produit cartésien des domaines de valeurs de ses différents nœuds. Elle est interprétée comme une conjonction de ses éléments.

Exemple 6.1 *Etant donné trois variables binaires V , P et C , pour le choix de la*

veste, celui du pantalon et celui de la chemise. De manière inconditionnelle, l'utilisateur préfère la couleur noire (V_n) à la couleur blanche (V_b) de la veste, et de même pour celle du pantalon ($P_n \succ P_b$). Si le pantalon et la veste sont de la même couleur, il préfère le gris (C_g) au blanc (C_b) pour la couleur de la chemise. En revanche, si le pantalon et la veste sont de couleurs différentes, il choisira la couleur blanche pour la chemise plutôt que la couleur grise. La figure 6.1 représente dans un CP-Net toutes ces informations qui décrivent les préférences de cette personne. On a $Dom(V) = \{V_n, V_b\}$, $Dom(P) = \{P_n, P_b\}$ et $Dom(C) = \{C_g, C_b\}$.

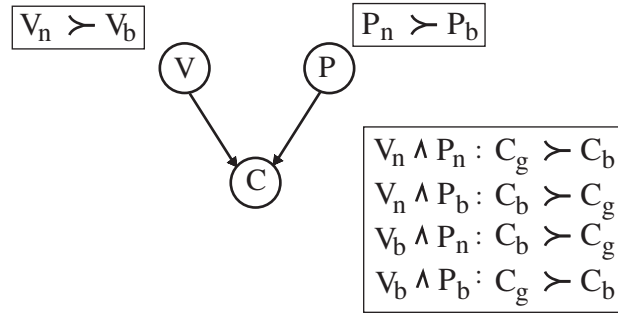


FIG. 6.1 – Exemple d'un CP-Net.

6.4 La logique QCL

La *logique du choix qualitatif* (QCL) [BBD04], est une extension de la logique propositionnelle. La partie non-standard de QCL est un nouveau connecteur $\overrightarrow{\times}$, nommé *disjonction ordonnée*. Intuitivement, si a et b sont deux alternatives représentées comme deux formules propositionnelles, alors nos préférences sur ces deux alternatives sont codées par " $a \overrightarrow{\times} b$ " qui signifie : " a si c'est possible, mais si a est impossible alors au moins b ".

Le langage QCL contient principalement trois types de formules : des formules propositionnelles permettant de représenter les connaissances d'un agent, des formules de choix de base pour représenter les préférences de structures simples et finalement, des formules de choix générales qui représentent les préférences de structures plus générales et plus complexes.

Les formules de choix de base (BCF)

Soit PS un ensemble de symboles propositionnels (ou atomes) et $PROP_{PS}$ un ensemble de formules propositionnelles obtenues à partir de PS et des connecteurs habituels de la logique propositionnelle (\neg, \wedge, \vee).

Une formule de choix de base correspond à la disjonction ordonnée de deux ou

plusieurs formules propositionnelles. Soit a_1, a_2, \dots, a_n , n formules propositionnelles de $PROP_{PS}$, la formule de choix de base $a_1 \overrightarrow{\times} a_2 \overrightarrow{\times} \dots \overrightarrow{\times} a_n$ est utilisée pour exprimer et ordonner une liste d'alternatives : certaines a_i sont vraies, a_1 est préférée en premier, mais si ce n'est pas possible alors a_2 , si cela n'est pas possible alors a_3 , et ainsi de suite.

Les formules de choix générales (GCF)

Une formule de choix générale représente toute formule qui peut être obtenue de PS en utilisant les connecteurs $\overrightarrow{\times}$, \wedge , \vee , \neg . Le langage composé de formules de choix générales est noté par GCF_{PS} .

Exemple 6.2 *Supposons que deux agents expriment leurs préférences au sujet de leurs loisirs. Le premier agent préfère fortement 'aller à la chasse' plutôt que 'faire une promenade à cheval' ou encore moins 'aller à la pêche'. Concernant le second agent, en plus qu'il préfère 'aller à la chasse' que 'se promener à cheval', il préfère aussi 'être accompagné' que 'd'être seul'.*

Les préférences du premier agent sont représentées par la formule :

"aller à la chasse $\overrightarrow{\times}$ faire une promenade à cheval $\overrightarrow{\times}$ aller à la pêche"

qui est une formule de choix de base, et celles du deuxième agent sont représentées par la formule :

"(aller à la chasse $\overrightarrow{\times}$ faire une promenade à cheval) \wedge (être accompagné $\overrightarrow{\times}$ être seul)".

6.5 Révisions et modifications de la logique QCL

6.5.1 Les variantes de la logique QCL

La logique QCL présente plusieurs limites dans le sens où sa relation d'inférence et le traitement de la négation ne sont pas satisfaisantes [BS08a]. Pour cerner les limites de la logique QCL , l'auteur de [Sed08] propose une nouvelle définition de la négation des préférences sans modifier la définition de la conjonction et la disjonction. Cette nouvelle logique s'appelle *Logique du Choix Qualitatif Minimale*, notée $MQCL$. Pour pouvoir exprimer d'autres types de préférences, [Sed08] propose, aussi, deux nouvelles logiques $PQCL$ [BS07a] (*Logique du Choix Qualitatif Prioritaire*) et $QCL+$ [BS08a] (*Logique du Choix Qualitatif Positive*). Les logiques $PQCL$ et $QCL+$ se caractérisent par une nouvelle définition de la négation qui est la même que $MQCL$ et des nouvelles définitions de la conjonction et de la disjonction. Dans le cadre de la logique $PQCL$, la conjonction (ou la disjonction) des préférences est

prioritaire, dans le cadre de $QCL+$, la conjonction (ou la disjonction) des préférences est positive.

6.5.2 La Logique du Choix Qualitatif Minimale $MQCL$

Les trois logiques $MQCL$, $PQCL$ et $QCL+$ se basent sur le même langage QCL (même vocabulaire et connecteurs). Elles partagent les mêmes propriétés comme la définition de la négation des préférences. Par contre, chacune des ces trois logiques se caractérise par une sémantique bien particulière, en ce qui concerne le type des préférences qu'elle permet de représenter.

La logique QCL ne fait pas de distinction entre la disjonction standard et la disjonction ordonnée lorsque la négation est présente dans les préférences. C'est à dire que par exemple :

$$\neg(a_1 \overrightarrow{\times} a_2 \overrightarrow{\times} \dots \overrightarrow{\times} a_n) \equiv \neg a_1 \vee \neg a_2 \vee \dots \vee \neg a_n.$$

Dans le cadre de la logique QCL , la négation d'une formule de choix de base ou de choix générale est équivalente à une formule propositionnelle, où la disjonction ordonnée est remplacée par la disjonction classique. Ce résultat implique qu'il n'est plus question d'ordonner les alternatives entre elles.

Dans le cadre de la logique $MQCL$ (et même de $PQCL$ et $QCL+$), une nouvelle définition de la négation a été donnée. Dans cette définition, si nous avons des préférences représentées par " $\neg(a_1 \overrightarrow{\times} a_2 \overrightarrow{\times} \dots \overrightarrow{\times} a_n)$ ", nous exigeons que $\neg a_1$ doit être considérée comme une première option, $\neg a_2$ doit être considérée comme une deuxième option, etc. Ainsi, la définition de la négation, donnée par [Sed08] est de la forme :

$$\neg(a_1 \overrightarrow{\times} a_2 \overrightarrow{\times} \dots \overrightarrow{\times} a_n) \equiv \neg a_1 \overrightarrow{\times} \neg a_2 \overrightarrow{\times} \dots \overrightarrow{\times} \neg a_n.$$

Par opposition à la logique QCL et dans le cadre des logiques $MQCL$, $PQCL$ et $QCL+$, la négation d'une formule garde la disjonction ordonnée et son degré de satisfaction est défini comme n'importe quelle formule de choix de base ou de choix générale.

Par exemple, dans le cadre de la logique QCL la négation d'une préférence de la forme "*il n'est pas vrai que je préfère aller à la pêche à ne pas y aller*" est codée par " $\neg(pêche \overrightarrow{\times} \neg pêche)$ ". Cette formule est équivalente à la contradiction suivante " $\neg pêche \wedge pêche$ ". Lorsqu'on applique la définition de la négation de la logique $MQCL$, on n'obtient jamais une contradiction, et la négation de cette préférence sera exprimée par "*je préfère ne pas aller à la pêche plutôt d'y aller*" et codée par " $\neg pêche \overrightarrow{\times} pêche$ ".

La relation d'inférence dans la logique $MQCL$ se montre plus satisfaisante que l'inférence QCL . En effet, la relation d'inférence de la logique $MQCL$ se caractérise

par une fonction de normalisation, notée par $\mathcal{N}_{\mathcal{MQL}}$. Cette fonction de normalisation permet la transformation, d'une manière équivalente, tout ensemble de préférences représentées par des formules de choix générales en un ensemble de préférences en formes normales représentées par des formules de choix de base.

6.5.3 La Logique du Choix Qualitatif Prioritaire \mathcal{PQL}

La logique \mathcal{PQL} [BS07a] est adaptée à la représentations des préférences prioritaires, c'est-à-dire des préférences ayant différents niveaux de priorité. Le choix pour ce type de logique est motivé par le fait que les préférences des agents sont certainement différentes au niveau de l'importance qu'elles procurent. Pour un problème donné, les agents expriment souvent plusieurs préférences sur un grand nombre d'alternatives afin de pouvoir satisfaire au minimum une des préférences souhaitées même si toutes les préférences exprimées ne sont pas totalement souhaitables.

Soit l'exemple suivant dans lequel un responsable informatique doit acheter un antivirus pour protéger le réseau informatique de la société qui l'a recruté. Ses préférences sont exprimées comme suit :

- Il préfère un antivirus qui fonctionne en réseau plutôt que par poste.
- Il préfère acquérir un antivirus libre qu'avec une licence.
- Il préfère un antivirus avec Pare-feu.
- Il préfère un antivirus avec une mise à jour fréquente plutôt que occasionnelle.
- Il préfère un antivirus qui occupe peu d'espace en mémoire centrale.

Il est clair que même si toutes ces préférences exprimées ne peuvent pas être satisfaites, ce responsable préfère quand même acheter l'antivirus vu les risques causées par la non protection contre les virus informatiques. Ainsi, la logique \mathcal{PQL} permet de donner des priorités aux préférences exprimées. Pour ces préférences, il est par exemple plus intéressant de considérer la préférence "*Il préfère un antivirus qui fonctionne en réseau plutôt que par poste*" prioritaire à la préférence "*Il préfère acquérir un antivirus libre qu'avec licence*". C'est à dire que le fait de protéger le réseau informatique par un antivirus (en réseau ou par poste) peut faire abstraction sur le type de l'antivirus acquis (libre ou avec licence).

L'inférence \mathcal{PQL} concernant les formules de choix de base est totalement équivalente à la relation d'inférence \mathcal{QL} . Cependant, lorsque les préférences sont représentées par des formules de choix générales, la relation d'inférence \mathcal{PQL} se montre plus satisfaisante que l'inférence \mathcal{QL} . En effet, elle suit le même principe que l'inférence \mathcal{QL} et se caractérise par de nouvelles définitions de la négation (celle de \mathcal{MQL}), conjonction et disjonction qui sont utiles pour l'agrégation des préférences des utilisateurs ayant différents niveaux de priorité.

Les principales caractéristiques de la logique \mathcal{PQL} sont les suivantes :

- Le langage de \mathcal{PQCL} est le même que celui de \mathcal{QCL} .
- La sémantique de toute formule est basée sur son degré de satisfaction dans un modèle particulier.
- La négation doit être aussi similaire que possible que celle de la logique propositionnelle. En particulier, la double négation d'une formule doit permettre de retrouver la formule originale.
- Les préférences prioritaires : la logique \mathcal{PQCL} doit être adéquate au traitement des préférences prioritaires codées par une conjonction (resp. disjonction) prioritaire. Si un agent exprime deux préférences p et q telle que p est plus importante que q , alors la logique \mathcal{PQCL} permet de coder ce type de préférences par : $p \wedge q$. Par exemple, pour exprimer que la préférence "*Il préfère un antivirus qui fonctionne en réseau plutôt que par poste*" est prioritaire par rapport à la préférence "*Il préfère acquérir un antivirus libre qu'avec licence*", on le codifie par la formule de choix générale : " $(\text{réseau} \xrightarrow{\times} \text{monoposte}) \wedge (\text{libre} \xrightarrow{\times} \text{licence})$ ".
- Relation d'inférence : la relation d'inférence de la logique \mathcal{PQCL} peut être construite de deux manières différentes mais équivalentes. La première méthode se base sur des règles d'inférence qui définissent le degré de satisfaction de chaque formule étant donnée une interprétation. La deuxième méthode est basée sur la fonction de normalisation qui permet de transformer tout ensemble de préférences représentées par des formules de choix générales en un ensemble de préférences simplifiées par des formules de choix de base.

6.5.4 La Logique du Choix Qualitatif Positive $\mathcal{QCL}+$

La Logique du Choix Qualitatif Positive $\mathcal{QCL}+$ [BS08a] est appropriée pour le traitement des préférences positives. La logique $\mathcal{QCL}+$ partage de nombreuses caractéristiques que la logique \mathcal{PQCL} . Elle est basée sur le même langage \mathcal{QCL} dans le sens où les préférences sont représentées par des formules de choix de base ou de choix générales et les connaissances sont représentées par des formules propositionnelles. La sémantique de toute formule dans le cadre de $\mathcal{QCL}+$ est également basée sur son degré de satisfaction dans un modèle particulier I . La négation dans le cadre de la logique $\mathcal{QCL}+$ est similaire à celle de \mathcal{PQCL} , par conséquent, en $\mathcal{QCL}+$ la double négation d'une formule est équivalente à la formule originale. Toutefois, la définition de la conjonction (resp. disjonction) n'est pas prioritaire. En fait, dans le cadre de la logique $\mathcal{QCL}+$, la conjonction et la disjonction définies sont plus adéquates pour la modélisation des préférences positives. Les préférences sont dites positives lorsqu'elles n'excluent pas de solutions dans la mesure où uniquement ce qui souhaitable est exprimé.

$\mathcal{QCL}+$ est caractérisée par une fonction de normalisation, notée $\mathcal{N}_{\mathcal{QCL}+}$, qui partage certaines caractéristiques de la fonction de normalisation de la logique \mathcal{PQCL} , telles que la décomposabilité par rapport à la négation, conjonction, disjonction et la disjonction ordonnée.

La conjonction et la disjonction dans le cadre de la logique $\mathcal{QCL}+$ sont symétriques. En effet, soient ϕ_1 et ϕ_2 sont deux formules dans \mathcal{GCFPS} , nous avons :

$$\phi_1 \wedge \phi_2 \equiv \phi_2 \wedge \phi_1 \text{ et } \phi_1 \vee \phi_2 \equiv \phi_2 \vee \phi_1.$$

La définition de la conjonction (resp. disjonction) dans $\mathcal{QCL}+$ diffère légèrement de celle utilisée dans le cadre de la logique \mathcal{PQCL} . Si on parle du niveau d'importance, dans le cadre de la logique $\mathcal{QCL}+$, les deux préférences ϕ_1 , ϕ_2 ont la même importance dans la mesure où tout ce qui est exprimé est souhaitable, c'est-à-dire que pour les préférences de la forme $\phi_1 \wedge \phi_2$, ϕ_1 est autant préférée que ϕ_2 . Par ailleurs, l'importance s'impose toujours au niveau de ϕ_1 (c-à-d. si $\phi_1 = (a_1 \vec{\times} a_2 \vec{\times} \dots \vec{\times} a_n)$ alors l'option a_1 est préférée à l'option a_2 , etc.) et également au niveau de ϕ_2 (c-à-d. si $\phi_2 = (b_1 \vec{\times} b_2 \vec{\times} \dots \vec{\times} b_n)$ alors l'option b_1 est préférée à l'option b_2 , etc.). Concernant la logique $\mathcal{QCL}+$, toutes les solutions (même celles qui falsifient les préférences) sont plus ou moins possibles, par contre pour \mathcal{PQCL} , la solution possible concerne une solution qui satisfait toutes les préférences d'un agent avec un certain degré.

6.6 Mesure de possibilité garantie et représentation logique

En théorie des possibilités, il existe une autre mesure appelée *mesure de possibilité garantie* [BK03,DHP00], notée par Δ .

6.6.1 Mesure de possibilité garantie Δ

Étant donné une distribution de possibilités π , la mesure de possibilité garantie d'une formule φ , notée $\Delta(\varphi)$, estime le niveau de cohérence de tous les modèles de φ avec les informations représentées par la distribution de possibilités π . Explicitement nous avons :

$$\Delta(\varphi) = \min\{\pi(\omega) : \omega \models \varphi \text{ et } \omega \in \Omega\}. \quad (6.1)$$

$\Delta(\varphi)$ estime le degré minimal pour lequel la formule φ est cohérente avec l'information disponible encodée par la distribution π . En clair, nous avons la relation suivante : $\forall \varphi, \Delta(\varphi) \leq \Pi(\varphi)$, où $\Pi(\varphi)$ exprime le degré de possibilité de φ . Cependant, aucune relation ne relie le degré de possibilité minimal $\Delta(\varphi)$ avec le degré de

nécessité $N(\varphi)$ (à l'exception de la relation $\Delta(\varphi) \leq 1 - N(\neg\varphi)$).

La mesure de possibilité garantie a été définie pour la représentation des préférences strictes [BK01,BDP01] de la forme $\Delta(\varphi) > \Pi(\psi)$, ce qui signifie que toute interprétation satisfaisant φ est préférée à n'importe quelle interprétation satisfaisant ψ . Les travaux de [BDKP02,BDKP06] ont également fait usage de la mesure de possibilité garantie dans la représentation bipolaire des préférences (positives et négatives). Cette dernière représentation sera brièvement exposée dans la prochaine section. Enfin, dans la logique possibiliste, cette mesure de possibilité garantie a aussi servi à la définition de bases de connaissances garanties.

6.6.2 Bases de connaissances possibilistes garanties

Une base de connaissances possibilistes basée sur la mesure de possibilité Δ (ou base de connaissance Δ), notée par \mathfrak{R} , est l'ensemble de formules pondérées suivant : $\mathfrak{R} = \{[\varphi_i, \alpha_i] : i = 1..n\}$, avec α_i correspondant à la limite inférieure de la mesure de possibilité garantie associé à φ_i , c'est à dire que $\Delta(\varphi_i) \geq \alpha_i$. Selon [BK03], il existe une unique distribution de possibilités associée à la base de connaissance Δ définie comme suit :

$$\forall \omega \in \Omega, \quad \pi_{\mathfrak{R}}(\omega) = \begin{cases} 0 & \text{si } \forall [\varphi_i, \alpha_i] \in \mathfrak{R}, \omega \not\models \varphi_i \\ \max\{\alpha_i : \forall [\varphi_i, \alpha_i] \in \mathfrak{R}, \omega \models \varphi_i\} & \text{sinon.} \end{cases} \quad (6.2)$$

Le degré de possibilité d'une interprétation ω correspond au plus grand des α_i dont la formule associée φ_i est satisfaite par ω (ω est modèle de φ_i). Si une interprétation falsifie toutes les formules alors son degré de possibilité sera égal à 0 (rien ne renforce l'idée que cette interprétation est quelque peu garantie). Dans le cas où les formules représentent des préférences, $\pi_{\mathfrak{R}}(\omega) = 0$ signifie que ω est une solution inacceptable vu qu'elle ne répond à aucune préférence.

Il est intéressant à signaler que la distribution de possibilités $\pi_{\mathfrak{R}}$ associée à la base \mathfrak{R} correspond au dual de la distribution de possibilités π_B associée à la base de connaissances possibiliste B basée sur la mesure de nécessité N . En effet, lorsqu'on considère une interprétation ω et en travaillant sur une base possibiliste B , la distribution de possibilités π_B tente de trouver les formules falsifiées par ω , ayant le plus grand poids ; tandis que, en travaillant sur une base \mathfrak{R} , la distribution $\pi_{\mathfrak{R}}$ cherche les formules satisfaites par ω ayant le plus grand poids. Les auteurs de l'article [BK03] ont montré qu'il existe une équivalence syntaxique entre une base de connaissances basée sur Δ et une autre basée sur N (degré de nécessité). Plus généralement, les auteurs ont montré que chaque notion utilisée dans la logique possibiliste standard possède sa contrepartie dans les bases de connaissances possibilistes basées sur Δ . L'exemple suivant illustre comment générer une distribution de possibilités $\pi_{\mathfrak{R}}$ à

partir d'une base de connaissances \mathfrak{R} .

Exemple 6.3 Soit $\mathfrak{R} = \{[bg \wedge sl \wedge cl, 1], [bg \wedge \neg sl \wedge cl, 0.7], [bg \wedge sl \wedge \neg cl, 0.4], [bb \wedge cl, 0.95]\}$ une base de connaissances basée sur la mesure garantie Δ . L'univers Ω comporte quatre variables binaires dont la signification est la suivante : 'bg' désigne faire une baignade en mer, 'sl' désigne le temps est ensoleillé, 'cl' désigne la mer est calme et 'bb' désigne faire une balade en bateau.

Chaque formule de la base \mathfrak{R} s'interprète de la manière suivante : la première, qui signifie faire une baignade en mer dans un temps ensoleillé et une mer calme, a une mesure de possibilité garantie = 1 ; la deuxième, qui signifie faire une baignade en temps couvert et dans une mer calme, a une mesure de possibilité garantie au minimum égal à 0.7 ; la troisième, qui signifie faire une baignade en temps ensoleillé mais dans une mer agitée, a pour mesure de possibilité garantie au minimum égal à 0.4 ; la dernière formule, qui signifie faire une balade en bateau dans une mer calme, a pour mesure de possibilité garantie au minimum égal à 0.95. La distribution de possibilités $\pi_{\mathfrak{R}}$ associé à la base \mathfrak{R} se calcule, selon la formule (6.2), comme l'indique le tableau suivant :

| ω | $\pi_{\mathfrak{R}}(\omega)$ |
|--|------------------------------|
| $\neg bg \wedge \neg sl \wedge \neg cl \wedge \neg bb$ | 0 |
| $\neg bg \wedge \neg sl \wedge \neg cl \wedge bb$ | 0 |
| $\neg bg \wedge \neg sl \wedge cl \wedge \neg bb$ | 0 |
| $\neg bg \wedge \neg sl \wedge cl \wedge bb$ | 0.95 |
| $\neg bg \wedge sl \wedge \neg cl \wedge \neg bb$ | 0 |
| $\neg bg \wedge sl \wedge \neg cl \wedge bb$ | 0 |
| $\neg bg \wedge sl \wedge cl \wedge \neg bb$ | 0 |
| $\neg bg \wedge sl \wedge cl \wedge bb$ | 0.95 |
| $bg \wedge \neg sl \wedge \neg cl \wedge \neg bb$ | 0 |
| $bg \wedge \neg sl \wedge \neg cl \wedge bb$ | 0 |
| $bg \wedge \neg sl \wedge cl \wedge \neg bb$ | 0.75 |
| $bg \wedge \neg sl \wedge cl \wedge bb$ | 0.95 |
| $bg \wedge sl \wedge \neg cl \wedge \neg bb$ | 0.4 |
| $bg \wedge sl \wedge \neg cl \wedge bb$ | 0.4 |
| $bg \wedge sl \wedge cl \wedge \neg bb$ | 1 |
| $bg \wedge sl \wedge cl \wedge bb$ | 1 |

TAB. 6.1 – Distribution de possibilités $\pi_{\mathfrak{R}}$ associée à la base \mathfrak{R}

L'interprétation $\omega = bg \wedge sl \wedge cl \wedge \neg bb$ satisfait uniquement la première formule ;

son degré de possibilité est donc égal à 1. Tandis que l'interprétation $\omega = bg \wedge sl \wedge cl \wedge bb$ satisfait la première et la dernière formule ; son degré de possibilité prend le max entre 1 et 0.95 ce qui vaut 1.

Les interprétations exprimant un temps couvert (non ensoleillé) et une mer agitée (non calme) falsifient toutes les formules de la base \mathfrak{R} , leur degré de possibilité est égal à 0. De même, les interprétations indiquant aucune baignade et aucune balade en bateau falsifient à leur tour toute formule de \mathfrak{R} . Enfin l'interprétation $\omega = \neg bg \wedge sl \wedge \neg cl \wedge bb$ traduisant pas de baignade et une balade en bateau en mer agitée falsifie toutes les formules, son $\pi_{\mathfrak{R}}(\omega) = 0$.

6.7 Représentation logique des préférences

Dans la littérature, Il existe deux familles pour la modélisation des préférences [Sed08] :

- La première famille est dite représentation quantitative ou numérique des préférences où à chaque alternative est affectée une ou plusieurs valeurs numériques. Cette représentation se base aussi sur la notion de fonction d'utilité.
- La deuxième famille est appelée représentation qualitative ou logique des préférences dans laquelle une relation binaire R est utilisée pour comparer les alternatives.

Dans cette section on se restreindra uniquement à la deuxième famille de modélisation ; plus précisément, on s'intéressera à la représentation des préférences basée sur la théorie des possibilités appelée la logique pondérée.

La théorie des possibilités offre différents formats pour la représentation des préférences. Plusieurs travaux de recherche ont été menés pour la représentation des préférences dans le cadre possibiliste ; parmi eux nous citons les travaux [DP95, DGPZ98, BDKP02, DP92b]. Dans la représentation des préférences bipolaires [BDKP02], les préférences d'un agent sont représentées à l'aide de deux bases de connaissances possibilistes : une pour représenter les rejets de l'agent et l'autre pour exprimer ces préférences positives.

Représentation des préférences bipolaires

Nous résumons brièvement dans ce qui suit, l'approche proposée dans [BDKP02]. Lorsqu'un agent exprime ce qui est plus ou moins acceptable, cela traduit ses rejets et quand il exprime ce qu'il veut atteindre, il s'agit de ses buts positifs. Ainsi, les rejets et les buts positifs d'un agent peuvent être représentés à l'aide de deux bases possibilistes distinctes.

Représentation logique des rejets

L'ensemble des rejets d'un agent, noté R , est composé de formules logiques auxquelles est associé un réel $a_i \in [0, 1]$ mesurant le niveau de rejet de l'agent pour une formule. Ainsi, $R = \{\mathcal{R}(\varphi_i) = a_i : i = 1, \dots, n \text{ et } a_i \neq 0\}$ où $\mathcal{R}(\varphi_i) = a_i$ signifie que l'agent rejette la formule φ_i avec le poids a_i .

On associe à l'ensemble R des rejets d'un agent, une distribution de possibilités π_R définie sur un ensemble des interprétations ω . $\pi_R(\omega)$ mesure le degré d'acceptation de ω par un agent vu ces préférences. Par exemple, $\pi_R(\omega) = 1$ signifie que ω est totalement accepté par l'agent, tandis que $\pi_R(\omega) = 0$ signifie que ω est rejeté. L'interprétation ω est considérée plus rejetée que l'interprétation ω' si $\pi_R(\omega) < \pi_R(\omega')$. La distribution de possibilités π_R est définie comme suit :

$$\forall \omega \in \Omega, \quad \pi_R(\omega) = 1 - \max\{a_i : \omega \models \varphi_i, \mathcal{R}(\varphi_i) = a_i \in R\}. \quad (6.3)$$

Exemple 6.4 Soient $\varphi_1 = bg \wedge \neg sl \wedge \neg cl$, $\varphi_2 = bg \wedge sl \wedge \neg cl$ et $\varphi_3 = bg \wedge \neg sl \wedge cl$ trois formules représentant les rejets exprimés par un agent tel que la base des rejets $R = \{\mathcal{R}(\varphi_1) = 1, \mathcal{R}(\varphi_2) = 0.75, \mathcal{R}(\varphi_3) = 0.5\}$. Cette base a pour signification : l'agent rejette totalement une baignade à la mer (bg) sous un temps couvert, non ensoleillé, (sl) et dans une mer agitée, non calme, (cl). Il rejette moins fortement une baignade ensoleillée dans une mer agitée. Enfin il rejette partiellement une baignade non ensoleillée dans une mer calme. La distribution de possibilités associée à R est donnée dans le tableau 6.2.

| ω | $\pi_R(\omega)$ |
|---|-----------------|
| $\neg bg \wedge \neg sl \wedge \neg cl$ | 1 |
| $\neg bg \wedge \neg sl \wedge cl$ | 1 |
| $\neg bg \wedge sl \wedge \neg cl$ | 1 |
| $\neg bg \wedge sl \wedge cl$ | 1 |
| $bg \wedge \neg sl \wedge \neg cl$ | 0 |
| $bg \wedge \neg sl \wedge cl$ | 0.5 |
| $bg \wedge sl \wedge \neg cl$ | 0.25 |
| $bg \wedge sl \wedge cl$ | 1 |

TAB. 6.2 – Distribution de possibilités π_R associée à l'ensemble des rejets R

Représentation logique des buts positifs

Les buts positifs B , exprimés par un agent, peuvent être également représentés à l'aide d'une base logique à laquelle est associée une distribution de possibilités

notée π_B . Par convention, $\pi_B(\omega) = 1$ signifie que l'agent est totalement satisfait, $\pi_B(\omega) = 0$ signifie que l'agent est indifférent par rapport à ω . L'interprétation ω est dite plus satisfaisante que l'interprétation ω' si $\pi_B(\omega) > \pi_B(\omega')$. Pour fournir la distribution de possibilités π_B , les préférences de l'agent sont représentées par des formules pondérées en termes de mesure de possibilité garantie Δ . Ainsi, les buts de l'agent sont représentés par une base $B = \{[\psi_j, b_j] : j = 1, \dots, m\}$ où ψ_j représente un but ou une préférence de l'agent et b_j le degré de possibilité garantie minimale de ψ_j . Intuitivement, toute interprétation ω qui satisfait ψ_j est considérée comme satisfaisante avec un degré au moins égal à b_j . La distribution de possibilités π_B associée à l'ensemble des buts positifs B est définie de la manière suivante :

$$\forall \omega \in \Omega, \quad \pi_B(\omega) = \begin{cases} 0 & \text{si } \forall [\psi_j, b_j] \in B, \omega \not\models \psi_j \\ \max\{b_j : \forall [\psi_j, b_j] \in B, \omega \models \psi_j\} & \text{sinon.} \end{cases} \quad (6.4)$$

Exemple 6.5 Soit l'ensemble des buts énoncés par un agent exprimé par la base $B = \{[bg \wedge sl \wedge cl, 1], [bg \wedge \neg sl \wedge cl, 0.6]\}$. La première formule signifie que l'agent sera totalement satisfait par sa baignade si le temps est ensoleillé et la mer est calme. En revanche, il sera moins satisfait si le temps est couvert mais la mer est quant même calme. La distribution de possibilités associée à B est donnée dans le tableau 6.3.

| ω | | | | | $\pi_B(\omega)$ |
|-----------|----------|-----------|----------|-----------|-----------------|
| $\neg bg$ | \wedge | $\neg sl$ | \wedge | $\neg cl$ | 0 |
| $\neg bg$ | \wedge | $\neg sl$ | \wedge | cl | 0 |
| $\neg bg$ | \wedge | sl | \wedge | $\neg cl$ | 0 |
| $\neg bg$ | \wedge | sl | \wedge | cl | 0 |
| bg | \wedge | $\neg sl$ | \wedge | $\neg cl$ | 0 |
| bg | \wedge | $\neg sl$ | \wedge | cl | 0.6 |
| bg | \wedge | sl | \wedge | $\neg cl$ | 0 |
| bg | \wedge | sl | \wedge | cl | 1 |

TAB. 6.3 – Distribution de possibilités π_B associée à l'ensemble des buts B

La représentation bipolaire des préférences est faite d'une manière indépendante, c'est à dire que les représentations des rejets et des buts positifs ne sont pas duales. L'explication donnée à la même valeur dans les deux distributions diffère. Par exemple, lorsque $\pi_R(\omega) = 1$ cela signifie que ω falsifie toutes les formules de l'ensemble R alors que $\pi_B(\omega) = 1$ signifie que ω satisfait totalement au moins une préférence.

6.8 Réseaux possibilistes garantis

Cette section fournit les fondements de base des réseaux possibilistes basés sur la mesure de possibilité garantie Δ . Avant d'aborder ces réseaux, nous introduisons cette section par la définition du conditionnement possibiliste basé sur la mesure Δ .

6.8.1 Conditionnement possibiliste basé sur la mesure Δ

Le conditionnement possibiliste [BMCM02, BABD⁺02] est une notion cruciale qui consiste à réviser les croyances sur les informations dont on dispose. Le conditionnement possibiliste consiste à modifier les connaissances initiales, exprimées par $\pi(\omega)$ et encodées par une distribution de possibilités π , suite à l'arrivée d'une nouvelle pièce d'information $\varphi \subseteq \Omega$. En conséquence, chaque $\pi(\omega)$ sera remplacée par $\pi(\omega|\varphi)$. Le conditionnement possibiliste basé sur la mesure Δ est défini comme suit :

Définition 6.1 *Étant donné une distribution de possibilités π , le conditionnement possibiliste basé sur la mesure Δ est défini par :*

$$\Delta(\omega|\varphi) = \min\{\pi_m(\omega|\varphi) : \omega \models \varphi\}, \quad (6.5)$$

où $\pi_m(\omega|\varphi)$ est le conditionnement possibiliste basée sur l'opérateur \min [DP04b] dont on rappelle la définition (voir l'équation 1.29) :

$$\pi_m(\omega|\varphi) = \begin{cases} 1 & \text{si } \pi(\omega) = \Pi(\varphi) \text{ et } \omega \models \varphi \\ \pi(\omega) & \text{si } \pi(\omega) < \Pi(\varphi) \text{ et } \omega \models \varphi \\ 0 & \text{sinon.} \end{cases} \quad (6.6)$$

La distribution de possibilités conditionnées $\Delta(\omega|\varphi)$ garantie des valeurs minimales de possibilité pour chaque interprétation ω . L'exemple suivant illustre comment obtenir cette distribution de possibilités conditionnées à partir de la définition précédente :

Exemple 6.6 *Considérons de nouveau, la distribution de possibilités $\pi_{\mathfrak{R}}$ dont les valeurs sont données dans le tableau 6.1. Supposons qu'une nouvelle information matérialisée par $\varphi = bg \wedge cl$ est parvenue. Pour chaque $\omega \in \Omega$, la distribution initiale $\pi_{\mathfrak{R}}(\omega)$ sera remplacée par le conditionnement possibiliste $\Delta(\omega|\varphi)$. Calculons, tout d'abord, le degré de possibilité de φ :*

$$\Pi(\varphi) = \max\{\pi_{\mathfrak{R}}(\omega) : \omega \models \varphi\} = \max\{0.75, 0.95, 1, 1\} = 1.$$

Le tableau 6.4 montre les valeurs de la distribution de possibilités conditionnelles basée sur le \min , $\pi_m(\omega|\varphi)$ obtenues à partir de la formule (6.6). Dans le même tableau

| ω | $\pi_m(\omega \varphi)$ | $\Delta(\omega \varphi)$ |
|--|-------------------------|--------------------------|
| $\neg bg \wedge \neg sl \wedge \neg cl \wedge \neg bb$ | 0 | 0 |
| $\neg bg \wedge \neg sl \wedge \neg cl \wedge bb$ | 0 | 0 |
| $\neg bg \wedge \neg sl \wedge cl \wedge \neg bb$ | 0 | 0 |
| $\neg bg \wedge \neg sl \wedge cl \wedge bb$ | 0 | 0 |
| $\neg bg \wedge sl \wedge \neg cl \wedge \neg bb$ | 0 | 0 |
| $\neg bg \wedge sl \wedge \neg cl \wedge bb$ | 0 | 0 |
| $\neg bg \wedge sl \wedge cl \wedge \neg bb$ | 0 | 0 |
| $\neg bg \wedge sl \wedge cl \wedge bb$ | 0 | 0 |
| $bg \wedge \neg sl \wedge \neg cl \wedge \neg bb$ | 0 | 0 |
| $bg \wedge \neg sl \wedge \neg cl \wedge bb$ | 0 | 0 |
| $bg \wedge \neg sl \wedge cl \wedge \neg bb$ | 0.75 | 0.75 |
| $bg \wedge \neg sl \wedge cl \wedge bb$ | 0.95 | 0.75 |
| $bg \wedge sl \wedge \neg cl \wedge \neg bb$ | 0 | 0 |
| $bg \wedge sl \wedge \neg cl \wedge bb$ | 0 | 0 |
| $bg \wedge sl \wedge cl \wedge \neg bb$ | 1 | 0.75 |
| $bg \wedge sl \wedge cl \wedge bb$ | 1 | 0.75 |

TAB. 6.4 – Distributions de possibilités conditionnelles $\pi_m(\omega|\varphi)$ et $\Delta(\omega|\varphi)$

sont données les valeurs de la distribution de possibilités conditionnelles $\Delta(\omega|\varphi)$ issues de la formule (6.5). $\forall \omega \models \varphi$, $\Delta(\omega|\varphi) = \min\{0.75, 0.95, 1, 1\} = 0.75$.

La définition du conditionnement possibiliste basée sur Δ sera à la base de la définition des réseaux possibilistes basés sur la mesure Δ .

6.8.2 Réseaux possibilistes garantis

Les réseaux possibilistes basés sur la mesure Δ sont définis de la même manière que les réseaux probabilistes et possibilistes standards [ABOY09] :

Définition 6.2 *Un réseau possibiliste garanti est caractérisé par deux composantes :*

- *Une composante graphique : Un réseau possibiliste garanti Δ est un graphe acyclique direct (DAG) dans lequel les nœuds correspondent aux variables et les arcs entre les nœuds traduisent les relations d'influence (ou de dépendance) entre ces variables.*
- *Une composante numérique représentant l'incertitude en chaque nœud. Dans les réseaux possibilistes basés sur Δ , l'incertitude est évaluée par la mesure de possibilité garantie Δ au lieu des distributions de possibilités π utilisées dans les réseaux*

possibilistes standards.

Plus précisément, un réseau possibiliste basé sur la mesure de possibilité garantie Δ noté par \mathcal{T} et défini sur un ensemble de variables V est un DAG dans lequel l'incertitude est mesurée, en chaque nœud $A_j \in V$, par :

- $\Delta_{\mathcal{T}}(a_i)$ si A_j est un nœud racine ($U_{A_j} = \emptyset$)
- $\Delta_{\mathcal{T}}(a_i | U_{A_j})$ si $U_{A_j} \neq \emptyset$,

avec U_{A_j} est l'ensemble des parents du nœud A_j et $a_i \in D_{A_j}$.

Explicitement, $\Delta_{\mathcal{T}}(a_i)$ correspond au degré de possibilité minimal associé à l'instance a_i du nœud A_j .

Exemple 6.7 La figure 6.2 illustre l'exemple d'un réseau possibiliste \mathcal{T} basé sur la mesure Δ et composé de cinq nœuds représentant respectivement les cinq variables binaires suivantes : 'SB' pour saison des baignades, 'SL' pour temps ensoleillé, 'CL' pour mer calme, 'BG' pour baignade en mer et 'BB' pour balade en bateau.

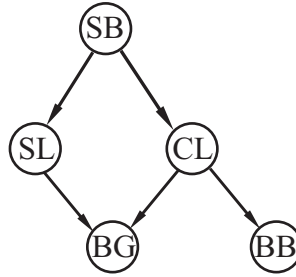


FIG. 6.2 – Exemple d'un réseau possibiliste garanti \mathcal{T} basé sur Δ

Les relations entre les nœuds du réseau s'expliquent ainsi : selon que nous sommes ou pas en saison des baignades 'SB', l'état du temps 'SL' (ensoleillé ou couvert) et de la mer 'CL' (calme ou agitée) en dépendent directement. La décision de faire une baignade en mer 'BG' dépend des états du temps 'SL' et de la mer 'CL' ensemble. Enfin, une balade en bateau 'BB' est conditionnée par l'état de la mer 'CL'.

En ce qui concerne la composante numérique associée à ce réseau possibiliste Δ , les tableaux 6.5 et 6.6 fournissent les distributions de possibilités conditionnelles initiales garanties de chaque nœud dans le contexte de ses parents.

Il est à souligner que, contrairement aux réseaux probabilistes ou aux réseaux possibilistes standards, les distributions de possibilités garanties ne sont pas nécessairement normalisées. C'est à dire que les degrés de possibilités garantis associés à un réseau \mathcal{T} basé sur Δ vérifient les inégalités suivantes :

- $\max\{\Delta_{\mathcal{T}}(a_i) : a_i \in D_{A_j}\} \leq 1$ si $U_{A_j} = \emptyset$,

| SB | $\Delta_{\mathcal{T}}(SB)$ | $SL \quad SB$ | $\Delta_{\mathcal{T}}(SL SB)$ | $CL \quad SB$ | $\Delta_{\mathcal{T}}(CL SB)$ | $BB \quad CL$ | $\Delta_{\mathcal{T}}(BB CL)$ |
|-----------|----------------------------|-------------------------|-------------------------------|-------------------------|-------------------------------|-------------------------|-------------------------------|
| sb | 0.40 | $sl \quad sb$ | 0.60 | $cl \quad sb$ | 0.70 | $bb \quad cl$ | 0.66 |
| $\neg sb$ | 0.25 | $sl \quad \neg sb$ | 0.15 | $cl \quad \neg sb$ | 0.33 | $bb \quad \neg cl$ | 0.12 |
| | | $\neg sl \quad sb$ | 0.10 | $\neg cl \quad sb$ | 0.50 | $\neg bb \quad cl$ | 0.25 |
| | | $\neg sl \quad \neg sb$ | 0.55 | $\neg cl \quad \neg sb$ | 0.60 | $\neg bb \quad \neg cl$ | 0.45 |

TAB. 6.5 – Distributions initiales de possibilités conditionnelles associées à \mathcal{T}

| BG | SL | CL | $\Delta_{\mathcal{T}}(BG SL, CL)$ |
|-----------|-----------|-----------|-----------------------------------|
| bg | sl | cl | 0.90 |
| bg | sl | $\neg cl$ | 0.40 |
| bg | $\neg sl$ | cl | 0.50 |
| bg | $\neg sl$ | $\neg cl$ | 0.15 |
| $\neg bg$ | sl | cl | 0.20 |
| $\neg bg$ | sl | $\neg cl$ | 0.30 |
| $\neg bg$ | $\neg sl$ | cl | 0.40 |
| $\neg bg$ | $\neg sl$ | $\neg cl$ | 0.60 |

TAB. 6.6 – Distributions initiales de possibilités conditionnelles associées à \mathcal{T}

$$- \max\{\Delta_{\mathcal{T}}(a_i|U_{A_j}) : a_i \in D_{A_j}\} \leq 1 \text{ si } U_{A_j} \neq \emptyset,$$

avec U_{A_j} l'ensemble des parents du nœud A_j .

Par exemple, pour le nœud SB de l'exemple précédent $\Delta_{\mathcal{T}}(sb) = 0.40$ et $\Delta_{\mathcal{T}}(\neg sb) = 0.25$ signifient que, respectivement, l'état sb est garanti au minimum avec un degré égal à 0.40 et que l'état $\neg sb$ est garanti, quant à lui, à un degré au minimum égal à 0.25.

Règle de chaînage garantie

La distribution de possibilités jointe associée au réseau possibiliste garanti basé sur Δ se définit par la règle de chaînage suivante [ABOY09] :

Définition 6.3 *Étant donné un réseau possibiliste \mathcal{T} basé sur Δ . La distribution de possibilités jointe, définie sur un ensemble de variables $V = \{A_1, A_2, \dots, A_n\}$, peut s'exprimer comme étant le maximum des n degrés de possibilités conditionnelles initiales garanties associés à chaque nœud du réseau. Explicitement, la distribution de possibilités jointe garantie, noté par $\pi_{\mathcal{T}}$, est obtenue par la règle de chaînage garantie basée sur Δ , comme suit :*

$$\pi_{\mathcal{T}}(A_1, A_2, \dots, A_n) = \max \{ \Delta_{\mathcal{T}}(A_i | U_{A_i}) : i = 1, \dots, n \} \quad (6.7)$$

En regardant de près, nous remarquons qu'il existe une similitude entre les équations (6.7) et (6.2). En effet, ces deux équations utilisent le même opérateur max pour définir les distributions de possibilités jointes garanties. Nous verrons dans la prochaine section qu'il est possible de transformer un réseau possibiliste garanti en base possibiliste basée sur la mesure Δ .

Cette règle de chaînage, définie plus haut, provient des relations d'indépendance induites par la structure du réseau \mathcal{T} et par les degrés de possibilités conditionnelles initiales.

Exemple 6.8 *Considérons de nouveau le réseau possibiliste \mathcal{T} de la figure 6.2 et les degrés de possibilités garanties initiales donnés par les tableaux 6.5 et 6.6. La distribution de possibilités jointe $\pi_{\mathcal{T}}$ est illustrée dans le tableau 6.7. À titre d'exemple, nous effectuons le calcul suivant :*

$$\begin{aligned}\Delta_{\mathcal{T}}(sb, sl, \neg cl, bg, \neg bb) &= \max\{\Delta_{\mathcal{T}}(sb), \Delta_{\mathcal{T}}(sl|sb), \Delta_{\mathcal{T}}(\neg cl|sb), \Delta_{\mathcal{T}}(bg|sl, \neg cl), \Delta_{\mathcal{T}}(\neg bb|\neg cl)\} \\ &= \max\{0.40, 0.60, 0.50, 0.40, 0.45\} \\ &= 0.60.\end{aligned}$$

6.9 Réseaux possibilistes garantis et préférences

Dans cette section, nous montrons qu'il est possible de représenter les préférences, exprimées par un utilisateur, à travers les réseaux possibilistes garantis. Pour cela, nous allons considérer les deux cas de préférences : inconditionnelles et conditionnelles.

6.9.1 Préférences basiques inconditionnelles :

Dans ce contexte un utilisateur peut formuler et ordonner une liste d'alternatives exprimant ses préférences pour un sujet donné et cela sans aucune condition préalable sur le choix des préférences. Considérons le cas de figure suivant : supposons qu'un utilisateur exprime ses préférences concernant la couleur qu'il aime. Les différentes alternatives (couleurs) à considérer sont : le *bleu* (B), le *vert* (V), le *jaune* (J) et le *noir* (N). Notons que cet utilisateur préfère ces couleurs dans leur ordre de citation.

Pour modéliser cette situation avec les réseaux possibilistes garantis, nous considérons un réseau \mathcal{T} composé d'un seul nœud (racine) représentant la variable *Couleur* et dont les instances sont les alternatives exclusives : *bleu* (B), *vert* (V), *jaune* (J) et *noir* (N).

| SB | SL | CL | BG | BB | $\pi_{\mathcal{T}}$ |
|-----------|-----------|-----------|-----------|-----------|---------------------|
| sb | sl | cl | bg | bb | 0.9 |
| sb | sl | cl | bg | $\neg bb$ | 0.9 |
| sb | sl | cl | $\neg bg$ | bb | 0.7 |
| sb | sl | cl | $\neg bg$ | $\neg bb$ | 0.7 |
| sb | sl | $\neg cl$ | bg | bb | 0.6 |
| sb | sl | $\neg cl$ | bg | $\neg bb$ | 0.6 |
| sb | sl | $\neg cl$ | $\neg bg$ | bb | 0.6 |
| sb | sl | $\neg cl$ | $\neg bg$ | $\neg bb$ | 0.6 |
| sb | $\neg sl$ | cl | bg | bb | 0.7 |
| sb | $\neg sl$ | cl | bg | $\neg bb$ | 0.7 |
| sb | $\neg sl$ | cl | $\neg bg$ | bb | 0.7 |
| sb | $\neg sl$ | cl | $\neg bg$ | $\neg bb$ | 0.7 |
| sb | $\neg sl$ | $\neg cl$ | bg | bb | 0.5 |
| sb | $\neg sl$ | $\neg cl$ | bg | $\neg bb$ | 0.5 |
| sb | $\neg sl$ | $\neg cl$ | $\neg bg$ | bb | 0.6 |
| sb | $\neg sl$ | $\neg cl$ | $\neg bg$ | $\neg bb$ | 0.6 |
| $\neg sb$ | sl | cl | bg | bb | 0.9 |
| $\neg sb$ | sl | cl | bg | $\neg bb$ | 0.9 |
| $\neg sb$ | sl | cl | $\neg bg$ | bb | 0.66 |
| $\neg sb$ | sl | cl | $\neg bg$ | $\neg bb$ | 0.33 |
| $\neg sb$ | sl | $\neg cl$ | bg | bb | 0.6 |
| $\neg sb$ | sl | $\neg cl$ | bg | $\neg bb$ | 0.6 |
| $\neg sb$ | sl | $\neg cl$ | $\neg bg$ | bb | 0.6 |
| $\neg sb$ | sl | $\neg cl$ | $\neg bg$ | $\neg bb$ | 0.6 |
| $\neg sb$ | $\neg sl$ | cl | bg | bb | 0.66 |
| $\neg sb$ | $\neg sl$ | cl | bg | $\neg bb$ | 0.55 |
| $\neg sb$ | $\neg sl$ | cl | $\neg bg$ | bb | 0.66 |
| $\neg sb$ | $\neg sl$ | cl | $\neg bg$ | $\neg bb$ | 0.55 |
| $\neg sb$ | $\neg sl$ | $\neg cl$ | bg | bb | 0.6 |
| $\neg sb$ | $\neg sl$ | $\neg cl$ | bg | $\neg bb$ | 0.6 |
| $\neg sb$ | $\neg sl$ | $\neg cl$ | $\neg bg$ | bb | 0.6 |
| $\neg sb$ | $\neg sl$ | $\neg cl$ | $\neg bg$ | $\neg bb$ | 0.6 |

TAB. 6.7 – Distribution de possibilités jointe de $\pi_{\mathcal{T}}$

La distribution initiale de possibilités garantie exprime l'ordre de préférence entre ces 4 alternatives comme le montre le tableau 6.8. (Notons que $\alpha \in]0, 1[$).

| <i>Couleur</i> | $\Delta_{\mathcal{T}}(\textit{Couleur})$ |
|----------------|--|
| <i>B</i> | α^1 |
| <i>V</i> | α^2 |
| <i>J</i> | α^3 |
| <i>N</i> | α^4 |

TAB. 6.8 – Distribution initiale de possibilités garantie associée à \mathcal{T}

Il est naturel que l'alternative (l'instance) la plus préférée *Couleur* = *B* possède le degré de possibilité garantie le plus élevé α^1 . Ainsi, le fait que $\Delta_{\mathcal{T}}(B) > \Delta_{\mathcal{T}}(V) > \Delta_{\mathcal{T}}(J) > \Delta_{\mathcal{T}}(N)$ s'interprète de la manière suivante : la couleur *Bleue* est préférée à la couleur *verte* qui elle-même préférée à la couleur *jaune*, elle-même préférée au *noir*.

Une autre façon qui permet de représenter et d'ordonner les préférences sur les couleurs est celle qui utilise les formules de choix de base (*BCF*) issues de la logique du choix qualitatif *QCL*. En effet, en utilisant le connecteur de disjonction ordonnée $\overrightarrow{\times}$, nous pouvons ordonner les alternatives des couleurs en exprimant la formule *BCF* suivante : $B \overrightarrow{\times} V \overrightarrow{\times} J \overrightarrow{\times} N$. Ceci veut dire que le bleu est préféré en premier mais si c'est pas possible alors le vert, si cela n'est pas possible alors le jaune , si cela n'est pas possible alors le noir.

En plus de cette formule *BCF*, d'autres formules logiques interviennent pour exprimer l'exclusivité entre les différentes couleurs (une couleur ne peut pas être jaune et verte à la fois). Aussi, la formule logique $\neg B \vee \neg V$ exprime qui si la couleur est bleue alors elle ne peut pas être verte et inversement.

D'où l'ensemble K de toutes les formules exprimant les préférences de l'utilisateur sur les couleurs : $K = \{B \overrightarrow{\times} V \overrightarrow{\times} J \overrightarrow{\times} N, B \vee V \vee J \vee N, \neg B \vee \neg V, \neg B \vee \neg J, \neg B \vee \neg N, \neg V \vee \neg J, \neg V \vee \neg N, \neg J \vee \neg N\}$.

Pour ce cas particulier de préférences basiques inconditionnelles, nous constatons que les réseaux possibilistes garantis sont aussi représentatifs que les formules de choix de base (*BCF*), si ce n'est pas mieux.

Toujours pour ce même cas, la remarque suivante permet d'établir une équivalence entre la relation d'inférence [BBD04] de la formule *BCF* et les valeurs de la distri-

bution jointe de possibilités garantie $\pi_{\mathcal{T}}$ associée au réseau garanti T :

Soit $\pi_{\mathcal{T}}$ la distribution jointe de possibilités garantie associée au réseau garanti T . Soit C la seule variable constituant Ω et ayant n instances c_1, c_2, \dots, c_n tel que $\forall i = 1..n, \Delta(c_i) = \alpha^i$. Soit ω une interprétation appartenant à Ω , alors :

$$\pi_{\mathcal{T}}(\omega) = \alpha^i \text{ si et seulement si } \omega \models_i c_1 \overrightarrow{\times} c_2 \overrightarrow{\times} \dots \overrightarrow{\times} c_n.$$

La vérification de cette équivalence peut se faire en justifiant les deux sens de l'implication :

L'implication \Rightarrow :

Etant donné que Ω ne contient que le singleton $\{C\}$ ayant n instances différentes, $\pi_{\mathcal{T}}(\omega) = \alpha^i$ signifie que $\omega \models c_i$ et $\omega \not\models c_j \forall j \neq i$.

$$\Rightarrow \omega \models c_1 \vee \dots \vee c_n \text{ avec } 1 \leq i \leq n \text{ et } \omega \not\models c_j, 1 \leq j < i$$

$$\Rightarrow \omega \models_i c_1 \overrightarrow{\times} c_2 \overrightarrow{\times} \dots \overrightarrow{\times} c_n.$$

L'implication \Leftarrow :

$$\omega \models_i c_1 \overrightarrow{\times} c_2 \overrightarrow{\times} \dots \overrightarrow{\times} c_n \Rightarrow \omega \models c_1 \vee \dots \vee c_n \text{ et } i = \min\{j : \omega \models c_j\}$$

$$\Rightarrow \omega \models c_1 \vee \dots \vee c_n \text{ et } \omega \models c_i \text{ et } \omega \not\models c_j \forall j \neq i$$

$$\Rightarrow \omega \models c_i \Rightarrow \pi_{\mathcal{T}}(\omega) = \alpha^i \text{ puisque } \Delta_{\mathcal{T}}(c_i) = \alpha^i.$$

6.9.2 Préférences conditionnelles

Le contexte de préférences conditionnelles survient lorsqu'un utilisateur exprime sa préférence sur un sujet, étant donné la présence d'un ou plusieurs de ses choix de préférences préalablement fait sur d'autres thèmes liés à ce sujet.

Considérons l'exemple suivant : supposons qu'un utilisateur est confronté à une série d'alternatives représentant différents types de voiture (*Peugeot*, *Renault*, *Fiat* et *autre*). L'utilisateur classe ces alternatives en allant du type de voiture le plus préféré vers le type le moins préféré (il préférera alors *Peugeot* à *Renault* à *Fiat* à toute *autre* marque). Ces préférences sont naturellement inconditionnelles. Supposons maintenant que ce même utilisateur est amené à définir la couleur $\{\text{bleu}, \text{vert}, \text{jaune et noir}\}$ du type de la voiture qui lui sera attribuée selon ses préférences. Pour chaque préférence de type de voiture, l'utilisateur indique ses préférences de couleur pour ce type. Il dira alors : "Si ma première préférence de type de voiture est satisfaite alors voici mes préférences sur sa couleur. Si ma deuxième préférence est satisfaite alors voici mes préférences sur sa couleur...". Les préférences de couleur sont donc conditionnelles au type de la voiture.

Afin de modéliser cette association de préférences par des réseaux possibilistes garantis, nous considérons deux nœuds, l'un pour exprimer la variable *type de voi-*

ture (ayant 4 instances : *Peugeot*, *Renault*, *Fiat* et *autre*) et l'autre pour exprimer sa couleur (ayant aussi 4 instances *bleu*, *vert*, *jaune* et *noir*). Un arc dirigé de la première variable vers la deuxième indique la dépendance du choix des couleurs par rapport au type de voiture. La figure 6.3 illustre ce réseau possibiliste garanti \mathcal{T} . Les distributions initiales conditionnelles de possibilités garanties $\Delta_{\mathcal{T}}$ sont présentées dans le tableau 6.9. (Pour une raison de clarté, nous ne présenterons qu'une partie de la distribution initiale conditionnelle de la variable couleur. $\alpha, \beta \in]0, 1[$).



FIG. 6.3 – Réseau possibiliste garanti.

| <i>type voiture</i> | $\Delta_{\mathcal{T}}(\textit{type voiture})$ | <i>couleur</i> | $\Delta_{\mathcal{T}}(\textit{couleur} \textit{type voiture} = \textit{Peugeot})$ |
|---------------------|---|----------------|---|
| <i>Peugeot</i> | β^1 | <i>B</i> | α^1 |
| <i>Renault</i> | β^2 | <i>V</i> | α^2 |
| <i>Fiat</i> | β^3 | <i>J</i> | α^3 |
| <i>autre</i> | β^4 | <i>N</i> | α^4 |

TAB. 6.9 – Distributions initiales conditionnelles de possibilités garanties

Par exemple, si la préférence sur le type de voiture *Peugeot* est satisfaite alors les préférences de l'utilisateur sur les couleurs de ce type seront exprimées et ordonnées ($\alpha^1 > \alpha^2 > \alpha^3 > \alpha^4$).

Par ailleurs, il est possible de représenter le même cas de préférences avec les CP-Nets. sauf que ces derniers sont beaucoup plus appropriés aux cas où les variables sont binaires. Ce qui n'est pas le contexte de notre cas.

En utilisant les formules de choix de base (*BCF*) de la logique \mathcal{QCL} , il est possible de modéliser et d'ordonner les alternatives correspondant aux préférences conditionnelles d'un utilisateur. En effet, l'exemple étudié peut s'exprimer par les formules suivantes (constituant l'ensemble des formules K) :

- $Peugeot \overrightarrow{\times} Renault \overrightarrow{\times} Fiat \overrightarrow{\times} autre$ pour exprimer les préférences inconditionnelles sur les types de voiture.
- $Peugeot \Rightarrow B \overrightarrow{\times} V \overrightarrow{\times} J \overrightarrow{\times} N$ pour exprimer les préférences conditionnelles sur les couleurs sachant que le type de voiture est *Peugeot*.
- $Renault \Rightarrow N \overrightarrow{\times} V \overrightarrow{\times} B \overrightarrow{\times} J$ pour exprimer les préférences conditionnelles sur les couleurs sachant que le type de voiture est *Renault*.
- $Fiat \Rightarrow J \overrightarrow{\times} B \overrightarrow{\times} V \overrightarrow{\times} N$ pour exprimer les préférences conditionnelles sur les couleurs sachant que le type de voiture est *Fiat*.
- $autre \Rightarrow B \overrightarrow{\times} N \overrightarrow{\times} V \overrightarrow{\times} J$ pour exprimer les préférences conditionnelles sur les couleurs sachant que le type de voiture est *autre*.

On peut ramener les quatre dernières formules en *BCF*, par exemple :

$$Peugeot \Rightarrow B \overrightarrow{\times} V \overrightarrow{\times} J \overrightarrow{\times} N \text{ est équivalente à } (Peugeot \Rightarrow B) \overrightarrow{\times} (Peugeot \Rightarrow V) \overrightarrow{\times} (Peugeot \Rightarrow J) \overrightarrow{\times} (Peugeot \Rightarrow N)$$

où encore :

$$(\neg Peugeot \vee B) \overrightarrow{\times} (\neg Peugeot \vee V) \overrightarrow{\times} (\neg Peugeot \vee J) \overrightarrow{\times} (\neg Peugeot \vee N).$$

Sans oublier qu'il faut ajouter toutes les formules logiques exprimant d'une part l'exclusion entre les types de voiture et d'autre part l'exclusion entre les couleurs.

Notre conclusion est que les réseaux possibilistes garantis sont les outils les plus adéquat pour représenter et ordonner les préférences. En effet, les réseaux possibilistes garantis se distinguent par rapport à la logique *QCL* par la simplicité de représentation : ils emploient clairement une composante graphique pour illustrer les préférences conditionnelles et une composante numérique pour ordonner ces préférences. Alors que la logique *QCL* nécessite un grand nombre de formules *BCF* pour représenter le même cas de préférences (pour l'exemple précédent de préférences conditionnelles, l'ensemble *K* contient au moins 19 formules *BCF*).

Un deuxième avantage mis en évidence par l'utilisation des réseaux possibilistes garantis pour la représentation des préférences est la diversité des états de celles-ci. En effet, une variable (nœud) du réseau garanti, exprimant une préférence, peut avoir une cardinalité supérieur à deux (plus que binaire). Cette propriété naturelle peut être un vrai handicap dans d'autres modèles de représentation des préférences notamment les CP-Nets.

6.9.3 Un exemple illustratif

Nous présentons ci-dessous un exemple de réseau possibiliste garanti représentant des préférences utilisateur. La figure 6.4 illustre un réseau garanti \mathcal{T} formé de six nœuds dont chacun représente, selon le cas, des préférences inconditionnelles ou conditionnelles sur les loisirs.

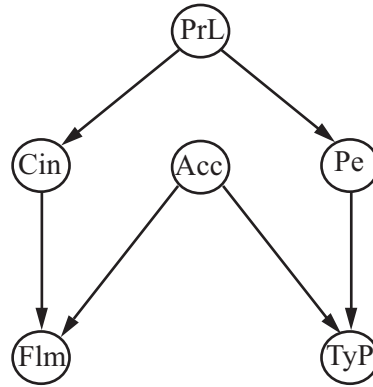


FIG. 6.4 – Exemple de réseau possibiliste garanti représentant des préférences.

Le premier nœud *PrL* exprime les préférences inconditionnelles sur le choix de la période pendant laquelle une personne préfère se détendre. Cette personne peut choisir entre l'une des deux périodes : *en semaine* ou pendant le *week – end*. Le deuxième nœud *Cin* exprime le loisir d'aller au cinéma pour voir un film pendant la première, la deuxième ou la troisième séance de présentation. La préférence d'une séance plutôt qu'une autre est conditionnée par le choix de la période de loisir. Le troisième nœud *Pe* représente le loisir d'aller à la pêche soit le matin, l'après-midi ou le soir. De même, le choix du moment pour aller à la pêche est conditionné par la période choisie. Le quatrième nœud *Acc* indique, inconditionnellement, si la personne préfère être accompagnée (ou pas) pendant l'un de ces deux loisirs. Le cinquième nœud *Flm* désigne les catégories des films (comique, d'action, science-fiction) que la personne peut visionner sachant qu'elle a choisi la séance de présentation et a exprimé sa préférence quant à être accompagné ou non. Le sixième nœud *TyP* exprime les modes de pêche (à la canne, en bateau) qu'une personne préfère pratiquer conditionnellement au moment de pêche et du fait qu'il soit accompagné ou pas. Le tableau 6.10 montre les distributions des possibilités garantis conditionnelles associées au réseau garanti \mathcal{T} . Ces distributions révèlent l'ordre des préférences d'un agent relativement aux alternatives de chaque variable conditionnellement aux préférences déjà exprimées. Par exemple, il est lisible que cet agent, préfère se détendre le week-end plutôt qu'au cours de la semaine ($\Delta_{\mathcal{T}}(we) > \Delta_{\mathcal{T}}(se)$). Aussi, cet agent préfère, s'il a choisi de se détendre en semaine, d'aller au cinéma pour voir un film à la deuxième séance plutôt que à la troisième et à la première séance ($\Delta_{\mathcal{T}}(s_2|se) > \Delta_{\mathcal{T}}(s_3|se) > \Delta_{\mathcal{T}}(s_1|se)$).

| PrL | $\Delta_{\mathcal{T}}(PrL)$ | Cin | $\Delta_{\mathcal{T}}(Cin PrL = se)$ | $\Delta_{\mathcal{T}}(Cin PrL = we)$ |
|-------|-----------------------------|-------|--------------------------------------|--------------------------------------|
| se | 0.4 | s_1 | 0.1 | 0.3 |
| we | 0.7 | s_2 | 0.8 | 0.7 |
| | | s_3 | 0.4 | 0.9 |

| Acc | $\Delta_{\mathcal{T}}(Acc)$ | Pe | $\Delta_{\mathcal{T}}(Pe PrL = se)$ | $\Delta_{\mathcal{T}}(Pe PrL = we)$ |
|------------|-----------------------------|------|-------------------------------------|-------------------------------------|
| acc | 0.75 | ma | 0.1 | 0.3 |
| $\neg acc$ | 0.4 | ap | 0.4 | 0.9 |
| | | sr | 0.8 | 0.7 |

| Flm | Cin | Acc | $\Delta_{\mathcal{T}}(Flm Cin, Acc)$ | Typ | Pe | Acc | $\Delta_{\mathcal{T}}(Typ Pe, Acc)$ |
|-------|-------|------------|--------------------------------------|-------|------|------------|-------------------------------------|
| com | s_1 | acc | 0.4 | ca | ma | acc | 0.4 |
| com | s_1 | $\neg acc$ | 0.1 | ca | ma | $\neg acc$ | 0.9 |
| com | s_2 | acc | 0.9 | ca | ap | acc | 0.6 |
| com | s_2 | $\neg acc$ | 0.6 | ca | ap | $\neg acc$ | 0.8 |
| com | s_3 | acc | 0.3 | ca | sr | acc | 0.5 |
| com | s_3 | $\neg acc$ | 0.15 | ca | sr | $\neg acc$ | 0.2 |
| act | s_1 | acc | 0.3 | bt | ma | acc | 0.85 |
| act | s_1 | $\neg acc$ | 0.1 | bt | ma | $\neg acc$ | 0.5 |
| act | s_2 | acc | 0.6 | bt | ap | acc | 1 |
| act | s_2 | $\neg acc$ | 0.7 | bt | ap | $\neg acc$ | 0.3 |
| act | s_3 | acc | 0.7 | bt | sr | acc | 0.15 |
| act | s_3 | $\neg acc$ | 0.9 | bt | sr | $\neg acc$ | 0.05 |

TAB. 6.10 – Distributions de possibilités garantis conditionnelles initiales

6.10 Représentation logique des réseaux garantis Δ

Dans cette section, nous montrons qu'un réseau possibiliste garanti peut se transformer, en temps polynomial, en une base logique possibiliste basée sur la mesure Δ . Le but de cette transformation est de profiter des avantages de la représentation logique des bases possibilistes garanties afin de tirer des conclusions plausibles.

Commençons par expliquer comment obtenir une base possibiliste garantie $\mathfrak{R}_{\mathcal{T}}$ équivalente à un réseau possibiliste garanti \mathcal{T} .

Soit A_j une variable et $\Delta_{\mathcal{T}}(a_i|U_{A_j})$ le degré de possibilité garanti local associé au nœud A_j . La base possibiliste $\mathfrak{R}_{\mathcal{T}}$ basée sur la mesure Δ est construite de la manière

suivante :

- Si A_j est un nœud racine ($U_{A_j} = \emptyset$), alors ajouter, pour chaque instance a_i de A_j , une formule pondérée $[a_i, \Delta_{\mathcal{T}}(a_i)]$ à la base $\mathfrak{R}_{\mathcal{T}}$.
- Si A_j admet des parents ($U_{A_j} \neq \emptyset$), alors ajouter, pour chaque instance a_i sachant chaque instance de ses parents, une formule pondérée $[a_i \wedge u_{A_j}, \Delta_{\mathcal{T}}(a_i|u_{A_j})]$ à la base $\mathfrak{R}_{\mathcal{T}}$.

Nous remarquons que le conditionnement sur l'ensemble U_{A_j} des parents du nœud A_j est remplacé par une conjonction entre les éléments de U_{A_j} . Formellement, la base $\mathfrak{R}_{\mathcal{T}}$ s'obtient par la formule suivante :

$$\mathfrak{R}_{\mathcal{T}} = \{[a_i \wedge u_{A_j}, \alpha] : \alpha = \Delta_{\mathcal{T}}(a_i|u_{A_j}) \in \mathcal{T}\}. \quad (6.8)$$

Exemple 6.9 *Considérons le réseau possibiliste garanti \mathcal{T} de la figure 6.2 et ses distributions de possibilités garanties initiales données par les tableaux 6.5 et 6.6. La base logique possibiliste basée sur Δ associé à \mathcal{T} est obtenue en appliquant la formule (6.8) :*

$$\begin{aligned} \mathfrak{R}_{\mathcal{T}} = \{ & [sb, 0.4], [\neg sb, 0.25], [sl \wedge sb, 0.6], [sl \wedge \neg sb, 0.15], [\neg sl \wedge sb, 0.1], [\neg sl \wedge \\ & \neg sb, 0.55], [cl \wedge sb, 0.7], [cl \wedge \neg sb, 0.33], [\neg cl \wedge sb, 0.5], [\neg cl \wedge \neg sb, 0.6], [bg \wedge sl \wedge \\ & cl, 0.9], [bg \wedge sl \wedge \neg cl, 0.4], [bg \wedge \neg sl \wedge cl, 0.5], [bg \wedge \neg sl \wedge \neg cl, 0.15], [\neg bg \wedge sl \wedge \\ & cl, 0.2], [\neg bg \wedge sl \wedge \neg cl, 0.3], [\neg bg \wedge \neg sl \wedge cl, 0.4], [\neg bg \wedge \neg sl \wedge \neg cl, 0.6], [bb \wedge cl, 0.66], [bb \wedge \\ & \neg cl, 0.12], [\neg bb \wedge cl, 0.25], [\neg bb \wedge \neg cl, 0.45]\}. \end{aligned}$$

Par exemple, en utilisant la formule (6.2) de distribution de possibilités associée à une base \mathfrak{R} , nous calculons :

$$\pi_{\mathfrak{R}}(sb, sl, \neg cl, bg, \neg bb) = \max\{0.40, 0.60, 0.50, 0.40, 0.45\} = 0.60$$

Ce qui donne le même résultat que l'exemple 6.6.

Proposition 6.1 *Soient \mathcal{T} un réseau possibiliste garanti et \mathfrak{R} une base possibiliste basée sur la mesure Δ obtenue à partir de \mathcal{T} . Si \mathcal{T} et \mathfrak{R} génèrent la même distribution de possibilités (c-à-d que $\pi_{\mathcal{T}} = \pi_{\mathfrak{R}}$), alors ils sont syntaxiquement équivalents.*

La preuve de cette proposition est la suivante :

soit $\omega = a_1, a_2, \dots, a_n = a_1 \wedge a_2 \wedge \dots \wedge a_n \in \Omega$.

Par définition, nous avons :

Proposition 6.2 *Soient \mathcal{T} un réseau possibiliste garanti, muni de sa distribution $\pi_{\mathcal{T}}$, et \mathfrak{R} une base possibiliste garantie, munie de sa distribution $\pi_{\mathfrak{R}}$, représentant l'équivalent syntaxique de \mathcal{T} ($\pi_{\mathcal{T}} = \pi_{\mathfrak{R}}$). Si $\pi_{\mathfrak{R}}$ est unique alors $\pi_{\mathcal{T}}$ l'est aussi.*

$$\begin{aligned}
\pi_{\mathcal{T}}(\omega) &= \pi_{\mathcal{T}}(a_1, a_2, \dots, a_n) \\
&= \max\{\Delta_{\mathcal{T}}(a_i | u_{A_j}) : j = 1, \dots, n\} \\
&= \max\{\Delta_{\mathcal{T}}(a_i \wedge u_{A_j}) : j = 1, \dots, n \text{ and } a_i \wedge u_{A_j} \subseteq \omega\} \\
&= \pi_{\mathcal{R}}(\omega). \blacksquare
\end{aligned}$$

6.11 Équivalence syntactique entre réseaux possibilistes garantis et réseaux qualitatifs

Après son équivalence logique avec les bases possibilistes garanties, nous allons montrer, dans cette section, que les réseaux possibilistes garantis peuvent aussi s'encoder en réseaux possibilistes standards basés sur des mesures possibilistes. Plus précisément, il s'agit de transformer notre réseau garanti basé sur Δ en un réseau possibiliste qualitatif. Pour éviter l'implémentation de nouveaux algorithmes de propagation destinés aux réseaux garantis, la transformation de réseaux possibilistes garantis en réseaux qualitatifs est importante afin de ré-utiliser les algorithmes d'inférence développés pour les réseaux basés sur le *min*, largement abordés dans la littérature [BAB05, BABK03, BS07b] (voir chapitre 3) ou développés dans le chapitre 4 de ce mémoire.

Par la suite, nous rappelons juste la définition de la règle de chaînage pour les réseaux possibilistes qualitatifs (sous-section 2.5.2) :

$$\pi_{G_m}(A_1, A_2, \dots, A_n) = \min\{\pi(a_i | U_{A_i}) : i = 1, 2, \dots, n\} \quad (6.9)$$

La proposition suivante précise comment est transformé un réseau possibiliste garanti en réseau possibiliste qualitatif.

Proposition 6.3 *Soit l'ensemble des variables $V = \{A_1, A_2, \dots, A_n\}$. Soit \mathcal{T} un réseau possibiliste garanti. Soit ΠG_m un réseau possibiliste basé sur le *min* et associé à \mathcal{T} de la manière suivante : Pour chaque nœud $A_i \in V$, $\pi(A_i | U_{A_i})$ est défini comme étant égal à $1 - \Delta_{\mathcal{T}}(A_i | U_{A_i})$. Alors,*

$$\forall A_i \in V, \pi_{\mathcal{T}}(A_1, A_2, \dots, A_n) = 1 - \pi_{G_m}(A_1, A_2, \dots, A_n) \quad (6.10)$$

La preuve de cette proposition est donnée comme suit :

$$\begin{aligned}
\pi_{\mathcal{T}}(A_1, A_2, \dots, A_n) &= \max\{\Delta_{\mathcal{T}}(A_i | U_{A_i}) : i = 1, \dots, n\} \\
&= \max\{1 - \pi(A_i | U_{A_i}) : i = 1, \dots, n\} \\
&= 1 - \min\{\pi(A_i | U_{A_i}) : i = 1, \dots, n\} \\
&= 1 - \pi_{G_m}(A_1, A_2, \dots, A_n)
\end{aligned}$$

De même, on peut écrire pour chaque variable A_i :

$$\begin{aligned}\Delta_{\mathcal{T}}(A_i) &= \min_{V \setminus \{A_i\}} \{\pi_{\mathcal{T}}(A_1, \dots, A_{i-1}, A_{i+1}, \dots, A_n)\} \\ &= \min_{V \setminus \{A_i\}} \{1 - \pi_{G_m}(A_1, \dots, A_{i-1}, A_{i+1}, \dots, A_n)\} \\ &= 1 - \max_{V \setminus \{A_i\}} \{\pi_{G_m}(A_1, \dots, A_{i-1}, A_{i+1}, \dots, A_n)\} \\ &= 1 - \pi(A_i) \quad \blacksquare\end{aligned}$$

Cette proposition signifie que chaque degré de possibilité garanti d'un événement quelconque peut être exprimé en fonction linéaire du degré de possibilité de cet événement. Ce résultat signifie aussi qu'il est possible de ré-utiliser les algorithmes de propagation définis pour les réseaux possibilistes qualitatifs afin de les mettre en œuvre sur les réseaux possibilités garantis. Afin d'appliquer l'inférence dans un réseau possibiliste garanti Δ , nous proposons la démarche suivante :

- (i) Transformer le réseau possibiliste garanti \mathcal{T} en réseau possibiliste qualitatif équivalent ΠG_m selon la proposition 6.3. Cette transformation n'affecte que la composante numérique du réseau garanti. La composante graphique des deux réseaux demeure identique.
- (ii) Sur le réseau qualitatif ΠG_m obtenu, appliquer l'un des algorithmes d'inférence pour déterminer les distributions de possibilités a posteriori des variables d'intérêt.
- (iii) En utilisant le résultat de la proposition 6.3, retrouver les degrés de possibilités garantis $\Delta_{\mathcal{T}}$, des variables d'intérêt, à partir de π .

L'exemple suivant applique la démarche précédente pour déterminer les degrés de possibilités garantis a posteriori suite à l'observation d'une évidence dans un réseau garanti \mathcal{T} . Notons que pour l'étape (ii), nous avons choisi d'appliquer l'algorithme Loopy Belief Propagation possibiliste qualitatif proposée dans le chapitre 4.

Exemple 6.10 *Considérons le même réseau garanti \mathcal{T} de l'exemple 6.7. Supposons qu'on a observé la variable SL à l'état sl (Le temps est ensoleillé). Nous allons déterminer pour chacune des autres variables, quel état est préféré conditionnellement à $SL = sl$.*

L'application de l'étape (i) de la démarche précédente fournit les tableaux 6.11 et 6.12 montrant les distributions de possibilités conditionnelles initiales associées au réseau qualitatif ΠG_m .

Il est à noter que le résultat obtenu est une distribution qui n'est pas normalisée ($\nexists \omega$ tel que $\pi_{G_m}(\omega) = 1$). Ceci n'affecte en rien le déroulement et l'issue des calculs

effectués pendant la propagation comme l'a montré l'auteur de [Ben06].

| SB | $\pi_{G_m}(SB)$ | SL SB | $\pi_{G_m}(SL SB)$ | CL SB | $\pi_{G_m}(CL SB)$ | BB CL | $\pi_{G_m}(BB CL)$ |
|-----------|-----------------|---------------------|--------------------|---------------------|--------------------|---------------------|--------------------|
| sb | 0.60 | sl sb | 0.40 | cl sb | 0.30 | bb cl | 0.34 |
| $\neg sb$ | 0.75 | sl $\neg sb$ | 0.85 | cl $\neg sb$ | 0.67 | bb $\neg cl$ | 0.88 |
| | | $\neg sl$ sb | 0.90 | $\neg cl$ sb | 0.50 | $\neg bb$ cl | 0.75 |
| | | $\neg sl$ $\neg sb$ | 0.45 | $\neg cl$ $\neg sb$ | 0.40 | $\neg bb$ $\neg cl$ | 0.55 |

TAB. 6.11 – Distributions initiales de possibilités conditionnelles associées à ΠG_m

| BG | SL | CL | $\pi_{G_m}(BG SL, CL)$ |
|-----------|-----------|-----------|------------------------|
| bg | sl | cl | 0.10 |
| bg | sl | $\neg cl$ | 0.60 |
| bg | $\neg sl$ | cl | 0.50 |
| bg | $\neg sl$ | $\neg cl$ | 0.85 |
| $\neg bg$ | sl | cl | 0.80 |
| $\neg bg$ | sl | $\neg cl$ | 0.70 |
| $\neg bg$ | $\neg sl$ | cl | 0.60 |
| $\neg bg$ | $\neg sl$ | $\neg cl$ | 0.40 |

TAB. 6.12 – Distributions initiales de possibilités conditionnelles associées à ΠG_m

Afin de propager l'évidence observée $SL = sl$, nous allons appliquer LBP possibiliste qualitatif. Le tableau 6.13 présente les degrés de possibilités a posteriori des variables non observés. La troisième colonne du même tableau donne les valeurs des degrés de possibilités garantis obtenus suite à l'étape (iii).

Les valeurs des degrés de possibilités garantis $\Delta_{\mathcal{T}}$ obtenues montrent que, sachant le temps est ensoleillé $SL = sl$, une personne préfère se baigner $\Delta_{\mathcal{T}}(BG = bg|SL = sl) = 0.6$ plutôt que non $\Delta_{\mathcal{T}}(BG = \neg bg|SL = sl) = 0.33$, ou bien préfère faire une balade en bateau $\Delta_{\mathcal{T}}(BB = bb|SL = sl) = 0.6$ plutôt que de rester sur terre $\Delta_{\mathcal{T}}(BB = \neg bb|SL = sl) = 0.33$.

6.12 Conclusion

Ce dernier chapitre comporte deux volets : le premier concerne la définition d'un réseau possibiliste garanti Δ . L'incertitude au niveau de chaque nœud est mesurée

| <i>Variable non observé</i> | $\pi_{G_m}(. SL = sl)$ | $\Delta_T(. SL = sl)$ |
|-----------------------------|------------------------|-----------------------|
| $SB = sb$ | 0.40 | 0.60 |
| $SB = \neg sb$ | 0.67 | 0.33 |
| $CL = cl$ | 0.67 | 0.33 |
| $CL = \neg cl$ | 0.40 | 0.60 |
| $BG = bg$ | 0.40 | 0.60 |
| $BG = \neg bg$ | 0.67 | 0.33 |
| $BB = bb$ | 0.40 | 0.60 |
| $BB = \neg bb$ | 0.67 | 0.33 |

TAB. 6.13 – Les degrés de possibilités a posteriori des variables non observés

par le degré de possibilité minimal Δ associé à chaque instance de ce nœud. Nous avons aussi donné une représentation logique équivalente à ce réseau possibiliste Δ sous la forme d'une base possibiliste garantie. Afin de résoudre le problème d'inférence au niveau des réseaux Δ , nous avons établi une équivalence syntactique entre ce réseau garanti et un réseau possibiliste qualitatif. Ceci nous a permis de ré-utiliser les algorithmes de propagation existants pour les réseaux garantis et éviter ainsi la définition d'une nouvelle méthode de propagation pour ces réseaux. Le second volet concerne la modélisation des préférences utilisateurs. Nous avons montré que les réseaux possibilistes garantis peuvent être un format adéquat pour représenter et ordonner les préférences utilisateur. Une comparaison entre un modèle de représentation des préférences à base des réseaux possibilistes garantis et deux autres modèles récents à savoir les CP-Nets et la logique *QCL* a permis de confirmer les apports et les avantages du modèle possibiliste garanti.

Conclusion générale

Les travaux présentés dans cette thèse se situent dans le cadre de la représentation, le traitement et l'interprétation de l'incertitude en utilisant des modèles graphiques. Nous nous sommes intéressés, plus précisément, d'une part, au problème d'inférence dans les réseaux possibilistes et d'autre part, à l'utilisation de la mesure de possibilité garantie dans ces réseaux.

Dans un premier temps, nous avons vu que les modèles graphiques (particulièrement les réseaux probabilistes et les possibilistes) sont des outils puissants pour la représentation et le raisonnement avec des connaissances incertaines et complexes. Par leur composante graphique, ces modèles offrent un environnement intuitif et modulaire pour la représentation des relations de dépendance entre les variables constituant le réseau. À travers leur composante numérique, ils expriment l'incertitude dans les données modélisées permettant ainsi de faire des raisonnements, d'appliquer l'inférence, d'extraire des connaissances ou de réaliser une fusion des données.

Ensuite, nous avons largement abordé le problème d'inférence, notion cruciale pour la propagation de l'évidence, dans les réseaux probabilistes et possibilistes. L'état de l'art présenté a montré que la plupart des algorithmes de propagation exacts et approximatifs sont considérés comme des problèmes NP-difficiles. Dans le cas où ces réseaux admettent des structures complexes, l'inférence exacte s'avère inapplicable et les algorithmes approximatifs apparaissent comme des alternatives aux algorithmes exacts.

Dans cette thèse, nous avons proposé une adaptation de l'algorithme approximatif *Loopy Belief Propagation* aux réseaux possibilistes quantitatifs et qualitatifs. Cette proposition se justifie du fait que la plupart des algorithmes de propagation possibilistes existants sont des adaptations directes des versions probabilistes. L'implémentation de l'algorithme adapté s'est réalisée sous *Bayes Net Toolbox*, une bibliothèque open source de fonctions Matlab. Les expérimentations, de notre adaptation, réalisées sur différentes structures graphiques quantitatives et qualitatives montrent que lorsque LBP possibiliste converge, il fournit des bonnes approxima-

tions des valeurs calculées par un algorithme exact.

Plus précisément, LBP possibiliste appliqué sur des réseaux qualitatifs de structures de type QMR-DT converge lentement mais avec des taux de convergence dépassant les 90%. quant aux valeurs générées sur ce type de réseau, l'expérimentation montre que la majorité des exécutions fournissent plus de 80% des valeurs exactes. ces résultats sont légèrement moins performants sur les réseaux aléatoires et de type QMR-DT quantitatifs.

Comme deuxième contribution, nous avons proposé la définition des réseaux possibilistes *garantis* basés sur la mesure Δ déjà utilisée pour définir les bases de connaissances garanties. Ces bases ont servi pour la représentation logique des préférences utilisateur. Le réseau possibiliste garanti ne diffère des autres réseaux possibilistes standards que par sa composante numérique. En effet, l'incertitude au niveau de chaque nœud du réseau garanti est quantifiée par la mesure Δ au lieu de la mesure de possibilité Π .

Nous avons montré que ces réseaux possibilistes garantis sont capables de modéliser et d'ordonner les préférences utilisateur. Une comparaison entre le modèle basé sur ces réseaux garantis et les modèles existants (CP-Nets et la logique *QCL*) a permis de confirmer les apports et les avantages du modèle possibiliste garanti.

Par ailleurs et afin de ré-utiliser les algorithmes de propagation existants, nous avons établi une équivalence syntactique linéaire entre les réseaux possibilistes basés sur Δ et les réseaux possibilistes qualitatifs.

Parmi les perspectives les plus intéressantes, on notera la prise en compte des notions de causalité et d'intervention [BS08b] dans les réseaux possibilistes garantis comme celles utilisées dans les réseaux possibilistes standards et hybrides [BS07b]. Les liens dans un réseau causal possibiliste garanti représenteront non seulement des relations de dépendance mais aussi des relations causales. La notion d'intervention, qui facilite l'identification des relations causales entre les différents éléments d'un système, sera prise en compte en introduisant l'opérateur *do* dans les réseaux causaux possibilistes garantis de la même manière que dans les réseaux possibilistes hybrides.

La deuxième perspective de ce travail concerne la mise en place d'un outil graphique, interactif et développé sous Matlab qui :

- Offre la possibilité aux utilisateurs de créer la composante graphique des réseaux possibilistes standards, hybrides et garantis, puis d'associer des distributions de possibilités conditionnelles initiales aux nœuds,
- Permet le choix d'un algorithme de propagation possibiliste (exact ou approxi-

- matif) pour qu'il soit appliqué sur le réseau crée,
- Donne la possibilité de comparer, visuellement, les résultats des différentes inférences de différents points de vue : temps d'exécution, exactitude des valeurs calculées, étude de convergence...

Bibliographie

- [ABOY08] A. Ajroud, S. Benferhat, M. N. Omri, and H. Youssef. An approximate propagation algorithm for product-based possibilistic networks. In *Tenth International Conference on Enterprise Information Systems (ICEIS'08)*, pages 321–326, Barcelona, Spain, 2008.
- [ABOY09] A. Ajroud, S. Benferhat, M. N. Omri, and H. Youssef. On the use of guaranteed possibility measures in possibilistic networks. In *Twenty-Second International Florida Artificial Intelligence Research Society Conference (Flairs'09)*, pages 517–522. AAAI-Press, 2009.
- [AOBY07] A. Ajroud, M. N. Omri, S. Benferhat, and H. Youssef. Loopy belief propagation in bayesian networks : origin and possibilistic perspectives. In *International Conference on Computing and e-Systems (TIGERA'07)*, 2007.
- [AOBY08] A. Ajroud, M. N. Omri, S. Benferhat, and H. Youssef. A note regarding ‘loopy belief propagation’ convergence in probabilistic and possibilistic networks. In *International Conference on Computing and e-Systems (TIGERA'08)*, pages 1–9, Hammamet, Tunisia, 2008.
- [AOBY09] A. Ajroud, M. N. Omri, S. Benferhat, and H. Youssef. Review of some uncertain reasoning methods. In *5th International Symposium on Distance Education : Technology, Pedagogy, Web and Networks*, pages 70–75, Hammam Sousse, Tunisia, 2009.
- [APBM10] L. Apffel, G. Palma, I. Bloch, and S. Muller. Fuzzy segmentation of masses in digital breast tomosynthesis images based on dynamic programming. In *2nd International Conference on Imaging Theory and Applications*, pages 7–13, 2010.
- [BA02] N. Ben Amor. *Qualitative possibilistic graphical models : from independence to propagation algorithms*. PhD thesis, Université de Tunis, 2002.

-
- [BAB05] N. Ben Amor and S. Benferhat. Graphoid properties of qualitative possibilistic independence. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 13(1) :59–96, 2005.
 - [BABD⁺00] N. Ben Amor, S. Benferhat, D. Dubois, H. Geffner, and Prade H. Independence in qualitative uncertainty frameworks. In *Proc.7th Int.Conf.on Principles of Knowledge Representation and Reasoning (KR'2000)*, pages 235–246, San Francisco, CA, USA, 2000. Morgan Kaufmann Publishers, Inc.
 - [BABD⁺02] N. Ben Amor, S. Benferhat, D. Dubois, K. Mellouli, and H. Prade. A theoretical framework for possibilistic independence in weakly ordered setting. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(2) :117–155, 2002.
 - [BABK03] N. Ben Amor, S. Benferhat, and M. Khaled. Anytime propagation algorithm for min-based possibilistic graphs. *Soft Comput.*, 8(2) :150–161, 2003.
 - [BABM00] N. Ben Amor, S. Benferhat, and K. Mellouli. Possibilistic independence vs qualitative independence. In *International Conference on Uncertainty Processing and Management of Uncertainty (IM-PU'2000)*, pages 349–356, Madrid, Spain, 2000.
 - [BABM01a] N. Ben Amor, S. Benferhat, and K. Mellouli. Some issues on the coherence of min-based possibilistic causal networks. In *Proc.Joint 9th IFSA World Congress and 20th NAFIPS International Conference*, pages 604–609, Vancouver, Canada, juillet 2001.
 - [BABM01b] N. Ben Amor, S. Benferhat, and K. Mellouli. A two-steps algorithm for min-based possibilistic causal networks. In *Proceedings of the 6th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'01)*, pages 266–277, Toulouse, France, 2001. Springer.
 - [BABM02] N. Ben Amor, S. Benferhat, and K. Mellouli. Qualitative possibilistic independence based on plausibility relations. In B. Bouchon-Meunier, J. Gutiérrez-Rios, L. Magdalena, and Ronald R. Yager, editors, *Technologies for Constructing Intelligent Systems*, volume 90, pages 31–44. Warsaw, Poland, 2002.
 - [BABMS07] N. Ben Amor, S. Benferhat, K. Mellouli, and S. Smaoui. Inférence dans les réseaux possibilistes basés sur le conditionnement ordinal. *Modèles graphiques non probabilistes(RSTI- RIA)*, pages 489–519, 2007.

-
- [Bal79] J. F. Baldwin. A new approach to approximate reasoning using a fuzzy logic. *Fuzzy Sets and Systems*, 2 :309–325, 1979.
 - [BBD04] G. Brewka, S. Benferhat, and L. B. Daniel. Qualitative Choice Logic. *Artificial Intelligence Journal(AIJ)*, 157(1-2) :203–237, 2004.
 - [BBD05] A. Brini, M. Boughanem, and D. Dubois. A model for information retrieval based on possibilistic networks. In *String Processing and Information Retrieval (SPIRE 2005)*, pages 271–282. LNCS, Springer Verlag, 2005.
 - [BBD09] M. Boughanem, A. Brini, and D. Dubois. Possibilistic networks for information retrieval. *International Journal of Approximate Reasoning*, 50 :957–968, 2009.
 - [BBHP99] C. Boutilier, R. I. Brafman, H. H. Hoos, and D. Poole. Reasoning with Conditional Ceteris Paribus Preference Statements. In *Proceedings of the 15th Conference on Uncertainty in Artificial Intelligence (UAI’99)*, pages 71–80, 1999.
 - [BDGP99] S. Benferhat, D. Dubois, L. Garcia, and H. Prade. Possibilistic logic bases and possibilistic graphs. In *Proceedings of the 15th Conference on Uncertainty in Artificial Intelligence (UAI’99)*, pages 57–64, Stockholm, Sweden, 1999.
 - [BDGP02] 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(21) :35–173, 2002.
 - [BDKP02] S. Benferhat, D. Dubois, S. Kaci, and H. Prade. Bipolar representation and fusion of preferences in the possibilistic logic framework. In *Proceedings of 8 th International Conference on Principles of Knowledge Representation and Reasoning KR’2002*, pages 421–432, Toulouse, France, 2002. Morgan Kaufmann Publishers, Inc.
 - [BDKP06] S. Benferhat, D. Dubois, S. Kaci, and H. Prade. Bipolar possibility theory in preference modeling : Representation, fusion and optimal solutions. *Information Fusion*, 7 :135–150, 2006.
 - [BDP97] S. Benferhat, D. Dubois, and H. Prade. *From semantic to syntactic approaches to Information combination in possibilistic logic*, pages 141–151. Physica Verlag, 1997.
 - [BDP01] S. Benferhat, D. Dubois, and H. Prade. Towards a possibilistic logic handling of preferences. *Applied Intelligence*, 14 :303–317, 2001.

-
- [Bea03] M.J. Beal. *Variational Algorithms for Approximate Bayesian Inference*. PhD thesis, University College London, 2003.
 - [Ben06] S. Benferhat. Merging possibilistic networks. In *17th European Conference on Artificial Intelligence (ECAI'06)*, pages 332–336, Riva del Garda, Italy, 2006. IOS Press.
 - [BG02] L. Biacino and G. Gerla. Fuzzy logic, continuity and effectiveness. *Archive for Mathematical Logic*, 41(7), 2002.
 - [BGeRK00] C. Borgelt, J. Gebhardt, and et R. Kruse. *Possibilistic Graphical Models*, pages 51–68. CISM Courses and Lectures 408, 2000.
 - [BGK98] C. Borgelt, J. Gebhardt, and R. Kruse. Possibilistic graphical models. In *Proceedings of International School for the Synthesis of Expert Knowledge (ISSEK'98)*, pages 51–68, Udine, Italy, 1998.
 - [BGT93] C. Berrou, A. Glavieux, and P. Thitimajshima. Near shannon limit error-correcting coding and decoding : Turbo codes. In *Proc. 1993 IEEE International Conference on Communications*, pages 1064–1070, Geneva, Switzerland, 1993.
 - [Bis06] C.M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.
 - [BK01] S. Benferhat and S. Kaci. A possibilistic logic handling of strong preferences. In *Proceedings Joint 9th IFSA Congress and 20th NA-FIPS International Conference*, pages 962–967, Vancouver, Canada, 2001.
 - [BK02] C. Borgelt and R. Kruse. *Graphical Models : Methods for Data Analysis and Mining*. New York, NY, USA, 2002.
 - [BK03] S. Benferhat and S. Kaci. Logical representation and fusion of prioritized information based on guaranteed possibility measures. *Artificial Intelligence*, 148 :291–333, 2003.
 - [BKZ96] M. Berthod, Y. Kato, and J. Zerubia. Bayesian image classification using markov random fields. *Image and Vision Computing*, 14(4) :285–295, 1996.
 - [BL95] J.-M. Boucher and P. Lena. Unsupervised bayesian classification, application to the forest of paimpont (brittany). *Photo Interprétation*, 32(1994/4, 1995/1-2) :79–81, 1995.
 - [BMCM02] B. Bouchon-Meunier, G. Coletti, and C. Marsala. Independence and possibilistic conditioning. *Annals of Mathematics and Artificial Intelligence*, 35(1-4) :107–123, 2002.

-
- [BNG04] O. Ben Naceur and C. Gonzales. Une unification des algorithmes d'inférence de Pearl et de Jensen. *Revue d'Intelligence Artificielle*, 18(2) :229–260, 2004.
 - [Bol02] B. Bollobás. *Modern Graph Theory*. Graduate Texts in Mathematics. Springer-Verlag New York Inc, 2002.
 - [BP00] L. Bruzzone and D. F. Prieto. Automatic analysis of the difference image for unsupervised change detection. *IEEE Trans. on Geoscience and Remote Sensing*, 38(3) :1171–1182, 2000.
 - [BP06] S. Benferhat and H. Prade. Compiling possibilistic knowledge bases. In Anna Perini Paolo Traverso Gerhard Brewka, Silvia Co-radeschi, editor, *17th European Conference on Artificial Intelligence (ECAI'06)*, pages 337–341, Riva del Garda, Italy, 2006. IOS Press.
 - [BS07a] S. Benferhat and K. Sedki. A revised qualitative choice logic for handling prioritized preferences. In *Ninth European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'07)*, pages 635–647, 2007.
 - [BS07b] S. Benferhat and S. Smaoui. Hybrid possibilistic networks. *International Journal of Approximate Reasoning*, 44(3) :224–243, 2007.
 - [BS08a] S. Benferhat and K. Sedki. Two alternatives for handling preferences in qualitative choice logic. *Fuzzy Sets and Systems Journal*, 159(15) :1889–1912, 2008.
 - [BS08b] S. Benferhat and S. Smaoui. Quantitative possibilistic networks : Handling interventions and ascribing causality. In *7th Mexican International Conference on Artificial Intelligence (MICAI'08)*, pages 720–731, Mexico, 2008. Springer, LNCS.
 - [BSCC89] I. A. Beinlich, H. J. Suermondt, R. M. Chavez, and G. F. Cooper. The ALARM Monitoring System : A Case Study with Two Probabilistic Inference Techniques for Belief Networks. In *Second European Conference on Artificial Intelligence in Medicine*, pages 247–256. Springer-Verlag, Berlin, 1989.
 - [BT08a] S. Benferhat and K. Tabia. Classification features for detecting server-side and client-side web attacks. In *23rd International Security Conference (SEC'08)*, pages 729–733, Milan, Italy, 2008. Springer LNCS.
 - [BT08b] S. Benferhat and K. Tabia. New schemes for anomaly score aggregation and thresholding. In *International Conference on Security and*

- Cryptography (SECRYPT'08)*, pages 21–28, Porto, Portugal, 2008. Springer.
- [BTN07] S. Barrat, S. Tabbone, and P. Nourrissier. A Bayesian classifier for symbol recognition. In *Seventh International Workshop on Graphics Recognition (GREC'07)*, Curitiba, Brésil, 2007.
- [CD00] J. Cheng and M. J. Druzdzel. Ais-bn : An adaptive importance sampling algorithm for evidential reasoning in large bayesian networks. *Journal of Artificial Intelligence Research (JAIR)*, 13 :155–188, 2000.
- [CD03] Y. Changhe and M. J. Druzdzel. An importance sampling algorithm based on evidence pre-propagation. In *In Proceedings of the Nineteenth Annual Conference on Uncertainty in Artificial Intelligence*, pages 624–631. Morgan Kaufmann Publishers, 2003.
- [CDLS03] R. Cowell, A. Dawid, S. Lauritzen, and D. Spiegelhalter. *Probabilistic Networks and Expert Systems*. Statistics for Engineering and Information Science. Springer, 2003.
- [CGH96a] E. Castillo, J. M. Gutiérrez, and A. S. Hadi. A new method for symbolic inference in bayesian networks. *Networks*, 28 :31–43, 1996.
- [CGH96b] E. Castillo, J. M. Gutiérrez, and Ali S. Hadi. *Expert Systems and Probabilistic Network Models*. Springer, New York, NY, USA, erste edition, 1996.
- [Cha06] Y. Changhe. *Importance Sampling for Bayesian Networks : Principles, Algorithms, and Performance*. PhD thesis, University of Pittsburgh, 2006.
- [CMT05] O. Cappé, E. Moulines, and R. Tobias. *Inference in Hidden Markov Models*. Springer, 2005.
- [Coo90] G. F. Cooper. The computational complexity of probabilistic inference using bayesian belief networks (research note). *Artif. Intell.*, 42(2-3) :393–405, 1990.
- [Coo99] G.F. Cooper. An overview of the representation and discovery of causal relationships using bayesian networks. In C. Glymour and G.F. Cooper, editors, *Computation, Causation, and Discovery*, pages 3–62. AAAI Press and MIT Press, 1999.
- [D'A93] B. D'Ambrosio. Incremental probabilistic inference. In *Proceedings of the Proceedings of the Ninth Conference Annual Conference on Uncertainty in Artificial Intelligence (UAI-93)*, pages 301–308, San Francisco, CA, USA, 1993. Morgan Kaufmann.

-
- [Dar00] A. Darwiche. A differential approach to inference in bayesian networks. In *Proceedings of Uncertainty in Artificial Intelligence*, pages 123–132, 2000.
 - [DC97] G. De Cooman. Possibility theory III : Possibilistic independence. *International Journal of General Systems*, 25(4) :353–371, 1997.
 - [dCFLH03] L. M. de Campos, J. M. Fernández-Luna, and J. F. Huete. The BNR model : foundations and performance of a bayesian network-based retrieval model. *International Journal of Approximate Reasoning*, 34 :265–285, 2003.
 - [dCH99a] L. M. de Campos and Juan F. Huete. Independence concepts in possibility theory : part I. *Fuzzy Sets Systems*, 103(1) :127–152, 1999.
 - [dCH99b] L. M. de Campos and Juan F. Huete. Independence concepts in possibility theory : part II. *Fuzzy Sets Systems*, 103(3) :487–505, 1999.
 - [DCK94] G. De Cooman and E. E. Kerre. A new approach to possibilistic independence. In *Proceedings of the Third IEEE International Conference on Fuzzy Systems*, pages 1446–1451, 1994.
 - [Dec98] R. Dechter. Bucket elimination : a unifying framework for structure-driven inference. *Artificial Intelligence*, 1998.
 - [Dem68] A.P. Dempster. A generalization of bayesian inference. *Journal of the Royal Statistical Society*, 30 :205–247, 1968.
 - [DFnDCHP94] D. Dubois, L. Fariñas Del Cerro, A. Herzig, and H. Prade. An ordinal view of independence, plausible reasoning and belief revision. In Ramon Lopez de Mantaras and David Poole, editors, *Proc. of the 10th Conference on Uncertainty in Artificial Intelligence (UAI'94)*, pages 195–203, Seattle, USA, 1994. Morgan Kaufmann Publishers.
 - [DGPZ98] D. Dubois, L. Godo, H. Prade, and A. Zapico. Making decision in a qualitative setting : from decision under uncertainty to case-based decision. In *6th Int. Conf. Principles of Knowledge Representation and Reasoning*, pages 1924–1930, San Francisco, CA, USA, 1998.
 - [DHP00] D. Dubois, P. Hájek, and H. Prade. Knowledge-driven versus data-driven logics. *J. of Logic, Lang. and Inf.*, 9(1) :65–89, 2000.
 - [DL93] P. Dagum and M. Luby. Approximating probabilistic inference in bayesian belief networks is NP-hard. *Artif. Intell.*, 60(1) :141–153, 1993.

-
- [DLN⁺90] F. M. Donini, M. Lenzerini, D. Nardi, F. Pirri, and M. Schaerf. Non-monotonic reasoning. *Artificial Intelligence Review* 4, pages 163–210, 1990.
 - [DLP94] D. Dubois, J. Lang, and H. Prade. Possibilistic logic. In *Handbook on Logic in Artificial Intelligence and Logic Programming*, pages 439–513. Oxford University press, 1994.
 - [DM83] T. Dietterich and R. Michalski. *A comparative review of selected methods of learning from examples*, pages 41–81. Morgan Kaufmann, R. michalski, J. carbonell, and T. mitchell edition, 1983.
 - [DMP98] D. Dubois, S. Moral, and H. Prade. Belief change rules in ordinal and numerical uncertainty theories. In D. Dubois and H. Prade, editors, *Belief Change*, pages 311–392. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.
 - [DP80] D. Dubois and H. Prade. *Fuzzy sets and systems : theory and applications*. Academic Press, New York, NY, USA, 1980.
 - [DP88] D. Dubois and H. Prade. *Possibility theory : An approach to computerized, Processing of uncertainty*. New York, NY, USA, 1988.
 - [DP91] D. Dubois and H. Prade. Epistemic entrenchment and possibilistic logic. *Artificial Intelligence*, 50(2) :223–239, 1991.
 - [DP92a] D. Dubois and H. Prade. Combination of fuzzy information in the framework of possibility theory. In R.C. Abidi and M.A. Gonzalez, editors, *Data Fusion in Robotics and Machine Intelligence*, pages 481–505. Academic Press, 1992. DP189.
 - [DP92b] D. Dubois and H. Prade. Possibility theory as a basis for preference propagation in automated reasoning. In *Proc. of the 1st IEEE Inter. Conf. on Fuzzy Systems 1992 (FUZZ-IEEE'92), San Diego, Ca.*, pages 821–832, San Diego, CA, USA, 1992.
 - [DP95] D. Dubois and H. Prade. Possibility theory as a basis for qualitative decision theory. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, pages 1924–1930, San Francisco, CA, USA, 1995. Morgan Kaufmann.
 - [DP98] D. Dubois and H. Prade. Possibility theory : qualitative and quantitative aspects. In *Quantified Representation of Uncertainty and Imprecision. D. M. Gabbay and P. Smets P., editors*, volume 1, pages 169–226. KLUWER ACADEMIC PUBLISHERS, The Netherlands, 1998. Handbook of Defeasible Reasoning and Uncertainty Management Systems.

-
- [DP04a] D. Dubois and H. Prade. On the use of aggregation operations in information fusion processes. *Fuzzy Sets and Systems*, 142 :143–161, 2004.
- [DP04b] D. Dubois and H. Prade. Possibilistic logic : a retrospective and prospective view. *Fuzzy Sets and Systems*, 144 :3–23, 2004.
- [Dur96] R. Durrett. *Probability Theory and Examples 2th edition*. Duxbury Press edition, 1996.
- [EEZA07] B. Elayeb, F. Evrard, M. Zaghdoud, and M. Ben Ahmed. SARI-POD : A Multiagent Possibilistic System For Web Information Retrieval. In *The 2007 International Conference on Information and Knowledge Engineering (IKE'07)*, pages 72–78, 2007.
- [EEZA08] B. Elayeb, F. Evrard, M. Zaghdoud, and M. Ben Ahmed. Qualitative approach for an intelligent possibilistic web information retrieval using multiagent system. In *The IEEE - 6th International Conference in Human System Learning (IEEE-ICHSL6)*, pages 2–13, 2008.
- [EEZA09] B. Elayeb, F. Evrard, M. Zaghdoud, and M. Ben Ahmed. Towards an intelligent possibilistic web information retrieval using multiagent system. *The International Journal of Interactive Technology and Smart Education (ITSE)*, 6(1) :40–59, 2009.
- [EM02] Y. Ephraim and N. Merhav. Hidden markov processes. *IEEE Transactions on Information Theory*, 48(6) :1518–1569, 2002.
- [ES95] K. Ezawa and T. Schuermann. Fraud/uncollectible debt detection using a bayesian network based learning system : A rare binary outcome with mixed data structures. In *Proceedings of the Eleventh Conference Annual Conference on Uncertainty in Artificial Intelligence (UAI-95)*, pages 157–166, San Francisco, CA, USA, 1995. Morgan Kaufmann.
- [Eva04] T. Evangelista. *Les IDS - Les Systèmes de détection d'intrusions informatiques*. Dunod, 2004.
- [Fab96] P. Fabiani. *Représentation Dynamique de l'Incertain et stratégie de Prise d'Information pour un Système Autonome en Environnement Evolutif*. PhD thesis, Ecole Nationale Supérieure de l'Aéronautique et de l'Espace, Toulouse, 1996.
- [Fay00] U. Fayyad. Knowledge discovery in databases : An overview. *Relational Data Mining*, pages 28–45, 2000.

-
- [FC89] R. Fung and K.-C. Chang. Weighing and integrating evidence for stochastic simulation in bayesian networks. In *Uncertainty in Artificial Intelligence 5*, pages 209–219, New York, NY, USA, 1989. Elsevier Science Publishing Company, Inc.
 - [FdF94] R. Fung and B. del Favero. Backward simulation in bayesian networks. In *Proceedings of the Tenth Annual Conference on Uncertainty in Artificial Intelligence (UAI-94)*, pages 227–234, San Francisco, CA, USA, 1994. Morgan Kaufmann Publishers.
 - [FDP⁺03] R. Fjortoft, Y. Delignon, W. Pieczynski, M. Sigelle, and F. Tupin. Unsupervised classification of radar images using hidden markov chains and hidden markov random fields. *IEEE Trans. on Geoscience and Remote Sensing*, 41(3) :675–686, 2003.
 - [FGG97] N. Friedman, D. Geiger, and M. Goldszmidt. Bayesian network classifiers. *Machine Learning*, 29(2–3) :131–163, 1997.
 - [FH99] N. Friedman and Joseph Y. Halpern. Belief revision : A critique. *Journal of Logic, Language and Information*, 8(4) :401–420, 1999.
 - [Fin07] G. A. Fink. *Markov Models for Pattern Recognition : From Theory to Applications*. Springer, 2007.
 - [Fon94] P. Fonck. *Réseaux d’inférence pour le raisonnement possibiliste*. PhD thesis, Université de Liège, Faculté des Sciences, 1994.
 - [Fon97] P. Fonck. A comparative study of possibilistic conditional independence and lack of interaction. *International Journal of Approximate Reasoning*, 16 :149–171, 1997.
 - [FPSM91] W.J. Frawley, G. Piatetsky-Shapiro, and C. Matheus. Knowledge discovery in databases : An overview. In G. Piatetsky-Shapiro and W. J. Frawley, editors, *Knowledge Discovery In Databases*, pages 1–30. AAAI Press and MIT Press, Cambridge, MA, USA, 1991.
 - [Fri87] M. Frisse. Searching for information in a hypertext medical handbook. In *Proceedings of the ACM conference on Hypertext (HYPERTEXT ’87)*, pages 57–66, 1987.
 - [GG84] S. Geman and D. Geman. Stochastic relaxations, gibbs distributions and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6) :721–742, 1984.
 - [GH96] D. Geiger and D. Heckerman. Knowledge representation and inference in similarity networks and bayesian multinets. *Artificial Intelligence*, 82(1-2) :45–74, 1996.

- [GH02] H. Guo and W. Hsu. A survey of algorithms for realtime bayesian network inference. In *Joint Workshop on Real-Time Decision Support and Diagnosis Systems (AAAI/KDD/UAI-2002)*, pages 1–12, 2002.
- [GK97] J. Gebhardt and R. Kruse. Background and perspectives of possibilistic graphical models. In *ECSQARU/FAPR'97 : Proceedings of the First International Joint Conference on Qualitative and Quantitative Practical Reasoning*, pages 108–121, London, UK, 1997. Springer-Verlag.
- [GRS96] W.R. Gilks, S. Richardson, and D.J. Spiegelhalter. *Markov Chain Monte Carlo in Practice*. Chapman and Hall/CRC, 1996.
- [GVP89] D. Geiger, T. Verma, and J. Pearl. d-separation : From theorems to algorithms. In *Proceedings of the Fifth Annual Conference on Uncertainty in Artificial Intelligence (UAI'89)*, pages 139–148, Amsterdam, The Netherlands, 1989. Elsevier Science Publishing Company, Inc.
- [GVP90] D. Geiger, T. Verma, and J. Pearl. Identifying independence in bayesian networks. *Networks*, 20 :507–534, 1990.
- [GWP88] A. Ginsberg, S.M. Weiss, and P. Politakis. Automatic knowledge base refinement for classification systems. *Artificial Intelligence*, 35(2) :197–226, 1988.
- [H95] P. Hájek. Fuzzy logic and arithmetical hierarchy. *Fuzzy Sets and Systems*, 3(8), 1995.
- [Has70] W. K. Hastings. Monte carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1) :97–109, 1970.
- [HB95] E. Horvitz and M. Barry. Display of information for time-critical decision making. In *Proceedings of the Eleventh Conference Annual Conference on Uncertainty in Artificial Intelligence (UAI-95)*, pages 296–305, San Francisco, CA, USA, 1995. Morgan Kaufmann.
- [Hen88] M. Henrion. Propagating uncertainty in bayesian networks by probabilistic logic sampling. In *Uncertainty in Artificial Intelligence 2*, pages 149–163, New York, NY, USA, 1988. Elsevier Science Publishing Company, Inc.
- [Hen90] M. Henrion. An Introduction to Algorithms for Inference in Belief Nets. In *Proceedings of the Fifth Annual Conference on Uncertainty in Artificial Intelligence (UAI'89)*, pages 129–138, 1990.

-
- [His78] E. Hisdal. Conditional possibilities independence and non interaction. *Fuzzy Sets and Systems*, 1 :283–297, 1978.
 - [HJ84] R. A. Howard and E. M. James. Influence diagrams. In R. Howard and J. Matheson, editors, *Readings on the Principles and Applications of Decision Analysis*, pages 721–762. Strategic Decision Group, Menlo Park, CA, USA, 1984.
 - [HSC89] E. Horvitz, Henri J. Suermondt, and Gregory F. Cooper. Bounded conditioning : Flexible inference for decisions under scarce resources. In *Association for Uncertainty in Artificial Intelligence*, 1989.
 - [Jen96] F. V. Jensen. *An Introduction to Bayesian Networks*. UCL press, London, UK, 1996.
 - [Jen01] F.V. Jensen. *Bayesian networks and decision graphs*. Springer-Verlag, New York, NY, USA, 2001.
 - [JJ99] T. S. Jaakkola and Michael I. Jordan. Variational probabilistic inference and the qmr-dt network. *Journal of Artificial Intelligence Research*, 10 :291–322, 1999.
 - [JLO90] F. V. Jensen, S. L. Lauritzen, and K. G. Olesen. Bayesian updating in causal probabilistic networks by local computations. *Computational Statistics Quarterly*, 5(4) :269–282, 1990.
 - [JZA07] F. Jemili, M. Zaghdoud, and M. Ben Ahmed. A framework for an adaptive intrusion detection system using bayesian network. In *Proceedings of IEEE International Conference on Intelligence and Security Informatics (ISI 2007)*, pages 66–70, 2007.
 - [JZA09] F. Jemili, M. Zaghdoud, and M. Ben Ahmed. Hybrid intrusion detection and prediction multiagent system. *International Journal of Computer Science and Information Security*, 5(1) :62–71, 2009.
 - [KF88] G. J. Klir and T. A. Folger. *Fuzzy sets, uncertainty, and information*. Prentice Hall, 1988.
 - [Kja94] U. Kjaerulff. Reduction of computational complexity in bayesian networks through removal of weak dependences. In *In Proc. Tenth Conf. on Uncertainty in Artificial Intelligence*, pages 374–382. Morgan Kaufmann, 1994.
 - [Kle84] E. P. Klement. Operations on fuzzy sets : an axiomatic approach. *Information Sciences*, 27 :221–232, 1984.
 - [Kli99] G. J. Klir. On fuzzy-set interpretation of possibility theory. *Fuzzy Sets and Systems*, 108(3) :263–273, 1999.

- [Kod95] Y. Kodratoff. Technical and scientific issues of kdd (or : Is kdd a science?). In *Proceedings of the 6th International Conference on Algorithmic Learning Theory (ALT'95)*, pages 261–265, London, UK, 1995. Springer-Verlag.
- [KP83] J. H. Kim and J. Pearl. A computational model for causal and diagnostic reasoning in inference systems. In *Proceedings of the Eighth international joint conference on Artificial intelligence (IJCAI'83)*, pages 190–193, San Francisco, CA, USA, 1983. Morgan Kaufmann Publishers Inc.
- [Lau96] S. L. Lauritzen. *Graphical Models*. Oxford : Clarendon Press, 1996.
- [LD94] Z. Li and B. D'Ambrosio. Efficient inference in bayes nets as a combinatorial optimization problem. *International Journal of Approximate Reasoning*, 11(1) :55–81, 1994.
- [LS90] S. L. Lauritzen and D. J. Spiegelhalter. Local computations with probabilities on graphical structures and their application to expert systems. *Readings in uncertain reasoning*, pages 415–448, 1990.
- [Mac03] D. J. C. MacKay. *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, 2003.
- [MBK⁺09] A. Makarenko, A. Brooks, T. Kaupp, H.F. Durrant-Whyte, and F. Dellaert. Decentralized data fusion : a graphical model approach. In *Proceedings of the 12th International Conference on Information Fusion*, pages 545–554, Seattle, WA, USA, 2009.
- [MMC98] R. J. McEliece, D. J. C. MacKay, and J. F. Cheng. Turbo decoding as an instance of pearl's belief propagation algorithm. *IEEE Journal on Selected Areas in Communications*, 16(2) :140–152, 1998.
- [MRS08] C. D. Manning, P. Raghavan, and H. Schütze. *Introduction to Information Retrieval*. Cambridge University Press, 2008.
- [Mur01] K. Murphy. The bayes net toolbox for matlab. *Computing Science and Statistics : Proceedings of Interface*, 33, 2001.
- [MWJ99] K. Murphy, Y. Weiss, and M. Jordan. Loopy belief propagation for approximate inference : An empirical study. In *Proceedings of the Fifteenth Annual Conference on Uncertainty in Artificial Intelligence (UAI-99)*, pages 467–475, San Francisco, CA, USA, 1999. Morgan Kaufmann Publishers.
- [NAAB08] O. Nempont, J. Atif, E. Angelini, and I. Bloch. A new fuzzy connectivity class application to structural recognition in images. In *14th*

- International Conference on Discrete Geometry for Computer Imagery*, pages 446–457, 2008.
- [Nov06] V. Novak. Fuzzy sets as a special mathematical model of vagueness phenomenon. In *Computational Intelligence, Theory and Applications*, pages 683–690. Heidelberg : Springer, 2006.
- [Nov09] V. Novak. Fuzzy set. In *Encyclopedia of Database Systems*, pages 1209–1210. 2009.
- [NWL⁺07] P. Naïm, P. Willemin, P. Leray, O. Pourret, and B. Anna. *Réseaux bayésiens*. 3e édition. Eyrolles edition, 2007.
- [Pea86] J. Pearl. Fusion, propagation, and structuring in belief networks. *Journal of Artificial Intelligence*, 29 :241–288, 1986.
- [Pea88] J. Pearl. *Probabilistic Reasoning in Intelligent Systems : Networks of Plausible Inference*. Morgan Kaufmann Publishers, Inc., San Mateo, CA, USA, 1988.
- [PGA87] D. Poole, R. Goebel, and R. Aleliunas. Theorist’ a logical reasoning system for defaults and diagnosis. In N.J. Cercone and G. McCalla, editors, *The Knowledge Frontier : Essays in the Representation of Knowledge*, pages 331–352, New York, NY, USA, 1987. Springer Verlag.
- [Pie03] W. Pieczynski. Markov models in image processing. *Traitement du Signal*, 20(3) :255–278, 2003.
- [Poo93] D. Poole. Average-case analysis of a search algorithm for estimating prior and posterior probabilities in bayesian networks with extreme probabilities. In *Proceeding of the thirteenth International Joint Conference on Artificial Intelligence*, pages 606–612, 1993.
- [Poo98] D. Poole. Context-specific approximation in probabilistic inference. In *In Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 447–454, 1998.
- [PS91] M. A. Peot and R. D. Shachter. Fusion and propagation with multiple observations in belief networks. *Artificial Intelligence*, 48 :299–318, 1991.
- [Rab90] L. R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. pages 267–296, 1990.
- [RJ87] G. Rebane and P. Judea. The recovery of causal poly-trees from statistical data. In *Proceedings of the Uncertainty in Artificial Intelligence 3 Annual Conference on Uncertainty in Artificial Intelligence*

- (*UAI'87*), pages 175–182, Amsterdam, The Netherlands, 1987. Elsevier Science.
- [Rub81] R. Y. Rubinstein. *Simulation and the Monte Carlo Method*. John Wiley and Sons, 1981.
- [Sap02] G. Saporta. Data fusion and data grafting. *Computational Statistics and Data Analysis*, 38(4) :465–473, 2002.
- [SDD90] R. D. Shachter, B. D'Ambrosio, and B. DelFabero. Symbolic probabilistic inference in belief networks. In *Proceedings of the Eighth National Conference on Artificial Intelligence*, pages 126–131, 1990.
- [Sed08] K. Sedki. *Raisonnement sous incertitude et en présence de préférences : Application à la détection d'intrusions et à la corrélation d'alertes*. PhD thesis, Université d'Artois, 2008.
- [Sha76] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, NJ, 1976.
- [Sha86a] R. D. Shachter. Intelligent probabilistic inference. In L. N. Kanal and J. F. Lemmer, editors, *Uncertainty in Artificial Intelligence*, pages 371–382. Amsterdam. The Netherlands, 1986.
- [Sha86b] R.D. Shachter. Evaluating influence diagrams. *Operations Research*, 34 :871–882, 1986.
- [Sha88] R.D. Shachter. Probabilistic inference and influence diagrams. *Operations Research*, 36 :589–604, 1988.
- [She97] M. Ross Sheldon. *Introduction To Probability Models 6th edition*. Elsevier, Academic Press edition, 1997.
- [Shi00] R. Shirey. *Internet Security Glossary*. 2000.
- [SK94] Ph. Smets and R. Kennes. The transferable belief model. *Artificial Intelligence*, 66(2) :191–234, 1994.
- [Sma07] S. Smaoui. *Réseaux possibilistes hybrides : Représentation des interventions et algorithmes*. PhD thesis, Université d'Artois, 2007.
- [SMH⁺91] M. Shwe, B. Middleton, D. Heckerman, M. Henrion, E. Horvitz, H. Lehmann, and G. Cooper. Probabilistic diagnosis using a reformulation of the internist- 1/qmr knowledge base i. the probabilistic model and inference algorithms. *Methods of Information in Medicine*, 30 :241–255, 1991.
- [SP89] R. D. Shachter and M.A. Peot. Simulation approaches to general probabilistic inference on belief networks. In *Uncertainty in Artificial*

- Intelligence 5*, pages 221–231, New York, NY, USA, 1989. Elsevier Science Publishing Company, Inc.
- [Spo88a] W. Spohn. A general non-probabilistic theory of inductive reasoning. In *Proceedings of the Fourth Annual Conference on Uncertainty in Artificial Intelligence (UAI'88)*, pages 149–158, New York, NY, USA, 1988. Elsevier Science Publishing Company, Inc.
- [Spo88b] W. Spohn. Ordinal conditional functions : A dynamic theory of epistemic states. In W. L. Harper and B. Skyrms, editors, *Causation in Decision, Belief Change and Statistics*, pages 105–134. 1988.
- [Sri02] R. Srinivasan. *Importance sampling - Applications in communications and detection*. Springer-Verlag, 2002.
- [SRNC⁺00] I. Silva, B. A. Ribeiro-Neto, P. Calado, E. de Moura, and N. Ziviani. Link-based and content-based evidential information in a belief network model. In *Proc. of the International ACM-SIGIR Conference*, pages 96–103, 2000.
- [TC89] H. Turtle and W. Croft. Inference networks for document retrieval. In *Proceedings of the 13th annual international ACM SIGIR Conference*, pages 1–24, 1989.
- [TIT03] K. Tanaka, J. Inoue, and D. M. Titterington. Loopy belief propagation and probabilistic image processing. In *XIII Proceedings of the 2003 IEEE Signal Processing Society Workshop*, pages 329–338. IEEE Computer Society Press, 2003.
- [Tor03] V. Torra. *Information fusion in data mining*, volume 123 of *Studies in fuzziness and soft computing*. Springer, 2003.
- [TS92] K. Thulasiraman and M. N. S. Swamy. *Graphs : Theory and Algorithms*. Wiley-Interscience, 1992.
- [Van03] B. Vantaggi. Conditional independence structures and graphical models. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 11(5) :545–571, 2003.
- [Wei00] Y. Weiss. Correctness of local probability propagation in graphical models with loops. *Neural Computation*, 12(1) :1–41, 2000.
- [WIL94] M. P. Wellman and Chao lin Liu. State-space abstraction for anytime evaluation of probabilistic networks. In *Proceedings of the Tenth Conference on Uncertainty in Artificial Intelligence*, pages 567–574. Morgan Kaufmann, 1994.

- [YFW03] J. S. Yedidia, William T. Freeman, and Y. Weiss. Understanding belief propagation and its generalizations. *Exploring artificial intelligence in the new millennium*, pages 239–269, 2003.
- [YPSPL05] C. Yuguo, D. Persi, Holmes Susan P., and Jun S. Liu. Sequential monte carlo methods for statistical analysis of tables. *Journal of the American Statistical Association*, 100 :109–120, 2005.
- [Zad65] L. A. Zadeh. Fuzzy sets. *Information and Control*, 8 :338–353, 1965.
- [Zad75] L. A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning - I. *Information Sciences*, 8(3) :199–249, 1975.
- [Zad78] L. A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1 :3–28, 1978.
- [Zha04] H. Zhang. The optimality of naïve bayes. In *17th International Florida Artificial Intelligence Research Society Conference (Flairs'04)*, Florida, USA, 2004.
- [ZP94] N. Zhang and D. Poole. A simple approach to bayesian network computations. In *Proceedings of the tenth Canadian Conference on Artificial Intelligence*, pages 171–178, 1994.
- [ZR00] D.A. Zighed and R. Rakotomalala. *Graphes d'Induction - Apprentissage et Data Mining*. Hermès, 2000.

Annexe A

Propriétés de la logique possibiliste

La logique possibiliste se caractérise par un ensemble de propriétés dont nous citons ce qui est en relation directe avec notre travail.

Une distribution de possibilités $\pi_{\mathcal{B}}$ n'est pas toujours normalisée. Il existe deux degrés complémentaires représentant, respectivement, la cohérence et l'incohérence d'une base possibiliste \mathcal{B} :

- Le degré de cohérence de \mathcal{B} , noté $Cons(\mathcal{B})$, indique à quel degré une interprétation est possible dans \mathcal{B} :

$$Cons(\mathcal{B}) = \max_{\omega \in \Omega} \pi_{\mathcal{B}}(\omega).$$

- Le degré d'incohérence de \mathcal{B} , noté $Inc(\mathcal{B})$, est le complément du degré de cohérence $Cons(\mathcal{B})$ qui peut être défini comme suit :

$$Inc(\mathcal{B}) = \max\{\alpha_i : \mathcal{B}_{\geq \alpha_i} \models \perp\},$$

avec $\mathcal{B}_{\geq \alpha_i}$ correspond à l'ensemble des formules possibilistes ayant un poids supérieur ou égal à α_i dans \mathcal{B} et \perp dénote les contradictions.

$Inc(\mathcal{B}) = 0$ signifie que \mathcal{B} est cohérente, par contre une valeur de $Inc(\mathcal{B}) = 1$ signifie que \mathcal{B} est incohérente. $0 < Inc(\mathcal{B}) < 1$ signifie que \mathcal{B} est partiellement cohérente.

Exemple 1.11 Soient \mathcal{B}_1 , \mathcal{B}_2 et \mathcal{B}_3 trois bases de connaissances possibilistes définies comme suit :

- $\mathcal{B}_1 = \{(a \vee \neg b, 0.7), (\neg a \vee b, 0.5), (a \vee b, 0.4)\}$,
- $\mathcal{B}_2 = \{(a, 0.4), (b, 0.5), (\neg a \vee \neg b, 0.8)\}$,
- $\mathcal{B}_3 = \{(a \wedge \neg b, 0.6), (\neg a \wedge b, 0.5), (\neg a \wedge \neg b, 0.4)\}$.

Le tableau 1.8 donne les distributions de possibilités associées à chacune des bases \mathcal{B}_1 , \mathcal{B}_2 et \mathcal{B}_3 , ainsi que leur degré de cohérence et d'incohérence.

D'après les valeurs des degrés d'incohérence, nous pouvons dire que \mathcal{B}_1 est cohérente, \mathcal{B}_2 est partiellement cohérente tandis que \mathcal{B}_3 est incohérente.

Définition 1.1 Deux bases de connaissances possibilistes \mathcal{B} et \mathcal{B}' sont dites équivalentes si et seulement si elles ont les mêmes distributions de possibilités associées.

| A | B | $\pi_{\mathcal{B}_1}$ | $\pi_{\mathcal{B}_2}$ | $\pi_{\mathcal{B}_3}$ |
|-----------------------|----------|-----------------------|-----------------------|-----------------------|
| $\neg a$ | $\neg b$ | 1 | 0.2 | 0 |
| $\neg a$ | b | 0.5 | 0.5 | 0 |
| a | $\neg b$ | 0.3 | 0.6 | 0 |
| a | b | 0.6 | 0.5 | 0 |
| $Cons(\mathcal{B}_i)$ | | 1 | 0.6 | 0 |
| $Inc(\mathcal{B}_i)$ | | 0 | 0.4 | 1 |

TAB. 14 – Distributions de possibilités associées aux bases \mathcal{B}_1 , \mathcal{B}_2 et \mathcal{B}_3

Formellement :

\mathcal{B} et \mathcal{B}' sont équivalentes ssi $\forall \omega \in \Omega, \pi_{\mathcal{B}}(\omega) = \pi_{\mathcal{B}'}(\omega)$.

Définition 1.2 Une distribution de possibilités π sur Ω satisfait une formule possibiliste (ϕ, α) si et seulement si la mesure de nécessité induite par π satisfait $N(\phi) \geq \alpha$.

Exemple 1.12 Soit $V = \{A, B, C\}$ un ensemble de trois variables binaires. La distribution de possibilités π associée à l'univers Ω est donnée par la table 1.9.

| A | B | C | π |
|----------|----------|----------|-------|
| $\neg a$ | $\neg b$ | $\neg c$ | 0.3 |
| $\neg a$ | $\neg b$ | c | 0.5 |
| $\neg a$ | b | $\neg c$ | 0.2 |
| $\neg a$ | b | c | 0.7 |
| a | $\neg b$ | $\neg c$ | 0.4 |
| a | $\neg b$ | c | 0.5 |
| a | b | $\neg c$ | 0.6 |
| a | b | c | 0.1 |

TAB. 15 – Distribution de possibilités π associée à Ω

Soit la formule possibiliste $(\neg a \wedge \neg b, 0.4)$. Pour que la distribution de possibilités π satisfasse cette formule, il faut que le degré de nécessité $N(\neg a \wedge \neg b)$, déterminé à partir de π , soit supérieur ou égal à 0.4. Ainsi en calculant :

$$N(\neg a \wedge \neg b) = \min(\Pi(\neg a), \Pi(\neg b)) = \min(0.7, 0.5) = 0.5 \geq 0.4.$$

Nous concluons que la distribution π satisfait la formule $\neg a \wedge \neg b$.

Définition 1.3 Une distribution de possibilités π sur Ω satisfait une base de connaissances possibiliste \mathcal{B} si et seulement si elle satisfait toutes les formules possibilistes

dans cette base.

La subsomption est une propriété qui permet de simplifier les bases de connaissances possibilistes en supprimant (ou ajoutant) quelques formules sans perte d'information. La subsomption est définie comme suit :

Définition 1.4 Soit (ϕ, α) une formule possibiliste dans \mathcal{B} . (ϕ, α) est subsumée par \mathcal{B} si les bases possibilistes \mathcal{B} et $\mathcal{B} \setminus \{(\phi, \alpha)\}$ sont équivalentes.

Les formules subsumées représentent des formules qui peuvent être supprimées ou ajoutées d'une base sans modifier la distribution des possibilités qui lui est associée.

Exemple 1.13 Soit la base possibiliste $\mathcal{B} = \{(a \vee \neg b, 0.7), (\neg a \vee b, 0.5), (a \vee b, 0.4), (b, 0.3)\}$. La formule $(b, 0.3)$ est subsumée de \mathcal{B} . Ce qui signifie que \mathcal{B} et $\mathcal{B} \setminus \{(b, 0.3)\}$ sont équivalentes et produisent la même distribution de possibilités. En effet le tableau 1.10 montre que cette formule n'a pas d'influence directe sur les valeurs de la distribution de possibilités associée à \mathcal{B} .

| ω | $\omega \models \phi_1$ | $\omega \models \phi_2$ | $\omega \models \phi_3$ | $\omega \models \phi_4$ | $\pi_{\mathcal{B}}(\omega)$ |
|----------------|-------------------------|-------------------------|-------------------------|-------------------------|------------------------------|
| ab | <i>oui</i> | <i>oui</i> | <i>oui</i> | <i>oui</i> | 1 |
| $a\neg b$ | <i>oui</i> | <i>non</i> | <i>oui</i> | <i>non</i> | $1 - \max\{0.5, 0.3\} = 0.5$ |
| $\neg ab$ | <i>non</i> | <i>oui</i> | <i>oui</i> | <i>oui</i> | $1 - \max\{0.7\} = 0.3$ |
| $\neg a\neg b$ | <i>oui</i> | <i>oui</i> | <i>non</i> | <i>non</i> | $1 - \max\{0.4, 0.3\} = 0.6$ |

TAB. 16 – Distribution de possibilités $\pi_{\mathcal{B}}$ obtenue à partir de la base \mathcal{B}

La fusion des connaissances possibilistes

Les principaux opérateurs de fusion dans le cadre possibiliste sont les conjonctions et les disjonctions. La fusion par conjonction est caractérisée par l'utilisation de l'opérateur *minimum* et est appropriée lorsque les informations disponibles sont de nature qualitatives. L'opérateur *produit* remplace le *minimum* dans le cas du raisonnement quantitatif. La disjonction idempotente se caractérise par l'utilisation de l'opérateur *maximum*.

La fusion des distributions de possibilités [DP04a, DP92a] consiste à combiner des informations incertaines multi-sources représentées par n distributions de possibilité $\{\pi_1, \dots, \pi_n\}$ en une nouvelle distribution de possibilités π^{\oplus} . Selon l'opérateur mis en jeu, la nouvelle distribution sera définie par :

$$- \forall \omega, \pi^{\min}(\omega) = \min_{i=1..n} \pi_i(\omega).$$

- $\forall \omega, \pi^*(\omega) = \prod_{i=1..n} \pi_i(\omega).$
- $\forall \omega, \pi^{max}(\omega) = \max_{i=1..n} \pi_i(\omega).$

Dans ce qui suit on s'intéresse en particulier à la fusion des distributions de possibilités utilisant l'opérateur *produit*.

La logique possibiliste permet une fusion des données au niveau syntaxique [BDP97]. En effet, lorsque l'information détenue est sous une forme syntaxique, la fusion consiste à combiner n bases de connaissances possibilistes $\{\mathcal{B}_i : i = 1, \dots, n\}$ de n sources en une nouvelle base de connaissances possibiliste en utilisant l'un des opérateurs de fusion (dans notre cas c'est l'opérateur *produit*). La fusion des connaissances possibilistes par l'opérateur *produit* consiste à calculer sa contrepartie syntaxique en partant des connaissances initiales comme le montre la proposition suivante :

Proposition 1.1 *Soient $\mathcal{B}_1 = \{(\phi_i, \alpha_i), i = 1, \dots, n\}$ et $\mathcal{B}_2 = \{(\psi_j, \beta_j), j = 1, \dots, m\}$ deux bases de connaissances possibilistes et soient π_1 et π_2 leurs distributions de possibilités associées. La base de connaissances possibilistes \mathcal{B}^* résultante de la fusion des bases \mathcal{B}_1 et \mathcal{B}_2 et dont la distribution de possibilités associée est $\pi^* = \pi_1 * \pi_2$ est donnée par :*

$$\mathcal{B}^* = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \{(\phi_i \vee \psi_j, \alpha_i + \beta_j - \alpha_i * \beta_j) : (\phi_i, \alpha_i) \in \mathcal{B}_1 \text{ et } (\psi_j, \beta_j) \in \mathcal{B}_2\}. \quad (11)$$

Exemple 1.14 *Soient les bases de connaissances possibilistes :*

- $\mathcal{B}_1 = \{(a \vee b, 0.6), (\neg a \vee b, 0.5)\},$
- $\mathcal{B}_2 = \{(b, 0.3), (\neg a \vee \neg b, 0.7)\}.$

Leurs distributions de possibilités associées π_1 et π_2 sont données dans le tableau 1.11. La fusion des deux bases \mathcal{B}_1 et \mathcal{B}_2 , par l'équivalent syntaxique de l'opérateur produit, donne lieu à une base résultante \mathcal{B}^ composée comme suit (après élimination de la formule subsumée $(b, 0.3)$) : $\mathcal{B}^* = \{(\neg a \vee \neg b, 0.7), (a \vee b, 0.72), (\neg a \vee b, 0.65)\}.$*

| A | B | $\pi_1(A, B)$ | A | B | $\pi_2(A, B)$ |
|----------|----------|---------------|----------|----------|---------------|
| a | b | 1 | a | b | 0.3 |
| a | $\neg b$ | 0.5 | a | $\neg b$ | 0.7 |
| $\neg a$ | b | 1 | $\neg a$ | b | 1 |
| $\neg a$ | $\neg b$ | 0.4 | $\neg a$ | $\neg b$ | 0.7 |

TAB. 17 – Distributions de possibilités π_1 et π_2 associées à \mathcal{B}_1 et \mathcal{B}_2

La distribution de possibilités $\pi^ = \pi_1 * \pi_2$ associée à la base combinée \mathcal{B}^* est donnée par le tableau 1.12. Notons qu'on peut retrouver cette même distribution de possibilités à partir de la formule 1.38.*

| A | B | $\pi^*(A, B)$ |
|----------|----------|---------------|
| a | b | 0.3 |
| a | $\neg b$ | 0.35 |
| $\neg a$ | b | 1 |
| $\neg a$ | $\neg b$ | 0.28 |

TAB. 18 – Distribution de possibilités résultante $\pi^* = \pi_1 * \pi_2$ associée à \mathcal{B}^*

Annexe B

Rappels sur la théorie des graphes

Nous commençons cette section par un bref rappel sur le vocabulaire et les notions de base utilisés dans la théorie des graphes [CGH96b, Bol02].

Soient $V = \{A_1, A_2, \dots, A_n\}$ un ensemble fini de variables et E un sous-ensemble du produit cartésien $V \times V$. Nous citons ci-dessous les définitions essentielles pour la compréhension de la suite de cette thèse :

- le couple (V, E) , noté par G définit un graphe où V constitue l'ensemble des sommets, appelé aussi nœuds, du graphe et E l'ensemble des liens entre les nœuds du graphe.
- Un sous-graphe d'un graphe G est un graphe $G' = (V', E')$ tel que $V' \subseteq V$ et $E' \subseteq E$.
- Un lien relie deux nœuds A_i et A_j . Un lien non orienté s'appelle arête. Un arc est un lien orienté de A_i vers A_j noté $A_i \rightarrow A_j$. A_i est appelé origine de l'arc et A_j l'extrémité.
- Un graphe constitué d'arcs est appelé graphe orienté. Un graphe constitué d'arêtes est appelé graphe non orienté.
- S'il existe un arc entre A_i et A_j , alors A_i est le parent de A_j et A_j est le fils de A_i .

L'ensemble des parents d'un nœud A_i est noté U_{A_i} ou plus simplement U_i .

L'ensemble des fils d'un nœud A_i est noté Y_{A_i} ou plus simplement Y_i .

- Une racine est un nœud qui n'a pas de parents.
- Une feuille est un nœud qui n'a pas d'enfants.
- Deux nœuds sont adjacents s'ils sont liés par un arc ou une arête.
- Un graphe est complet si deux nœuds quelconques sont reliés dans au moins une direction.
- Un sous-graphe complet, maximal et non inclu dans un autre sous-graphe est appelé clique.
- Un chemin est une suite de nœuds tel qu'il existe dans le graphe un arc entre chaque paire de nœuds successifs $(A_k, A_{k+1}) : k = i, i + 1, \dots, j - 1$.

Un chemin d'origine A_i et d'extrémité A_j est noté $A_i \mapsto A_j$. Si $A_j = A_i$ alors le chemin est appelé cycle.

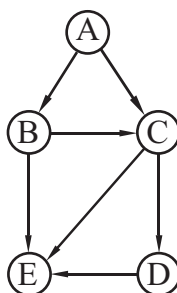
- Un graphe est connecté (où dit connexe) s'il existe un chemin possible entre chaque paire de nœuds.
- Une chaîne est un chemin dans lequel on ne tient pas compte de l'orientation des arcs. Une chaîne de A_i vers A_j est noté $A_i - A_j$. lorsque $A_j = A_i$ la chaîne est appelé boucle. Une boucle est un cycle non orienté.
- $Des(A_i)$ désigne l'ensemble des descendants de A_i ; cela signifie qu'il existe un chemin $A_i \mapsto A_k$ pour tout A_k appartenant à $Des(A_i)$.
- $Anc(A_i)$ désigne l'ensemble des ancêtres de A_i ; cela signifie qu'il existe un chemin $A_k \mapsto A_i$ pour tout A_k appartenant à $Anc(A_i)$.

Définition 2.1 Les graphes orientés ne possédant pas de cycles sont appelés Graphes Acycliques Orientés (en anglais *Directed Acyclic Graphs* ou *DAG*) [TS92].

Définition 2.2 Un DAG simplement connecté [KP83] est un graphe où entre deux nœuds il existe au plus un chemin. Un polyarbre [RJ87] est un DAG connecté sans boucles.

Définition 2.3 Un DAG à connexions multiples est un DAG où entre deux nœuds il existe plus d'un chemin. C'est un DAG avec boucles.

Exemple 2.1 Considérons le graphe orienté G de la figure 2.1 composé de cinq nœuds A , B , C , D et E .



Exemple d'un DAG.

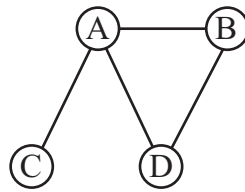
Les nœuds C , D et E forment un sous-graphe du graphe G . Par exemple, il est clair que d'après l'arc $A \rightarrow B$; A est le parent du nœud B qui est lui même le fils du nœud A .

Pour le nœud B , l'ensemble des parents est $U_B = \{A\}$ et l'ensemble des fils est $Y_B = \{C, E\}$. A est un nœud racine ($U_A = \emptyset$) et E est un nœud feuille ($Y_E = \emptyset$).

L'ensemble des ancêtres du nœud D est $Anc(D) = \{A, B, C\}$. L'ensemble des descendants du nœud A est $Des(A) = \{B, C, D, E\}$.

Il existe plusieurs chemins dans ce graphe : par exemple, les chemins $ABCDE$ et ACE . De plus, ce graphe est connexe ne contenant pas de cycle mais il contient plusieurs boucles (par exemple $BEDCB$) : il est donc un DAG à connexions multiples. Si au lieu de l'arc $C \rightarrow E$ on avait l'arc $E \rightarrow C$, dans ce cas on aurait un cycle $CDEC$ et par suite ce graphe ne serait pas un DAG.

La figure 2.2 montre un exemple d'un graphe non orienté. Dans ce graphe il existe deux cliques. Le premier formé par les nœuds A et C : ces deux nœuds sont adjacents et forment le plus grand sous-graphe complet. Le sous-graphe formé par les nœuds A et B n'est pas une clique car il est inclus dans la clique ABD (le plus grand sous-graphe complet contenant les nœuds A, B et D).



Exemple d'un graphe non orienté contenant deux cliques.